



Thèse

2016

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Spatial Data Infrastructures for addressing environmental challenges: stocktaking, capacity building, implementation and assessment

Guigoz, Yaniss

How to cite

GUIGOZ, Yaniss. Spatial Data Infrastructures for addressing environmental challenges: stocktaking, capacity building, implementation and assessment. Doctoral Thesis, 2016. doi: 10.13097/archive-ouverte/unige:88107

This publication URL: <https://archive-ouverte.unige.ch/unige:88107>

Publication DOI: [10.13097/archive-ouverte/unige:88107](https://doi.org/10.13097/archive-ouverte/unige:88107)

UNIVERSITE DE GENEVE
Département F.-A. Forel des sciences de
l'environnement et de l'eau

FACULTE DES SCIENCES
Professeur A. Lehmann

INSTITUT DES SCIENCES DE
L'ENVIRONNEMENT
Docteur G. Giuliani

Spatial Data Infrastructures for addressing environmental challenges: Stocktaking, Capacity Building, Implementation and Assessment

THESE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention sciences de
l'environnement

par

Yaniss GUIGOZ

de

Genève (GE)

Thèse n° 4943

GENEVE

Atelier d'impression ReproMail

2016

La Faculté des sciences, sur le préavis de Monsieur A. LEHMANN, professeur associé et directeur de thèse (Institut F.-A. Forel, Institut des sciences de l'environnement), Monsieur G. GIULIANI, docteur et codirecteur de thèse (Institut des sciences de l'environnement), Monsieur H. DAO, professeur titulaire (Département de géographie et environnement, Faculté des sciences de la société), Monsieur D. BÉROD, docteur (Hydrology Division, World Meteorological Organization, Genève, Suisse) et Madame M. A. BROVELLI, professeure (Como Campus, École polytechnique de Milan, Italie), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 24 juin 2016

Thèse - 4943 -



Le Doyen

TABLE OF CONTENTS

TABLE OF CONTENTS	3
ABSTRACT.....	7
RESUME.....	10
ACKNOWLEDGEMENTS.....	13
REMERCIEMENTS.....	15
1 INTRODUCTION	17
1.1 Background and fundamental concepts	17
1.1.1 Data	20
1.1.2 SDI.....	24
1.1.3 Capacity Building.....	29
1.2 Research questions	34
1.3 Thesis structure, contributing papers and projects	37
1.3.1 Structure of the thesis	37
1.3.2 List of contributing projects	38
1.3.3 List of contributing research papers.....	39
2 SDI STOCKTAKING.....	41
2.1 Introduction on SDI stocktaking.....	41
2.2 Spatial Data Infrastructures in Africa: a gap analysis	45
2.2.1 Abstract	45
Keywords	45
2.2.2 Introduction.....	46
2.2.3 Methodology.....	49
2.2.4 Results.....	51
2.2.5 Discussion and Perspectives	52
2.2.6 Conclusions.....	58
2.2.7 Acknowledgments	60
2.2.8 Supplementary notes.....	60

Access mechanism	68
2.2.9 Supplementary Tables	74
2.2.10 Supplementary Figures.....	79
2.3 Chapter key outcomes	88
3 CAPACITY BUILDING	89
3.1 Introduction to capacity building	89
3.2 Bringing GEOSS services into practice: a capacity building resource on Spatial Data Infrastructures (SDI).....	92
3.2.1 Abstract	92
3.2.2 Introduction.....	92
3.2.3 Objectives:	95
3.2.4 Workshop material.....	96
3.2.5 Dissemination	101
3.2.6 Lessons learned and recommendations:	102
3.2.7 Conclusions and future perspectives:	105
3.2.8 Supplementary material	107
3.3 Leading the way toward an environmental National Spatial Data Infrastructure in Armenia.....	111
3.3.1 Abstract	111
3.3.2 Introduction.....	112
3.3.3 Environmental status of Armenia and data sharing activities	112
3.3.4 Potential and benefits of an Armenian nSDI	114
3.3.5 Applying the EGIDA methodology to Armenia.....	116
3.3.6 Conclusion	119
3.4 Chapter key outcomes	120
4 IMPLEMENTATION	122
4.1 Introduction on implementation.....	122
4.2 SCOPED-W: SCAlable Online Platform for extracting Environmental Data and Water-related model outputs (co-lead, submitted)	126
4.2.1 Abstract	126

4.2.2	Introduction.....	126
4.2.3	SWAT input and output.....	130
4.2.4	SCOPED-W: architecture and implementation.....	131
4.2.5	Discussion.....	137
4.2.6	Conclusions.....	141
4.3	Enabling Discovery of African Geospatial Resources (co-author, published) 142	
4.3.1	Abstract:	143
4.3.2	Introduction:	143
4.3.3	Methodology, implementation and preliminary results	145
4.3.4	Discussion:	152
4.3.5	Conclusions and perspectives:	154
4.4	Facilitating the production of ISO-compliant metadata of geospatial datasets (co-author, published).....	156
4.4.1	Highlights.....	156
4.4.2	Abstract	156
4.4.3	Introduction.....	157
4.4.4	Methodology.....	158
4.4.5	Implementation	159
4.4.6	Discussion.....	162
4.4.7	Conclusions & perspectives:.....	163
4.5	Chapter key outcomes	164
5	ASSESSMENT.....	166
5.1	Introduction to SDI assessment	166
5.2	EGAL	168
5.2.1	Abstract	169
5.2.2	Introduction.....	169
5.2.3	Methodology.....	172
5.2.4	Results.....	185
5.2.5	Discussion.....	186

5.2.6	Conclusions	191
5.2.7	Annexes	192
5.3	The EOPOWER impact assessment framework	199
5.3.1	Impact assessment of the “Bringing GEOSS Services into Practice” workshop	201
5.3.2	Earth Observation promotion activities and their impact in Armenia 212	
5.4	Chapter key outcomes	219
6	CONCLUSIONS	221
6.1	Answers to research questions	221
6.2	Contributions, innovations, and their relevance to scientific and practical advancements	234
6.3	Recommendations and perspectives	236
6.4	Final conclusion	239
	ANNEXES	242
	List of figures	242
	List of acronyms	243
	REFERENCES	246

ABSTRACT

Environmental challenges resulting in particular from climate change are becoming increasingly worrying. In order to study and address these issues, political decisions are needed that must be supported by solid scientific evidences. Data forms the building block on top of which a whole process leading to informed political decisions could be built. Despite the evident importance of data, many technical, institutional, political, legal or social barriers still hinder an optimal data use in the decision-making process in many parts of the World. In order to address these barriers, a general framework called “Spatial Data Infrastructure” (SDI) has been conceptualized and developed during the last twenty years. SDI aims at integrating data, technological, institutional and individual components in a concerted approach for improving data availability and its discovery, access, use, creation, collection, exchange and sharing. This should eventually benefit the various stakeholders who need quality data in specific geographic areas.

Notwithstanding the promising benefits of SDIs and their acceptance in numerous places as an essential infrastructure in a modern society, their real impact and societal penetration do not seem to meet the expectations according to certain authors. Given the importance of SDIs for improving data accessibility and hence for addressing environmental challenges, it is necessary to understand in a realistic manner the status of their implementation, their societal or environmental impact, and the impact of some SDI specific activities on the whole SDI. This should enable to properly target the efforts, justify expenses and fundraising. But the dynamic and multi-dimensional nature of SDIs makes these tasks complicated. In addition, the development of an SDI and the visibility of its impact on a society are a long-term process that requires long-term vision, commitment and sustainable funding. Moreover, the resources needed to establish, maintain or monitor an SDI are proportional to its geographic scale, which means that certain actions that can be performed at a local level SDI in much detail cannot be done with the same level of detail for a global SDI.

It is therefore difficult to have a single solution or procedure that would fit any SDI without taking into account its socio-political context. This is due to the central role of human aspects in SDI and reflected in the institutional and people SDI components. It is therefore primordial to include the human component in any SDI activity, for example by systematically organizing capacity building activities. The underlying hypothesis of this thesis consists in saying that capacity building is the key element to address both tangible and intangible barriers to the success of SDI and that it should be the primary element of any SDI strategy. The main aim of this research is therefore to examine the role of capacity building for addressing the various barriers to SDI implementation.

For this purpose, the various elements constituting the stages in the SDI cycle are examined: (1) the stocktaking evaluation needed before any action can be planned; (2) the possible capacity building actions that can be taken; (3) the SDI implementation actions possible; (4) the assessment of the SDI components and of the actions taken.

We first examine the existing SDI assessment methodologies, which have been developed with different approaches depending on the purpose of the assessment. Some of these approaches might focus on a specific geographic scale (e.g. well suited for performing a national level assessment), others on a specific theme (e.g. cadastre). In any case these approaches use specific indicators, some of which might be similar depending on the approach. The diversity of these approaches reflects the complex nature of SDIs. Recognizing the multiple SDI realities, this study discusses the “multi-view assessment framework” that has been developed by some authors as an attempt to group these approaches through a general framework. Given the limitations it presents to the need of assessing the SDI implementation status in Africa that we have described in this thesis, we propose a specific methodology for executing a rapid assessment of SDI at a continental level with limited resources. This assessment not only shows the general score of Africa in terms of SDI implementation but also highlights the worldwide scarcity of SDI monitoring data. This opens the door for discussions on opportunities to improve this situation, which should clearly be addressed through the institutional component, at several politico-administrative levels and shows the crucial human role in this process.

We then demonstrate the benefits of capacity building for SDI through two different and complementary approaches: (1) the role played by some specific capacity building material for individual empowerment and (2) the general underlying role of capacity building in a specific national use case in Armenia. The key aspect of capacity building in SDI makes it necessary to carefully plan it and anticipate several success enablers. Among these success enablers, there is an overwhelming need to tailor capacity building activities, to the participants. The awareness raising part of capacity building, is of particular importance to demonstrate the challenges to decision-makers.

Three technological solutions for SDI improved implementation are then presented and discussed. They all propose a reduction of the technological complexity for the data providers and users while transferring it to another level managed by IT specialists. The solutions proposed largely rely on the possibilities offered by the concept of data and metadata interoperability that is implemented through standards. Additionally, the “open” emerging trend for software, data and standards is discussed and presented as a necessary enabler of the interoperability concept that would otherwise be more difficult to put in place. We believe that the different solutions proposed all contribute to a better SDI implementation for an increased data use and production. These solutions primarily address the infrastructure component of SDI but also have a strong influence on the “people” SDI component.

The last chapter of the thesis discusses the importance of assessment at various levels of SDI. We distinguish two main types: (1) assessment of the SDI. It can be an assessment of the whole SDI status (stocktaking) as discussed in the second chapter, or assessment of the status of some specific SDI components (e.g. SDI geoportals); (2) assessment of the impact of SDI. It can be the global impact of SDI on a society where it is implemented, but it can also be the impact on the whole SDI of some specific actions taken for implementing or improving the SDI. This is for example the case when capacity building or technological actions

taken need to be assessed. Two assessment methodologies are proposed in this chapter, one for each main type of assessment: a first methodology is proposed for measuring the quality of geoportals and a second one for measuring the impact of Earth Observation solutions, using the capacity building material and activities described in the chapter on capacity building as a use case.

We conclude by highlighting some keywords that recur regularly in this thesis, such as: leadership, vision, network, commitment, coordination, simplicity, complexity reduction, interoperability, and customization. We argue that capacity building is the solution to address these various elements that are critical to successful SDI. We also state that awareness raising and capacity building at the level of individuals are key steps because they have a decisive influence on the other capacity building aspects, individuals being common denominators to all SDI components. We add that a balance needs to be found between the necessary technical uniformity and the cultural plurality inherent to mankind, on the model of a system of systems. Finally, this study describes the cyclic aspect of SDI process that is composed of: stocktaking – capacity building – implementation and assessment. Any SDI process should regularly pass through these stages for a durable improvement. We also describe similar sub-cycles for capacity building and SDI implementation that also need to continuously improve their performance for the benefit of the whole system.

RESUME

Les défis environnementaux, résultant en particulier des changements climatiques, se font de plus en plus menaçants. Afin d'étudier et de prendre des mesures pour résoudre ces problèmes, des décisions politiques sont nécessaires, qui doivent pouvoir s'appuyer sur de solides preuves scientifiques. Les données constituent la base sur laquelle tout un processus menant à des décisions politiques avec preuves à l'appui peut se construire. Malgré l'importance manifeste des données, de nombreuses barrières techniques, institutionnelles, politiques, légales ou sociales empêchent encore leur utilisation optimale dans un processus de prise de décision dans de nombreuses parties du Monde. Afin de réduire ces obstacles, un cadre général dénommé "Spatial Data Infrastructures" (SDI) ou "Infrastructures de Données Géospatiales" en français, a été conceptualisé et développé durant les vingt dernières années. Les SDIs visent à intégrer les données avec les composantes technologiques, institutionnelles et individuelles dans une approche concertée visant à améliorer la disponibilité des données et leur découverte, accès, utilisation, création, collecte, échange et partage. Cette approche devrait au final bénéficier aux diverses parties prenantes qui ont besoin de données de qualité pour des zones géographiques déterminées.

Malgré les bénéfices potentiels prometteurs des SDIs ainsi que leur large acceptation en tant qu'infrastructure essentielle dans une société moderne, leur impact réel et leur pénétration sociétale ne semblent pas être à la hauteur des attentes selon certains auteurs. Etant donné l'importance des SDIs pour améliorer l'accès aux données et donc pour répondre aux défis environnementaux, il est nécessaire de comprendre de manière réaliste le statut de leur implémentation, leur impact sociétal ou environnemental, et l'impact de certaines de leurs activités spécifiques sur l'ensemble. Ceci devait permettre de cibler les efforts de manière appropriée, justifier les dépenses et la récolte de fonds nécessaires. Mais la nature dynamique et multidimensionnelle des SDIs rend ces tâches compliquées. De plus, le développement d'un SDI et la visibilité de son impact sur la société sont un processus de longue haleine qui demande une vision, un engagement et un financement de long terme. Il faut également tenir compte du fait que les ressources nécessaires à l'établissement, au maintien ou au suivi d'un SDI sont proportionnelles à son échelle géographique, ce qui signifie que certaines actions qui peuvent être effectuées de manière détaillée dans un SDI local ne peuvent pas forcément l'être avec le même niveau de détail pour un SDI global.

Par conséquent, il est difficile d'avoir une solution ou une procédure unique qui conviendrait à n'importe quel SDI, indépendamment de son contexte politico-social. Ceci est dû au côté "humain" prépondérant dans un SDI et reflété dans ses composants institutionnels et individuels. Il est dès lors primordial d'inclure la composante humaine dans toute activité liée aux SDI, par exemple en organisant systématiquement des activités de renforcement des capacités. L'hypothèse sous-jacente de ce travail de recherche consiste à dire que le renforcement des capacités est l'élément clé permettant de répondre à la fois aux barrières tangibles et intangibles qui empêchent le succès d'un SDI, et que ce renforcement

des capacités devrait être l'élément central de toute stratégie liée aux SDIs. Le but principal de cette recherche est donc d'examiner le rôle du renforcement des capacités pour répondre aux différentes barrières à l'implémentation des SDIs.

Pour cela, les éléments constituant les différents stades dans le cycle des SDIs sont examinés: (1) l'évaluation de la situation (stocktaking) qui est un prérequis à la planification de toute action; (2) les actions possibles de renforcement des capacités qui peuvent être prises; (3) les actions possibles d'implémentation des SDIs; (4) l'évaluation des composants SDI et des actions prises.

Nous examinons tout d'abord les méthodologies existantes d'évaluation des SDIs, qui ont été développées dans des approches différentes en fonction de la raison de l'évaluation. Certaines de ces approches peuvent privilégier une échelle géographique particulière (par exemple elles peuvent être particulièrement bien adaptées à une évaluation au niveau national), d'autres peuvent être plus spécifiques à un thème (ex: le cadastre). Dans tous les cas ces approches utilisent des indicateurs spécifiques, dont certains peuvent être identiques suivant l'approche. La diversité de ces approches reflète la nature complexe des SDIs. Reconnaissant la réalité multiple des SDIs, ce travail de recherche discute le cadre d'évaluation multi-perspectives (multi-view assessment framework) qui a été développé par certains auteurs comme une essai de regroupement de ces approches au travers d'une méthodologie générale. Etant donné les limitations que ce cadre présente au besoin d'évaluer le statut d'implémentation des SDIs en Afrique que nous décrivons dans cette thèse, nous proposons une méthodologie spécifique pour effectuer une évaluation rapide de SDIs à n niveau continental avec des ressources limitées. Cette évaluation montre non seulement le score général de l'Afrique en termes d'implémentation des SDIs, mais souligne également la rareté des données de suivi SDI au niveau global. Ceci ouvre la porte à des discussions sur les opportunités d'amélioration de la situation, qui devrait être résolue à travers la composante institutionnelle à plusieurs niveaux politico-administratifs, et démontre une fois encore l'importance du côté humain dans ce processus.

Nous démontrons ensuite les bénéfices du renforcement des capacités pour le SDI à travers deux approches différentes et complémentaires: (1) le rôle joué par du matériel spécifique de renforcement des capacités pour un renforcement des individus et (2) le rôle sous-jacent du renforcement des capacités dans le cas d'étude spécifique de l'Arménie. Le rôle clé du renforcement des capacités dans les SDI nécessite une planification attentive et une anticipation de plusieurs activateurs de succès. Parmi ces activateurs, il y a la nécessité d'adapter les activités de renforcement des capacités aux participants. La partie de sensibilisation revêt une importance toute particulière pour faire prendre conscience des défis aux décideurs.

Trois solutions technologiques pour une implémentation améliorée des SDI sont ensuite présentées et discutées. Elles proposent toutes une réduction de la complexité technologique pour les fournisseurs et utilisateurs de données en la transférant à un autre niveau géré par les spécialistes en information et technologie. Les solutions proposées reposent en grande partie sur les possibilités offertes par le concept d'interopérabilité des données et

métadonnées, qui est implémenté à travers les standards. De plus, l'émergence de la tendance "ouverte" pour les logiciels, les données et les standards est discutée et présentée comme un activateur nécessaire au concept d'interopérabilité, qui serait autrement plus difficile à mettre en place. Nous croyons que les différentes solutions proposées contribuent à une meilleure implémentation des SDIs pour une utilisation et production accrue de données. Ces solutions répondent en premier lieu à la composante "infrastructure" des SDIs mais ont aussi une forte influence sur la composante des "individus".

Le dernier chapitre de cette thèse discute l'importance de pouvoir faire des évaluations à plusieurs niveaux des SDIs. Nous en distinguons deux catégories principales: (1) l'évaluation d'un SDI. Cette évaluation peut consister en une évaluation du statut global d'un SDI (stocktaking) comme discuté dans le second chapitre, ou bien en une évaluation du statut de certaines composantes spécifiques d'un SDI (par exemple les géoportails); (2) l'évaluation de l'impact d'un SDI. Ceci peut consister en une évaluation de l'impact global que peut avoir un SDI sur une société où il est implémenté, mais il peut s'agir également de l'impact de certaines actions spécifiques prises pour implémenter ou améliorer un SDI. C'est par exemple le cas lorsque des actions de renforcement des capacités ou d'amélioration technologique doivent être évaluées. Deux méthodes d'évaluation sont proposées dans ce chapitre, une pour chaque type principal d'évaluation: une première méthodologie est proposée pour mesurer la qualité des géoportails et une seconde pour mesurer l'impact de solutions liées à l'Observation de la Terre, en utilisant comme cas d'étude le matériel de renforcement des capacités ainsi que les activités décrites dans le chapitre sur le renforcement des capacités.

Nous concluons en soulignant certains mots clé qui reviennent régulièrement dans cette thèse, tels que: leadership, vision, réseau, engagement, coordination, simplicité, réduction de la complexité, interopérabilité et personnalisation. Nous affirmons que le renforcement des capacités est la réponse à ces différents éléments qui sont essentiels au succès d'un SDI. Nous déclarons aussi que la sensibilisation et le renforcement des capacités au niveau des individus sont des étapes clé car elles ont une influence décisive sur les autres aspects du renforcement des capacités, de par le fait que les individus sont le dénominateur commun à toutes les composantes SDI. Nous ajoutons qu'un équilibre doit être trouvé entre l'uniformité technique nécessaire et la pluralité culturelle inhérente à l'humanité, sur le modèle d'une approche privilégiant les systèmes de systèmes. Finalement, cette étude décrit l'aspect cyclique des SDIs, qui est composé de: état des lieux (stocktaking) – renforcement des capacités – implémentation et évaluation. Chaque processus d'implémentation de SDI doit passer à intervalle régulier à travers ces stades pour une amélioration permanente. Nous décrivons également des sous-cycles similaires pour le renforcement des capacités et l'implémentation des SDIs, qui doivent en permanence améliorer leur performance pour le bénéfice du système entier.

ACKNOWLEDGEMENTS

For many years I have been lucky to be part of several projects involving SDI and capacity building activities of the highest scientific and technical interest. This gave me the opportunity to not only deepen my knowledge and develop new skills in the domain, but also to travel to several parts of the World and meet many people, some of which have become friends. My participation in these projects is the result of different factors conjunction that probably not happen many times in life and I wish to express my gratitude to all those who have been part of this conjunction.

I would like to warmly thank first of all Prof. Anthony Lehmann, director of the enviroSPACE laboratory and director of this thesis, for accepting to dedicate the necessary time and energy to supervise my work and for giving me the chance to be part of a team. His positive mindset, encouraging approach, humanity, friendship, ability to gather the pieces of the puzzle for consistency of the big picture have been precious contributions to my work for several years and a permanent motivation towards self improvement.

I would like to equally thank Dr. Grégory Giuliani, co-director of this thesis, for his invaluable role in the conjunction of factors and contribution to the present work. His leading role in SDI development and implementation, his motivation and team spirit gave me the opportunity to broaden my knowledge and develop new skills. This was even more pleasant with friendship, which allowed sharing wonderful times especially during the numerous work related trips in Africa or Europe, that started with the AFROMAISON project back in 2011 and continued with many more projects since then.

I would like to also deeply thank Dr. Nicolas Ray for his precious role in many aspects of this work, his friendship and wise advices. He gave me the opportunity to be part of several projects essential for this work, such as ARPEGEO, ECOARM2ERA, EOPOWER and IASON. His presence, encouragements, trust and belief in human qualities contributed to gain self-confidence and share the best of myself.

I also cordially thank Dr. Pierre Lacroix for all the collaborative work achieved together in the different projects. His hard and thorough work, commitment, altruism and perseverance have been major assets to overcome difficulties encountered in a team spirit embellished by humor and friendship.

The scientific, professional and human qualities of the SDI team collaborators just mentioned reflect the same spirit present among the larger teams of colleagues at UNEP/GRID-Geneva and ISE that I am lucky to work or have worked with for several years. I would like to also thank them for their daily collaborative spirit and their flexibility that benefits us all: Karin Allenbach, Andrea de Bono, Jean-Philippe Richard, Bruno Chatenoux, Cédric Gampert, Pascal Peduzzi, Hy Dao, Stefan Schwarzer, Stéphane Kluser, Diana Rizzolio, Ron Witt, Jean-Michel Jaquet, Denisa Rodila, Martin Lacayo, Martin Schlaepfer and Marc Fasel.

Obtaining a project supposes a lot of preparatory teamwork for an uncertain result. I would like to warmly thank all my colleagues who have played a role in obtaining the various projects contributing to this thesis. I would also like to acknowledge the European Commission “seventh framework programme” that funded the enviroGRIDS (Grant Agreement n°226740), AFROMAISON (Grant Agreement n°266379), EOPOWER (Grant Agreement n° 603500), IASON (Grant Agreement n°603534) and EcoArm2ERA (Grant Agreement n°294924) projects, as well as the Swiss National Science Foundation that funded the ARPEGEO project through the “Scientific co-operation between Eastern Europe and Switzerland” (SCOPES) programme 2009-2012.

Finally, on a more personal level, I would like to thank my friends and family for their support, encouragement and patience during these past years. I would like to thank in particular: my wife Amita for sharing with me the necessary sacrifices by taking on her shoulders the daily domestic and educational work for a long period; my children Anisha, Nitya and Nikhil for accepting to have less of my presence temporarily; my mother Michelle and my late father Emile who have spared no sacrifice for my education; other members of my family who have supported me morally and materially, and taught me endurance through mountaineering.

I dedicate this thesis to my children Anisha, Nitya and Nikhil, who represent the hope we must place in the young generation for a better Earth, and to whom we shall devise our knowledge through capacity building.

REMERCIEMENTS

Depuis quelques années j'ai eu la chance de participer à plusieurs projets d'intérêt de par leur niveau scientifique et technique, incluant des activités d'Infrastructures de Données Spatiales ainsi que de renforcement des capacités. Ceci m'a donné la possibilité non seulement d'approfondir mes connaissances et de développer de nouvelles compétences dans le domaine, mais également de voyager dans différents endroits du monde et d'y côtoyer de nombreuses personnes, dont certaines sont devenues des amis. Ma participation à ces projets est le résultat de la conjonction de plusieurs facteurs qui ne se produisent probablement pas plusieurs fois dans une vie, et je souhaite exprimer ma gratitude à toutes les personnes qui ont participé à cette conjonction.

Je voudrais tout d'abord remercier chaleureusement le Professeur Anthony Lehmann, directeur du laboratoire enviroSPACE et directeur de cette thèse, pour avoir accepté de dédier le temps et l'énergie nécessaires à la supervision de mon travail et pour m'avoir donné la chance de faire partie d'une équipe. Sa positivité, ses encouragements, son humanité, son amitié, sa faculté de réunir les pièces du puzzle pour la cohérence de l'ensemble ont été des précieuses contributions à mon travail depuis plusieurs années et une motivation permanente à une amélioration personnelle.

J'aimerais remercier de la même manière le Dr. Grégory Giuliani, co-directeur de cette thèse, pour son rôle inestimable dans cette conjonction de facteurs et pour sa contribution au travail présent. Son rôle clé dans le développement et l'implémentation des SDI, sa motivation et son esprit d'équipe m'ont donné l'opportunité d'élargir mes connaissances et de développer de nouvelles compétences. Ces aspects professionnels agrémentés d'amitié ont permis de partager de merveilleux moments, en particulier durant les nombreux voyages professionnels en Afrique et en Europe, qui ont débuté avec le projet AFROMAISON en 2011 et ont continué avec de nombreux autres projets depuis lors.

Je voudrais également remercier sincèrement le Dr. Nicolas Ray pour son précieux rôle dans de nombreux aspects de ce travail, pour son amitié et ses conseils avisés. Il m'a donné l'opportunité de prendre part à de nombreux projets essentiels pour ce travail tels que ARPEGEO, ECOARM2ERA, EOPOWER et IASON. Sa présence, ses encouragements, sa confiance et sa foi dans les qualités humaines ont contribué à forger une plus grande confiance personnelle et à partager le meilleur de moi-même.

Je voudrais aussi remercier cordialement le Dr. Pierre Lacroix pour tout le travail collaboratif effectué ensemble à travers différents projets. Son ardeur à la tâche, sa rigueur, son dévouement, son altruisme et sa persévérance ont été des atouts majeurs pour surmonter les difficultés rencontrées dans un esprit d'équipe, agrémenté d'humour et d'amitié.

Les qualités scientifiques, professionnelles et humaines de l'équipe des collaborateurs SDI mentionnés reflètent le même état d'esprit présent parmi les équipes plus larges de collègues au PNUE/GRID-Genève et à l'ISE, avec lesquels j'ai la chance de travailler ou d'avoir travaillé depuis plusieurs années. Je

voudrais également les remercier pour leur esprit collaboratif et leur flexibilité qui nous bénéficient à tous: Karin Allenbach, Andrea de Bono, Jean-Philippe Richard, Bruno Chatenoux, Cédric Gampert, Pascal Peduzzi, Hy Dao, Stefan Schwarzer, Stéphane Kluser, Diana Rizzolio, Ron Witt, Jean-Michel Jaquet, Denisa Rodila, Martin Lacayo, Martin Schlaepfer et Marc Fasel.

Obtenir un projet requiert un travail d'équipe préparatoire considérable pour un résultat incertain. J'aimerais remercier chaleureusement tous mes collègues qui ont joué un rôle dans l'obtention des nombreux projets qui ont contribué à cette thèse. J'aimerais également remercier le septième programme cadre de la Commission Européenne qui a financé les projets enviroGRIDS (Grant Agreement n°226740), AFROMAISON (Grant Agreement n°266379), EOPOWER (Grant Agreement n° 603500), IASON (Grant Agreement n°603534) and EcoArm2ERA (Grant Agreement n°294924), ainsi que le Fonds National Suisse qui a financé le projet ARPEGEO à travers le programme SCOPES (coopération scientifique entre l'Europe de l'Est et la Suisse) 2009-2012.

Finalement, à un niveau plus personnel, j'aimerais remercier profondément mes amis et ma famille pour leur support, leurs encouragements et leur patience durant ces dernières années. J'aimerais remercier en particulier: ma femme Amita pour partager avec moi les sacrifices nécessaires en prenant sur ses épaules le travail domestique et éducationnel familial durant une si longue période; mes enfants Anisha, Nitya et Nikhil pour accepter ma présence réduite de manière temporaire; ma mère Michelle et feu mon père Emile qui n'ont reculé devant aucun sacrifice pour mon éducation; d'autres membres de ma famille qui m'ont soutenu moralement et matériellement, et qui m'ont appris l'endurance à travers l'alpinisme.

Je dédie cette thèse à mes enfants Anisha, Nitya et Nikhil, qui représentent l'espoir que nous devons placer dans la nouvelle génération pour une meilleure planète, et à qui nous pouvons léguer notre connaissance à travers un renforcement des capacités.

1 INTRODUCTION

1.1 Background and fundamental concepts

The last decades witnessed an unprecedented record number of extreme environmental events mostly attributable to climate change (IPCC, 2014a). The rhythm of global warming is even accelerating since most of the warmest years occurred in the past ten years¹ with dramatic consequences for example on ice sheets². This led to a situation where natural disasters generate huge economic losses estimated between US\$2.5 and US\$24.2 trillion (Dietz et al., 2016), forcing the insurance-industry to sourcing and applying climate data in their decision-making process (Bell-Pasht and Krechowicz, 2015). It is now admitted that more than half of the observed increase in global average surface temperature is of anthropogenic nature (IPCC, 2014a). This situation needs to be addressed but this is a long term and complex fight that challenges all actors of the society worldwide, from the general public to politicians, through industry, science and politics, both in public and private sectors. The challenges consist not only in finding technological solutions to mitigate greenhouse gases emissions, but also in political agreements at various scales, from local to global, between the concerned players. Only political actors can issue binding regulations that must be applied at relevant scale. Nowadays, most of the crucial environmental challenges (e.g. global warming, deforestation) faced by humanity need to be addressed internationally given the global aspect of these issues. As an illustration, the notion of “planetary boundaries” has been developed (Rockström et al., 2009) that defines nine planetary boundaries within which humanity can operate safely. They concern: climate change, ocean acidification, stratospheric ozone depletion, atmospheric aerosol loading, biogeochemical flows, global freshwater use, land-system change, rate of biodiversity loss and chemical pollution. Three of these planetary boundaries (climate change, rate of biodiversity loss, global nitrogen cycle) have already been transgressed, which shows the emergency to act internationally.

According to Ryabinin (2015), our political systems are suitable to address national problems but much less for global issues. The UN and international organizations fill this important gap in governance and this is why such international negotiations take place under the auspices of the United Nations. This started in 1972 with the Stockholm UN Conference on Human Environment³, followed by the major Rio Earth summit in 1992 and consecutive decennial Earth Summits (Johannesburg in 2002, Rio again in 2012). Additionally, yearly UN conferences known as Conferences of the Parties (CoP) have been taking place since 1995 to assess progresses for addressing climate change⁴. Despite immediate results not always meeting the expectations, these global conferences represent a unique opportunity of engaging and maintaining

¹ http://geodata.grid.unep.ch/extras/graph_hottest_years.php

² http://geodata.grid.unep.ch/extras/arctic_sea_ice_graph_min-max.php

³

https://en.wikipedia.org/wiki/United_Nations_Conference_on_the_Human_Environment

⁴ https://en.wikipedia.org/wiki/United_Nations_Climate_Change_conference

dialogue between governments and other environmental players and finding agreements, which would probably not take place at the global scale otherwise. This is a long-term process, whose results range from simple declarations to binding treaties depending on the years, with sometimes concrete noteworthy action agreements such as the United Nations Framework Convention on Climate Change (UNFCCC), the Agenda21, the Kyoto protocol or the latest Paris COP21 agreement⁵.

Political decisions sustaining such treaties are always difficult to obtain, as they are at the crossroads of various economic and political interests, often divergent. In any case, solid scientific evidences are needed to sustain environmental negotiations at the political level and environmental models play a crucial role as they have an impact on decisions of public authorities (Maué et al., 2011). This shows that the translation of results from monitoring and research into policies is of vital importance (Ryabinin, 2015). Unfortunately, much of the scientific outputs (e.g. models, measurements) remain beyond the understanding of end-users and can hence not be integrated into policies (Bell-Pasht et al., 2015). Nevertheless scientific and political worlds must be bridged as « it is the only way to drive global sustainable development that delivers social inclusion, environmental sustainability and economic prosperity »⁶. Consequently, it is also the role of scientists to convey simple and clear messages to politicians, which is precisely the target of synthesis reports or summaries for policymakers (e.g. the IPCC climate change 2014 synthesis report (IPCC, 2014a) and several summaries for policymakers (IPCC, 2013, 2014b, c) ; the GEOSS water strategy executive summary (GEO, 2013b) ; the UNEP Global Environment Outlook (GEO) summary for policymakers (UNEP, 2012) ; the Millenium Ecosystem Assessment various synthesis reports⁷ ; the UN Global Assessment Report on Disaster Risk Reduction (GAR) pocket version (UNISDR, 2015b) or factsheet⁸).

These simplified messages are outcomes of complex analyses based on models requiring quantities of data, including their spatial location as it is estimated that over 80% of governmental data has a locational basis (Forbes, 2016; Rajabifard, 2002).

Data can be considered here as a building block on top of which a whole process leading eventually to decisions is built, as schematized in the widely recognized Decision-Information-Knowledge-Wisdom (DIKW) pyramid model (Ackoff, 2010). An alternative version of this pyramid, containing an additional “Decisions” level is proposed in Figure 1. In this pyramid, (1) data are products of observation but need to be useable for being useful; (2) to this end, data is given a context (description allowing to answer questions like who, what, where, or when) that (3) makes it useable, interpretable by experts with the necessary know-how; at this stage, raw data has become useful knowledge thanks to its

⁵ https://en.wikipedia.org/wiki/Paris_Agreement

⁶ <http://www.nature.com/news/interdisciplinarity-how-to-catalyse-collaboration-1.18343#/ref-link-1>

⁷ <http://www.millenniumassessment.org/en/Synthesis.html>

⁸ http://www.preventionweb.net/english/hyogo/gar/2015/en/gar-pdf/GAR15_at_a_glance_EN.pdf

description and interpretation, and allows to answer the questions “how” and “why”. This knowledge (4) can in turn be given an additional value by increasing its effectiveness, integrating it with other knowledge to form a superior level of knowledge called wisdom. Wisdom can then transform into (5) decision if action is effectively taken based on wisdom.

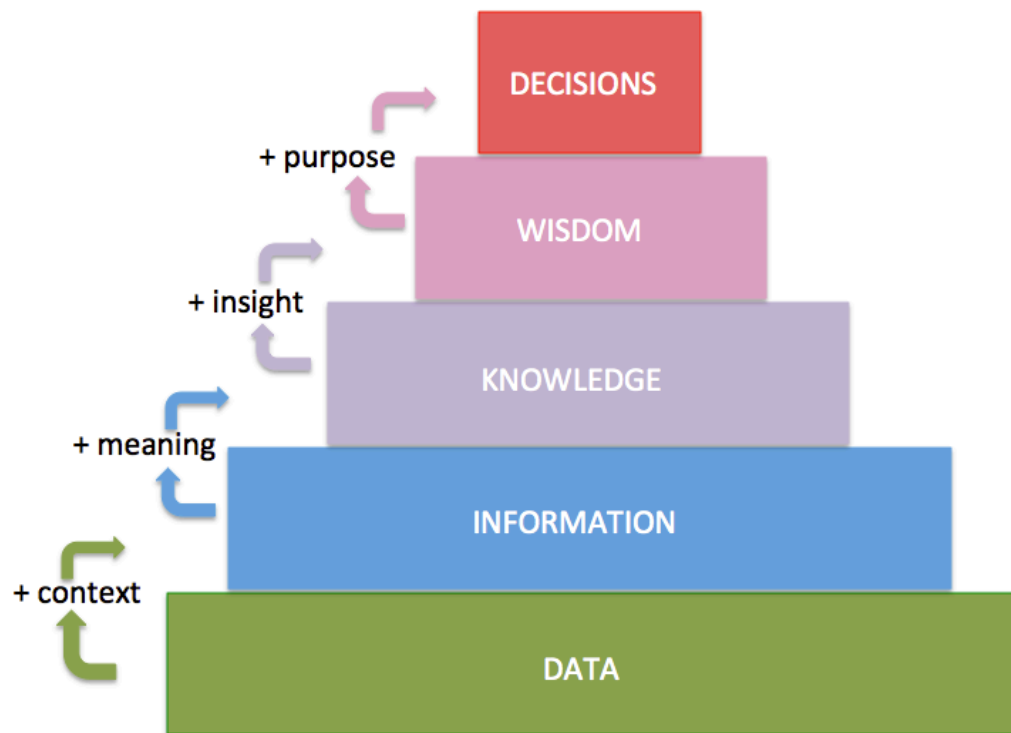


Figure 1: the Decision-Information-Knowledge-Wisdom (DIKW) pyramid (based on Ackoff, 2010)

Despite large acceptance of this data to wisdom and decisions vision, critics have been formulated (Weinberger, 2010) and discussed (Rowley, 2007). This debate is out of the scope in this thesis and we consider here this hierarchical and sequential vision of the data being the building block of a pyramid leading to informed decisions as experienced in the daily work of environmental sciences. This is illustrated by the 80% estimate of all business decisions that involve geographic data (Erskine et al., 2013).

In addition to the importance of this fundamental data block, Giuliani (2011) showed that many other blocks such as interoperability, standards, communication, education, etc. are also required to reach the perfect pyramidal vision of improved data access and geoprocessing for informed decisions. These different blocks will be discussed in this thesis but a short section about data at this stage is necessary hereafter given its fundamental role in the whole decisional process that affects in the end the quality of environmental models on which decisions are based. We are conscious of the fact that despite all evidences that scientists can bring, political decisions might be biased by superior political

or economic motivations. But this should not stop the willingness to address environmental challenges through evidences and solutions as robust as possible.

1.1.1 Data

1.1.1.1 Data concepts

Several terms linked to data such as spatial data, geographic data, geospatial data or core data are often used in an interchangeable manner so we would like first of all to clarify some of these often used terms and the meaning we give them in our purpose, starting from their common ground: data.

Data can be defined as “any collection of related facts arranged in a particular format; often, the basic elements of information that are produced, stored, or processed by a computer” (ESRI, 2016) or “symbols that represent properties of objects, events and their environment. They are the products of observation. But are of no use until they are in a useable (i.e. relevant) form” (Rowley, 2007). This definition matches with the DIKW pyramid, also considering data as the building block for the next steps. From these definitions, we can retain that data are properties of observed objects or events, which need to be processed for being useful. “Many disciplines ‘spatialize’ data by constructing visualisations of non-spatial material in the form of maps or graphics” (Golledge et al., 2008); in the environmental domain, the observed objects or events all have a meaningful spatial (and temporal) dimension, which means that when talking about data, this implicitly means spatial data. Besides, environmental sciences targeting the terrestrial system (geo) in the broad sense (including atmospheric phenomena for example), the word “geospatial” becomes equivalent to “spatial” in our purpose. Finally, even if “geographic” data is a subset of spatial data, “geographic”, “spatial” and “geospatial” data are often used interchangeably⁹ but we will stick to “data” or “spatial data” interchangeably in this thesis that “relate to a location on the Earth” (Dessers et al., 2015).

Apart from their spatial dimension, data are often mentioned with additional attributes such as “fundamental”, “core”, “environmental”, “big” or “open” that allow to narrow down their typology for a given context. These attributes are non-exclusive as fundamental data can also be open data for example. Even if an exhaustive list of such data typology is out of scope here, it is however necessary to give a brief definition of some of these terms regularly associated with data, keeping in mind their relevance in environmental sciences:

1.1.1.1.1 Fundamental data

Van Loenen (2006) distinguishes two categories of datasets: framework datasets (that correspond to fundamental data) and thematic datasets. Fundamental data sets can be defined as *“the minimum primary sets of data that cannot be derived from other data sets, and that are required to spatially represent phenomena,*

⁹ <http://gis.stackexchange.com/questions/34733/spatial-data-geodata-geographic-data-geospatial-data>

objects, or themes important for the realization of economic, social, and environmental benefits” (UNECA, 2007b). To highlight their importance, Rajabifard (2001a) states that governments and organizations within each nation should agree on which fundamental datasets are required to meet their common interests. This is in line with Nebert (2005) who defines them as *“dataset for which several government agencies, regional groups and/or industry groups require a comparable national coverage in order to achieve their corporate objectives and responsibilities”*. Several years after these calls, the situation remains unsatisfactory as the United Nations Initiative on Global Geospatial Information Management (UN-GGIM) reiterated at its fifth session (August 2015) the urgent need for a set of global fundamental geospatial data themes for consistently measuring, monitoring and managing sustainable development processes (Rajabifard, 2016). An example of the fundamental datasets for Africa is available in section 2.2.10.1.

A parallel can be drawn with the concept of Essential Variables (EVs) that consists in a set of critical domain-specific variables, identified based on criteria of relevance, feasibility and cost effectiveness¹⁰. Essential Variables have been or are in the process of being identified in several scientific domains, such as the Essential Climate Variables (ECVs) for the climate community (Bojinski et al., 2014); the Essential Ocean Variables (EOVs) for the ocean community; Essential Biodiversity Variables (EBVs) for the communities dealing with biodiversity (Pereira et al., 2013). Essential Variables are critical for specific communities whereas fundamental datasets are more generic and necessary for effective and efficient decision making and development across all communities of a given area.

1.1.1.1.2 Big data

Big data is *“data requiring high management capabilities characterized by the 3Vs: Volume, Velocity and Variety”* (Laney, 2001). It has been extended later to three other characteristics: Veracity (need of documenting data quality and uncertainty), Value (need of filtering data to have valuable information) and Visualization (need of presenting complex data and information in an effective way) (Nativi et al., 2015). This recent concept is of particular importance in the environmental domain as new technologies result in growing size and variety of Earth Observation data that are at the same time great opportunities but pose great challenges to scientists and information technology experts (Nativi et al., 2015). One of the biggest challenge is to structure big data more efficiently so that it can become smart data (Spiegel, 2015).

1.1.1.1.3 Open data

Open data is *“data that can be used, reused and redistributed without restriction other than (perhaps) the requirement to attribution or share-alike.”*¹¹ The concept

¹⁰ http://www.wmo.int/pages/prog/gcos/documents/bams_ECV_article.pdf

¹¹ <http://www.isitopendata.org/guide/>

of open data goes in line with the growing trend towards various aspects of “openness” that are promoted by the Open Knowledge Foundation (Molloy, 2011). The principles of open data have been adopted by the Group on Earth Observations (GEO) and the many economic, societal, educational, political and scientific benefits of open data value are summarized by the GEO secretariat (CODATA, 2015). The Global Open Data Index¹² or the Open Data for Africa¹³ project are self-speaking successful illustrations of open data.

1.1.1.1.4 Linked data

The W3C¹⁴ defines linked data as a “*collection of interrelated datasets on the Web*”. This interrelation is made possible by a standardized description (e.g. RDF or SPARQL) of data following four rules (Berners-Lee, 2006). If successfully described, data can be semantically queried and connected, becoming part of the web of data or semantic web, giving a chance to connect data from diverse domains and hence opening the way to new types of applications (Bizer et al., 2009).

1.1.1.1.5 Metadata

Metadata is a “*structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities*” (Ma, 2006). In short it is a description of a data/resource, that is fundamental for an efficient and effective data/resource discovery mechanism (Giuliani et al., 2016). Metadata must hence always be created for any produced dataset or other resource, as it allows potential dataset/resource users to discover it through all the necessary information at disposal: its creation date, extent, description, spatial reference system, owner or condition of use.

1.1.1.2 Data issues

The concepts discussed in the previous section highlight the importance of data in building up a society increasingly based on knowledge for ensuring quality decisions at the other side of the pyramid.

However, Giuliani (2011) identified several barriers that prevent an optimal discovery, access or use of data. A non-exhaustive list of these barriers is: cost, lack of interoperability, constraining legal aspects, lack of metadata, lack of communication. All these barriers have deeper causes that can be technical, institutional, political, legal or social and result in waste of time and money and duplication of efforts, to the detriment of environmental research. Karpouzoglou (2016) argues that some of the barriers discussed are manifestations of power struggles that can create additional barriers.

¹² <http://index.okfn.org/place/>

¹³ <http://www.afdb.org/en/knowledge/statistics/open-data-for-africa/>

¹⁴ <https://www.w3.org/standards/semanticweb/data>

Besides these structural issues, new challenges arise with evolution of technology (e.g. big data, linked data) that require new technological architectures for storing, sharing and analyzing data, but also new institutional frameworks to integrate this evolution that demands more sharing and more collaboration (e.g. for distributed computing).

1.1.1.3 Solutions to data issues

Solutions are then needed to try solving or at least lowering the barriers to an efficient data use, but the willingness, possibility and capability of removing the barriers need to be distinguished (Mazzetti et al., 2013). Based on the causes of the barriers mentioned in the previous section, we can distinguish three categories of solutions to data issues: (1) technical solutions, (2) institutional solutions and (3) individual solutions.

1.1.1.3.1 Technical solutions

Technical solutions are needed to solve primarily the issue of data interoperability, which is *“the ability of two or more systems or components to exchange information and to use the information that has been exchanged”* (Geraci, 1991). Giuliani (2011) showed that interoperability allows to take advantage of the exchange possibilities offered by the Internet, and gives users the ability to find, access, and use data. Data interoperability is made possible by standards that are *“agreed specification of rules and guidelines about how to implement software interfaces and data encodings”* (Open Geospatial Consortium, 2015). The Open Geospatial Consortium (OGC)¹⁵ is the main international organization responsible for making and maintaining open geospatial standards allowing geographic data interoperability. It works in close collaboration with the International Organization for Standardization (ISO)¹⁶ to ensure coherence between all the aspects of the standards.

Interoperability makes it possible to interconnect various data through standards, which opens the way to new solutions such as direct machine-to-machine communication in view of simplifying actions requested by end-users. Several papers linked to technological solutions to data issues will be presented in the fourth chapter of this thesis.

1.1.1.3.2 Institutional solutions

Institutional solutions are needed to solve the political and legal aspects of data issues as it is for example the role of national institutions to define a data policy, to enforce it through laws, and to create the necessary institutional arrangements for these solutions to happen.

¹⁵ <http://www.opengeospatial.org/>

¹⁶ <http://www.iso.org>

At the international level, there are currently no binding laws as could occur at national level, except in the particular case of the European Union where the INSPIRE directive (European Commission, 2007) requests the EU countries to align on a common data policy rules.

Apart from the legal aspect of data policies, an intense effort of coordination is also needed at all geographical scales in order not to duplicate efforts and to concentrate energies towards a mutual beneficial solution. At local/national levels, this can be solved through a well-defined institutional arrangement but at international level, this requires an international coordinated approach. To this end, the Group on Earth Observations (GEO) has been established in 2005 as a *“voluntary partnership of governments and organizations that envisions a future wherein decisions and actions for the benefit of humankind are informed by coordinated, comprehensive and sustained Earth observations and information”*¹⁷. GEO concentrates its efforts in building the “Global Earth Observation System of Systems” (GEOSS), a *“set of coordinated, independent Earth observation, information and processing systems that interact and provide access to diverse information for a broad range of users in both public and private sectors, and links together existing and planned observing systems around the world”*¹⁸.

1.1.1.3.3 Individual solutions

Solutions are also needed at the level of individuals, as they are eventually the data producers or users and hence the ones able to influence solutions to data issues. The data solutions at the individual level require first of all awareness of the existence of these issues and understanding of related consequences. It also requires skills for managing data and related tools, knowledge to take advantage of standardized data and opportunities they provide.

Without capabilities of individuals, the technological and institutional solutions to data issues become meaningless as their potential is not understood or used since individuals are the real data providers and users, and institutions are a construct of individuals. This raises the issue of Capacity Building that is central in this thesis, as well as a need to integrate these different categories of solutions into a single framework to set up the ideal conditions for lowering barriers towards a better data flow, which is the topic of the next section dedicated to Spatial Data Infrastructures (SDI).

1.1.2 SDI

1.1.2.1 SDI definitions

The need for a coordinated surveying and mapping approach has been existing for decades, and even back to the middle of the 19th century in the USA (Robinson, 2008) but the term Spatial Data Infrastructure emerged for the first time in 1993 (ESRI, 2010) from the U.S. National Research Council. Several

¹⁷ <https://www.earthobservations.org>

¹⁸ <http://www.earthobservations.org/geoss.php>

definitions of SDI exist depending on the discipline or the administrative/political levels (Chan et al., 2001):

- “The term “Spatial Data Infrastructure” (SDI) is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data” (Nebert, 2005)
- “Spatial data infrastructures (SDIs) facilitate the collection, maintenance, dissemination, and use of spatial information “ (Kok and van Loenen, 2005)
- “Spatial Data Infrastructure (SDI) is an initiative intended to create an environment in which all stakeholders can co-operate with each other and interact with technology, to better achieve their objectives at different political/administrative levels. “ (Chan et al., 2001)
- “Framework of technologies, policies, and institutional arrangements that together facilitate the creation, exchange, and use of geospatial data and related information resources across an information-sharing community.” (ESRI, 2010)

There is no agreement among specialists on a unique SDI definition, its components and their relationships (Crompvoets et al., 2008). We can however retain from these various definitions that SDIs are more than just data, but a whole environment similar to roads, railways and electricity distribution that supports sustainable development and in particular economic development, environmental management and social stability (Rajabifard et al., 2001a). SDIs provide the spatial dimension in the relationship of people to land and the location component supporting the systems to manage modern societies (Williamson et al., 2003). They indeed consist in a framework that allows integrating data, technology, policy and institutions with objective to improve geospatial data availability, discovery, access, use, creation, collection, exchange and sharing. This explicitly addresses the technological and institutional levels to data solutions presented in the previous section, and indirectly raises the importance of the individual level that is central to these different aspects.

1.1.2.2 SDI characteristics

It is commonly admitted (Giuliani and Peduzzi, 2011a; Giuliani et al., 2013d; Rajabifard, 2002) that SDIs are made of five principal components: (1) data, (2) people, (3) access network, (4) policy and (5) standards as represented in Figure 2.

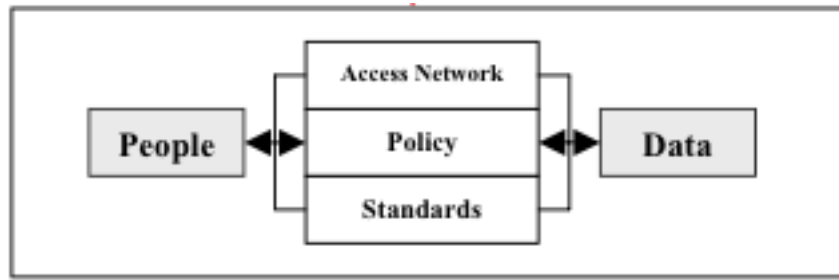


Figure 2: Nature and relations between SDI components (source: Rajabifard, 2002)

Rajabifard (2002) argues that all decisions require data. Since people are key to decision-making, and relationship of people to data become increasingly complex, he highlights the central supporting role of some other components (access network, policy and standards) to facilitate the interaction between people and data for governance. He considers the access network as the means by which the datasets are made accessible to the community. These means require a technical infrastructure (e.g. servers, Internet access, agreed-on standards) but also a policy that is determined within an institutional network. We can hence distinguish four main categories in a SDI: (1) People; (2) Data; (3) Infrastructure and (4) Institutions (Figure 3). However, much of the technology needed already exists and the success will depend on an institutional and cultural willingness to share data out of ones immediate work group (Williamson et al., 2006). For Cooper et al. (2013), the building and maintenance of the organizational structures (institutional aspect) are key to ensure the existence of an SDI. Kok (2005) defines five crucial organizational aspects for the development of a national SDI: (1) a vision; (2) leadership; (3) a communication strategy; (4) coherence and (5) intention of the geographic community to initiate new innovations.

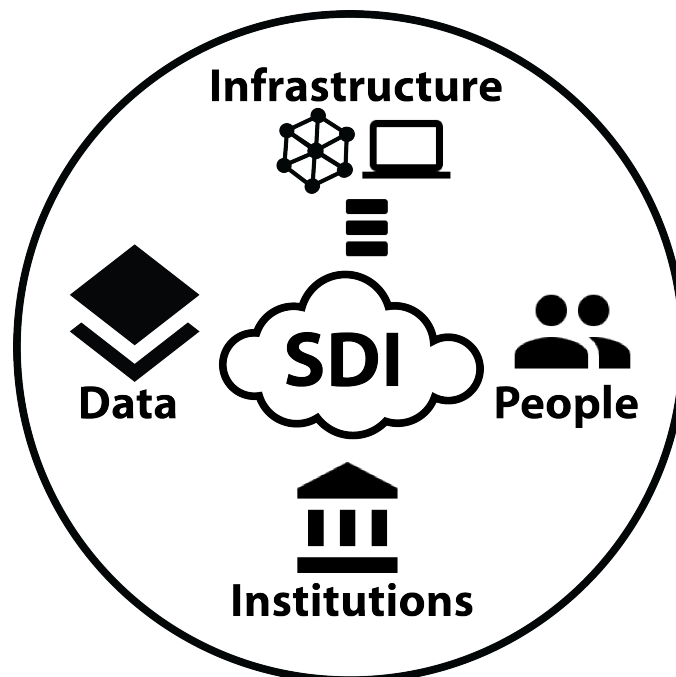


Figure 3: SDI components

Rajabifard (2002) also states that SDIs are characterized by a systemic dynamicity due to the rapidly evolving nature of their technological component, requiring the other components to permanently adapt. As an illustration, permanent technological progress allows more data processing through improved hardware or distributed computing, as well as data crowdsourcing through new devices. This requires new skills from the people component and new data policies for managing the legal aspect and flow of this increasing amount of data (e.g. open data policy) from the institutional component.

The dynamic and multi-dimensional nature of the SDIs (Chan et al., 2001; Kok et al., 2005) makes their understanding, approach and monitoring a complex task. For example Chan (2001) argues that giving a definition of SDI depends on the various perspectives or views of the different users and their interests. This shows once again the importance of the individual (people) component, which gives a meaning to the SDI concept. Consequently, the development of an SDI is by nature a long-term project (Najar et al., 2007b; Williamson et al., 2006) which needs long term investment, that develops gradually with different stages (Kok et al., 2005), addressing in priority the most pressing issues such as collecting and sharing data before including the political aspect. Moreover, there is no fit-for-all solution or uniform approach to successfully set up a SDI (Van Orshoven, 2003).

Rajabifard (2004) describes the hierarchical characteristic of SDIs, that range from corporate and local levels to the global level, which forms an environment allowing decision-makers of any geographical level to use data from the other levels (Rajabifard et al., 1999), creating a dynamic and hierarchical relationship between each level (Figure 4). Rajabifard also distinguishes between two possible views of the nature of this SDI hierarchy: (1) the umbrella view, in which SDI of a higher level encompasses all the SDI components of the levels below, and (2) the building block view where each lower building block supports the higher level block by providing it with necessary lower level data. The umbrella view suggests that the five SDI components are in place at any geographical level to use and share the data of the lower level whereas the building block view suggests that each level is able to provide the superior level with the spatial data needed.

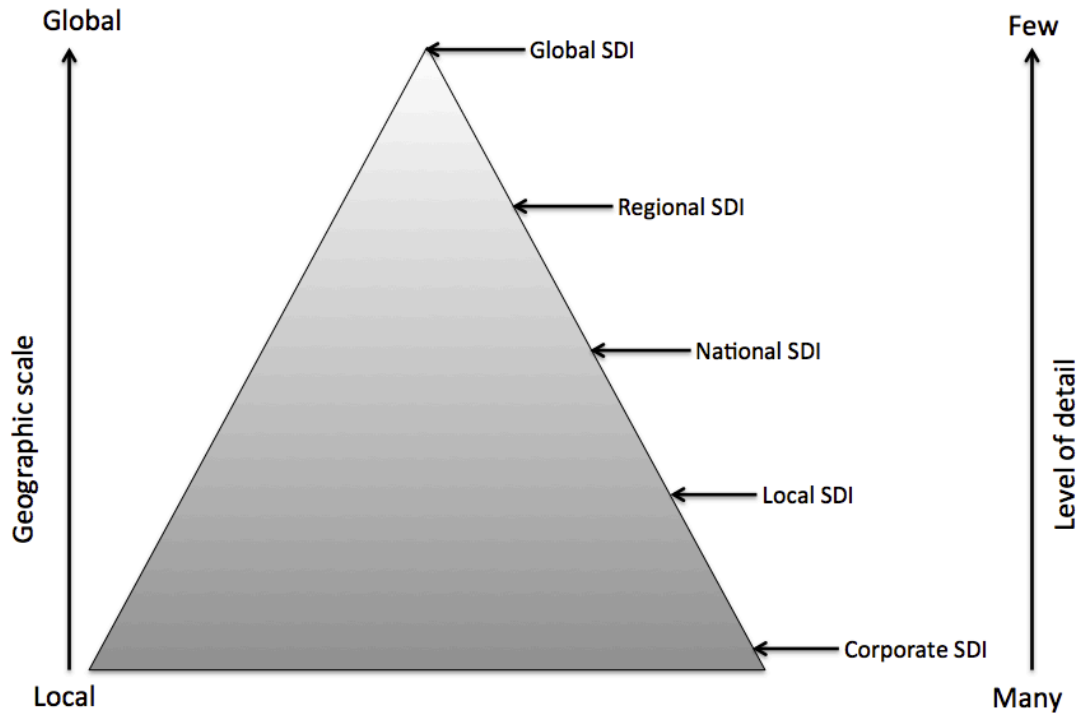


Figure 4: SDI hierarchy (based on Rajabifard, 1999)

Williamson (2006) argues that it is at the sub-national level that the majority of relevant data (collection of land taxes, land use planning, infrastructure development) for people is produced. But at the same time other authors (Najar et al., 2007b; Norris, 2015) highlight the special significance of the national level through National Spatial Data Infrastructures (NSDI) as even if data are produced at sub-national level, it is the national government that can add a stamp of authority to data and services and provide authoritative data. Additionally, this is also at the national level that juridical, political and administrative national decisions are taken and supersede the local levels. In any case, a SDI should result from a coordinated vision at all relevant levels, motivated by the perspective of ready access to spatial data to support decision-making (Rajabifard et al., 2004) and supported by funding models to guarantee ongoing SDI development and maintenance.

Based on these goals and characteristics, SDIs should help avoiding fragmentation, gaps in availability of Geographic information, duplication of data collection and problems of identifying, accessing or using available data (Van Orshoven, 2003). In the information society where we live, such a framework is definitely needed to support the daily management of the various interconnected components of the society, eager of quality data for quality decisions. This should benefit stakeholders by maximizing the value of their investments in geoprocessing systems and data, hence allowing significant financial savings in data gathering resulting in economic, social and environmental benefits (Manso Callejo and Castelein, 2010). These potential benefits have led to the development of SDIs in many countries, generating benefits in many fields of the public sector and beyond (Longhorn and Blakemore, 2007) and led to

acceptance of SDIs as an essential infrastructure in a modern society, necessary to spatially enable it (Crompvoets et al., 2008).

Despite this momentum of enthusiasm and potential benefits, some authors (Dessers et al., 2015; Díaz Sánchez et al., 2012; Nedovic-Budic et al., 2004) argue that the real impact of SDIs is not up to the expectations because they often stay out of reach and have failed to achieve the desired level of impact and penetration. This raises the question of the causes of this shift between expected benefits on one side and the real SDI implementation and efficient data sharing on another side. Several barriers are mentioned by different authors (Craglia, 2010; GEO Secretariat, 2015; Sebake and Coetzee, 2013; Van Orshoven, 2003; Williamson et al., 2003) and concern all SDI components. They can be technical (e.g. data inconsistency or incompatibility, lack of data documentation, lack of interoperability), institutional (e.g. incompatible mandates or level of expertise between institutions, data policy issues, lack of funding) or social (e.g. cultural resistance to innovation or to sharing, lack of motivation or commitment, language). Addressing these various types of barriers towards successful SDI implementations requires a better case-by-case understanding for suggesting tailored solutions based on experience and best practices.

1.1.3 Capacity Building

Best practices regarding spatial data approach and use, in particular for SDI, have been built up based on years of experts' learning and experiment through research and projects. All this accumulated knowledge needs to be brought into practice and adopted worldwide for proper global implementation. It demands transfer of this knowledge from the SDI expert practitioners to the targeted societies, in order to build or re-enforce their capacity with the best practices experienced. However, this requires to understand the specific needs of the various societies and social systems in which an SDI operates (Rajabifard et al., 2004) as many barriers, including socio-economic or cultural ones might prevent an efficient SDI framework adoption. The notion of capacity building is then key and requires some clarifications.

1.1.3.1 Capacity Building concepts

Williamson (2006; 2003) states that capacity building is a complex issue whose definition depends on the user and the context. He proposes to see it as a methodology aiming to provide a sustainable outcome through assessing and addressing a whole range of relevant issues and interrelationships, that will eventually give a system, an organization or a person the power to perform and produce properly. The Group on Earth Observations (GEO) bases its approach of capacity building on a UN one that encompasses human, scientific, technological, organizational and institutional resources and capabilities with the goal to enhance the abilities of stakeholders to evaluate and address crucial questions related to policy choices and different options for development (GEO secretariat, 2006).

We can retain from these definitions that capacity building is a process targeting several types of resources and capacities, to re-enforce the target by giving it the ability to independently evaluate, decide, produce or perform. The target becomes hence less dependent thanks to knowledge, which is an asset in power games. Eade (1997) confirms that capacity building is concerned with social and political relationships and cannot be viewed in isolation from the social, economic and political environment of the society, making it critical to understand the context for a successful capacity building. Williamson's definition also distinguishes two essential steps in capacity building: a first one that consists in assessing the situation and a second one for addressing it. This is confirmed by Rajabifard (2004) who clearly states that capacity building has two components: 1) capacity assessment, which is the analysis of desired capacities against existing capacities (UNDP, 2009) and 2) capacity development, which is the process through which individuals, organizations and societies obtain, strengthen and maintain capabilities (UNDP, 2009).

Williamson (2003) distinguishes three levels for building capacity:

- The societal level, that is for example an entire country or society;
- The entity or organizational level, that might be a government, an organization;
- The individual level.

Among these levels, he considers the individual level as the most critical as it needs to function efficiently and effectively within the two upper ones: the organization and the society.

Besides the capacity building levels, Williamson (2006) also distinguishes several capacity factors that are important for a successful SDI implementation: technological, human and financial capacity.

We can then distinguish some specific elements in the capacity building concept: firstly, a capacity building process, that is made chronologically of the selection of the target community(ies), an assessment of their capacity and then a formal building of the missing or weak capacities highlighted in the assessment. This building of capacity consists in a series of actions such as awareness raising, promotion, education and training. Rajabifard (2001a) highlights the importance of awareness raising at the beginning of the capacity building process by the necessity for stakeholders, from politicians to technical people, to understand the potential and advantages of Geographic information, SDIs, and cooperation, that might otherwise not be diffused outside the geospatial community (Giuliani, 2011).

These actions are supported by capacity building instruments or resources that consist in any human, organizational, methodological or technical elements supporting knowledge development. As an illustration, the "Global Earth Observation CApacity Building" (GEOCAB)¹⁹ portal has been developed in the frame of EU FP7 projects with purpose of providing a central entry point for

¹⁹ <http://www.geocab.org/>

capacity building resources linked to Earth Observations. The types of capacity building resources available in the portal, corresponding to what we consider as instruments of capacity building are the following:

- **Organization:** institution, person, project, whose purpose is related to the use of data derived from Earth Observation
- **Reference:** reference project or initiative that has mobilized capacity building resources in Earth Observation
- **Access program:** agreement for accessing to EO images or derived products
- **Document:** technical or scientific document or tutorial on EO data or software use
- **Service:** online service allowing to discover, view, access or process EO data
- **EO product:** satellite images or derived data
- **Software:** software or library for cataloguing, analyzing, processing, visualizing and laying out geospatial data
- **Training**
- **Marketing toolkit:** methodological tool to support the development of projects based on the use of EO data
- **Event:** workshop, conference, training session related to the use of EO data

These various types of capacity building resources highlight the importance of access to information, that is a key element of capacity building, which requires the most appropriate combination of resources depending on the targeted community(ies) or capacity building level.

1.1.3.2 GEO Capacity Building activities

Rajabifard (2004) argued that the international community needs to pay more attention to capacity building. Echoing this statement, the Group on Earth Observations (GEO) that coordinates Earth Observation activities at the global level, established a capacity building working group (foundational task CD-01) that developed a capacity building strategy in 2006 (GEO secretariat, 2006) and also has a dedicated capacity building section on its website²⁰. This strategy is partly based on a preliminary survey that revealed weaknesses in Earth observation capacities, mainly in developing countries, such as limited access to capacity building resources, lack of e-science infrastructure for Earth Observation education and training, inefficient connectivity, lack of awareness about the value of Earth observations, etc.

These elements guided the GEO capacity building strategy through the establishment of a vision, some guiding principles, objectives and actions.

GEO (2006) defines three target levels of capacity building:

²⁰ <http://www.earthobservations.org/cb.php>

- The **human** level: it aims at giving individuals the necessary skills for Earth Observations management
- The **institutional** level: it aims at ensuring that the institutional aspect linked to Earth Observations (e.g. policies) support Earth Observations activities
- The **infrastructure** level: it aims at having a solid infrastructure (e.g. servers, access network) to support Earth Observations activities

The GEO Capacity building levels can be seen as horizontal levels across the society, targeting the necessary elements of society for building an efficient SDI. These horizontal levels are complementary with the ones of Williamson that are vertical and aim at re-enforcing Capacity Building at each level of the society. This complementary approach ensures that the concerned individuals, organizations or societies have the required knowledge for all the SDI aspects (Figure 5).

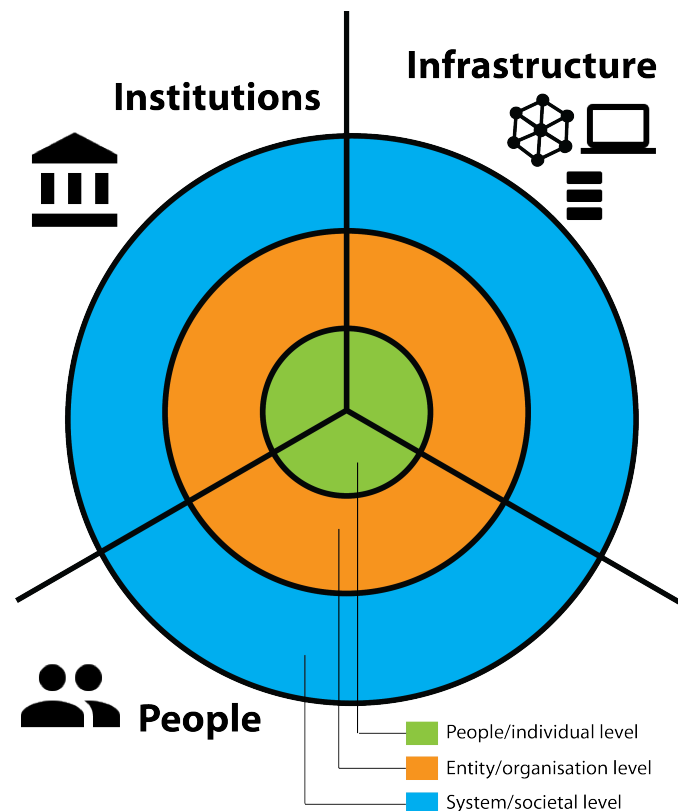


Figure 5: Capacity Building levels (based on Williamson, 2003 and the GEO Capacity Building strategy, 2006)

Additionally, the three Capacity Building levels defined by GEO show a close matching with the SDI components described in section 1.2.2, as illustrated in Figure 6:

- the GEO human level corresponds to the “people” SDI component;
- the GEO institutional level corresponds to the “institutions” SDI component;
- the infrastructure level is found in both capacity building and SDI

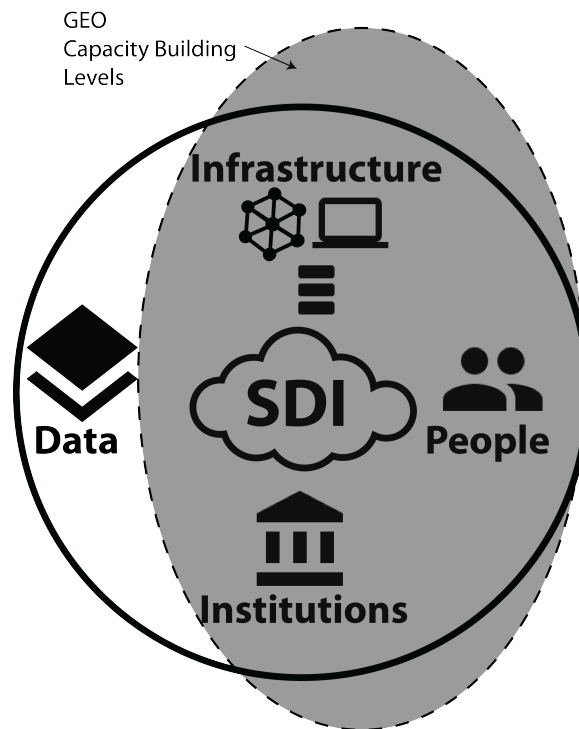


Figure 6: SDI components and GEO Capacity Building levels

This suggests that the GEO capacity building approach can easily be brought into practice for SDI based on similar entities types targeted. Rajabifard (2004) specifies that since the adoption of Agenda 21 in 1992, there has been an increased emphasis on building human and technological capacity to access and use available spatial data to support decision-making.

The combination of the coordinating and capacity building roles of GEO in Earth Observation is also favorable to the promotion of dialogue between stakeholders or even adverse parties by gathering the parties and giving them access to available information, knowledge and simulations. Despite a non-binding role, such an “informed” dialogue can help to reconcile positions and envisage common solutions.

Finally, the global coordination and capacity building efforts linked to Earth Observations are materialized through the GEOCAB portal mentioned prior, that is jointly developed by the European Commission and the Group on Earth Observations to promote and provide a central access to capacity building material.

1.1.3.3 Capacity Building added value for SDIs

Given its empowering role, capacity building represents an unavoidable challenge to re-enforce SDIs worldwide and is perceived as a central component, as much required as other SDI elements such as SDI vision, integration of spatial datasets or financial support by several authors ((Rajabifard et al., 2004; Williamson et al., 2003). The importance of Capacity Building is illustrated by its clear mention in the seventeenth UN Sustainable Development Goal (data, monitoring and accountability) for its support to developing countries to

increase significantly the availability data with relevant characteristics in national contexts (Rajabifard, 2016).

Masser (2008) states that a key challenge is to develop a SDI that will serve the majority of the society that is not spatially aware. Williamson (2006) adds that this should be done by providing not only access to spatial information, but also to additional resources such as business goals, strategies, processes, operations and models. That would contextualize information and allow SDIs to progress from information age to knowledge age, contributing to make science and data more transparent and accessible, empowering citizens and increasing public confidence in scientific enquiry (Karpouzoglou et al., 2016). The capacity building process, mainly through its awareness raising action, seems a perfectly fit-for-purpose element to address such a challenge.

Capacity building can bring a lot of advantages for the benefit of SDIs:

- More autonomy and less external dependence in SDI related fields for individuals (not only the spatially aware elites), institutions and society in general depending on the level(s) targeted in the capacity building process;
- Societal advantages in a SDI committed society where spatial data supports daily needs (e.g. meteorological, mobility, business applications based on spatial data);
- Individual/institutional/societal participation to global SDI efforts through standardized data sharing allowing a better use of environmental data;

All these elements seem to give capacity building a prominent role in the SDI implementation, which is precisely what we want to explore in more details in this thesis.

1.2 Research questions

We have seen in the previous sections that Spatial Data Infrastructures are the recognized framework to address spatial data sharing issues, which is necessary to tackle current environmental challenges. SDI shall then ideally be recognized worldwide and at each geographic level as the solution to data sharing issues. This recognition is becoming a reality with many SDI related initiatives, projects or institutions being created at various geographic levels. We can mention a few: the Global Spatial Data Infrastructure (GSDI) association²¹ or the Group on Earth Observations (GEO)²² at global level; the INSPIRE directive at regional European

²¹ <http://gsdiassociation.org/>

²² <https://www.earthobservations.org>

level or the AfriGEOSS²³ initiative at the African level; many National Spatial Data Infrastructures at country level or subnational initiatives²⁴.

However, their real implementation status is not clear since establishing and maintaining a complete Spatial Data Infrastructure, with the necessary technological, political and individual arrangements is a lengthy process that might require deep societal changes. The barriers to these changes can be tangible, such as the adoption of technological standards for interoperability and easily measurable, but some other intangible barriers like sharing willingness or SDI vision are much more difficult to monitor. In both cases, we argue that the capacity building process is unavoidable to properly address tangible and intangible barriers towards a full SDI implementation at various geographic levels. This affirmation results from the central role of individuals in both SDI implementation and capacity building processes. In SDI implementation, even though several components (technological, institutional, humans) must be addressed, the individuals are also the ones influencing and forming the political and technological levels. In building capacity, the building target is always an individual or a group of individuals that has then a direct effect on their technical capacity and institutional or societal influence, that will in turn influence the data component. Building capacity of this “human” element should then always be the primary action of any sustainable and successful SDI implementation.

Consequently, the aim of this research is to:

Examine the role of capacity building in successful SDI implementation by lifting tangible and intangible barriers.

The associated research questions are the following:

1) What are the existing methodologies to evaluate SDI implementation?

In order to find out if and where SDI capacity building is needed, it is necessary to evaluate the SDI implementation status. It is only with such information that it becomes possible to determine what actions need to be taken, which might include capacity building. To this end, we want to know if methodologies exist for assessing SDI implementation status, and if these can be applied without distinction at all geographic levels.

2) What are the innovative solutions to lift the barriers of SDI implementation?

After identifying the SDI implementation status in given areas, it is necessary to determine what types of actions need to be undertaken to improve SDI

²³ <http://www.earthobservations.org/afrigeoss.php>

²⁴ many of them linked from the GSDI website:

<http://gsdiassociation.org/index.php/publications/sdi-links.html>

implementation, that vary a lot depending on the barriers encountered. In some cases technological solutions might be needed while in some other cases institutional solutions are required. For example, technical barriers might be a brake (e.g. in the case of scientists not experts in GIS) to some people, preventing them to further share data they have produced, due to the complexity of the GIS workflows. In such cases, technological barriers could be lowered through automation, reducing in parallel the need for Capacity Building while improving at the same time data use and sharing. Through this question, we will focus on the technological barriers, but other types of barriers will also be discussed.

3) How to measure the impact of CB activities on Spatial Data Infrastructures?

Determining the SDI implementation status and innovative actions to be undertaken to improve gaps is core for progressing to a better SDI implementation. It is however necessary to be able to measure the impact of the actions taken, in particular the capacity building actions. This is particularly important for justifying the benefits of the actions undertaken and potentially obtain funding for continuing such actions. Being for example able to measure the institutional or societal uptake resulting from capacity building efforts in SDI related field is essential for better targeting future capacity building efforts.

1.3 Thesis structure, contributing papers and projects

1.3.1 Structure of the thesis

This thesis is structured in six chapters addressing the research questions defined in section 1.2.

Chapter 1 sets the scene by describing the reasons why data are of crucial importance for addressing environmental challenges. It then introduces some fundamental concepts underlying the thesis: data concepts, issues and possible solutions; Spatial Data Infrastructure definition and characteristics; capacity building concepts and international implementation. It finally describes the aim of this research and the associated research questions to address it.

Chapter 2 focuses on the first major action to perform for implementing or improving an SDI, which consists in performing an inventory of the situation or stocktaking. It explores the existing methodologies, discusses their difficulties and limitations, and proposes a new approach for assessing the SDI situation at a continental level with limited financial and human resources. It also discusses the findings discovered in applying this methodology to Africa and potential solutions to improve the situation.

Chapter 3 discusses the first type of actions that shall be taken once the SDI situation has been clarified: capacity building activities. Using successful use cases, it aims at demonstrating the importance of this type of actions all along the SDI implementation process, and in particular at the beginning through awareness raising activities.

Chapter 4 focuses on the implementation aspect of SDI through their technical side. It recognizes that technical complexity is a key issue that must be addressed in any effort towards a wider data production and use. In an attempt to contribute to this effort, it proposes three different technical solutions that rely on the principle of interoperability: a first one that proposes a customized access to data; another one that demonstrates the feasibility and benefit for users of harmonizing heterogeneous data; and a last one that tackles the metadata issue through simplification of the process.

Chapter 5 aims at demonstrating the importance of assessment at various levels for SDI. It distinguishes between assessment of the whole SDI or some of its components, and assessment of the impact that SDI as a whole can have, or impact that some actions can have on the whole SDI. Two assessments methods are presented in this chapter: one for an SDI component and one for the impact of activities on the SDI with concrete use cases.

Chapter 6 concludes this research in answering the three research questions, making some recommendations and opening the way to some perspectives of research. It ends this research work with a final conclusion containing a condensed reasoning based on the various findings of this thesis.

1.3.2 List of contributing projects

This research benefited directly or indirectly from six international projects (Figure 7) supported by the European Commission “seventh Framework Programme” (FP7) as well as the Swiss National Science Foundation (SNSF) “Scientific co-operation between Eastern Europe and Switzerland” (SCOPES).

The EU FP7 enviroGRIDS project (<http://www.envirogrids.net/>) was a four years project (2009-2013) that aimed at assessing water resources in the Black Sea catchment in the past, present and future based on different development scenarios. It developed many datasets compatible with the European INSPIRE Directive on spatial data sharing across Europe, that are now available, along with their metadata, on the enviroGRIDS platform (<http://blacksea.grid.unep.ch/>). It promoted the use of web-based services to share and process large amounts of key environmental information, and organized several capacity building events in the domain.

The EU FP7 AFROMAISON project (<http://www.afromaison.net/>) was a three years project (2011-2014) that aimed at putting into practice the concept of integrated natural resources management (INRM) at meso-scale by providing a practical approach and tools that can be applied in a variety of environmental and socio-economic conditions. Several tools were developed, including a specific AFROMAISON SDI²⁵ aiming at improving exchange of information between the various partners and stakeholders. Many capacity building activities were also performed during the project, including activities aiming at re-enforcing GIS and SDI knowledge among partners.

The EU FP7 EOPOWER project (<http://www.eopower.eu/>) was a two years project (2013-2015) that built on the results of previous projects (e.g. to valorize former datasets, success stories or tools developed) and aimed at creating the conditions for sustainable economic development through the increased use of Earth Observation products and services for environmental applications. This serves the higher goal of effective use of Earth Observation for decision-making and management of economic and sustainable processes. Synergies were made with other projects such as EU FP7 IASON and EU FP7 EcoArm2ERA, giving the opportunity to organize several joint capacity building events.

The EU FP7 IASON project (<http://www.iason-fp7.eu>) was a two years project (2013-2015) that aimed at establishing a permanent and sustainable network of scientific and non-scientific institutions, stakeholders and private sector enterprises belonging to the European Union and Mediterranean and Black Sea regions countries. It targeted three main domains of activities for which Earth Observation could play a significant role: actions to address climate change; research and innovation to improve resource efficiency; raw material management. The IASON « Permanent Networking Facility » (<http://iason-fp7.eu/pnf/>), which is an online directory gathering regional partners and their expertise, is one of the most visible achievements of the project. Many synergies

25

http://www.afromaison.net/index.php?option=com_content&view=article&id=73&Itemid=182

were established with the parallel EU FP7 EOPOWER project, including joint capacity building events. Furthermore, it sought to uptake results of previous projects (e.g. enviroGRIDS).

The EU FP7 EcoArm2ERA project (<http://www.ecoarm2era.eu/>) was a three years project (2011-2014) that aimed at reinforcing the cooperation capacities of Armenia's leading research institute in environmental research and ecology: the Armenian Center for Ecological-Noosphere Studies (CENS). This was done by defining and promoting a development strategy for improving the institution's capacities, visibility and competitiveness; developing a strategic partnership with other European research institutions; building the competencies needed by Armenian researchers to participate in EU FP7 and H2020 programs. Synergies were established with the EU FP7 EOPOWER project to use Armenia as a pilot for the application, assessment and consolidation of the EGIDA methodology. This methodology aims to re-enforce institutional capacity building through a set of best practices and guidelines for a sustainable contribution to the Global Earth Observation System of Systems (GEOSS).

The SNSF SCOPES ARPEGEO project (<http://www.arpegeo.sci.am/>) was a two years project (2011-2013) designed to establish a tripartite institutional partnership between the University of Geneva (UNIGE), the Armenian Center for Ecological-Noosphere Studies (CENS) and the Armenian Institute for Informatics and Automation Problems (IIAP). It aimed at building capacities at CENS and IIAP regarding the management, processing and sharing of geospatial data in the environmental domain. The joined efforts of the partners resulted in a successful deployment of the first environmental data sharing and interoperability services in Armenia, which strengthened the national capacities of geospatial data sharing, increased the visibility and national position of CENS as an expert in environmental research, and expanded their regional and international networks in this field.

1.3.3 List of contributing research papers

Guigoz, Y., Giuliani, G., Nonguierma, A., Lehmann, A., Mlisa, A. and Ray, N. (2016). "Spatial Data Infrastructures in Africa: a gap analysis." *Journal of Environmental Informatics*.

Guigoz Y./Lacroix P., Rouholahnejad E., Ray N. and Giuliani G. (2016). SCOPED-W: Scalable Online Platform for extracting Environmental Data and Water-related model outputs. *Transactions in GIS, in press*.

Guigoz, Y., Lacroix, P., Ray, N., Lehmann, A., Dao, H., Lacayo, M. and Giuliani, G. "EGAL: a methodology for Environmental Geoportals Assessment and Label", submitted in *International Journal of Spatial Data Infrastructures Research*.

Giuliani, G., Papeschi, F., Mlisa, A., Lacroix, P. M. A., Santoro, M., Nonguierma, A., Cools, J. and Guigoz, Y. (2015). "Enabling Discovery of African Geospatial Resources." *South-Eastern European Journal Issue of Earth Observation and Geomatics* 4(1S): 1-16.

Giuliani G., Guigoz Y., Lacroix P., Ray N., Lehmann A., (2016) Facilitating the production of ISO-compliant metadata of geospatial datasets. *International Journal of Applied*

Giuliani G., Lacroix P., Guigoz Y., Roncella R., Bigagli L., Santoro M., Mazzetti P., Nativi S., Ray N., Lehmann A., Bringing GEOSS services into practice: a capacity building resource on spatial data infrastructures (SDI). *Transactions in GIS*, in press.

Asmaryan S., Saghatelyan A., Astsatryan H., Bigagli L., Mazzetti P., Nativi S., Guigoz Y., Lacroix P., Giuliani G., Ray N. (2014) Leading the way toward an environmental National Spatial Data Infrastructure in Armenia. *South-Eastern European Earth Observation and Geomatics* 3:52-62.

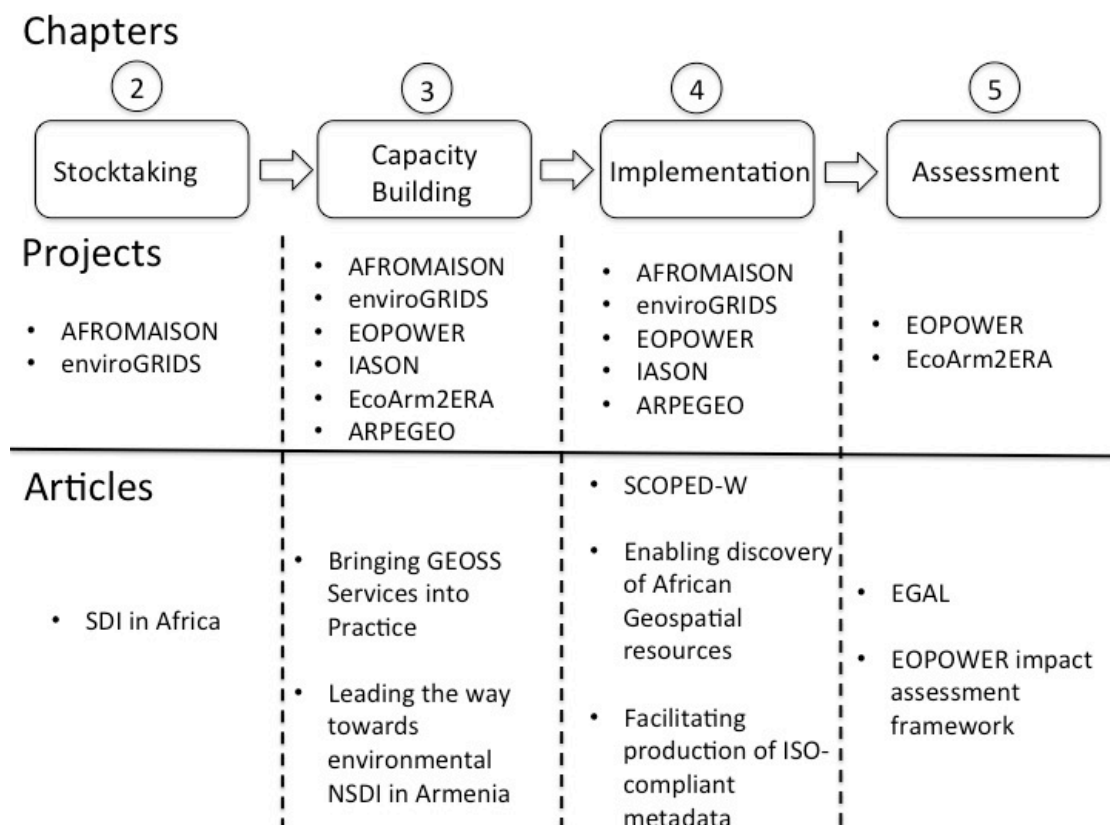


Figure 7: Synthesis of projects and articles contribution to the chapters of the thesis

2 SDI STOCKTAKING

2.1 Introduction on SDI stocktaking

SDIs are now recognized as an essential infrastructure in modern societies (Crompvoets et al., 2008) given the benefits they can bring if successfully implemented. The definition of an ideal SDI depends from the perspective and can consist in the satisfaction of end users through a SDI fitting their needs or the involvement of all relevant stakeholders in the SDI use (Grus et al., 2006). A successful SDI implementation can hence be seen as a proper functioning of the various SDI components (skilled people, committed institutions and laws, adequate technologies) ensuring a successful data workflow that will satisfy the society and include the relevant stakeholders. In order to evaluate the implementation level of an SDI, it is necessary to perform an assessment of their effectiveness and efficiency (Crompvoets, 2006) for several reasons (Giff and Crompvoets, 2008; Grus et al., 2006; Grus et al., 2007; Najjar et al., 2007b; Williamson et al., 2006):

- Need to determine if an SDI achieves its defined objectives
- Need to know how to improve the SDI performance in case the objectives are not achieved
- Need to concretely report an SDI status to justify the matching between the goals and the investment (cost/benefit relations)
- Need to attract new funding with precise defined goals whose achievement can be measured
- Understand the SDI potential impact on the geoinformation market
- Need to contribute significantly to increasing knowledge about the key qualities of SDI.
- Need to make short term results visible for easier getting political support

In order to address SDI assessment needs, several methodologies have been developed with different approaches focusing on specific aspects of SDI, that are summarized in (Grus et al., 2007). Some of these methodologies have been applied in selected use cases while others are only conceptual. But each of them only captures a partial aspect of SDIs that are of complex, dynamic and multifaceted nature. SDIs can be seen as a complex system that is more than the sum of its parts, and requires complex methods to properly reflect the reality. Despite the numerous existing methodologies, none of them seems to meet the requirements of practitioners (Grus et al., 2011; Nushi et al., 2015). This complexity is one of the main challenges that makes their assessment difficult, along with other obstacles such as: assessment data availability, reliability and objectivity; high dependence on focal points; language; lack of commitment of targeted stakeholders; assessment cost (Grus et al., 2006). These obstacles need to be taken into account in a SDI assessment framework. A unique and universal assessment method would require a single and globally accepted SDI definition, which is not the case (Chan et al., 2001). Consequently, an ideal SDI assessment framework should integrate this complexity by being able to: incorporate the different views and conceptions of SDI role and objectives; take into account the

predictable and unpredictable changes; describe the evolution of SDI in time instead of a single moment snapshot (Grus et al., 2007).

Based on these requirements, a framework based on Complex Adaptive Systems assessments has been developed, called the “Multi-View Assessment Framework” (Cromptvoets et al., 2008; Grus et al., 2011; Grus et al., 2007). It accepts multiple SDI realities, views and definitions instead of controlling and capturing complexity. It aims to guiding users towards the most suitable existing assessment approaches and methods based on their purpose (Figure 8).

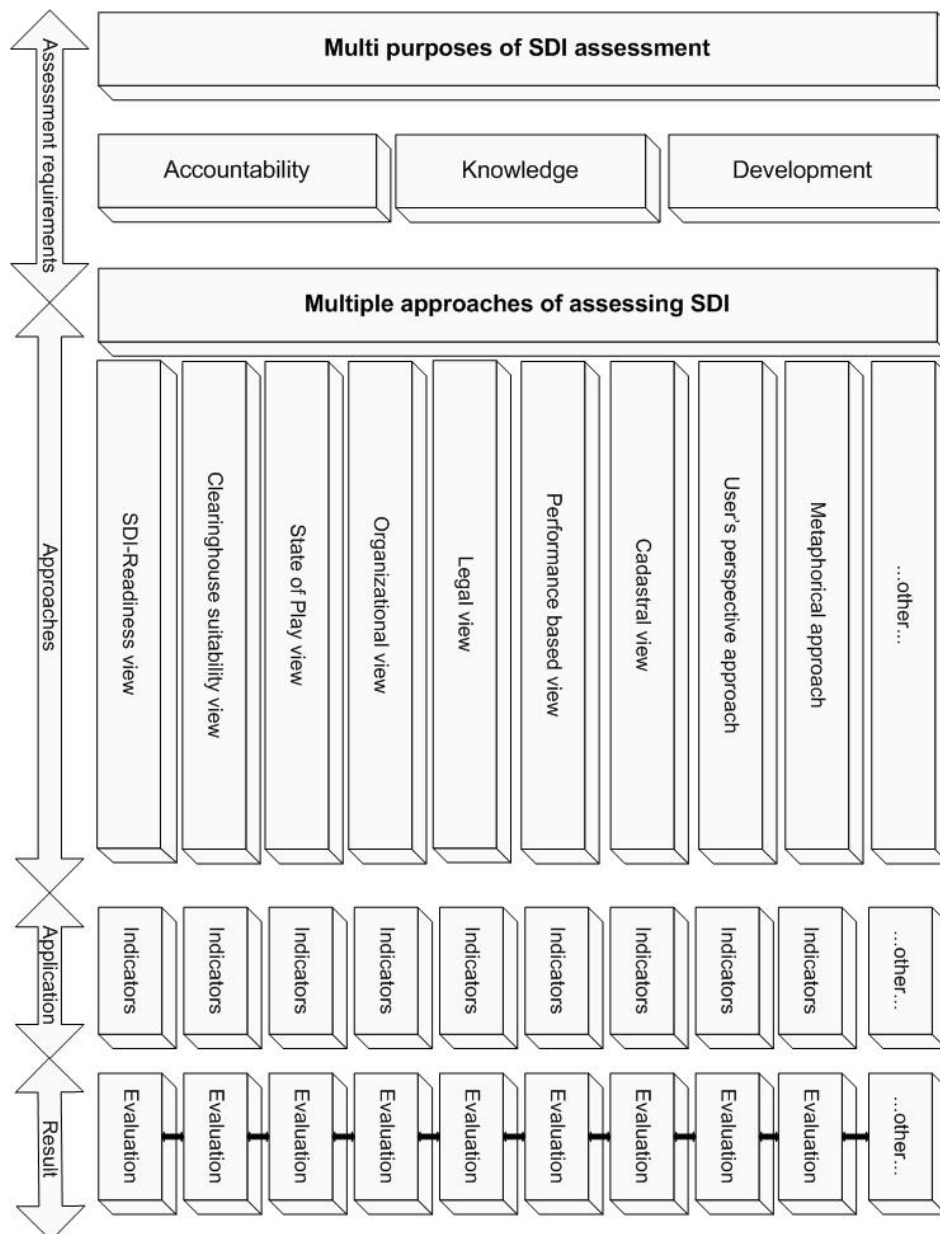


Figure 8: the Multi-View Assessment Framework (Grus et al., 2007)

The Multi-View Assessment Framework is four tiers:

- (1) the “purpose” of the assessment, that is made of three non exclusive classes:

- accountability: a SDI falls in this category when the purpose is to monitor its results to determine if they match the defined goals, or if a change in the SDI results in an increased use of the SDI (e.g. more people using data after the data policy change)
 - knowledge: this is when the purpose is to understand the mechanisms and forces behind a SDI, in view to improve it.
 - development: this is when the purpose is to monitor an SDI's development and recommend changes in case it does not develop as planned.
- (2) A selection of possible approaches, allowing assessing the SDI from different view points (e.g. organizational, technical, SDI-readiness). Each approach has particular goals and methods, and might better fit one or several purposes, as summarized in Grus et al. (2007) or Giff et al. (2008).
 - (3) The application part that consists in measuring the SDI. This is based on several methods proposed such as case studies, surveys, document analysis, or key informants as well as indicators to be defined by the user, such as performance indicators.
 - (4) The results part of the framework, that allows to evaluate the assessed SDI as well as the approach chosen, to make sure it is acceptable to the stakeholders.

Some examples of application of this multi-view framework exist (Grus et al., 2011; Nushi et al., 2012). However, despite the originality of this framework, there are several limitations to using this multi-view approach for the following reasons: (1) some of the approaches remain conceptual and not developed; (2) performing several approaches to capture multi aspects of the SDI requires a solid team of experts, with necessary resources to replicate the assessment at different periods; (3) most of the proposed approaches are better suited for assessing a small scale SDI (from local to national scale) given the resources needed, which is especially true if multi view assessments have to be performed.

It might be necessary to evaluate the SDI status of a whole region, for example a continent. In such a case, it becomes difficult to mobilize the necessary resources (time, funding and people) necessary for a traditional assessment using one or several of the approaches proposed in the multi-view framework. For large scale SDIs, the necessary level of detail in existing approaches can not be accommodated (Grus et al., 2007), making it necessary to develop an innovative approach, closer to a rapid assessment requiring less details. Despite loss of details, it gives the advantage to get a broad picture of the SDI implementation status with few resources. A further refinement in particular areas or countries using other traditional assessment approaches remains possible afterwards.

The conjunction of several factors such as African exposure to environmental challenges, ongoing projects and initiatives (Afromaison²⁶, AfriGEOSS²⁷), and good contact with some African SDI stakeholders made us consider Africa as an interesting use case where to develop such an innovative approach for measuring SDI implementation level. This is the aim of the first paper of this

²⁶ <http://www.afromaison.net/>

²⁷ <http://www.earthobservations.org/afrigeoss.php>

thesis, called “Spatial Data Infrastructures in Africa: a gap analysis” and published in the “Journal of Environmental Informatics”.

2.2 Spatial Data Infrastructures in Africa: a gap analysis

Yaniss Guigoz^{1,2}, Grégory Giuliani^{1,2}, André Nonguierma⁴, Anthony Lehmann^{2,3}, Andiswa Mlisa⁵, Nicolas Ray^{1,2}

¹ United Nations Environment Programme, Division of Early Warning and Assessment, Global Resource Information Database – Geneva, International Environment House, 11 chemin des Anémones, CH-1219 Châtelaine.

² University of Geneva, Institute for Environmental Sciences/enviroSPACE, 66 Boulevard Carl-Vogt, CH - 1205 Geneva, Switzerland.

³ University of Geneva, Forel Institute, Uni Carl-Vogt, CH-1211 Geneva 4, Switzerland

⁴ United Nations Economic Commission for Africa, Information, Science & Technology Division, Menelik II Ave., Addis Abeba, Ethiopia

⁵ GEO Secretariat, 7 bis avenue de la Paix, 1211 Geneva

2.2.1 Abstract

The need for spatially explicit thematic data is currently increasing in parallel to the development of observing, storing and processing capabilities. This requires an integrated data management structure in which human and institutional aspects play a key role as part of a Spatial Data Infrastructure (SDI).

We focus in this study on the African continent to evaluate the status of its SDI implementation. Because assessing SDI at a continental scale in a traditional way (i.e. following methods developed for national assessments) requires financial resources and mechanisms only affordable to developed countries (e.g. European Union), alternative ways have been explored based on fourteen key SDI indicators that were validated by SDI experts in a previous study. Data was collected for each African country through the African leading SDI institution (UN Economic Commission for Africa) and through Internet searches. We found relatively weak scores of the fourteen SDI indicators for African countries compared to the rest of the World, but with notable differences within Africa. We discuss the implication of the lack of information available on the Internet to assess SDI status in Africa. We conclude that it is necessary to improve statistical information in most African countries. This requires an agreed-on geospatial data structure and organization between concerned institutions that is only achievable through a shared global vision on geospatial data governance. To this end, we suggest a few quick wins and several new mechanisms that would enhance the flow of SDI statistical information and improve data management structure in Africa.

Keywords

SDI, Africa, Gap Analysis, monitoring, brokering

2.2.2 Introduction

Environmental issues cannot be solved solely at local, national, or regional scales without an integrated global approach (GEO, 2010a). Furthermore, it requires a knowledge integration from various natural and social sciences as exemplified by Integrated Natural Resources Management (INRM)²⁸ and Integrated Water Resources Management (IWRM) frameworks (Koch et al., 2013). Earth Observation (EO) and Geographic Information Systems (GIS) technologies have much evolved in the last decades to address this data challenge and to assist in environmental monitoring, modeling and analysis as demonstrated in climate change (Xia et al., 2014), ecosystem services (Yang and Yang, 2014), remote sensing (El-Askary et al., 2015) or water management (Su et al., 2013). They allow for example to answer the fundamental question on where to take action before any action is effectively taken. An interesting African example in the water domain is the ESA's TIGER initiative (ESA, 2013) that uses satellite observations to inform local authorities about the state of this vital resource. Another African example is the SERVIR-Africa (2014) project that monitors and forecasts ecological changes and responds to natural disasters. At the global scale, an increasing number of useful Earth Observation products exist: the NASA Shuttle Radar Topographic Mission (USGS, 2010), the ESA global land cover map (ESA, 2014), the FAO Global Soil Map, the UNISDR Global Risk Data Platform (Giuliani et al., 2011a), or the daily ice extent from the University of Illinois (2014). A downside of these products is the sheer amount of data produced, which requires growing storage and management capacities (Mazzetti et al., 2014).

Despite wide availability and use of EO and GIS tools, access to quality environmental and geospatial data remains the largest challenge for supporting decision-making. A plea was made at the Rio Conference twenty years ago for spatially-explicit data to address global environmental issues (Clarke, 1999). This is even more important today, as the integration of environmental data at various spatial and temporal scales is necessary to better understand our global system and take appropriate actions. If geospatial data are necessary for tackling many environmental challenges in developed countries, it is even more so in developing countries where high demographic rates is combined to massive rural exodus, water or power shortages. In emergency situations, quality data is also crucial as demonstrated after the 2010 Haiti earthquake (UNITAR, 2014). According to EIS-Africa, sustainable development of a particular village, city, province, or country requires access to data about the environment (EIS-Africa, 2002).

Data availability is the first step, but accessibility to this data (and associated metadata) by all stakeholders is what can really make a difference. Data must be shared to facilitate its integration with other datasets, and to produce integrated knowledge. But data production, access, use and dissemination are dependent on many other factors such as related laws, regulations, standards, infrastructure

²⁸ Insights and guidance for putting Integrated Natural Resources Management into practice in Africa at meso-scale. Based on the results of the AfroMaison Project. AfroMaison, 2014.

and human factors. This justified the development of Spatial Data Infrastructures (SDI) for efficient geospatial data workflow and management. A commonly accepted definition of an SDI is “*the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of, and access to, spatial data* » (Nebert, 2008). Furthermore, Rajabifard (2002), Giuliani and Peduzzi (2011a), and Giuliani *et al.* (2013d) distinguish the following five main SDI components : Data (geospatial data), People (human resources), Access network (networking technology), Policy (institutional framework) and Standards (technical standards).

2.2.2.1 Local, national and continental SDIs: availability and access

A SDI can be established at different levels, ranging from local to national, continental or global levels. A local SDI typically focuses on detailed datasets of small geographic extent for use at local scale. On the other side of the scale, a global SDI fosters global datasets, generally at lower resolution.

Rajabifard *et al.* (2000) introduced the notion of hierarchical relationship between these different geographic levels of SDI. Each of these different SDI levels has an influence on the upper ones. For example, a National SDI (NSDI) should theoretically provide access to contents of the local SDIs within the country. But this requires a well-organized architecture, both at technical and institutional levels, which is only possible through a National Policy. This is even more needed at higher levels such as continental or global. Setting up such policies at supranational levels is very challenging and requires strong political integration and willingness.

One of the best examples of such a continental SDI is the European INSPIRE directive (European Commission, 2007) that seeks to establish a spatial information infrastructure for the European Union to support environmental policies, and policies or activities that have potential impacts on the environment. The INSPIRE extent groups the European Union countries plus some other voluntary, European non-member states (e.g. Switzerland, Turkey) (European Commission, 2013b).

It is now well recognized that regional and national SDIs can greatly benefit users at different governmental levels. For any SDI implementation, having these five components correctly taken into consideration and adequately implemented creates favorable conditions for production of, access to, use and dissemination of geospatial information. In turn, these favorable conditions of data flow will feed the data needs for informed decisions in many sectors of the society and will be a real added-value for many societal building blocks (Masser, 1998), (Rajabifard *et al.*, 2000). According to the Federal Geographic Data Committee (FGDC, 2014), a NSDI allows reducing the duplication of effort among agencies, improve quality and reduce costs related to geographic information. It makes geographic data more accessible to the public, increases the benefits of using available data, and establishes key partnerships with states, counties, cities, tribal nations, academia and the private sector.

2.2.2.2 SDI assessment

The adoption of SDI principles and technologies is far from being equally acknowledged in all parts of the world. Countries such as USA, Australia, Canada, or Germany have been pioneers in adoption of SDI concepts (Monett and McLeod, 2013) with well advanced NSDI strategies, whilst most other countries are at various stages of adoption and implementation, with many still lacking tangible initiatives or institutions to lead the effort. It is therefore important to be able to map out national or regional performance in developing and implementing SDI capacities (Giff, 2006). Several authors (Delgado Fernandez et al., 2005; Eelderink, 2006; Steudler et al., 2008; van Loenen and van Rij, 2008; Vandenbroucke, 2009) have worked on developing various SDI assessment frameworks. A chosen methodology for SDI assessment highly depends on the geographic extent of the study. Measuring SDI at national or sub-national level is very different from measuring SDI at a continental level for which time and money constraints are a major limiting factor.

2.2.2.3 The case of Africa

Endowed with abundant and diversified natural resources, Africa requires appropriate tool to manage the resources, more so as it is facing major pressing issues such as climate change impacts and environmental stresses. The continent could particularly benefit from an integrated SDI implementation at national, regional and continental scales. Among other reasons that plead in favor of such underlying assumption include: (1) environmental pressure is increasing very fast and necessitates urgent solutions for which multi-disciplinary and transnational environmental data is needed; (2) technical infrastructures are emerging and the time is right to have SDI concepts and tools adopted; (3) there is a great opportunity from the proximity to Europe and its advanced SDI involvement (e.g., INSPIRE directive, participation in many projects promoting SDI particularly in Africa); (4) several SDI actors and initiatives already exist in Africa (see note SN1).

An assessment of the current status of SDI implementation in Africa and an analysis of where the gaps are at continental scale is therefore needed. The objectives of this study are then:

- i) to examine innovative way of performing SDI assessment, notably by making extensive use of Internet searches and using the existing networks of SDI actors;
- ii) to assess the SDI implementation in Africa through indicators and assessment variables that take into account all the SDI components;
- iii) to compare the African situation to the rest of the World.;
- iv) to suggest some ways for improving the SDI implementation and monitoring in Africa

2.2.3 Methodology

2.2.3.1 Review and choice of SDI assessment frameworks

There are currently five main existing SDI assessment frameworks: (1) the SDI Readiness Index (Delgado Fernandez et al., 2005), (2) the organizational maturity matrix (van Loenen et al., 2008), (3) the performance indicators (Steudler et al., 2008), (4) the INSPIRE & NSDI State of Play methodology (Vandenbroucke, 2009), and (5) the set of fourteen key indicators to assess NSDIs in developing countries (Eelderink, 2006).

The SDI readiness index is not appropriate for the African case as it aims at scoring an individual country, which is not meant for study and comparison at continental level. The theory behind organizational maturity matrix combines organizational indicators with development stages, which again does not make it appropriate for a quantitatively measurable continental study, as it would require considerable time to evaluate each country's development stage. Thirdly, Steudler's performance indicators is related to land administration and only gives a broad framework with some possible general indicators for evaluating SDIs, which is not appropriate at continental scale. The core of the INSPIRE & NSDI State of Play methodology is about collecting information from websites, documents and experts. Eelderink's fourteen key indicators for assessing NSDIs have been validated by experts; besides, this reasonably low number of indicators makes it possible to gather sufficient data to perform a meaningful study at the African scale. As the goal of a SDI continental assessment in Africa is similar to the one of INSPIRE, which is to monitor the implementation of INSPIRE (Vandenbroucke, 2009), we decided to follow this methodology in combination with Eelderink's fourteen indicators to assess NSDIs in developing countries. Eelderink groups these key indicators by SDI component, including an additional "Other" component (see Table ST1).

2.2.3.2 Targeted SDI levels

Even if we aim at assessing the continental status of SDI implementation in Africa, the reference level remains countries as most statistics are produced at this level. For the INSPIRE directive in the European Union, the national level is crucial for coordination and implementation (Vandenbroucke, 2010a) as National SDI have a full impact on all levels of the SDI hierarchy (global, regional, state/provincial and local) (Eelderink, 2006). With data for particular SDI indicators at country level, it becomes easy to (1) make continental statistics (e.g., mean, standard deviation), (2) compare the continent to other continents or to an ideal situation, (3) compare continental sub-regions with each other, and (4) compare countries with each other.

2.2.3.3 Selected African SDI champion institution

There is no formal SDI body at a continental level in Africa, as it is the case in the European Union with the infrastructure for Spatial Information in Europe (INSPIRE) or in the USA with the Federal Geographic Data Committee (Makanga and Smit, 2010). However, the UN Economic Commission for Africa (UNECA) plays a key role in Africa in terms of advancing SDI in the continent (Schwabe

and Govender, 2009a). For example, the *Committee on Development Information, Science and Technology* (CODIST) (UNECA, 2013) meeting, held every two years in Africa and organized by UNECA, allows African countries to officially gather around the SDI thematic, in order to advance the SDI agenda in Africa through the resolutions voted during the meeting. Besides CODIST, the ICT department of UNECA performs many actions to promote SDI across Africa (e.g. SDI information collection, conferences and workshops). UNECA is therefore the continental institution of choice to help collecting appropriate data on SDI in addition to information gathered on the Internet.

2.2.3.4 Sources of assessment data

Information provided by UNECA mainly comes from a survey among African countries performed in 2011 (see figure SF2). According to UNECA, 25 of the 53 contacted countries replied to this survey. We re-used UNECA's survey for the following reasons: (1) to our knowledge the UNECA questionnaire is the only existing one on SDI covering all African countries, with the same set of questions, which makes the answers comparable; (2) we consider the commitment of African countries' to UNECA to be better than to other organizations, which impacts the level of responsiveness to the questionnaire; (3) using a single survey should minimize misinterpretation of the meanings of used terminologies that may be different among surveyors.

In addition to the UNECA survey, Internet searches and online databases were used for the assessment.

2.2.3.5 Selected assessment variables for the 14 indicators

The proposed methodology consists in gathering information from UNECA and from various sources on the Internet at national level, and in selecting and organizing information based on fourteen key indicators defined by Eelderink (2006).

In order to perform quantitative SDI comparisons, it is necessary to gather values for each of these key indicators. This requires defining specific variables, or sometimes proxy variables, to measure the key indicators. The methodology and sources of information for measuring each of the 14 key SDI indicators can be found in the online supplemental material (see note SN2).

Among all the proposed variables, some can be considered as "variables of objectives" allowing to concretely measure the current status of an SDI component. The others are "variables of means" to reach the quantitative objectives in longer term by creating a favorable environment. For example, the number of people who attended the CODIST 2011 meeting is a countable variable of objective, while the socio-political stability index is a variable of means that helps in assessing the general SDI environment. The variables of means can be considered as input variables for creating a favorable SDI environment. The variables of objectives are the countable results of SDI status, objectives to measure and improve; these are considered as output variables.

The key indicators, assessment variables, type of variables and sources used can be visualized in table ST2.

2.2.4 Results

The results for each assessment variable can be visualized in an online table with reference sources at <http://africa-sdi.grid.unep.ch>. They are also summed up at the bottom of each indicator's column by continent when possible or for Africa only. The continental results are expressed either as a percentage or as a real value depending on the indicator. The detailed results description by indicator is available in the supplemental note SN3.

The SDI's "data" component results were very weak in Africa (11%) compared to the 87-100% data availability reported for Europe (Vandenbroucke, 2011).

In the "people" SDI component, the "capacity building" key indicator revealed a huge gap for GIS and SDI capacity building in Africa compared to the ideal situation where most countries of a continent should at least have a few institutions teaching GIS to increase local capacity. For the "willingness to share" indicator, neither information about existing data sharing policies nor any national geoportal could be found in African countries. Only memberships of international SDI initiatives could be evaluated, which places Africa in the same position as Asia and the Americas. All the variables measured for the "Human capital" and "SDI awareness" key indicators show that Africa is lagging behind the other continents. The latter revealed a low number of Internet users in Africa despite a growing but still low number of active mobile-broadband subscribers, and low attendance of African countries in key continental SDI meetings (CODIST).

Based on the measurements performed on the four key indicators of the "people" SDI component, Africa has the lowest value of all continents.

The "Access network" SDI component measured through the defined assessment variables of the "access mechanism" key indicator revealed that Africa is in crucial need of improving its information infrastructure and ICT services in order to properly implement SDI.

The "Policy" SDI component measurement in Africa revealed difficulties to track information on SDI funding policies or even on ICT expenditure for the "funding" key indicator. The key indicators "vision", "institutional arrangements", "leadership" and "socio-political stability" are also low compared to other continents. The supplemental figure SF4 shows the details by country of the political stability index, with Africa less socio-politically stable than other continents.

Regarding the "standards" SDI component, information we found for the indicator "metadata availability" for fundamental datasets in Africa was scarce, making it difficult to draw a general trend. In terms of "interoperability" indicator assessment, only a few countries reported national working groups on standards, but about half of the African countries are compatible with the African geodetic Reference Framework, which is encouraging but still low.

Unfortunately, no African institution was member of the leading geospatial standards institution (OGC) as of February 2013.

Finally, regarding the “Other” SDI component defined by Eelderink et al. (2008a) and associated “initiatives connected to SDI” indicator, the GEO membership of African countries is comparable with other continents, as well as the availability of African data in portals of international SDI initiatives such as GEOSS, EyeonEarth, One Geology or GBIF.

2.2.5 Discussion and Perspectives

2.2.5.1 Innovative way of performing SDI assessment

As a continental SDI implementation assessment is different from a national or sub-national assessment, a specific assessment framework is needed. This has so far only been done in Europe through the INSPIRE State of Play, the goal of which is to describe, analyze and assess the status of INSPIRE and NSDI implementation in 34 countries in Europe (Vandenbroucke, 2011). The assessment framework used in INSPIRE State of Play could not be directly transposed to Africa for the following reasons:

(1) The INSPIRE State of Play aims at monitoring on a permanent basis the implementation of the INSPIRE directive in Europe that addresses 34 spatial data themes needed for environmental applications (European Commission, 2014b). The INSPIRE infrastructure in the member states does not necessarily equate the National SDIs (e.g. these might also cover other sectors such as agriculture, spatial planning, or additional technological components) (Vandenbroucke, 2009). For Africa, we do not want to measure the implementation of a particular directive but we want to measure the general SDI components in each country. Even though we followed the business approach of the INSPIRE State of Play, we adapted it to a more NSDI-oriented approach. This is the reason why we completed it with the SDI assessment framework of Eelderink (2006).

(2) In the INSPIRE State of Play, the indicators shall be collected by each Member State on an annual basis and the results must be made public. This is not possible in the African context where there is no political integration such as the European Union and hence no binding directive such as INSPIRE to request indicator collection from member states.

(3) The resources available for this study did not allow a deep institutional analysis such as in the INSPIRE State of Play.

The assessment indicators proposed by Eelderink have the advantages of (1) targeting developing countries, (2) being validated by SDI experts based on case studies in Latin America, Asia and Africa, and (3) consisting in a reasonable number of indicators. For measuring these 14 indicators, Eelderink performed a detailed assessment on six case studies. This detailed assessment was possible given the low number of countries, but in the present assessment we want to

have a general overview of the SDI situation in 54 African countries. We performed then a lighter assessment of each country, based on assessment variables different from Eelderink but still relevant for populating the 14 indicators required.

The main bottleneck of such an assessment is the difficulty to obtain data for some assessment variables due to the large number of countries and heterogeneous availability of SDIs information, both on the Internet and in the SDI champion institution. At the same time, this opens the way to new synergies and mechanisms needed to improve the data availability, between the SDI continental institution and the concerned countries. This is where the political aspects come into play and it is then crucial to improve the SDI implementation status in Africa.

The assessment framework we have used is innovative in the sense that it combines a proven methodology for collecting information used at continental scale in Europe (INSPIRE State of Play) with SDI assessment variables validated by experts used at national level. Moreover, it is the first time that an SDI assessment at the African scale has been performed.

2.2.5.2 SDI assessment

The performed analysis highlighted several trends regarding SDI status in Africa: (1) Africa has the lowest ranking of all continents (except Oceania) in most key variables assessed when an international comparison is possible; (2) for the variables that are only assessed in Africa, most of the results show a very weak SDI status; (3) there is a serious difficulty in finding data about SDI key indicators and assessment variables in Africa on the Internet and even at UNECA, considered as a key institution in Africa for SDI-related information. This makes the SDI monitoring at a continental level not reliable yet, not only in Africa but also worldwide with notable regional differences (e.g. data for European Union countries are easier to find thanks to INSPIRE monitoring).

Getting such information at country level is already a challenge, as it requires a deep analysis, often dependent on the accuracy of national surveys or on the good-willingness of people in the governments providing information. At a continental level, it is even more challenging as a coordinating body (UNECA in the case of Africa, the European Commission through the INSPIRE directive in the European Union) needs to gather data of all member countries. There seems to be an obvious lack of a standardized mechanism that would allow for regular provision of SDI monitoring data. Such mechanisms exist in other fields like economy or health. For example, a lot of health-related data are available at the World Health Organization's website (WHO, 2013), allowing for countries monitoring and comparison. Similarly, the World Bank's website (World Bank, 2013) provides valuable economic indicators. This makes these domains' monitoring much convenient and allows for targeted improvement of the concerned indicators. In the case of WHO, this could be explained by the membership of all the countries and a well-structured organization, with a clear global mandate for health coordination role. Such a global mandate is delivered by the member states through the United Nations. SDI being interdisciplinary by nature, the coordination role is not clear yet in most places of the world. This is

being addressed more and more at national level through NSDIs but is still lacking at continental (except in the European Union with INSPIRE) or global levels. Unfortunately, this prevents an optimal targeting of SDI efforts in a specific area or continent.

The WHO example just mentioned and the successful European SDI through the INSPIRE directive both result from a political consensual decision (at UN level in the first case, at European level in the second case) leading to a successful mechanism of data production and monitoring. This necessitates a political will that remains to be strengthened across Africa and might be one of the causes that lead to a lack of a common SDI African vision. This translates into weak SDI commitments in African countries, even though UNECA plays its coordinating role in organizing SDI continental meetings such as CODIST or other SDI-related events. The weak attendance to the CODIST meetings (about 50% of African countries) confirms this lack of commitment.

However, there is hope that grouping of African countries such as the 15 countries within the Economic Community Of West African States (ECOWAS) may trigger a sub-continental organized environmental data and information infrastructure. This is already the case in the energy sector with the ECOWAS Observatory for Renewable Energy and Energy efficiency (ECOWREX) that is currently redeveloping its rich map viewer (ECOWREX, 2014).

Further initiatives such as AfriGEOSS (GEO, 2014a) may raise awareness on the benefits of SDI and promote national commitments for the implementation of SDIs in the continent.

This analysis also revealed important intra-African differences with some African countries always in the head group for all the assessed variables: South Africa, Algeria and Botswana. In some cases, Egypt and Tunisia have also relatively good scores. The score of these five countries for some of the assessed variables can be visualized in the table ST3 along with a deeper analysis of the intra-African differences (note SN4).

A combination of all the assessed variables into one single value of SDI status per African country would be interesting for comparison. The types of assessment variables being different by nature makes it difficult to combine them together to give a unique value by country. However, one type of variables can be combined to obtain an indicative value by country, which can then be mapped to better visualize the SDI status in Africa. This has been done for the “objective variables” (or output variables) and is available both on the online table for the numbers (<http://africa-sdi.grid.unep.ch>) and in the figure 1 for the map. It shows the same trends as the ones discussed just above, with South Africa, Botswana and Namibia outstanding, while Algeria, Nigeria and Madagascar perform quite well also.

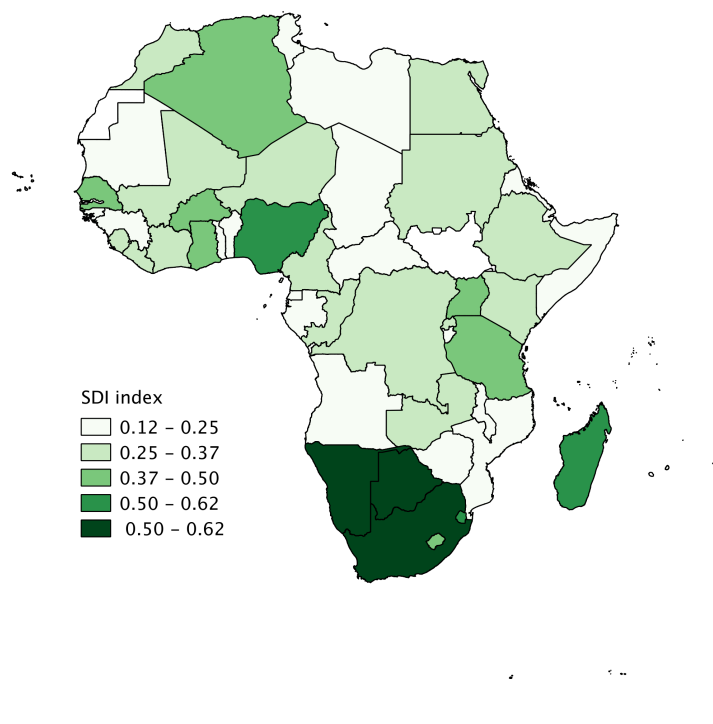


Figure 1: SDI index by country, based on "objective" or "output" variables
Note: the boundaries shown on this map do not imply official endorsement or acceptance by the authors

The methodology used for building the index is detailed in the supplementary note SN5.

Among the issues discussed so far, two main elements should be retained in terms of SDIs in Africa: (1) National SDIs need to be much improved in Africa and (2) there is a crucial need for better data to monitor the SDI status through assessment variables described in this article.

2.2.5.3 African situation to the rest of the World.

As shown in the results section and illustrated at http://africa-sdi.grid.unep.ch/SDI_variables_display_region.php, Africa scores less than other continents for most SDI-related variables measured, except the percentage of GDP spent in ICT and the official adoption of metadata standards. This necessitates a comparison with other regions where the SDI implementation status is much better, in order to find out if some elements would be transposable to Africa.

Europe is regularly mentioned as a successful example of a continental SDI implementation thanks to its federated approach legally bound through the INSPIRE directive. The most visible result of this enterprise is the European "INSPIRE geoportal" (European Commission, 2013a), gateway to European

member states services (data and metadata). Regular reviews and reporting of the SDI status like the INSPIRE State of Play allow maintaining and improving this mechanism, resulting in much SDI metainformation available. The strong European political commitment that led to the INSPIRE directive has been key. Getting such political signal in other parts of the World, in particular in Africa, requires political integration and common vision that are not consolidated enough yet.

Idrees (2012) compared approaches and strategies for NSDI implementations between the developed and developing World, and found out that political and ethnic polarity are visible key factors observed with the developing countries hindering speedy implementation of policies. He also states that successful and sustained SDI implementation will largely depend on the political will of the leaders. If this is already the case at national levels, it is even more so at a continental level. The African political body likely to initiate a continental political will in terms of SDI integration is the African Union. But even if this was on top of its agenda, it does not have the same powers as the European Union to pass a binding directive to its member states such as INSPIRE. This means that the European model cannot be directly transposed to Africa given the different political structure of the key institutions. Nevertheless, even if a strong political integration is lacking on the continent, alternatives exist to improve SDI and SDI monitoring with assets already available in Africa.

2.2.5.4 Suggestions for improving the SDI implementation and monitoring in Africa

First of all, in terms of standardized geospatial data availability, a system of systems approach instead of a federated approach would have the advantage of lowering the constraints on the side of the data producers. Indeed, in a system of systems approach, it is not necessary for the data producers to modify their standards in order to align with a mandatory standard as it is the case in the federated approach. This is the role of a mediator, a so-called broker, to perform the matching between the input standard available and the output standard needed. This brokering approach (Nativi and Bigagli, 2009a), (Nativi et al., 2012) and related GI-cat tool (ESSILab, 2014) have been notably adopted by the GEOSS and the Earthcube partnership (<http://earthcube.org>) as a backbone for their catalogue infrastructures. This approach and tool should ideally be adopted at a central African SDI institution if it is to become the continental reference system serving the African geospatial data in a harmonized way. The presence of UNECA to play such a role in Africa is a strength that needs to be used. This has been done in the framework of EU FP7 funded projects Afromaison (2011) and EOPOWER (2014). The GI-Cat has been implemented in Afromaison to broker existing African resources, while it has been customized and transmitted to UNECA in the frame of EOPOWER to become an African broker (Afromaison, 2013). This will also have the advantage to allow many organizations that informally contribute to SDI development even though they do not have the mandate (Makanga et al., 2010) to uptake their geospatial information.

In terms of SDI monitoring on the continent, UNECA is an asset as it is the recognized continental institution and already performs SDI monitoring surveys. These are currently weakened by the lack or imprecise answers of some countries. Two elements could be taken into consideration to improve the SDI monitoring survey from UNECA. The first one would be to set up an online database at the UNECA website, automatically populated by online forms submitted by National Mapping Agencies through an online survey form, which would allow an easier populating, management and sharing of basic SDI information. This could be started with at least a few variables and then extended to more variables later. The lack of motivation of stakeholders at national levels to fill such online surveys could be the main barrier, but overcoming this could be done through (1) an increased awareness of being an essential node of an integrated continental effort, (2) a kind competition with other countries to better position their own country in terms of SDI information provision, (3) the promises of the reduction in the number of dispersedly or redundant requests for information.

The second element to improve SDI monitoring in Africa is the example of “data flows” scoring for each member country of the European Environment Information and Observation Network (EIONET, 2012). On a simple webpage it is possible to see which country does better than the others in terms of environmental indicators reporting. When clicking on a given country, a graphical appreciation (with a given number of smileys) of progress is shown for each monitored indicator compared to the previous year. This allows to directly appreciating where to concentrate efforts and stimulate countries to improve significantly. A similar mechanism for SDI indicators and related assessment variables monitoring would be interesting to put in place at a continental level.

Such an assessment, as well as a list of national geoportals of the continent (Najar et al., 2007a) and other SDI relevant elements could be the components of a dedicated “SDI section” of the UNECA’s website, giving African SDI much more visibility and awareness to create a positive dynamism.

Another important asset of Africa is its participation in many projects and initiatives, past or on-going, that have a strong SDI or Earth Observation component, e.g. AEGOS (2011), Africover (FAO, 2013), EIS-Africa (2014), Geonetcast (GEO, 2013a), FEWS NET Africa Data Portal (USGS and USAID, 2013), SAFARI2000 (NASA, 2013), SERVIR-Africa (2014), SDI-Africa monthly newsletters (GSDI, 2012) and mailing list (GSDI, 2011). The recently launched AfriGEOSS initiative, developed in the GEO framework, aims to enhance Africa’s capacity for accessing, producing, using and managing Earth observations data and information. This will be achieved through implementing a coordination framework taking into account, national, regional and continental level. Under the coordination network periodic user needs and status of data access and products development will be undertaken at national and regional level, thereby informing the continental status. It is thereby expected that AfriGEOSS will add great value to the implementation and monitoring of SDI in Africa.

Moreover, the results of the search of African data in portals such as GEOSS, GBIF, and OneGeology showed that Africa is quite well connected to current SDI-

related initiatives, which is a positive push for complying to SDI international standards and therefore for improving African SDI initiatives. This African integration in international cooperation projects is an opportunity that allows the continent to remain in a dynamic of best practices in terms of SDI, as in the example of the African broker made possible by international projects. This should continue to be used for addressing the different issues in view of better implementation of SDI and SDI monitoring on the continent.

The concept of "open data" means that data should be freely available to everyone without restrictions (copyright, patents or control) (Wikipedia, 2015). This concept should also be further investigated to set up a better data flow and automatic mechanism of SDI statistics and key indicators. Indeed open data, not restricted by passwords, becomes more easily accessible by web services for direct parsing by machines without human intervention and hence give live access to data. This approach is for example used by the new UNEP monitoring system called "UNEP Live" (UNEP, 2014). A same live data flow for SDI statistics and key indicators would be a very valuable added value for SDI live monitoring.

Such open approach to data seems obvious but is unfortunately far from being the norm even though more and more open data initiatives exist. A very encouraging example of the fight against the lack of data is the "Open data for Africa" (African Development Bank, 2014) platform, financed by the African Development Bank (AfDB), that hosts open data for all 54 African countries since July 2013. Coupled with free and open source GIS technologies, this could really help better flow of data and less financial and institutional barriers.

In terms of capacity building, there is also a lot of work to do for GIS and SDI education. E-learning with initiatives such as UNIGIS international association (UNIGIS, 2014), the growing number of "Massive Open Online Courses" (MOOC) provided by universities worldwide, or material developed in various projects like "Bringing GEOSS services into practice" (Giuliani et al., 2014b) are solutions that could be more broadly used in the short to middle term to address the GIS education shortage existing in Africa.

2.2.6 Conclusions

2.2.6.1 Innovative way of performing SDI assessment

The first objective of this study was to examine an innovative way of performing an SDI assessment, at the continental scale. The methodology chosen and described addresses this objective since the assessment framework used combines a proven methodology for collecting information used at continental scale in Europe (INSPIRE State of Play) with SDI assessment variables validated by experts and usually used at national level. This methodology allows making use both of the Internet global search possibilities and a continental expertise (UNECA). Combined together, they provided necessary data for a continental SDI implementation measurement. The main limitation is due to the lack of data both on the Internet and at the SDI champion institution. This has to be solved at the political level, for example through the African Union, by a more binding directive on the model of the INSPIRE directive for the European Union. The

assessment variables proposed in this study are a first attempt for populating a suite of appropriate indicators and give a first overview of the African status of SDI implementation.

The second objective was to perform the evaluation of SDI status in Africa using assessment variables defined further to the methodology chosen. This revealed a weak status of SDI implementation in Africa, but suggestions and opportunities were also discussed to address this situation. This assessment also showed the difficulty in finding reliable data for measuring the assessment variables on the Internet or at the SDI key institution. This is true not only for Africa but also for other continents, making it necessary to have a better global mechanism for SDI implementation monitoring.

In order to improve the SDI implementation in Africa, the human components (policy, people) are key as they form the basis on which the technical components or physical infrastructure (data, access network), also essential to a successful SDI, can be efficiently used. We therefore recommend putting in place a proper political mechanism to support a continental SDI. UNECA already tries to do it through the CODIST meetings for several years but increased attendance and more commitment from African countries are needed. Besides, great opportunities lie in the cooperation among African and European institutions through for instance FP7 and H2020 framework projects.

The third objective was to compare the situation of Africa in terms of SDI status to the rest of the world. As demonstrated, a lot of efforts still need to be put in place to improve most of SDI components as well as a better SDI statistical information monitoring. This is true for Africa but also in other places of the World where regional SDI monitoring is strongly dependent on a regional integrated political vision. The success of the European SDI is mostly due to a strong political will translated through the INSPIRE directive. The discussion focused on the transposability of this European model to Africa and the conclusion is that this is not possible given the different nature of the key political institutions. Nevertheless, other solutions to improve SDI implementation in Africa have been proposed based on Africa's assets. It is important to reiterate such a survey again in a few years to see how the SDI implementation evolves in Africa compared to other continents.

The fourth objective of our study was to suggest some ways for improving the SDI implementation and monitoring in Africa. An important asset in Africa is the presence of UNECA that is already coordinating SDI activity on the continent as well as an initiative such as AfriGEOSS. UNECA and Africa in general would greatly benefit from a re-enforced mandate of UNECA in terms of continental SDI authoritative institution, as it is the key continental leader to implement SDI and SDI monitoring solutions. Other elements such as involvement of Africa in international projects with geographic components, open data growing political trend, and online capacity building are opportunities that are worth supporting and re-enforcing in Africa. All these elements are more dependent on international decisions (e.g. UN for UNECA, network of institutions for participation in international projects) than national or regional ones. Progresses

at the continental level should help overcoming the regional and national political barriers influencing SDI policy and status.

2.2.6.2 Final recommendations

Our final recommendations to improve SDI implementation in Africa are then to (1) obtain more political commitment to SDI from African governments, which might require more Earth Observation dissemination and promotion; (2) to reinforce the role of UNECA as the officially recognized continental SDI leader institution and give it the necessary resources to put in place a proper SDI implementation strategy in Africa; this could for example be done in the framework of the AfriGEOSS initiative; (3) to establish a proper African SDI online monitoring tool that would for example contain online SDI monitoring surveys, or a centralized visual mechanism for monitoring SDI status of African countries through online comparative maps, for example through a regional observatory that could be hosted at UNECA; this should create incentives for countries to perform better than the others; and (4) to establish a capacity building program on SDI at national and continental level.

2.2.7 Acknowledgments

This work was supported by the European Commission through the seventh research framework funded projects *AFROMAISON* (Grant agreement no 266379) and *EOPOWER* (Grant agreement no 603500). The views expressed in the paper are those of the authors and do not necessarily reflect the views of the institutions they belong to. The boundaries and names shown on the maps do not imply official endorsement or acceptance by the authors. We thank three anonymous reviewers for their helpful comments on an earlier version of this manuscript.

2.2.8 Supplementary notes

2.2.8.1 Supplementary Note S1: UN Economic Commission for Africa (ECA) historical role for geoinformation in Africa

In 2000, the UN Economic Commission for Africa (ECA) conducted a study on the “Future Orientation of Geoinformation in Africa.” The objective of the study was to “raise awareness of African governments and other sectors of society on the importance of geographic information in socio-economic development and to identify practical mechanisms to facilitate spatial data collection, access and use in the decision-making processes, both nationally and regionally, through a participatory approach.” The study recommended that all geoinformation

activities should be oriented towards developing SDIs, whose components utterly span all aspects of the production, management, dissemination and use of geoinformation. The study's report was received by the second session of the Committee on Development Information and became ECA's guiding principle for its advocacy work in the area of geoinformation.

As a result of ECA's and various partners' activities , awareness has been raised on the advantages of SDI to provide the spatial information needed for integrated economic and development planning. Prior to ECA's focus on SDI, several of the delegates to the geoinformation subcommittee of the first session of the Committee on Development Information (CODI, 1999) were opposed to the concept of SDI, and preferred to retain emphasis on mapping as a standalone activity. With the presentation of the report on the future orientation of geoinformation activities in Africa during the second session of the committee, only few delegates were still opposed to the shift of emphasis from mapping to SDI. By the third session of the committee (2003), the concept of SDI has been fully accepted and discussion shifted to the coordination of the various national SDI initiatives, as well as the need to create an African Regional SDI led by ECA.

2.2.8.2 Supplementary Note S2: Details on the methodology and sources of information for measuring each of the 14 key SDI indicators

1. **Availability of digital datasets:** in order to narrow down the definition of digital datasets, we decided to focus on fundamental datasets. Among existing definitions of fundamental datasets (UNECA, 2007b; Van Loenen and Kok, 2007), we decided to assess the availability of the digital format of the thirty fundamental datasets defined by UNECA at national level in Africa (see figure SF1). Most of these datasets have a defined scale for use at the national level and are ideally produced, maintained or coordinated by the country's National Mapping Agency (NMA). We therefore measured the availability of these datasets in NMAs through countries answer in UNECA's survey as the primary source of information.

Given countries' heterogeneous answers (e.g. "yes", or general description of the status) instead of each fundamental dataset's status, we attributed the value "yes" when the countries answered "yes" or when the explanation seemed to show they have the digital datasets available, and "no" otherwise. We have also considered as "no" the countries that did not answer the survey, assuming they would have positively answered if they had the requested information. The measurement of this variable is then strongly dependent on the quality of countries answers to the UNECA's survey and it is currently not possible to have a more appropriate level of detail (e.g. percentage of fundamental datasets available by country) than "fully available" or "not fully available" ("yes" or "no"). This methodology is also valid for the results of other UNECA's questions that are used for other indicators assessment.

2. **Capacity Building:** Capacity building is about developing the necessary skills in SDI and SDI-related fields such as GIS or remote sensing. An efficient mean to build capacity is the teaching at secondary and tertiary levels. Coetzee and Eksteen (2012) have performed an internet survey of "GIS education in tertiary institutions in Africa" in which they measure the number of tertiary institutions offering GIS teaching by African country. This forms an excellent variable to demonstrate countries commitment to national GIS/SDI capacity building, even though such data is subject to change rapidly (e.g. the case of Ethiopia mentioned in Gemedu (2012)). This allows to calculate the percentage of African countries teaching GIS and eventually compare it with the rest of the world. It should be noted that GIS education provides a skill set that does not necessarily mean it will only be used in the SDI field. However, we consider that an SDI expert has a GIS background, hence the importance of the GIS teaching variable.
3. **Willingness to share:** This key indicator is not easy to assess directly but remains essential to understand the government data sharing policy (Najar et al., 2007a). We retained the following proxy variables : (1) the existence of a national geospatial data sharing policy, (2) the existence of a national geoportal, and (3) the membership to international data sharing initiatives such as the « Group on Earth Observations » (GEO, 2014c). Despite no data

found for the first proxy variable, information on the other two proxy variables can help to quantify the willingness to share data, even though they also depend on the country's infrastructure or the availability of data.

4. **Human Capital:** As data on number of SDI experts by country or continent do not exist yet, proxy indicators are necessary to estimate the level of SDI skill. Firstly, we consider the "ICT skills" index (ITU, 2011) developed by the International Communication Union (ITU), combining adult literacy rate, gross secondary school enrolment and gross tertiary school enrolment as a proxy indicator given the importance of ICT skills for SDI development (e.g. for developing geoportals, processing geospatial data). Secondly, we consider the presence of GIS companies in a country as an indicator of GIS use. Environmental Sciences Research Institute (ESRI) being the commercial leader in GIS worldwide, we consider its number of offices (ESRI, 2012) as a second variable for the estimation of the SDI human capital in a country. Keeping in mind the growing importance of open source products and the leading role of the Open Source Geospatial Foundation (OSGeo) in open source geospatial software collaborative development (OSGeo, 2015), the number of OSGeo products users or representations in Africa would be a very interesting variable to assess the number of SDI experts. Unfortunately it has not been possible to find such recent and reliable information. Consequently, we stick to the variables "ICT skills index" and "ESRI offices number" to obtain a good approximation of Africa's situation in terms of SDI Human Capital.
5. **SDI Awareness:** Measuring a country's SDI awareness is necessary because it is correlated with the importance given to invest (Annoni et al., 2002). Quantifying SDI awareness in a population of potential end-users, especially at a continental scale, is not feasible without a proper large-scale survey that is currently not available, hence the need of proxy variables.. SDI being intrinsically linked to the use of Internet, the first variable is the status of Internet in the country because if there is no or limited Internet there cannot be an SDI (e.g. access to satellite images is compromised). Hence, we used the "percentage of Internet users" variable from ITU statistics (ITU, 2014). Workshops being the ideal mean to raise awareness, measuring the number of SDI workshops taking place in a given country or attended by a given country specialists over a year would show a country's interest and awareness of SDI. Tracking all SDI workshops and attendees in Africa over a year was not possible, and we chose instead to monitor the countries attending CODIST meetings occurring every two years at UNECA, for years 2007, 2009 and 2011.
6. **Access mechanism:** According to Najar et al. (2007a), a good access mechanism indicator is the inclusion of web services in current SDI initiatives such as a national geoportal, or a national participation to a regional geoportal, that generally publishes web services. But the possibility for a country to provide a good access mechanism is also dependent on the national ICT infrastructure. To this end, we used ITU's

« Digital Access Index » that is a compilation of indicators that describe a country's information infrastructure (Kozma and Wagner, 2005).

7. **Funding:** Assessing the “funding” indicator would ideally require assessing the funding of each SDI component. But given the complexity of this measurement, we focused on the most tangible part that consists in the technical part of the SDI, the Information and Communication Technologies (ICT). Although the percentage of GDP spent in SDI would be the ideal indicator, this information does not exist and we use instead the percentage of GDP spent in ICT. We also checked whether a national policy on SDI funding exists.
8. **Vision:** At national level, a National SDI (NSDI) is the necessary tool for bringing geospatial information management into practice. But this only works if there is a national vision to gather all necessary elements and set up the framework. Hence, we define here that a country has a SDI vision if SDI is found in the national legislation in relation to some concept of a NSDI.
9. **Institutional arrangements:** This indicator reveals the level of a country's institutional SDI integration. The integration of a national SDI strategy into other national strategies is indeed crucial to have an efficient NSDI. For example, when a country is collecting health data, it should also collect simultaneously geographic information (e.g. geographic coordinates) linked to this data. We used here UNECA surveyed information on the relations of SDI to some other National Policies.
10. **Leadership :** In order for a country's SDI agenda to progress, it needs good coordination by a national leader institution or group of institutions that can be referred to as SDI coordinating bodies. In theory, a country's NMA should play this role but in practice other institutions such as statistics offices can play a more important role. In any case, the existence of a formal SDI coordinating body at national level is the variable to be surveyed here.
11. **Socio-Political stability:** According to the SDI Cookbook (Nebert, 2008), the development of a SDI will rely on socio-political stability and legal context. We chose the year 2013 values of the World Bank's time series index “Political Stability and Absence of Violence/Terrorism” (World Bank, 2014) that reflects the perceptions of the likelihood that a government is destabilized or overthrown by unconstitutional or violent means for assessing the socio-political stability indicator, as it is important to take into account the latest political developments in the World, like the Arab springs that concern Africa.
12. **Metadata availability:** In order to optimize data use/production and avoid duplication, end users need to be able to know what data exist and their associated metadata. The establishment of web based metadata services is a sign of SDI success (World Bank, 2011). Data/metadata catalogs serve this purpose by informing in details about data they contain. The metadata (availability) key indicator encompasses three fundamental

questions in our view: (1) what proportion of the existing fundamental datasets have metadata?; (2) does the metadata follow a standard?; (3) is this metadata available somewhere on the Internet, ideally in a geoportal?

Being able to assess the percentage of fundamental datasets with metadata would be ideal but such information could not be found. Instead, we chose to find out whether a country: (1) has adopted a particular metadata standard for geospatial data, (2) has a national working group on metadata, and (3) has a metadata and clearinghouse gateway.

- 13. Interoperability :** This indicator aims at measuring the ability to share data, software and hardware across organizations and end users (Eelderink, 2006). It implies that standards have been chosen for interoperability of datasets and/or services. As a first variable to assess we chose to gather information on the existence of a Working Group on standards. Secondly, it exists an African Geodetic Reference Frame (AFREF) (RCMRD, 2014). To be compatible with the AFREF, an African country must have as many AFREF ground stations as possible, which is the second assessment variable. Finally, countries' membership of the « Open Geospatial Consortium » (OGC, 2013b), reference organism in terms of geospatial data standards, is the assessment variable to determine adherence to geospatial data standards.

Initiatives connected to SDI: this last indicator assesses countries SDI activities. For Africa, a few international SDI initiatives have been considered: (1) GEO Membership; then, African data available in the (2) Global Earth Observation System of Systems (GEOSS) geoportal (GEO, 2012); (3) Eye on Earth geoportal (EyeonEarth, 2012); (4) One Geology geoportal (OneGeology, 2012); (5) Global Biodiversity Information Facility (GBIF, 2012).

2.2.8.3 Supplementary Note S3: Details on the results for each of the assessment variables measured

2.2.8.3.1 Data

Only 11% of African countries (Algeria, Botswana, Lesotho, Madagascar, Namibia and South Africa) are reported by UNECA as having the complete set of fundamental digital data through countries' NMA. This makes the

“Data” SDI component very weak in Africa compared to the 87%-100% reported for Europe in the 2010 INSPIRE & NSDI State of Play. Some of these fundamental datasets such as land cover or geology can be freely obtained from the Internet for most countries. But these are only a small part of the fundamental datasets as defined by UNECA. For countries without complete availability of fundamental datasets, it was not possible to get more detailed information about the missing datasets.

2.2.8.3.2 People

2.2.8.3.2.1 Capacity Building

Coetzee and Eksteen (2012) showed that GIS education in tertiary institutions is present in 26 out of the 54 African countries (see figure SF3). But among these 26 countries, the absolute numbers range from one single institution in Tanzania up to 18 institutions in South Africa, with most countries having less than 5 tertiary institutions providing GIS teaching. Only 14% of the surveyed tertiary education institutions offer GIS education. With only one GIS-teaching institution for 10 million inhabitants in Africa, the only solution for most African wishing to study GIS is to go abroad or to take online courses. As of January 2013, only 3.6% of participants to Massive Online Open Courses (MOOCs) were located in Africa, which is behind all the continents except Oceania (North America: 35.2%; Europe: 28.2%; Asia: 21.4%; South America: 8.8%; Africa: 3.6%; Oceania: 2.8%) (Universities UK, 2013). This information concerns MOOCs in general, not the ones specialized in GIS, which could not be found.

This clearly shows a huge gap in terms of GIS and SDI capacity building compared to the ideal situation where most countries of a continent should at least have a few institutions (varying proportionally to a country's size) teaching GIS for building local capacity. One should also keep in mind that ICT use in schools is very much dependent on the national ICT infrastructure (Kozma et al., 2005) and Africa is the continent with the lowest value in the Digital Access Index measuring it.

2.2.8.3.2.2 Willingness to share

Information about existing data sharing policies in African countries could not be found, which needs to be addressed institutionally. The availability of a national geoportal is considered as a variable showing the willingness to share. Unlike all other continents, we were not able to find a single African national geoportal. finally, membership of an international SDI initiative can also indicate a willingness to share data; 41% of the African countries are members of the GEO, which is comparable to Asia and the Americas.

2.2.8.3.2.3 Human Capital

ITU's mean ICT skills index is the lowest in Africa (4.05), half the value of Europe. Tunisia has the highest index value in Africa with 6.94, but this is nearly equivalent to the lowest value for Europe (6.79). The lowest measured value of Africa (Niger with 1.44) is also the lowest in the World for countries with available data.

ESRI offices representations are found in only 11% of African countries, whereas this figure is larger than 50% (Americas: 57%, Asia: 61%; Europe: 58%) for all continents except Oceania (15%).

All the variables assessed for the "Human Capital" key indicator show that Africa is much behind the other continents.

2.2.8.3.2.4 SDI Awareness

The first variable defined by the authors to assess this indicator is the number of Internet users per 100 inhabitants (year 2012). Its percentage value is 13.52% for Africa, the lowest value compared to all other continents that range between 32.04% for Oceania, 35.68% for Asia, 45.05% for the Americas and nearly 70% for Europe. This makes the number of Internet users in Africa only a third of the global average. This is then a major issue in terms of SDI use and hence SDI awareness. Even if all the major SDI actors and users were already counted among the percentage of internet users in Africa, the low percentage of internet connected people still remains a barrier to national and continental SDI awareness raising as the general public or some decision-makers need to understand the advantages of SDI, which is done through the internet.

Problems related to Internet access through fixed lines in Africa are known but mobile solutions are considered more adapted both from an economic and cultural perspective (Gallagher, 2012). This could mean that smart phones are becoming real opportunities of development. But according to the latest ITU's statistics (ITU, 2014), Africa is also in last position for "Active mobile-broadband subscriptions" with 10.9 subscriptions per 100 inhabitant against 67.5 in Europe, 48 in the Americas and 20.6 in Asia. It will however be interesting to follow this indicator in the longer term as it is crucial that Africa addresses now the SDI

challenges as geospatial data issues are imbedded in the development of the smart phone industry.

Regarding the African countries attendance to the CODIST workshop (2007, 2009, 2011), around 50% of the African countries were represented (2007: 56%, 2009: 52%, 2011: 52%), which is quite a low number that does not seem to increase over the years.

With the last four indicators making up the assessment of the “people” component of SDI (capacity building, willingness to share, human capital, SDI awareness), one can see that this component is the lowest for all the indicators in Africa.

2.2.8.3.3 Access network

Access mechanism

An Internet search allowed us to find national geoportals or regional geoportals with national participation for some countries in the World, mainly in Europe and North America. In Europe, the INSPIRE geoportal (European Commission, 2013a) provides data for 20 European countries. Except Africa, every continent has at least one country with a geoportal. But no national or regional running geoportal could be found for any African country at the time of this study, even though UNECA is planning to create an “SDI-Africa” portal in the future. Except in the GSDD’s links webpage (GSDD, 2013) where some geoportals are listed by geographic scale (global, regional, national, local) but without date of last update of the listing, we could not find any other website listing national geoportals.

Regarding the Digital Access Index, Africa has the lowest mean value (0.2) of all the continents, just under Oceania and much lower than Europe (0.55) or the Americas (0.48). The highest value for Africa (Seychelles with 0.54) is even below the mean value of Europe (0.55). This clearly shows that Africa is in crucial need of improving its information infrastructure and ICT services in order to properly implement SDI.

2.2.8.3.4 Policy

2.2.8.3.4.1 Funding

The percentage of ICT expenditure relative to countries GDP is a piece of information that cannot easily be found. The most recent data found is for 2008, and only for some countries in the World. On the African continent, only data for some of the most IT-developed African countries (Algeria, Cameroon, Egypt, Kenya, Morocco, Nigeria, Senegal, South Africa and Tunisia) are available (Econstats, 2013), with an average of 6,7%. This number is high compared to the 6% average worldwide, but it is most probably not representative of the other African countries for which no data are available.

Policy on SDI funding could not be found for any of the African countries except South Africa that has a SDI Act, 2003. Hence, no general trend for this policy indicator can be derived. There is a serious need to be able to better track funding for SDI and even for ICT that is a crucial component of SDI.

2.2.8.3.4.2 Vision

UNECA provided the list of African countries that (1) have NSDI formally established (which means the NSDI concept has been accepted in the country but not necessarily put in practice yet) and (2) have working groups for its implementation. Eleven African countries out of 54 have an NSDI formally established and ten countries have SDI working groups established. It should be noted that some countries have SDI working groups established whereas a NSDI does not formally exist. Also, some countries have a NSDI formally established but without SDI working groups. Altogether about 20% of African countries have started a NSDI program, consequently to a national SDI vision. This figure is quite low compared to the European Union where all countries must comply to the INSPIRE directive.

2.2.8.3.4.3 Institutional arrangements

SDI integration in some national policies is a good indicator of institutional integration status for SDI. UNECA has surveyed the integration of SDI with National Information and Communication Infrastructure (NICI) and National Strategy for the Development of Statistics (NSDS). According to UNECA's figures, 24% of African countries, mainly the ones with a NSDI or SDI Working groups established, have such institutional arrangements and hence tend to SDI integration across various institutions. This number is of course a positive sign of SDI strategies emerging in African countries, but is still low at continental scale.

2.2.8.3.4.4 Leadership

According to the UNECA figures, complemented by information found in the Internet, 28% of the African countries have a SDI coordinating body, which is a good indicator of leadership. Again these countries are mainly the ones with a NSDI or SDI Working groups established. There are exceptions like Madagascar, Sudan, Tanzania or Zambia where SDI coordinating bodies exist but no official working groups or NSDI are formally established. Explanation for these particular cases needs further investigation from UNECA.

2.2.8.3.4.5 Socio-political stability

The World Bank political stability index ranges from approximately -2.5 (weak) to 2.5 (strong) in the World. In 2013 it is the lowest in Africa compared to the other continents, with a mean value of -0.58. By comparison, Europe's mean value is 0.56 and countries with a good SDI reputation have even higher values: Canada: 1.03; Australia: 1.02; Norway: 1.33. The online figure http://africa-sdi.grid.unep.ch/SDI_documents.html#Political_stability shows the details of the index for each country; it is also consistent with the index from the Economist (The Economist, 2009) that shows Africa less socio-politically stable than other continents.

2.2.8.3.5 Standards

2.2.8.3.5.1 Metadata (availability)

UNECA has contributed to the development of an African profile of the ISO standard for metadata. According to UNECA sources as well as Internet searches, only four African countries have adopted an official metadata standard: Botswana, Ethiopia, Nigeria and South Africa.

Regarding the existence of a metadata and clearing house gateways, UNECA records only three countries with such a gateway: Botswana, South Africa and Uganda²⁹. If these numbers remain valid, it means that the metadata from only 6% of African countries is discoverable on the Internet. An alternative would be to find national metadata in regional geoportals; but as already mentioned no such regional portal has been found in Africa.

According to UNECA, 9% of African countries (Botswana, Ethiopia, Mali, Senegal and Swaziland) have a metadata working group in place. However, one should be cautious about this number as it was gathered from a UNECA survey for which the responders had to indicate the names of their SDI sub-working groups. Only those groups with the proper quotation of “metadata working group” were counted, which makes this number not necessarily representative.

We conclude that information about metadata availability for fundamental datasets in Africa is very scarce and no general trend can hence be drawn from it.

2.2.8.3.5.2 Interoperability

For information about the existence of a national Working Group on standards, the only available information comes from UNECA’s survey that gathered the names of countries’ SDI sub-working groups. Only three African countries report the existence of a working group on standards.

However, the number of countries that are compatible with the African geodetic Reference Framework (which means countries that have at least one Global Position reference point) is monitored by UNECA and equals to twenty-five African countries (46% of the African countries) as shown in figure SF5 and in the online table.

According to the Open Geospatial Consortium (OGC) website (OGC, 2013a), no governmental, national, sub-national or local African institution is member of OGC (as of February 2013). In comparison, all the other regions of the World have at least one of their institutions represented in the OGC.

To summarize, the indicators for interoperability show that Africa needs to make strong progress as (1) it is not represented at all in the main geospatial data standardization consortium, (2) there is very few indication about work being done in African countries for standardization, and (3) too few African countries have a national geodetic reference compatible with the rest of the continent.

2.2.8.3.6 Other

2.2.8.3.6.1 Initiatives connected to SDI

One of the main current worldwide SDI initiatives is the Global Earth Observation System of Systems (GEOSS), led by the Group on Earth Observations (GEO). Europe is well represented in GEO with membership of 67% of its countries. Africa is in the average with 41% of its countries represented, equivalent to Americas and Asia.

²⁹ The corresponding websites were down when the authors tried to verify the information

In terms of availability of African data in the GEOSS portal, the number of occurrences when typing the word “Africa” in July 2012 comes at the 3rd rank compared to the other continents with 6,274 occurrences against: about 15,000 for Europe, 12,000 for the Americas, 4,200 for Asia and around 100 for Oceania.

When doing the same exercise 7,5 months later (first round: 9 Jul. 2012, 2nd round: 25 Feb. 2013), the number of occurrences for each continent had increased considerably: between about four times (Europe) to 94 times (Oceania). During the same period, Africa has lost one rank and came to the 4th place. However, African data remains much visible and active, as it has multiplied its entries by a factor of 6,4 during this period.

The ranking is the same for the Eye on Earth portal as of July 2012 with around 500 occurrences for Africa against 700 for the Americas, 500 for Europe, 200 for Asia and only 5 for Oceania. This experience could not be repeated in February 2013 as the interface has changed and the website has later been abandoned.

The same ranking is also observed on the “OneGeology” platform whereas Africa does even better on the “Global Biodiversity Information Facility” (GBIF) portal with the largest number of occurrences in datasets compared to the other continents.

Of interest to note is that most of these datasets are global in nature, they are not necessarily developed at continent level and not owned by the African institutions. National and regional datasets available in these portals is very limited.

2.2.8.4 Supplementary Note S4: Details on the intra-African differences

This analysis also revealed important intra-African differences with some African countries always in the head group for all the assessed variables: South Africa, Algeria and Botswana. In some cases Egypt and Tunisia have also relatively good scores. The score of these five countries for some of the assessed variables can be visualized in a the supplementary table ST3.

Not surprisingly, South Africa appears to be the leading country in terms of SDI development and implementation. This country is always in the top five countries for all variables, except for the socio-political stability index. Botswana does also well in some of the variables.

In northern Africa, Algeria is the regional leader, followed in some cases by Egypt and Tunisia. But the Egyptian and Tunisian good scoring seems due more to the technological aspects (ICT) than to proper SDI issues as can be seen with the SDI index by country described in supplemental note SN5 as well as the related Figure 1 in the main article, whereas Algeria gets a more constant good scoring for most of the SDI aspects. Tunisia and Egypt should then not be considered as part of the SDI head group. This might suggest that the lack of investments in Africa's ICT infrastructure is a real threat that might deepen the gap between African countries regarding access to data and information, necessary for sustainable development.

2.2.8.5 Supplementary Note S5: Details on the methodology used for building the SDI index by country

Most of the “objective variables” used for the index are binary (yes or no) and have hence been attributed 1 if yes, and 0 otherwise. Three of these objective variables were in absolute numbers (number of ESRI offices, number of institutions offering GIS teaching, Number of AFREF stations); they have been attributed 1 if the absolute number is at least 1, and 0 otherwise. This way, all the variables are on a common scale between 0 and 1.

About half of these objective variables result from the UNECA survey (see figure SF1) (variables: core fundamental datasets, NSDI formally established, SDI Working Groups established, Relation of SDI to other national policies, SDI coordinating body existing, National Working Group on metadata existing, Metadata & Clearinghouse existing, National Working Group on standards existing). As explained in the methodological part of this paper (section 2.5, under the “availability of digital datasets” point), we attributed the value “yes” when a country answered “yes” or when the explanation seemed to show they have the digital datasets available, and “no” otherwise. Besides, 28 countries have not replied to the survey, which might have an impact on the reality of the results. But we assume that if a country would fulfill one or several of the variables it would advertise it by replying to the survey; this means that we assume that no answer can be considered as a no and we therefore stick to a “yes” or “no” schema.

As the number of assessment variables is unequal between the different SDI pillars, we cannot simply sum up the value of every variable by country and divide it by the number of assessment variables to obtain a value by country. This would mean that each variable has a similar weight and would not take into account the number of variables by SDI pillar. For example, the weight of the SDI “data” pillar (with only one assessment variable) would be very weak compared to the seven objective variables of the “people” pillar. We have instead made an average of the assessment variables by SDI pillars. Then we have made an average of these SDI pillars averages to obtain the final value, considering that each pillar has the same weight and each assessment variable inside a same pillar has also the same weight as this map is simply indicative, to obtain a trend. Obtaining a weight for each variable and pillar would require a SDI experts committee agreement.

We finally represent these values by country on a map with 5 equal intervals categories.

2.2.9 Supplementary Tables

2.2.9.1 *Supplementary Table S1: Key variables for the assessment of NSDIs in developing countries*

SDI components	Key indicators
Data	1. Availability of digital datasets
People	2. Capacity building
	3. Willingness to share
	4. Human capital
	5. SDI awareness
Access network	6. Access mechanism
Policy	7. Funding
	8. Vision
	9. Institutional arrangements
	10. Leadership
	11. Socio-political stability
Standards	12. Metadata (availability)
	13. Interoperability
Other	14. Initiatives connected to SDI

Adapted from Eelderink et al., (2008a)

2.2.9.2 Supplementary Table S2: Assessment variables chosen for key indicators

SDI components	Key indicators	Assessment variables	Type (I = input O = Output)	Sources used
Data	Availability of digital datasets	Core Fundamental Datasets validated and available at the country NMA	O	UNECA internal data
People	Willingness to share	Existence of a data sharing policy	I	Not found
		Availability of a national geoportal	O	Various internet sources
		GEO membership	O	http://www.earthobservations.org/ag_members.shtml
	Human capital	ICT skills 2010	I	http://www.itu.int/ITU-ICT/publications/idi/material/2011/MIS_2011_without_annex_5.pdf
		ESRI Offices	O	http://www.esri.com/about-esri/offices
	Capacity building	Tertiary institutions offering GIS education	O	Data used for the article (Coetzee et al., 2012)
	SDI awareness	Internet users per 100 inhab	I	http://www.itu.int/ITU-ICT/dai/index.htm
		Mobile broadband users per 100 inhabitants	I	http://www.itu.int/en/ITU-Statistics/Pages/stat/default.aspx
		Countries participating to the CODIST workshops 2007/2009/2011	O	UNECA internal data
Access network	Access mechanism	Availability of a national geoportal	O	Various internet sources +

				http://inspire-geoportal.ec.europa.eu/
		Participation to a regional geoportal	0	Various internet sources
		Digital Access Index	I	http://www.itu.int/ITU-D/ict/dai/index.html
Policy	Funding	Existence of a SDI funding policy	I	Not found
		% of GDP spent in ICT, 2008	I	http://www.oafrica.com/statistics/world-databank-ict-expenditure-gdp/
	Vision	NSDI formally established	0	UNECA internal data
		SDI working groups established	0	UNECA internal data
	Institutional arrangements	Relations of SDI to other National Policies (NICI, NSDS, etc)	0	UNECA internal data
	Leadership	Countries with SDI coordinating bodies	0	UNECA internal data
	Socio-political stability	Political Stability Index	I	http://info.worldbank.org/governance/wgi/index.aspx#home
Standards	Metadata (availability)	Official Metadata Standard Adopted	0	Internet sources
		Existence of national WG on metadata	0	UNECA internal data
		Metadata & Clearinghouse gateway existing (=discovery)	0	UNECA internal data + internet sources
	Interoperability	Existence of national WG on standards	0	UNECA internal data + internet sources
		compatibility with the African geodetic reference	0	UNECA internal data
		OGC membership by country	0	http://www.opengeospatial.org/ogc/me

				members/report?sortBy=%27gov%27#Government-National
Other	Initiatives connected to SDI (country's activity)	GEO Membership	0	http://www.earthobservations.org/ag_members.shtml
		GEOSS portal search	-	http://www.geoportal.org
		Eye on Earth portal search	-	http://network.eyearth.org/home
		One Geology portal search	-	http://portal.onegeology.org
		GBIF portal search	-	http://data.gbif.org

2.2.9.3 Supplementary Table S3: Intra-African scoring of selected countries for some measurable variables

Scoring/Ranking among African countries	South Africa	Algeria	Botswana	Egypt	Tunisia
Core fundamental datasets validated	yes	yes	yes	no	no
GEO membership	yes	yes	no	yes	yes
ICT skills index	3rd	5th	7th	6th	1st
Institutions giving GIS teaching (Capacity Building)	18	11	0	6	0
Number of ESRI offices	1	0	0	1	1
Number of internet users/100 inhab.	4th	17th	9th	10th	5th
Digital Access index	3rd	10th	4th	7th	6th
NSDI established and/or SDI WG existing	yes	yes	yes	no	no
Integration of NSDI with NICI or NSDS	no	no	no	no	no
Existing SDI coordinating body	yes	no	yes	no	no
Socio-political stability ranking	34th	24th	6th	12th	5th
Compatibility with AFREF	yes	yes	yes	no	no

2.2.10 Supplementary Figures

2.2.10.1 Supplementary Figure S1: fundamental datasets for Africa

Table 7: Fundamental data sets for Africa

Level	Category	Data Theme	Data Set	Definition
0	Primary Reference	Geodetic Control Network	Geodetic control points	List of coordinates with information on the history of establishment of the network as well as network design in digital map/GIS format.
			Height datum	List of heights of primary height points in digital map/GIS form (vertical datum surface)
			Geoid model	Geoid-ellipsoid separations (heights at individual points) to convert from GPS observations to heights
I	Base geography	Rectified Imagery	Aerial photography	Aerial photography
			Satellite imagery	Satellite imagery
		Hypsography	Digital elevation model	Vertical distance from the earth's surface to a base defined by the adopted height datum
			Spot heights	Heights of peaks
		Hydrography	Bathymetry	Vertical distance of earth's surface from base defined by Lowest Astronomical Tide
			Coastline	The limit of land features usually at mean high water level.
II	Administration and spatial organisation	Boundaries	Natural water bodies	Location of watercourses, drainage network, and all inland water bodies (streams, rivers, canals, ponds, lakes, etc.)
			Governmental units	Limits of administrative and jurisdictional authority (International, national, sub-national boundaries, and local government areas)
			Populated places	Population centres including urban areas, towns, localities, and rural settlements
		Enumeration areas	Enumeration areas	Boundaries of areas delineated for the purpose of collecting demographic census information
			Place Names	Official and local names of places
		Geographic names	Feature Names	Official and local names of cultural and geographic features (including roads)
			Land Parcels/Cadastral	A consistent framework of land parcel/cadastral boundaries defined for land tenure purposes, referenced to a common datum
		Land Tenure	Land Tenure	Current, proposed and historical details of all tenures, e.g., details of ownership, vesting, and including traditional forms of land holding.
			Street Address	Unique Street Address of parcels/properties
			Postal or zip code zones	Boundaries of post code areas
III	Infrastructure	Transportation	Land use planning zones	Boundaries of areas of permitted/restricted land use defined by planning authorities (includes conservation areas, heritage sites, and restricted areas)
			Roads	Network of physical roads and carriageways
			Road centrelines	Centrelines of roads and carriageways
			Railways	Network of railway lines
			Airports and ports	Location of airports, sea ports, and navigation aids
		Structures	[Bridges and tunnels]	
			Power	Locations of trunk or national grid power line networks and major assets/installations, and sources
		Utilities and services	Telecommunications	Locations of trunk communication networks and major assets
			Land cover	Observed bio-physical cover over on the earth's surface ¹
			Soils	Boundaries and classifications of soil resources
III	Environmental Information	Natural environment	Geology	Boundaries and classification of geological units

1. FAO Land Cover Classification System — Classification concepts and user manual (Software version 2, Draft Version, Nov. 2004), p.7

2.2.10.2 Supplementary Figure S2: UNECA's questionnaire

CODIST.2 – Questionnaire for Country Reports

This questionnaire is designed to allow a fast and easy filling. The United Nations Economic Commission for Africa will be happy to receive any type of information that you would consider useful for the Second session of the Committee on Development Information, Science and Technology (CODIST.2).

Please return the questionnaire filled out to the following address:

Andre Nonguierma

United Nations Economic Commission for Africa (UN-ECA)

PO Box 3005 Addis Ababa (Ethiopia)

Fax: (251) - (251) - 115.510.512

E-mail: ANonguierma@uneca.org or EcaGeoinfo@uneca.org

Name of Country: _____

Details of person in your country to be contacted for further information or clarification, or to communicate relevant information:

	Primary Contact	Alternate Contact
Name:	_____	_____
Dept:	_____	_____
Phone No:	_____	_____
Fax No:	_____	_____
Email:	_____	_____

For each question, please select the option that is closest to the situation in your country

Status of the National Spatial Data Infrastructure (NSDI)

- ☐ The government has formally established of a national (geo) spatial data infrastructure/framework/policy.
When the NSDI/NGDI was formally established? _____
- ☐ The government has endorsed the concept but it has not been formally established.
When was the NSDI/NGDI concept endorsed? _____
- ☐ National stakeholders have endorsed the concept; government endorsement is now awaited.
When was the NSDI/NGDI discussed and endorsed by stakeholders? _____
- ☐ The concept is currently under discussion among stakeholders
- ☐ No immediate plan to implement SDI because the concept is not yet understood by stakeholders.

Status of the implementation of the NSDI

- ☐ Name of SDI Coordinating Body
Please specify when it became operational and formal citation _____
- ☐ Names of Working Groups (WG) or Sub-Committees of the SDI Committee
Please list the names of WG or Sub-Committees and their responsibilities _____
- ☐ Name of the Authority responsible for Geographical Names (if any)
Please specify when it became operational and formal citation _____
- ☐ Dates of Meetings of the SDI Committee and Working Groups since CODIST.1 (2009).
National, Local _____
- ☐ Progress in SDI implementation since CODIST.1 (2009).

Activities to date, Intermediate results, prospects, etc. _____

Status of National Information and Communication Infrastructure (NICI) process:

- ☐ Policy document has been approved and implemented.
When did it become operational (specify month and year)? _____
- ☐ Draft Policy document is in consultation phase
When is the consultation expected to be completed? _____
- ☐ National consultative workshop has been conducted and draft document is in preparation.
When will the draft expected to be completed? _____
- ☐ Not aware of status of NICI
If you choose this response, please skip next question _____
- ☐ Relations of NSDI to other National Policies (e-Government, National Statistic Development Strategies, etc...) process.
Standardisation, Interoperability, Data sharing, etc. _____

Status of core or fundamental datasets:

- ☐ Core or fundamental datasets have been agreed on and are now available in digital form
List _____
- ☐ Core or fundamental datasets have been agreed on and digitisation is in progress
Status of the digitization _____
- ☐ Core or fundamental datasets have been agreed on and are provided in analogue form
List _____
- ☐ Discussion is continuing on what to include in core or fundamental datasets
Status of the consensus reached so far: _____
- ☐ The issue has not yet been addressed
Any comments? _____
- ☐ Map Revisions being undertaken
Maps types, Progress _____
- ☐ New Mapping initiatives / Projects
List on-going project and initiatives _____

Status of metadata and clearinghouse:

- ☐ A national system exists for maintaining metadata of spatial data resources:

- ☐ The system is accessible through a clearinghouse gateway

URL: _____

- ☐ The system is web-based, but not accessible through a standard clearinghouse gateway

URL: _____

- ☐ The system is online and accessed through special client software

- ☐ The metadata are distributed on CDs

- ☐ The metadata are distributed in printed form

- ☐ The metadata are distributed in some other form

Please specify _____

- ☐ A national metadata system is being created but it is not yet accessible to users

Please specify when work on the system started: _____

Please further specify when it is expected to be available: _____

- ☐ A national metadata system has been agreed on, but work has not started to create it

- ☐ Discussion is going on about creating a national metadata system

- ☐ No immediate plans for a national metadata system

Global Navigation Satellite Systems (GNSS) Reference Stations:

Please provide Number of Reference Stations in the country: _____

Please indicate how many continuous operating reference stations : _____

Please attach sheet with the following information for each GNSS Reference Station:

- ☐ Site Name: _____
- ☐ City/Town: _____
- ☐ Country: _____
- ☐ Date Installed: _____
- ☐ Latitude (Northing): _____
- ☐ Longitude (Easting): _____
- ☐ Elevation (m, ellips.): _____
- ☐ Receiver Type: _____
- ☐ Antenna Type: _____
- ☐ Brief Description of Monument: _____

Human Resources and Capacity Building:

Please provide number of staff with the listed specialisation and qualification below:

- ☐ Staffing Complement
Photogrammetry: _____
Geodesy: _____
Cadastral: _____
GIS: _____
RS: _____
Databases: _____
- ☐ Number of:
Short Trainings _____
Workshops _____

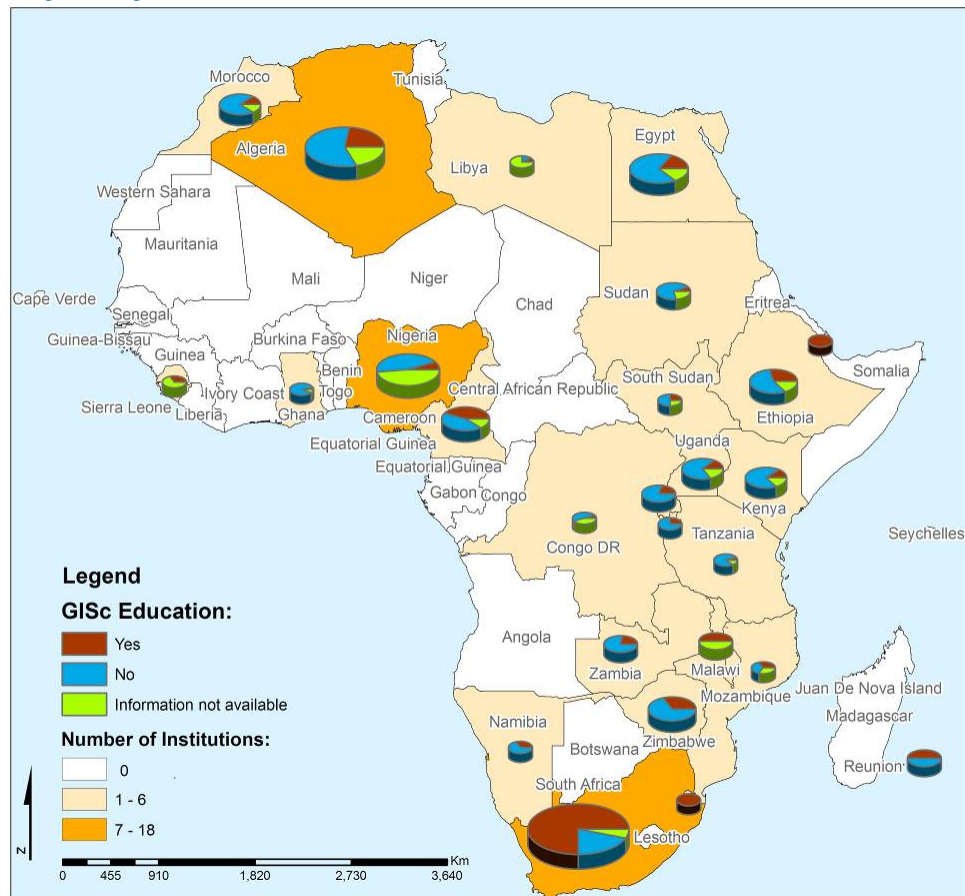
Please outline any major activities or events related to the acquisition and management of geoinformation resources that have occurred in the last five years (e.g., mapping campaigns, map digitisation, data harmonisation, coordination office or committee, capacity building, database creation, equipment acquisition).

Kindly also attach summary report of not more than 5 pages on activities of interest

- _____
- _____
- _____
- _____
- _____
- _____
- _____
- _____
- _____
- _____

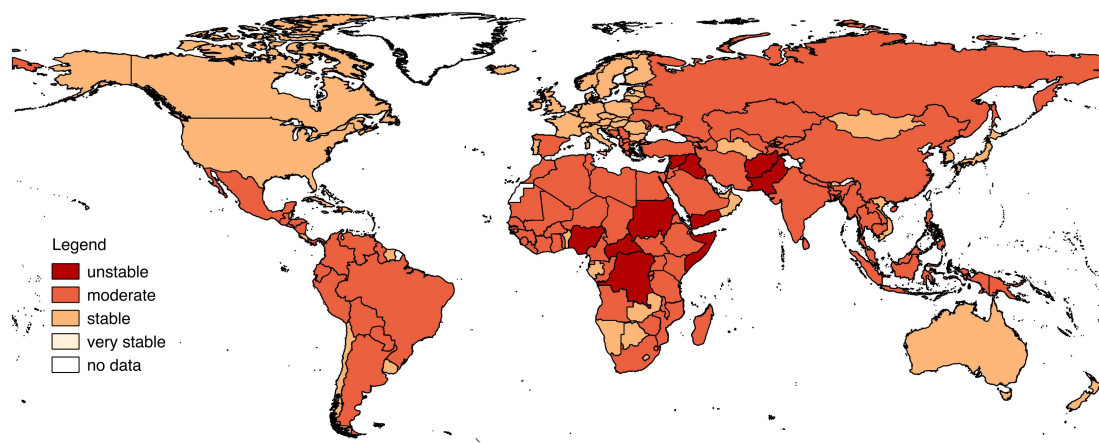
Any Other relevant information

2.2.10.3 Supplementary Figure S3: Tertiary institutions in Africa presenting some form of GIS education



Reproduced with permission from Coetzee et al., (2012).
 Note: the boundaries and names shown on this map do not imply official endorsement or acceptance by the authors.

2.2.10.4 Supplementary Figure S4: Map of the Political Stability Index, 2013



2.2.10.5 Supplementary Figure S5: AFREF stations 2012



2.3 Chapter key outcomes

- Several SDI assessment methodologies exist but different factors need to be taken into account before deciding which one(s) to use: (1) the SDI scale, (2) the resources available to perform the assessment, (3) the purpose of the assessment, (4) the commitment of the SDI stakeholders.
- Given the complexity of SDIs, complex systems are needed to properly assess the various facets of reality. The risk is that only a fragment of experts is able to perform complex assessments that require a lot of resources (e.g. multi-view framework) in time, expertise and funding.
- The geographic scale of the SDI is of particular importance for assessment as it drastically impacts the level of details available. A compromise shall then be found between the assessment's accuracy and the several factors to consider.
- Large-scale SDI assessments are possible but require a rapid assessment methodology providing a broad picture, which can further be completed by smaller scale assessments.
- In any SDI assessment, key respondents (people and/or institution) are essential to ensure a successful assessment.
- In a large-scale rapid assessment, data availability might be scarce, making it necessary to use proxy variables to assess some indicators.
- To address SDI monitoring data scarcity worldwide and particularly in Africa, new mechanisms need to be set up: 1) key SDI coordinating bodies should be designated at several SDI levels, at least at national and regional (e.g. continent) levels; 2) the SDI coordinating bodies should agree on a vision, objectives and an architecture to efficiently share and update agreed-on SDI monitoring variables; 3) tools and best practices should also be set up for rapidly improving efficiency (e.g. standardized online reporting tools, SDI dedicated sections).
- Most of these suggestions require a clear institutional mandate coming from a political supra-national consensus on the model of the INSPIRE directive arising from the European Commission. Such a political consensus requires political integration as well as awareness and funding. Moreover, commitment of all parties, especially the national respondents, is essential, and might be encouraged by incentives.
- Improvements should not be expected everywhere at once, but a leading group of countries might pave the way, for example through international initiatives, for giving directions towards a wider implementation of SDI and monitoring best practices. Such synergies might open the way to further integration, political or economic between participating countries.

3 CAPACITY BUILDING

3.1 Introduction to capacity building

The previous chapter on SDI stocktaking demonstrated the feasibility of a large-scale SDI implementation assessment and allowed to point out African weaknesses in all SDI components. In cases similar to the African one, it is then necessary to address these weaknesses on all fronts, but not necessarily at the same time since SDIs develop gradually addressing in priority the most pressing issues (Kok et al., 2005). Several kinds of actions can be taken to this end, such as infrastructure improvement (e.g. internet bandwidth or electricity supply), technical complexity reduction, or awareness and skills re-enforcement. We argue that the latter is the most crucial one and should come prior to the actions, as awareness at all levels (from the GIS technicians to the political leaders) can trigger the necessary impulsions to address all Capacity Building levels defined by GEO: infrastructure, institutional and human. Awareness and commitment among the stakeholders controlling funding resources is of particular importance (Nushi et al., 2015).

As defined in introduction, several general steps of the Capacity Building process can be differentiated: (1) selection of the target community(ies); (2) assessment of their capacity(ies); (3) formal capacity building, including awareness raising, promotion, education and training. In this general process, the methods and tools will vary depending on the audience type as capacity can not be built in the same way for the general public or for experienced scientists for example. Similarly, the SDI aspect to re-enforce will also determine the tools and duration of building capacity. For example, an institutional capacity building might be a very long process that lasts several years, by opposition to an individual skills re-enforcement that can be performed in a few weeks or months. These considerations will also affect the resources needed for the capacity building activities. Some commonly used tools of capacity building are the following: workshops, success stories, training activities, summer schools, Massive Online Open Courses (MOOC), webinars.

This chapter aims to demonstrate the benefits of capacity building through two different and complementary approaches: (1) the development and improvement of a capacity building material to re-enforce individual skills linked to SDI, which also have an influence on the infrastructure aspect given the software and hardware orientation proposed; (2) an example of an individual, institutional and societal capacity building in which individual skills have been re-enforced, institutional aspects have been improved, as well as the national SDI infrastructure re-enforced and optimized. Both approaches are non-exclusive as capacity developed through the first approach serves the purpose of the second approach.

The first approach is illustrated by the article called *“Bringing GEOSS services into practice: a capacity building resource on Spatial Data Infrastructure”*. It describes the “Bringing GEOSS Services Into Practice” (BGSIP) method that aims at building the capacity for trainees to understand the general SDI concepts and to master the necessary tools for the whole chain of data storing, publishing, documenting, processing, viewing, downloading, analyzing and sharing. This method follows

several principles that should be considered in any capacity building material, especially when addressing Spatial Data Infrastructures: (1) it addresses the language barrier by having part of the material translated into several languages; (2) it uses free and open source software, open data (GEO Secretariat, 2015) and open standards, in conformity with the GEO capacity building strategy (GEO secretariat, 2006); (3) it follows a train-the-trainer concept, aiming at building local capacity so that trainees become local/regional trainers; (4) it is modular and flexible, allowing to adapt to a broad range of audiences despite its best suitability to scientific audience with a technical background; (5) it uses a creative common license, giving a maximal flexibility to users, such as reproducing and modifying the material for their own audience. Such material is key to build individual capacity, which can in turn become capacity available to an institution or an infrastructure, benefiting in fine to a whole society.

The second approach is implemented in the paper called *“Leading the way toward an environmental National Spatial Data Infrastructure in Armenia”*. It describes the whole process that allowed Armenia to become successful and recognized internationally in Earth Observations and SDI domains. It demonstrates the major role played by capacity building activities at individual, institutional and infrastructural levels making Armenia a success story with its recent GEO membership (November 2014). The BGSIP method has been used as the main capacity building instrument for the individual level. The institutional level also benefited from key individuals who followed the BGSIP workshop and got re-enforced capacity and knowledge in SDI. It was complemented by another capacity building methodology targeting the institutional level, called the “EGIDA methodology” (Mazzetti et al., 2013; Plag et al., 2013), which has been used to develop the Armenian environmental National Spatial Data Infrastructure. EGIDA is a methodological set of practices and guidelines that aims at guiding national/regional science and technology communities to make a sustainable contribution to GEOSS and relevant European initiatives. The actions consist of two types of activities that run in parallel in different steps: networking activities to identify and address the relevant Science and Technology community(ies) and actors, and technical activities to guide the infrastructure development and align it with the GEO/GEOSS interoperability principles (Figure 9).

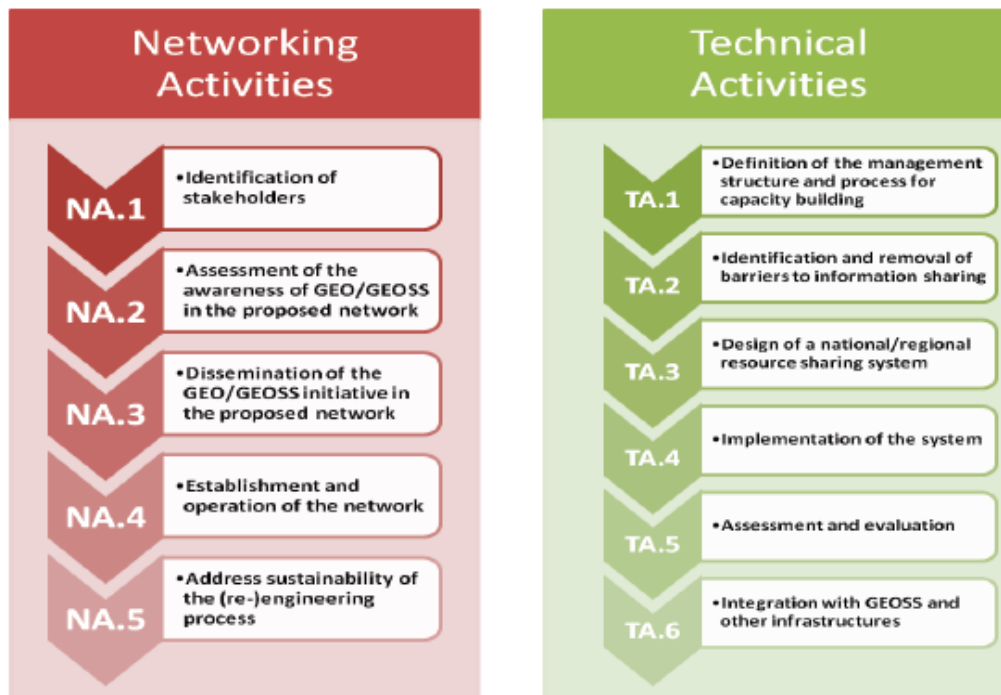


Figure 9: the EGIDA methodology activities (Mazzetti et al., 2013)

The importance of network of actors in the Earth Observation domain is reflected in the EGIDA methodology, as it is one of the building blocks for re-enforcing a country's capacity. Being able to identify the various actors of a network according to relevant criteria is essential and is implemented in the IASON permanent networking facility (PNF)³⁰. The PNF is an online tool containing a list of Earth Observation actors of certain geographic areas (Balkans, Caucasus, Mediterranean area) with sorting criteria such as their role in Earth Observation, the theme that address and the societal sector they are part of.

³⁰ <http://iason-fp7.eu/pnf/>

3.2 Bringing GEOSS services into practice: a capacity building resource on Spatial Data Infrastructures (SDI)

Gregory Giuliani^{1,2}, Pierre Lacroix^{1,2}, Yaniss Guigoz^{1,2}, Roberto Roncella⁴, Lorenzo Bigagli⁴, Mattia Santoro⁴, Paolo Mazzetti⁴, Stefano Nativi⁴, Nicolas Ray^{1,2}, Anthony Lehmann^{1,3}

Addresses:

¹University of Geneva, Institute for Environmental Sciences, EnviroSPACE Lab., CH-1211 Geneva 4, Switzerland

²Global Resource Information Database (GRID) – Geneva, International Environment House, 11 chemin des Anémones, CH-1219 Châtelaine, Switzerland

³University of Geneva, Forel Institute, CH-1211 Geneva 4, Switzerland

⁴Institute of Atmospheric Pollution Research, CNR, Area della Ricerca di Roma 1, Via Salaria Km 29,300 Monterotondo, (RM), Italy

3.2.1 Abstract

Data discoverability, accessibility, and integration are frequent barriers for scientists and a major obstacle for favorable results on environmental research. To tackle this issue, the Group on Earth Observations (GEO) is leading the development of the Global Earth Observation System of Systems (GEOSS), a voluntary effort that connects Earth Observation resources world-wide, acting as a gateway between producers and users of environmental data. GEO recognizes the importance of capacity building and education to reach large adoption, acceptance and commitment on data sharing principles to increase the capacity to access and use Earth Observations data.

This paper presents “*Bringing GEOSS services into practice*” (BGSIP), an integrated set of teaching material and software to facilitate the publication and use of environmental data through standardized discovery, view, download, and processing services, further facilitating the registration of data into GEOSS. So far, 520 participants in 10 countries have been trained using this material, leading to numerous of Spatial Data Infrastructures implementations and 1000 tutorial downloads. This workshop lowers the entry barriers for both data providers and users, facilitates the development of technical skills, and empowers people.

3.2.2 Introduction

Research in environmental sciences is currently hindered by various barriers that impede efficient data discovery and access (Craglia et al., 2012b; Giuliani et al., 2011b). Earth Observation (EO) data are an essential prerequisite to analyze

past, current, and future environmental trends (Beniston et al., 2012; Nativi et al., 2013a). EO can be considered as the collection of necessary data to measure and monitor Earth's physical, chemical and biological systems. They are usually acquired through remote sensing techniques (e.g., satellite, airborne systems) or field sensors (e.g., thermometer, wind gauge, ocean buoy, seismometer). The difficulties in efficiently accessing data are a major obstacle for success in research themes like climate change, biodiversity, or disasters management (Lehmann et al., 2014). To tackle this issue and facilitate environmental data discovery and access, it is essential to convince data holders to make their data available to a larger audience and to understand the benefits of data sharing (Charvat et al., 2013b; Myroshnychenko, 2010). Currently, one of the main challenges that authorities are facing worldwide is the coordination and effective use of the vast amount of environmental data that is generated (Giuliani et al., 2013d). To address this challenge, the concept of Spatial Data Infrastructure (SDI) has emerged (Craglia et al., 2010; Masser, 2005; Nebert, 2005). This framework aims at integrating data sources, systems, network linkages, standards and institutional issues to efficiently deliver environmental data from heterogeneous sources to the widest possible audience (Giuliani and Gorgan, 2013b; Giuliani et al., 2011b; Lehmann et al., 2014; Nebert, 2005). According to the Global SDI Association (GSDI), SDIs are *"an umbrella of policies, standards, and procedures under which organizations and technologies interact to foster more efficient use, management, and production of environmental data"* (Nebert, 2005) and have the ultimate objective to support easy access to and utilization of environmental data (e.g., discovery, visualization, evaluation, access).

SDIs depend on interoperability to significantly enhance data discovery and accessibility. Interoperability can be defined as the ability to exchange and make use of data and information between two or more components. The Open Geospatial Consortium (OGC) is leading the international effort to develop open interoperability standards for geospatial data and information, to enable data discovery (e.g. Catalog Service for the Web – CSW), data visualization (e.g. Web Map Service – WMS), data download (e.g. Web Feature Service – WFS & Web Coverage Service – WCS), and data processing (e.g. Web Processing Service – WPS). Indeed, there are many other ways of sharing environmental data (e.g., download of static files, physical support like DVD or USB keys); the preferred option depends on various factors such as Internet connectivity or computer performance (Giuliani et al., 2013a). However, the BGSIP approach focuses on data sharing through web services as this is an emerging technology supporting the major data sharing initiatives at national/regional/global scales, ensures access to up-to-date data sets, enables interoperability, and is well adapted in a environment with good Internet connectivity.

At the global scale, the leading effort to enhance coordination and provisioning of environmental data is represented by the Group on Earth Observations (GEO) that is developing the Global Earth Observation System of Systems (GEOSS) (GEO secretariat, 2005a). The aim of GEOSS is to enhance the relevance of EO for addressing environmental problems, and to offer full and open access to comprehensive information on and analyses of the environment. GEO adopted

Data Sharing Principles³¹ to expend data reuse through GEOSS and promote open data access to answer broad societal challenges (GEO, 2009). To achieve GEOSS vision and objectives, anyone who wishes to participate in GEO should endorse the Data Sharing Principles: *“(1) there will be full and open exchange of data, metadata and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation; (2) all shared data, metadata and products will be made available with minimum time delay and at minimum cost; and (3) all shared data, metadata, and products being free of charge or no more than cost of reproduction will be encouraged for research and education”*. GEO recognizes the importance of Capacity Building and education to reach large acceptance, adoption, and commitment on Data Sharing Principles to increase the capacity to access and use EO data (Donert, 2015b; Giuliani et al., 2013d). GEO’s capacity building definition follows the United Nations Conference on Environment and Development (UNCED) definition of including *“human, scientific, technological, organizational, and institutional resources and capabilities”* to *“enhance the abilities of stakeholders to evaluate and address crucial questions related to policy choices and different options for development”* (GEO secretariat, 2006). Within its current work plan, GEO has a dedicated task on Capacity Building (ID-02 Developing Institutional and Individual Capacity) that aims at implementing the vision stated in its strategy (European Commission, 2014a; GEO secretariat, 2006).

In its strategy, GEO defines three levels of capacity building: (1) human (e.g., education and training of individuals); (2) institutional (e.g., enhancing the understanding within organization and governments of the value of geospatial data to support decision-making); and (3) infrastructure (e.g., installing/configuring/managing of the needed technology). GEO also recommends demonstrating the benefits of data sharing through appropriate examples, best practices, and guidelines in order to strengthen: (1) existing observation systems; (2) capacities of decision-makers to use it; and (3) capacities of the general public to understand important environmental, social and economical issues at stake. Such initiatives can also give data providers the opportunity to become more visible and trustworthy nationally and internationally by participating in the effort of building GEOSS. The GEO capacity building strategy also identifies several issues that represent at the same time as many opportunities to improve the situation such as (Noort, 2011, 2012, 2013): limited access to capacity building resources; lack of e-science infrastructure for EO education and training; need for criteria and standards for EO capacity building; gaps between EO research and operational application; inefficient connectivity between providers and users of EO systems; need for cooperation within and between developed and developing countries and regions; lack of awareness about the value of EO among decision makers; and duplication of EO capacity building efforts.

Currently, there are many tutorials (e.g., OpenLayers³², GeoServer³³, PostGIS³⁴, QGIS³⁵) and workshops material (e.g., Boundlessgeo Workshops³⁶) that are

³¹ <https://www.earthobservations.org/dswg.php>

³² <http://openlayers.org/en/v3.9.0/doc/tutorials/>

available on the Internet explaining how to use OGC and ISO standards to publish, document, visualize, and process data. However, one of the major issues is that they focus on one or two topics or tools, but to our knowledge none of them gives a clear vision of both data provider and the consumer perspective. It becomes therefore more difficult for stakeholders to understand the benefits of concepts like data sharing and interoperability. This can potentially lead to a lack of interest and commitment, and finally act as a barrier to participation to initiatives such as GEOSS. The GEOSS tutorials are probably the best attempt to integrate these resources and are available from: http://wiki.ieee-earth.org/Documents/GEOSS_Tutorials. However, these tutorials tend to be outdated (i.e., not maintained and using old versions of software) and they do not cover all the aspects of providing and consuming data. Furthermore, learners are required to identify and go through several tutorials before being able to share and use interoperable services.

Recognizing these issues and opportunities, the “*Bringing GEOSS services into practice*” (BGSIP) workshop (Giuliani et al., 2014b) has been initiated in the framework of the European research project enviroGRIDS (Lehmann et al., 2015) as an integrated set of teaching material and software to give the necessary knowledge to efficiently share and use environmental data through web-based services. This material is simple to use and easy to deploy, facilitates the publication and use of data and metadata through discovery, view, download, and processing services, as well as the registration of data into GEO/GEOSS. This paper discusses the strategy put in place to achieve a coherent set of teaching material and its effective delivery specifically addressing the needs of GEO/GEOSS.

3.2.3 Objectives:

The main objective of the *BGSIP* workshop is to promote the GEO Data Sharing Principles by teaching how to configure, deploy, and use a set of open source software to set up a spatial data infrastructure. This general goal is supported by three specific objectives to assist trainees to learn (1) how to publish and share data and metadata using OGC and ISO standards, (2) how to register services into GEOSS, and (3) how to use services discovered through GEOSS in desktop and web-based GIS clients.

The main requirements to design the course are the following:

- Alternate theory and hands-on exercises;
- Simple start and intuitive user interface;
- Platform independent (e.g., Windows, Mac, Linux);

³³ <http://docs.geoserver.org/2.7.1/user/tutorials/index.html>

³⁴ <http://postgis.net/documentation/>

³⁵ <http://www.qgistutorials.com/en/>

³⁶ <http://workshops.boundlessgeo.com>

- Geospatial standards explained, taught, and used;
- Scalable infrastructure for both learning and production environments;
- Open to modification of all teaching material;
- Multilingual instructions;
- Ease of deployment in case participants want to rapidly set up an SDI in production environment.

This workshop concentrates mostly on the human and infrastructure levels of the GEO Capacity Building Strategy. The ultimate strategy behind this workshop is to train-the-trainer by enabling the participants to train other colleagues within their own institutions and to spread the word about data sharing. The institutional level is therefore also tackled by raising awareness internally on the benefits of data sharing.

This workshop cannot cover all topics related to data sharing. The primary objective is to give data providers and users the necessary knowledge to share and consume data through widely used interoperable standards. Once the necessary knowledge is acquired, it becomes possible for them to tackle other issues like data quality or to handle other types of data sources (e.g., sensors or crowdsourced data). In the long term, this will also enable their participation not only in GEO/GEOSS but also in other data sharing initiatives like the Infrastructure for Spatial Information in the European Community (INSPIRE) (European Commission 2007).

3.2.4 Workshop material

Based on the requirements presented in the previous section, the following teaching material has been produced and developed: (1) a tutorial documentation available as a PDF document and eBook (available on iTunes and Google Play), (2) introductory PowerPoint presentations on essential SDI concepts, and (3) a virtual machine (VM) with all software, documentation and data already installed. The following open source software (Table 1) and standards (Table 2) are used in the workshop.

Table 1: Software used in "Bringing GEOSS services into practice"

Name	Version	Website
<i>Provider/Server-side</i>		
PostgreSQL	9.3.1	http://www.postgresql.org
PostGIS	2.1	http://postgis.refractor.net
GeoServer	2.6.1	http://geoserver.org

GeoNetwork	2.10.4	http://geonetwork-opensource.org
PyWPS	3.2.2	http://pywps.wald.intevation.org
GI-cat	10.0.2	http://essi-lab.eu/do/view/GIcat
GI-axe	2.0.3	http://essi-lab.eu/do/view/GIaxe/
<i>Consumer/Client-side</i>		
OpenLayers	3.3.0	http://openlayers.org
QGIS	2.8	http://www.qgis.org

Table 2: Standards used in "Bringing GEOSS services into practice"

Name	Abbreviation	Reference
OGC Web Map Service	WMS	(Open Geospatial Consortium, 2006a)
OGC Web Feature Service	WFS	(Open Geospatial Consortium, 2005)
OGC Web Coverage Service	WCS	(Open Geospatial Consortium, 2006b)
OGC Keyhole Markup Language	KML	(Open Geospatial Consortium, 2009)
OGC Styled Layer Descriptor	SLD	(Open Geospatial Consortium, 2007e)
OGC Catalog Service for the Web	CSW	(Open Geospatial Consortium, 2007b)
OGC Web Processing Service	WPS	(Open Geospatial Consortium, 2007c)
ISO 19115		(ISO, 2014)

ISO 19139 - Geographic information -- Metadata -- XML schema implementa tion		http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557
ISO 19119 - Geographic information -- Services		http://www.iso.org/iso/catalogue_detail.htm?csnumber=39890

The choice has been made to use open source software and open standards because they are valuable resources to develop freely available educating and teaching material. This will help to freely disseminate this material to the widest audience possible allowing trainees to become trainers. Finally, they are efficient solutions to share data and to ensure a sustainable technology transfer by making accessible cost-effective and user-friendly solutions.

The approach taken is that each chapter starts with a question that is addressed and answered by interweaving theory and step-by-step practical exercises (Table 3).

Table 3: Structure of the workshop, with related tools and standards

Chapter	Title	Software	Standards
1	Concepts on SDI	N/A	N/A
2	How to store geospatial data?	PostgreSQL/PostGIS	N/A
3	How to publish geospatial data?	GeoServer	WMS, WFS, WCS, KML, SLD
4	How to document and search geospatial data?	GeoNetwork	CSW, ISO
5	How to process geospatial data?	PyWPS	WPS
6	How to view geospatial data?	OpenLayers, QGIS	WMS, KML
7	How to download geospatial data?	QGIS	WFS, WCS

8	How to analyze geospatial data?	QGIS	WPS
9	How to share geospatial data?	GI-cat, GI-axe	WMS, WFS, WCS, CSW, WPS

The material follows a structured order reflecting a general path from service providers to consumers (Figure 1). After an introduction on concepts on SDI, chapters 2 to 5 concentrate on the provider side explaining how to create a PostgreSQL/PostGIS geospatial database and load vector data; how to publish vector and raster data as OGC services with GeoServer; how to document data with ISO-compliant metadata and store it in a metadata catalog; and finally how to create a geoprocessing script to analyze geospatial data. Chapters 6 to 8 focus on the consumer side illustrating how to visualize data with WMS in various clients (e.g., OpenLayers, QGIS, Google Earth); how to download data with WFS and edit a vector layer; and how to consume a WPS service in GIS client for geospatial analysis. The final chapter discusses how to share these services and participate in GEOSS by registering resources in the GEO Discovery and Access Broker (Nativi et al., 2009a; Nativi et al., 2013a; Nativi et al., 2013b; Nativi et al., 2015).

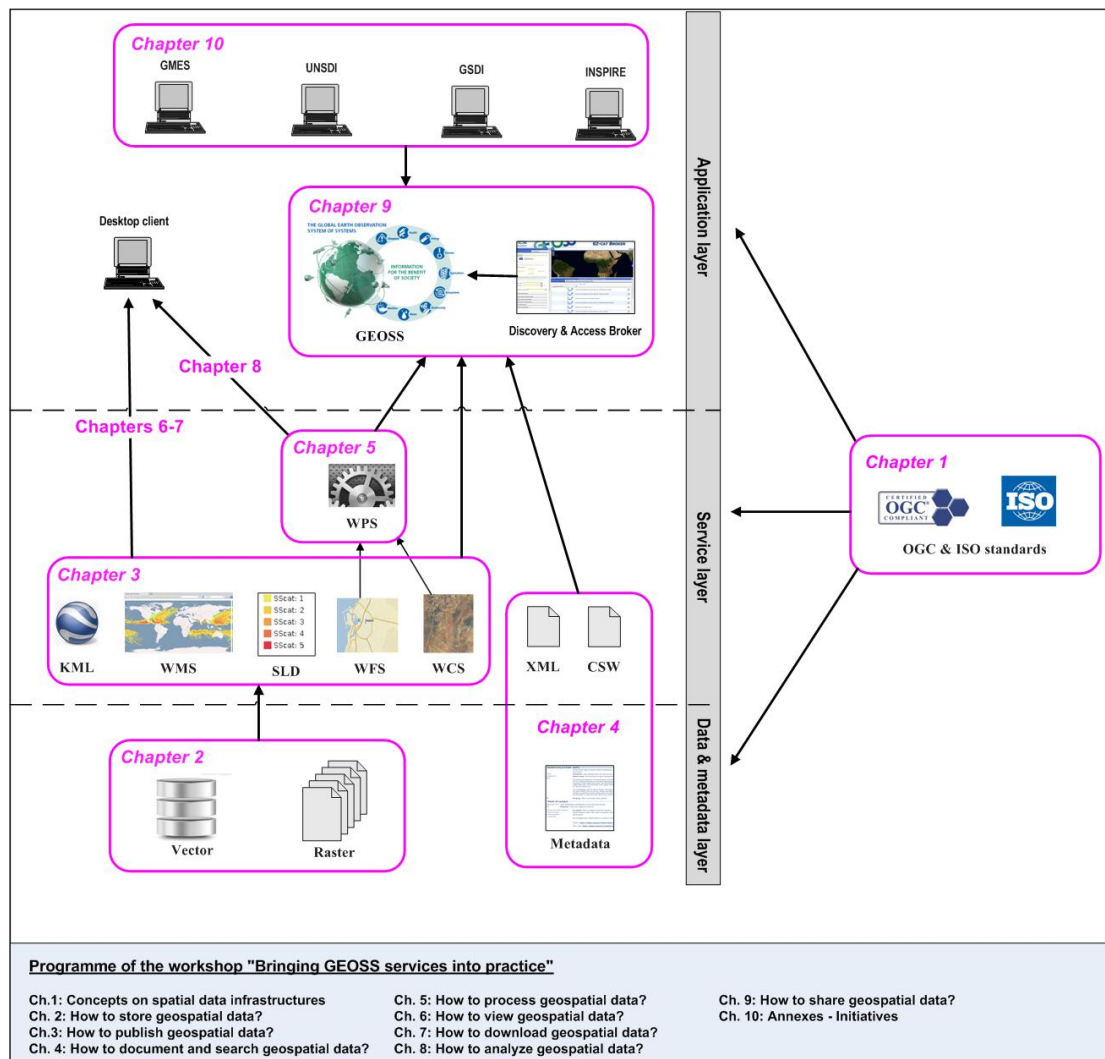


Figure 1: The general structure of the workshop

In the same spirit, the data used in this workshop are based on international (e.g., Global Risk Data Platform (Giuliani et al., 2011a), Transboundary Water Assessment Program)), regional (e.g., ECOWREX, project on renewable energy in West Africa) and project-related (e.g., enviroGRIDS (Giuliani et al., 2013b; Lehmann et al., 2014), IASON and EOPOWER on the use of EO for environmental applications)) environmental data platforms. These data sets cover the following themes: risk to natural hazards, water, renewable energy, hydrology, and climate. All these data sets are openly and freely available (i.e., no copyright issues) and are used in several chapters to create continuity within the workshop. The data used is neither voluminous nor spatially complex so as to be easily handled by beginners. The same reasons for using open source software apply to use the open data sets and in particular their rights to redistribute. Commercial data sets were not an option in our case because we wanted to freely redistribute data sets together with the software to handle them. This is in line with the free and open sharing spirit promoted by GEO/GEOSS and supported by the workshop.

All the material (i.e. software, data, tutorial) is integrated and provided in a dedicated and tailored Linux VM based on Oracle VirtualBox. This VM can be executed on Windows, MacOS, or Linux and requires a minimum of 4GB of memory and at least 20GB of disk space. Once installed users can directly start the tutorial. Furthermore, having a VM that is already preconfigured allows users to first train with SDI concepts and technologies and once they are familiarized they can deploy it as a production server. By using the VM users are able to save time on setup, focus on open standards, and later deploy it as a production environment and make adjustment if necessary. Therefore, this is an important incentive for the users. This material comes with a 200-slide PowerPoint presentation of the theoretical content of the workshop. This teaching material of the workshop is already available in seven languages: Arabic, Croatian, English, French, Russian, Serbian, Spanish and soon in Czech.

3.2.5 Dissemination

The material of the workshop can be downloaded on a dedicated website: <http://www.geossintopractice.org> and the tutorial is available as an interactive ebook on both iTunes/iBooks Store³⁷ and Google Play Books³⁸. The material is completely free of charge and is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. This type of license allows any user to freely share (e.g., copy and redistribute in any medium or format) and adapt (e.g., remix, transform and build upon) the material. However, users are obliged to give appropriate credit (e.g., Attribution), are not allowed to use this material for commercial purposes (e.g., NonCommercial) and if users remix, transform or build upon the material, they must distribute their contributions under the same license (i.e., ShareAlike).

The workshop itself can take various formats by teaching one, some, or all chapters, with an estimated maximum of 12 hours. This is complemented by videos presenting each chapter of the tutorial on a dedicated “*Bringing GEOSS services into practice*” YouTube channel³⁹. In terms of dissemination, the workshop has been taught 15 times between 2010 and 2015 to about 520 participants in Bulgaria, Georgia, Morocco, the Netherlands, Romania, Serbia, Switzerland, Tunisia and Turkey (the up-to-date agenda is available at: <http://goo.gl/IWBz8M>). At the University of Geneva, the workshop has been included as a course that lasted two days and students were further asked to put their knowledge into practice by developing and publishing a web application. In two other cases the last half-day of the workshop was dedicated to the integration of geospatial data into the SDI of the hosting institution. The workshop was also presented at international events (e.g., as a side event during

³⁷ <https://itunes.apple.com/us/book/bringing-geoss-services-into/id806182409>

³⁸

https://play.google.com/store/books/details/Gregory_Giuliani_Bringing_GEOSS_services_into_prac?id=Nv6nAgAAQBAJ

³⁹ <https://www.youtube.com/channel/UCfBVYFBQw1aEU7M1j9zbW7A>

the GEO-X⁴⁰ Plenary) and followed by discussions on how to assess its impact. The workshops have addressed diverse audiences ranging from students to teachers, scientists to policy makers, and NGOs and public employees to private companies. However, the workshop remains quite technical and is ideally suited to environmental scientists with technical background. In order to reach a broader audience, other formats of the workshop are planned, less technical and/or more thematic (e.g., in the field of water resources, climate change or mineral exploitation). In addition, more than 1'000 downloads of the workshop material have been recorded since March 2014. The tutorial is registered in GEOSS and in GEOCAB⁴¹, and is promoted by the GEO secretariat.

A questionnaire was also sent to 450 past attendees to help measure the impact of the workshop (Figure S1). While the number of respondents is low (e.g., only 50 people replied so far), 32 of them stated they have taught others what they learned at the workshop. Even if this does not mean they have reproduced the exact workshop, it shows that the workshop's principles can be reproduced. 75% of the respondents say they are in a position to teach this workshop in their organization. 65% say they have / are planning to share data through GEOSS. Finally, one third of the respondents set up an SDI in their organization.

These results show that the objectives of being simple, interoperable, flexible, scalable, and multilingual have been successfully achieved. In particular this can be also helpful for users of specific scientific community to participate to an initiative like GEOSS and using what is in GEOSS to develop tailored application targeting the needs of a specific community (e.g., Community portals (Cau et al., 2013; Gorgan et al., 2013a)).

3.2.6 Lessons learned and recommendations:

Since 2010, the experience gained and lessons learned in developing, implementing, and assessing the workshop to develop capacities of different user groups allow us to draw some recommendations for creating a capacity building resource on SDI, but also to highlight some limitations.

3.2.6.1 Benefits & Impacts

The workshop had an impact on research project partners, countries and institutions where the Bringing GEOSS services into practice workshop was taught (Giuliani et al., 2013d). At the human level, several partners of EU/FP7 projects decided to implement their own SDI. They all have stated that having participated in the workshop convinced them about the necessity to share data and the benefits of using web-based interoperable services. Many positive comments in the impact survey (e.g. *"I want to use GEOSS services in my PhD works"*; *"We hope to be able to collaborate with you in other research and education / training projects related to web-based data sharing"*) echo this sentiment. At the institutional level, institutions like the Commission on the Protection of the Black Sea Against Pollution (whose International Secretariat has been accepted as an Observer in GEO in November 2015) and the

⁴⁰ http://www.earthobservations.org/me_sevent.php?id=147

⁴¹ www.geocab.org

International Commission for the Protection of the Danube River (ICPDR) found that sharing data using OGC and ISO standards could bring several benefits for their assessment and reporting processes and prompted the effort to implement and/or upgrade their infrastructures. At the country level, these workshops helped to raise awareness about GEO/GEOSS: three countries (Armenia, Bulgaria, Georgia) where workshops were given became GEO members. Finally, the wide adoption of OGC standards among project partners facilitated the development and implementation of different software solutions and enabled communication between different computing infrastructures (Charvat et al., 2013b; Giuliani et al., 2013b; Giuliani et al., 2013c). This allowed to process large amounts of environmental data on distributed computing infrastructures to analyze high-resolution satellite images or execute large hydrological models (Bacu et al., 2013; Bektas et al., 2013; Gorgan et al., 2011; Gorgan et al., 2012a; Gorgan et al., 2012b; Mihon et al., 2013a; Mihon et al., 2013b).

Besides these general impacts, several other benefits both measurable directly (e.g., questionnaire sent to participants) and indirectly (e.g., following the work of some participants, institutions) can be featured. At the technological level, participants have registered several services into GEO/GEOSS. This demonstrated the increase of awareness and use of open standards and will ultimately facilitate discovery and access to hundreds of data sets. The use of Free and Open Source Software (FOSS) was really instrumental in building capacities and implementing data sharing solutions. In particular, these solutions are attractive for students, professionals, enterprises, and institutions that do not have the means to afford expensive commercial solutions. More importantly the open source software solutions used in the workshop have proven to be very reliable and efficient (Giuliani et al., 2013a). The fact that all the teaching material can be freely disseminated helps trainees to become trainers. This helps to lower entry barriers for both data providers and users, strengthens development of technical skills, and empowers people. This can potentially have a scientific and societal impact by increasing the number of data sets available; facilitating discovery and access to data; enabling decision-makers and the general public to access relevant, up-to-date, and scientifically sound environmental information; giving a sense of belonging to an active community; and possibly taking better political decisions. Finally, this workshop may have an impact on GIS educators and curricula because sharing and documenting data is part of the elementary scientific approach enhancing accountability, credibility, and reproducibility. It gives them access to a complete teaching resource presenting cutting-edge web-based technology and helping them to educate students on the benefits of data sharing.

3.2.6.2 Limitations

During this workshop several challenges have been identified. In addition to technological aspects (e.g., computer languages, computer performances) the main issues are related to institutional (e.g., political/cultural context, policies, organization, resources) and human (e.g., skills, knowledge) aspects.

Based on our own experience in giving the workshop together with collected feedbacks from participants, other limitations have been identified. Our experience shows that a one-day session with no more than 10 participants per

instructor is ideal for both teaching and learning. Alternating between talks, hands-on and videos proved to be a good way to capture and keep attention of the audience. However, the heterogeneity of participant's skills; the balance between the number of participants and instructors; and the fact that the workshops strongly depends on a fast (>2Mbps) Internet connectivity are important factors to take into consideration to ensure a good user experience. This is illustrated by a comment in the user survey feedback stating, *"Less demanding data would be easier, loading and processing the current data is quite slow"*. Based on user feedback, it seems that the chapter on storing data is too technical, and this is a notable barrier for some participants since it includes the first hands-on exercise. Furthermore, in terms of assessment of the workshop, the number of respondents was limited. Increasing the number of answers can help to improve feedbacks and draw more robust conclusions based on participants' comments. Finally the question of sustainability of the workshop is critical. We could evaluate that upgrading the full material (virtual machine with latest software versions, tutorial, presentations in one language) requires between two and three weeks of work. The recent integration of the workshop in courses at University of Geneva (2014) is positive but a successful dissemination beyond the academic sphere will require identifying key mechanisms (e.g. GEO processes), projects, and people. A possible solution is to integrate capacity building workshops into future projects, as was done in IASON, EOPOWER and ClimVar⁴².

3.2.6.3 Recommendations

Based on the experience acquired and the identified benefits, impacts, and limitations, there are several recommendations for developing a capacity building resource on SDI such as the BGSIP workshop:

- Promote the use of FOSS software and the development of freely available education and teaching material. This will facilitate reaching and disseminating teaching resources to the widest audience possible. This will also ensure a sustainable technology transfer by making available cost effective and user-friendly solutions.
- Massive learning solutions like MOOC are promising teaching solutions to better promote data sharing needs and solutions to a large audience. However, MOOC might not be appropriate for some technical hand-on exercises that would require heavy live interaction with the teacher.
- Capacity building activities should let users experiencing the benefits of data sharing through appropriate examples; by communicating best practices; and developing guidelines and policies. Altogether this will facilitate reaching agreement and endorsement on the use of new standards and enhance an "open and sharing spirit".
- Specific measurable goals must be defined during the planning phase of a workshop development. Such a strategy should consider various indicators (e.g. web material download statistics, user surveys, agreement

⁴² <http://www.globalclimateforum.org/index.php?id=127>

with some users for a long term impact follow-up, etc.). This will help to assess impact, get feedback, and better understand how the workshop has influenced their work.

- The targeted audience of such a Capacity Building resource must be well defined from the beginning, and tailored material ready depending on the audience (e.g. be prepared for technical issues such as connectivity, hardware or to languages issues that might require translators). The audience is usually heterogeneous and therefore the materials should not be too technical or too conceptual elements, because otherwise some participants may not be able to follow.
- In order to maximize the impact, it is essential to target and include key high-level institutional individuals when delivering such capacity building workshop. We found that having a pre-workshop high-level presentation on the benefits of data sharing would often facilitate the adoption of the technical choices later on.
- A long-term maintenance and upgrade mechanism must be also planned right from the beginning in order to ensure a quality and cutting-edge sustainability of the Capacity Building resource.
- A final recommendation is to clearly highlight the anticipated benefits of the workshop, such as: increased awareness of data sharing issues and solutions; complete technical understanding of the chain from data collection to data publication; improved technical skills; increased regional and national capacities with the potential increased development; enabling institutions and people to closely cooperate and share a common vision.

3.2.7 Conclusions and future perspectives:

In order to encourage data providers to be more “open” and facilitate access to their data, a long-term commitment to education, capacity building, and research is essential. The *“Bringing GEOSS services into practice”* teaching material is probably the first attempt to build capacity simultaneously for both data providers and data users. It facilitates the configuration, use, and deployment of a set of open source software to implement an SDI. It also explains how to publish and share data and metadata using OGC and ISO standards and how to register published services into GEOSS. Finally, it explains how to use services that are discoverable in GEOSS to develop tailored applications for specific purposes and/or a community of users. This material has been developed in the train-the-trainer spirit enabling the reuse and exchange of knowledge and new capacities that have been acquired. The main features of the workshop are simplicity, data interoperability, interoperability of material, flexibility, multilingual, and scalability. As a result, 520 participants in 10 countries have been trained using this material, leading to dozens of on going SDI implementations (Asmaryan et al., 2014; Astsatryan et al., 2012), and more than 1'000 downloads. The introduction material (PowerPoint presentations) is translated in several languages (Arabic, Croatian, English, French, Russian, Serbian, Spanish, Czech). This workshop lowers entry barriers for both resource

users and providers, facilitates the development of technical skills, and empowers people.

Experience has shown that this material is useful, but some chapters are probably too technical for beginner users. The authors are planning to develop a lighter and shorter version of this training material with less technology and simpler tools (e.g. GeoNode). Another option for further user targeting is to develop training material based on “*Bringing GEOSS services into practice*” but focusing on thematic issues such as the GEO/GEOSS Societal Benefits Areas (SBAs: Agriculture, Biodiversity, Climate, Disasters, Ecosystems, Energy, Health, Water, Weather). Finally, to complement the workshop other chapters can be added to take into account subjects that are not yet discussed in detail such as the use of sensors, the management of crowdsourced data, or the assessment and control of data quality. The latter is of particular importance and has been explored by several EU-funded projects that define data quality indicators (e.g., GeoViQua, QA4EO) (Díaz et al., 2012b) or deal with managing uncertainty (e.g., UncertWeb) (Bastin et al., 2013). With such substantive additional material, the development of a dedicated Massive Open Online Course (MOOC) would be possible, and greatly increase the EO capacity building range in the line with the GEO CB strategy.

Beyond overall capacity building of the technical aspects of SDI implementation, in some countries the workshop had a direct influence on national policies on Earth Observation and data sharing. For example, Georgia and Armenia recently became GEO members (in 2013 and 2014, respectively). Armenia was an early adopter of the technical and institutional recommendations of the workshop, notably with the creation of a Master-level course in SDI and computational technology. In 2014, the country also adopted its national resolution N136 aiming at building a national Spatial Data Infrastructure. The recent nomination (in November 2015) of Armenia as a member of the GEO Executive Committee is yet another success that can be, in part, attributed to the positive outcomes of the BGSIP workshop.

Although a few preliminary success stories have been discussed, such as with Armenia, the impacts and benefits of such a workshop typically continue to accrue over a longer period of time. These insights will be part of future articles in which the impacts of SDI capacity building activities will be studied in more detail, including sharing the lessons learned in each country, with the goal of ultimately empowering others to follow their lead.

The sustainability of the workshop material itself is also at stake. Indeed, the SDI community is dynamic and changes over time, so there is the risk that the workshop could become outdated. Therefore, it is important to maintain and regularly upgrade the workshop material with new software versions that would come out or with new standards that will emerge in the future. To this end, advantage will be taken of this workshop being also taught at the University of Geneva (e.g., Certificate of Geomatics⁴³). This will provide the necessary resources for maintaining it.

⁴³ <http://www.unige.ch/sig/enseignements/cgeom.html>

Besides, it is foreseen to enrich the course with practical exercises based on thematic data, in fields such as climate change, water resource and mineral exploration.

Some characteristics of BGSIP facilitate its sustainability. First of all, the tutorial has a modular structure making easy to update specific sections or even adding entirely new sections where required. Moreover, the tools for the hands-on part of the course are provided in a virtual machine that can be easily updated during the preparatory phases. These characteristics allow quite a fast update and tailoring of courses. Upgrading the VM and the tutorial requires a few person-weeks, a limited time and effort that can be easily allocated on most initiatives and projects budget. Since its creation in 2010 the workshop has been upgraded each year utilizing various project funding.

Moreover the *“Bringing GEOSS Services into practice”* course has been proposed by the EU-FP7 IASON project⁴⁴ as a best practice for actions aiming to reduce technical barriers to data sharing. The integration of the workshop as a guideline in the revised version of the EGIDA Methodology (Bigagli and Lipiarski, 2015a; Nativi et al., 2013c), a general methodological approach for the re-engineering of Earth Observation infrastructures, itself integrated in the EOPOWER Methodological Framework for Impact Assessment of Earth Observation for Environmental Applications⁴⁵, will give an increased visibility of the workshop. This might provide new opportunities for further developing or updating the workshop, contributing to its sustainability.

3.2.8 Supplementary material

⁴⁴ <http://iason-fp7.eu>

⁴⁵ <http://www.eopower.eu/?q=node/118>

* First name
* Last name
* Email
* Country Choose one of the following answers Please choose...
* Company/Organisation name
* Type of organization Check any that apply ☐ Private ☐ Government ☐ Education/Academic ☐ Not for profit ☐ Personal use ☐ Other:

*** Position/Job role**
Check any that apply

- ☐ Technical lead
☐ GIS professional
☐ Decision maker
☐ Manager
☐ Other:

*** Have you created a Spatial Data Infrastructure in the frame of your institution, including webservices?**
In case of "yes", please indicate the urls, endpoints, ... of your webservices

Choose one of the following answers

- ☐ Yes
☐ No

Please enter your comment here:

*** Have you shared or are you planning to share data of your institution through GEO/GEOSS (http://www.geoportal.org/web/guest/geo_home_stp)?**
In case of "no", please indicate if this is due to:

- institutional barriers,
- financial barriers,
- technical barriers,
- time constraint,
- other

Choose one of the following answers

- ☐ Yes
☐ No

Please enter your comment here:

*** In your institution, are you in a position to train others with material acquired during this workshop?**

☐ Yes ☐ No

Have you or are you planning to teach what you learnt during the workshop in your organization in order to improve data sharing?
In case of "no", please indicate if this is due to:

- institutional barriers,
- financial barriers,
- technical barriers,
- time constraint,
- other

Choose one of the following answers

☐ Yes ☐ No

Please enter your comment here:

Do you have any other comments or wishes for future editions of the workshop?

Figure S1: Online questionnaire used to assess the *Bringing GEOSS Services into Practice* workshop since 2010.

3.3 Leading the way toward an environmental National Spatial Data Infrastructure in Armenia

Asmaryan Sh.¹, Saghatelyan A.¹, Astsatryan, H.², Bigagli, L.³, Mazzetti P.³, Nativi S.³, Guigoz, Y.^{4,5}, Lacroix, P.^{4,5}, Giuliani, G.^{4,5}, Ray, N.^{4,5}

¹Center for Ecological-Noosphere Studies of the National Academy of Sciences of the Republic of Armenia, 68, Kh. Abovtan str., Yerevan, 0025, Armenia

²Institute for Informatics and Automation Problems of the National Academy of Sciences of the Republic of Armenia, 1, P. Sevak str., Yerevan, 0014, Armenia

³ National Research Council of Italy, Institute of Atmospheric Pollution Research, Earth and Space Science Informatics Laboratory, Area della Ricerca di Firenze, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy

⁴Institute for Environmental Sciences, EnviroSPACE Lab., University of Geneva, 7 route de Drize, CH 1227 Carouge / GE, Switzerland

⁵United Nations Environment Programme, Global Resource Information Database, 11 chemin des Anémones, CH 1219 Châtelaine/GE, Switzerland

3.3.1 Abstract

Once the most industrialized republic of the Soviet Union, Armenia inherited a dramatic ecological situation from the Soviet era. As the key national environmental academic entity, the Center for Ecological-Noosphere Studies (CENS) of the National Academy of Sciences of the Republic of Armenia has a strong national role in delivering authoritative environmental information and data sets. To enhance data sharing towards its stakeholders, CENS engaged in recent years in several international capacity building projects directed to the setting up of an environmental Spatial Data Infrastructure (SDI). These activities were successful in showing the potential of data sharing in Armenia, to gain visibility in the country and the South Caucasus region, and to start engaging in international voluntary partnerships such as the Group on Earth Observations (GEO). CENS now envisions to scale up its SDI infrastructure to an environmental national SDI (nSDI) in order to support a wider range of geospatial services. This paper discusses several aspects and challenges of the envisioned strategy. First, we present how the current components of the implemented SDI benefit the scientific and environmental communities in Armenia. Second, we examine how the EGIDA methodology can be applied to support the process of scaling up the infrastructure to become a nSDI, one of the pilot studies in the EU/FP7 EOPOWER project. Finally, we discuss the potential of future full-scale provision of geospatial services in Armenia and how these could benefit the various stakeholders involved in Armenia and in the South Caucasus region.

3.3.2 Introduction

The term Spatial Data Infrastructure (SDI) refers to technologies, policies and people supporting the sharing of geographic information throughout all levels of government, commercial and the non-profit sectors, academia and citizens (Giuliani, 2011). The goal of an SDI is to make geospatial information more accessible to the public, to improve quality of this information, to avoid duplication effort and to “establish key partnerships with states, counties, cities, tribal nations, academia and the private sector to increase data availability” (FGDC, 2013).

A SDI can be implemented at different geographical scales. Well-known examples of SDIs are the Global Earth Observation System of Systems (GEOSS) (GEO Secretariat, 2005b) and the United Nations Spatial Data Infrastructure (UNSDI) (Henricksen, 2007) at the global level, and the Infrastructure for Spatial Information in the European Community (INSPIRE) (European Commission, 2007) at the regional level. Examples of SDIs at the sub-national level are numerous (Riecken et al., 2003; Vandenbroucke et al., 2009). The example of Spain is singular as most provinces have built up their own SDI, e.g. IDEAndalucia, GeoEuskadi, Cartomur, IDECanarias and the SDI of Catalonia (Craglia and Campagna, 2009b; Garcia Almirall et al., 2008).

One of the first countries that implemented a National Spatial Data Infrastructure (nSDI) was the United States of America in the 1990s, under the impetus of the Federal Geographic Data Committee (FGDC). This initiative was engaged after President Clinton signed in 1994 the Executive Order 12906 that defined the nSDI as “the technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data” (Clinton, 1994). Many national initiatives have followed since that time (e.g., (Crompvoets et al., 2004; Rajabifard et al., 2001a)) like in Australia under the impulse of the Australian and New Zealand Land Information Council (ANZLIC), in Malaysia (Arshad and Hanifah, 2010) and in the Netherlands (Kok et al., 2005). The setting up of a nSDI is logically influenced by national specificities such as the political background, the technological state of progress and the environmental policy. In Japan for example the building of the nSDI was mainly driven by the concern for handling earthquake-related emergencies (Masser, 2005).

3.3.3 Environmental status of Armenia and data sharing activities

Armenia was part of the former Soviet Union and was one of the most industrialized Soviet republics. Large-scale industrial activities such as mining, chemical and electrical industry, and machinery led to a severe impact on the environment (Aleksandryan, 2006; Kakarekaa et al., 2004; Kurkjian, 2000; Kurkjian et al., 2002; Petrosyan et al., 2004; Saghatelyan, 2007; Saghatelyan and Sahakyan, 2007). After the Soviet breakdown, these industrial activities collapsed in Armenia, but some recovery occurred in the mid-1990s essentially due to the activities of several mining companies. In the meantime, the economic policy shifted towards a strong support to industrial development, which was accompanied with a lack of interest for associated environmental issues. As a result, mining-related industries such as dressing and metallurgical plants were permitted to operate without environmental regulations, and geo-exploration

and exploitation works in deposits were conducted disregarding nature protection norms. Consequently, the unfavorable ecological situation inherited from the Soviet industrial era substantially worsened.

Geographically, Armenia lays in the northern part of the South Caucasus and is a place of origin of two major water arteries: rivers Kura and Araks. All countries of the region sharing borders with Armenia use Kura-Araks catchments and share emerging environmental problems. Activities targeting environment research, awareness and conservation in Armenia are vital for the country's future and by extension for the South Caucasus region.

Among the few organizations dealing with environmental studies in Armenia, the Center for Ecological-Noosphere Studies (CENS) of the National Academy of Sciences of the Republic of Armenia is an active group with strong national leadership. CENS carries out environmental research activities on the complex assessments and modeling of ecological state of various environmental compartments (soil, water, plants) and develops scientific and methodical fundamentals of ecological expertise and optimization of natural resource management processes in the country. As a result of these activities, CENS has built and filled since early 1990 a large spatio-temporal registered environmental fieldwork database, together with a file-based multi-scale geodatabase for Armenia.

However, no centralized way of managing these environmental data and the national inventory of environmental data (emission, land-use, etc.) existed. To overcome this shortcoming, CENS together with University of Geneva and the Institute for Informatics and Automation Problems (IIAP, the leading Armenian research and technology development institute on Information and Communication Technologies) of the National Academy of Sciences of the Republic of Armenia initiated the creation of the national distributed processing capacities for environmental data sharing which was successfully implemented and deployed in the framework of the SNSF-SCOPES ARPEGEO ("**AR**menian distributed **P**rocessing capacities for **E**nvironmental **GEO**spatial data", <http://arpegeo.sci.am>) project. This 2-year project (2011-2013) enabled the deployment of the first environmental data sharing and interoperability services in Armenia, which strengthened the national capacities of geospatial data sharing, increased the visibility and national position of CENS as an expert in environmental research, and expended their regional and international networks in this field.

The ARPEGEO project strongly benefited from the existing firm foundation of the Armenian e-infrastructure that integrates networks, distributed computational and storage resources, experimental workbenches, data repositories, tools, instruments, and other operational support enabling national and global virtual research collaborations. The e-infrastructure is operated by IIAP and offers research data services and repositories enabling scientists from many disciplines to upload and share data in Armenia and beyond.

The main output of the ARPEGEO project was an environmental Spatial Data Infrastructure combining data resources, distributed computing platforms and computational services. A web portal of interoperable geoprocessing services

was developed to offer complex geoprocessing capabilities, and to hide low-level access mechanisms to computational resources by high-level graphical interfaces, making even non-GIS expert users capable of defining and executing distributed applications. The geospatial and environmental data sets and their associated metadata existing at CENS were integrated into the SDI by adopting the international standards from the Open Geospatial Consortium (Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS)) and from the International Organization for Standardization (ISO 19139, 19115) (Astsatryan et al., 2012). Innovative Web Processing Service (WPS) workflows were developed to connect to the Geographic Resources Analysis Support System (GRASS GIS, (Neteler et al., 2012)) in order to compute a set of vegetation indices on user-defined satellite images (Astsatryan et al., 2014; Astsatryan et al., 2015). These WPS workflows can access both grid and cloud resources of the Armenian National Grid Initiative (ArmNGI, <http://www.grid.am>). The ArmNGI environment is used in case of distributed processing of large amounts of spatial data with very complex calculations.

3.3.4 Potential and benefits of an Armenian nSDI

It is now recognized that “international collaboration is essential for exploiting the growing potential of Earth observations to support decision making in an increasingly complex and environmentally stressed world” (http://www.earthobservations.org/about_geo.shtml). Accordingly, an international effort has been going on for more than a decade to better coordinate earth observation globally, mainly through the Group on Earth Observations (GEO). GEO is a voluntary partnership of governments and international organizations that provides the framework necessary to this coordination. Ninety countries are currently members of GEO, giving them a frame to coordinate their strategies and investments in Earth Observation. Besides, “GEO continues to focus significant effort on building both human and technological capabilities” (GEO, 2014d).

Even though Armenia is not officially a GEO member yet (the letter of intent has been sent by Armenian authorities to the GEO Secretariat in August 2014), it has the maturity to become one as there is a growing awareness in the national scientific community, as well as a political understanding at the highest level, of the advantages of being part of this global coordination effort. GEO, through its Global Earth Observation System of Systems (GEOSS), targets nine Societal Benefit Areas (SBAs: disasters, health, energy, climate, water, weather, ecosystems, agriculture and biodiversity) that are key to ensure a proper development of countries and improve socio-economic needs of the country and the region. GEOSS “seeks to focus the attention of world leaders on existing and developing capabilities of GEO members and participating organizations, identifying and filling capability gaps and—very important—relating these observations to specific benefits for society” (Lautenbacher, 2006). Armenia could greatly benefit of this coordinated effort as: (1) it would improve the national data flow through the mandatory acceptance of open-data sharing principles for member states; (2) it would stimulate a better cooperation between the different state institutions responsible for national data to reach open-data sharing principles; (3) it would benefit from global contributions in Capacity Building for Earth Observation.

In parallel to Armenia's integration to international initiatives such as GEO, the country is also more and more involved in the European Research Area (ERA), for example through EU funded projects such as the FP7 EcoArm2ERA (<http://ecoarm2era.eu/>) project coordinated by CENS and the FP7 INARMERA (<http://www.inarmeraproject.am>) project coordinated by IIAP. The ERA is "a unified research area open to the world based on the Internal market, in which researchers, scientific knowledge and technology circulate freely." (see: http://ec.europa.eu/research/era/era_communication_en.htm). It aims at optimizing the European research as "if Member States manage their agenda in relative isolation, it is inevitable that several research teams across Europe will engage into similar projects." (European Commission, 2012). Being integrated to this framework will allow Armenia to (1) further enable scientific exchange and partnering between the Armenian researchers and colleagues from ERA through long term strategic partnerships; (2) give Armenian researchers more opportunities for joint research or projects collaborations with EU funded projects, thus addressing the national scientific brain drain problem (i.e., Armenian scientists leaving the country to find a job elsewhere); (3) improve the competencies needed by Armenian researchers and staff members to participate to the EU funded project calls.

Armenia's integration into international initiatives or the ERA will allow the country to enhance its visibility, both regionally and globally, and could potentially give it a regional role of expertise in environmental coordination efforts that could pave the way to a better political integration with neighboring countries. CENS and IIAP are already well versed in the GEO activities through participation in several GEO meetings, and both institutions have built capacities in OGC standards and SDI technologies through hosting of various related workshops, notably the "Bringing GEOSS services into practice" workshop (Giuliani et al., 2014b).

Initiating, developing and linking CENS SDI to distributed resources were the first essential steps. Now is the appropriate time to consider scaling up this SDI with the aim of becoming the authoritative environmental platform in Armenia, allowing both the Armenian data providers to register their data sets and services, and all Armenian stakeholders to benefit from the infrastructure by easily accessing its content in standardized ways. This vision is what we are referring to with the term *Armenian environmental nSDI*. It is a *national* SDI in the sense that it will ideally be recognized at national level to be the authoritative source for environmental data in Armenia, and consequently be supported by the appropriate national data policies (still to be developed). Note that we hereby restrict this vision to the environmental data and services, but the long-term vision would be to target a truly cross-discipline national SDI, incorporating cadastral, infrastructure, energy sectors, etc.

The benefits of the envisioned Armenian environmental nSDI would be numerous and could directly profit other spheres, for example:

- An effective Armenian environmental nSDI could positively influence similar initiatives initiated elsewhere in the South Caucasus region. For example the neighboring country Georgia, which was just accepted as

the 90th GEO member country, is still lacking a roadmap toward an environmental SDI. The close collaboration between many Georgian and Armenian researchers (notably through collaboration in the FP7 enviroGRIDS project; (Lehmann et al., 2015) is greatly facilitating exchange of good practices and experience, which could help Georgia to better frame the actions related to their nSDI.

- The actions associated with implementing the Armenian environmental nSDI will also facilitate the organizational and policy aspects of the stakeholders involved. This will be of great help to finalize the membership of Armenia in GEO and the related follow-up actions (e.g., establishment of a national GEO Committee).
- An Armenian environmental nSDI could facilitate the integration of other types of data and associated services. For example, integration of environmental data with geological data sets could make it possible to assess the risks of dangerous sites such as tailing repositories and various toxicants burial sites (Saghatelyan et al., 2013; Volfson et al., 2010).

3.3.5 Applying the EGIDA methodology to Armenia

The FP7 project EGIDA (Coordinating Earth and Environmental Cross-Disciplinary Project to Promote GEOSS; <http://www.egida-project.eu>) has produced a general methodological approach (Mazzetti et al., 2013) for implementing a (re-)engineering process of the existing Science and Technology infrastructures and systems, to be adopted at the national/regional level for a sustainable contribution to GEOSS and other relevant European initiatives. The EGIDA methodology has been adopted by the FP7 EOPOWER project (<http://www.eopower.eu>) to be applied to four pilot studies in which this methodology will be further improved. One of these pilot studies is Armenia, with the goal of "contributing to institutional Capacity Building in order to make Earth Observation resources optimally used towards sustainable development in Armenia". The realization of this pilot using the EGIDA methodology is therefore directly aligned with the vision of an Armenian environmental SDI. Applying the EGIDA methodology in this context will allow grounding CENS SDI into a fully operational tool that is effectively and sustainably linked to GEOSS, and this will therefore pave the way towards the realization of the Armenian environmental nSDI.

The EGIDA Methodology is based on a System of Systems approach, through the mobilization of resources made available from the participation in national, European and international initiatives and projects, hence it seemed applicable in the context of infrastructural and technological recommendations for Open Access to research data.

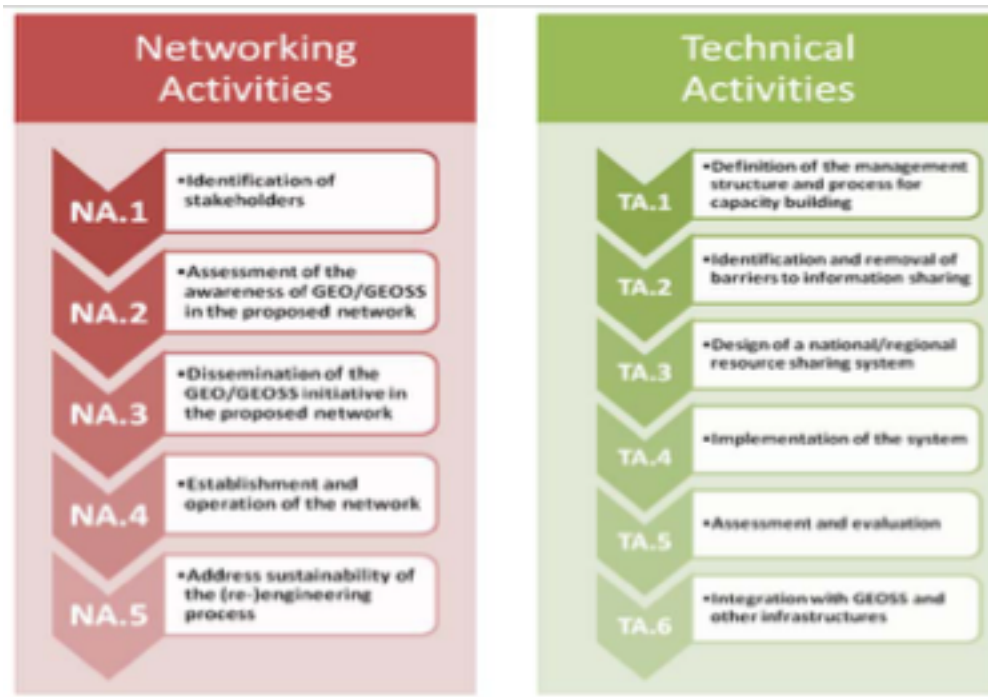


Figure 1 – Overview of the main networking and technical activities in the EGIDA methodology

As shown in Figure 1, the EGIDA Methodology defines two sets of activities running in parallel:

- Networking Activities: to identify and address the relevant Science and Technology community and actors (Community Engagement);
- Technical Activities: to guide the infrastructure development and align it with the GEO/GEOSS interoperability principles (Capacity Building).

For each activity several actions and sub-actions are defined, with related practices and guidelines derived from the design phase.

EGIDA defines three typical scenarios, depending of the scope of the planned action. The Armenia nSDI can be mainly related to EGIDA scenario S1, which is defined as:

S1) Regional/National Initiative Scenario: a national project aimed to deploy a national/regional infrastructure for sharing information relevant for GEO/GEOSS. Project partners may adopt the EGIDA methodology for an effective and efficient mobilization of resources, in order to design and develop the infrastructure making it a sustainable contribution to GEO/GEOSS (p. 13 in (Mazzetti et al., 2013)).

Both networking and technical activities described in the EGIDA Methodology could be applied and would benefit the implementation of the Armenian environmental nSDI.

Networking activities have the purpose of identifying and addressing the relevant science and technology communities and actors. Profiling and involving the stakeholders, for example by means of workshops, conferences, and thematic discussions, is crucial to build a network, as well as to reach out to potential new parties not previously involved. As workshops were already organized in Armenia by CENS, parts of the networking activities (e.g. the assessment of the awareness of GEO/GEOSS, and the dissemination of GEO/GEOSS initiative in the proposed network) are considered partly done, although future workshops (notably the EcoArm2ERA final event scheduled for October 2014 in Yerevan) will target other stakeholders and therefore enhance the networking activities. Previous work by CENS, and its existing contacts, will also facilitate the establishment and operation of the network, as formal/informal networking among key stakeholders is already in place.

Technical activities have the purpose to guide the infrastructure development and align it with the GEO/GEOSS interoperability principles (Capacity Building). As leadership is crucial in the implementation of a nSDI, a central coordination point will be established, similar to the GEO secretariat, which will be especially important if Armenia becomes a GEO member in the near future. To define the management structure and process for capacity building, we will identify the most relevant transversal areas of interest in the general environmental theme, for biodiversity, biochemistry, etc. Other technical aspects may be investigated (e.g. data policy). Due to time and resource limitations of the EOPOWER project, analysis activities such as the identification and removal of barriers to information sharing will not be performed. These may be implemented in the near future. Design activities will also be skipped, as the SDI already existing at CENS will be the node from which environmental nSDI and GEO membership will be established, by expanding it to also include other stakeholders' data. In fact, by experience from previous projects and close collaboration with Armenian partners, CENS appears as the most advanced center in the country in terms of SDI. As portal, metadata catalog and view/access services already exist, the implementation of the Armenia nSDI will focus on enabling advanced services (brokering, semantic discovery, etc.) and documentation. Integration with GEOSS should be straightforward, as CENS services are already registered with GEOSS.

As anticipated above, although recommended by the EGIDA methodology, and considered important for the positive outcome of the Armenia nSDI, we will not have the resources to carry out technical activities aiming at the identification and removal of the several types of barriers to information sharing that can arise (EGIDA considers behavioral, economical, legal, and technical barriers). These can be conceivably planned in the medium term, after the EOPOWER time frame. Concerning technical barriers, which could be perceived as an immediate hindering factor for the successful uptake of the Armenian nSDI, it is important to underline the *“technical barriers are related not to the will or possibility to share resources, but to the capability to do it. Some participants may be willing and authorized to share resources but are not able to do it.”* (p. 41 in (Mazzetti et al., 2013)). As recommended by the EGIDA methodology, *“technical barriers are removed through the availability of technical expertise and tools. These include the establishment of technical Task Forces, and training activities for individual capacity building such as workshops, summer schools, web lectures, etc.”*

In particular, EGIDA Guideline TA.2.1 will be beneficial to mitigate technical barriers:

“The existence and nature of obstacles to data sharing can be discovered and analysed through surveys and interviews. Members of the Stakeholders Network can provide information about behavioural, legal, technical and financial barriers to data sharing.” (p. 42 in (Mazzetti et al., 2013)).

3.3.6 Conclusion

At the heart of the South Caucasus region, Armenia is currently suffering from several environmental problems. Some of them could be addressed and mitigated through improved environmental modeling, forecasting and analysis. Data availability and integration, and the connection to international data sharing initiatives such as GEO could greatly help in this regard. Leading the way to an environmental national SDI appears therefore as a priority for the key national players (CENS and IIAP) that deliver environmental data and associated services for already a long time. The recent involvement of CENS and IIAP in several major international European projects and initiatives targeting the improvement of data sharing is a very timely and ideal situation to envision a long leap forward for Armenia. The EGIDA methodology is also well suited to scale-up the existing SDI infrastructure and to improve the Armenian network of stakeholders. The experience that will be gained through this process will in turn help to improve the EGIDA methodology, which will benefit other countries in the region and beyond.

3.4 Chapter key outcomes

- Capacity Building is an essential element to improve a SDI as it influences the individual, institutional and infrastructure components
- Capacity Building activities should be tailored to the audience to be efficient. For example, awareness raising activities (e.g. events, success stories) will better suit decision-makers whereas technical skills re-enforcement activities (e.g. workshops, training) will better fit GIS technicians.
- Combined events convening heterogeneous audiences split in different parts (e.g. one part for decision-makers and another part for SDI professionals) are a successful option.
- Various barriers can arise during capacity building activities, that might partly be addressed by best practices: (1) language issues: propose multi-lingual material or live translation, at least partial; (2) hardware issues (e.g. inadequate laptops performance, keyboards, internet connectivity): organize technical equipment for the workshop if possible; (3) heterogeneous skills in the attendance: define pre-requisites for attending the workshop and have modular material that can easily be accommodated to the audience.
- When planning Capacity Building activities, it is recommended to anticipate several issues: (1) measurable goals (e.g. attendees satisfaction, attendees SDI implementation through surveys) to assess the impact of the Capacity Building activities, which might be required for further funding; (2) long-term maintenance and upgrade of the material; (3) local long-term uptake of the Capacity Building activities by using a train-the-trainers approach, which requires committed key individuals in the process right from the beginning.
- Free and open source material (e.g. software, tutorials) should be promoted and used, as well as open standards and open data. This will have a considerable influence on the cost, interoperability and local uptake.
- When Building Capacity, it is important to provide attendees with a global picture of the various implicated issues beyond the technical ones. This is the role of awareness raising to set the scene and demonstrate the whole context of data sharing and benefits, which might influence the intangible factors linked to SDI (e.g. behaviors) and promote dialogue.
- The train-the-trainer concept is important to consider in a capacity building process as it gives the opportunity for local trainees to become in turn local, regional or national trainers. This can trigger a bottom-up

dynamic with understanding of local specificities that might have a regional impact.

- When building capacity at institutional or societal levels, it is important to identify and build a network of stakeholders (e.g. the PNF) that might be formal or informal.
- Leadership (individual and institutional) for coordinating such a network is essential and highlights the importance of key individuals that are committed on the long term and aware of the potential benefits.
- Capacity Building activities through a network can have a multiplier effect through increased activity of newly created committed communities creating new forms of cooperation.
- Institutional or Societal Capacity Building is a long-term process that is made of a set of Capacity Building activities and requires more resources than individual Capacity Building.
- International collaboration proved to be an essential element for Capacity Building in Armenia as it brings the necessary knowledge and resources (e.g. material, funding).
- To be successful, Capacity Building activities must remain simple, illustrative, positive (by showing potential benefits, success stories), customized (e.g. by using local examples) and flexible (e.g. modularity of workshops that can be adapted to various audiences). These elements are important so that Capacity Building activities are not only preserves of experts, but open to a large audience that will eventually contribute data.
- A successful national capacity building strategy targeting GEO/GEOSS principles will in fine be beneficial both for the country and GEOSS. This is illustrated with Armenia where the whole Capacity Building activities (e.g. BGSIP, EGIDA) were oriented towards GEOSS, and succeeded in making Armenia a new GEO member. Both GEO and Armenia benefit now from each other.
- The benefits of the successful process for Armenia are numerous: better national data flow, intra-country improved cooperation, benefit from global contributions in Capacity Building through GEO, more scientific exchange and partnering for international projects preventing brain drain to other countries, improved national skills.
- A country having successfully achieved its objectives can have a positive regional influence by becoming a trainer and undertaking in turn similar Capacity Building initiatives through collaboration, exchange of good practice and experience.

4 IMPLEMENTATION

4.1 Introduction on implementation

The chapter on stocktaking showed that SDI implementation is weak in Africa compared to other regions of the world due to technological, institutional, economic, social or political barriers. Developing countries in particular have to face financial and technological barriers on data sharing (Nativi et al., 2015). Some solutions have already been suggested to address some of these barriers but we want to focus in this chapter on solutions that might help to address technological and indirectly individual barriers to SDI implementation. This type of solutions is based on reduction of the technological complexity for the user while transferring it to another level that will not impact the user, which should lift some technological and individual barriers.

The fundamental concept underlying this approach is called interoperability, which is *“the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units”*⁴⁶. Interoperability can only be achieved through standardization, allowing the different parties to have the same understanding of the information to exchange. This is the role of the standards, which are technical documents describing all the necessary elements for a successful interoperability, qualified of “key enablers of interoperability” by the Global Spatial Data Infrastructure (GSDI) (Open Geospatial Consortium, 2004). In order to thoroughly exploit the vast amount of information and knowledge available worldwide, it is necessary to connect data, metadata and models over the Internet. This is possible by making them interoperable through standards.

The main organizations responsible for standardization in the geospatial fields are:

- The International Organization for Standardization (ISO)
- The Open Geospatial Consortium (OGC)

ISO and OGC are complementary in the sense that ISO describes the conceptual components of a standard and the relationships between these components whereas OGC focuses on publishing the specifications of implementation (Coetzee, 2011). But there are now several OGC standards that have also been adopted as ISO standards. The OGC developed many open geospatial standards⁴⁷, including web services interfaces that are “software systems designed to support interoperable machine-to-machine interaction over a network”⁴⁸. These web services are essential to implement the data, metadata and models interoperable connection over the Internet. The most commonly used OGC web services are:

⁴⁶ <https://www.iso.org/obp/ui/#iso:std:iso:19119:ed-1:v1:en>

⁴⁷ <http://www.opengeospatial.org/standards>

⁴⁸ <https://www.w3.org/TR/ws-gloss/#defs>

- Web Maps Service (WMS) (Open Geospatial Consortium, 2006a) allowing spatial data visualization through its portrayal representation as an image;
- Web Feature Service (WFS) (Open Geospatial Consortium, 2005) allowing vector data access;
- Web Coverage Service (WCS) (Open Geospatial Consortium, 2006b) allowing raster data access;
- Web Processing Service (WPS) (Open Geospatial Consortium, 2007c) facilitating the publishing of geospatial processes, that are algorithms or models for spatially referenced data;
- Catalogue service for the (CSW) (Open Geospatial Consortium, 2007b) allowing to publish and access digital catalogues of metadata.
- Sensor Observation Service (SOS) (Open Geospatial Consortium, 2007d)

Technological complexity, which might be a barrier to a wider data use and production, can be addressed in different ways using geospatial standardized web services. This chapter discusses innovative ways of reducing complexity in the geospatial domain through three different approaches: (1) data access customization; (2) heterogeneous data discovery and access; (3) metadata semi-automatic generation.

The first approach, data access customization, is user centric as it recognizes that users might be overwhelmed with the large amounts of spatial data available on the Internet, the difficulty to access it and the time required to process it so that it fits their needs in terms of format, scale, extent. For example, Bell-Pasht et al. (2015) or Hewitt et al. (2015) illustrate this by the advantages that companies or decision-makers would get from tailored climate information data and services for longer-term decisions and planning, early warning of potential hazards, and climate variability and change adaptation and mitigation. This might make a particular thematic data set of interest fastidious to discover, visualize, access and process. This might also be counter-productive in the sense that users might be discouraged to look for the datasets and maybe re-create such datasets themselves when this is possible, causing hence redundancy.

The first step to minimize these issues consists in promoting thematic geoportals, which allow an organization or a community of information users and providers to aggregate and share content and create consensus (Maguire and Longley, 2005). The next step consists in giving users the necessary tools to reduce as much as possible their effort in discovering, accessing and processing data. For example, giving user the possibility to interactively choose the dataset to download as well as the desired extent through a dedicated data extractor participates in a set of actions to: reduce the necessary download time, which is useful in a low bandwidth context; bypass the dataset extent processing, which saves time and necessary desktop skills for the user. These benefits for the user can only encourage him to access and use such spatial data, and potentially share findings and further data with the thematic community in return. The first article of the chapter, *“SCOPED-W: Scalable Online Platform for extracting Environmental Data and Water-related model outputs”*, proposes such an interactive approach

targeting the water community through a Soil and Water Assessment Tool (SWAT) use case dedicated platform.

The second approach recognizes the existence of heterogeneous spatial data formats and aims at promoting approaches and tools for discovering and accessing such data, which is of utmost importance in multidisciplinary contexts. The underlying concept of this approach is based on the brokering concept (Nativi et al., 2009a; Nativi et al., 2009b; Nativi et al., 2012). It starts from the ascertainment that even if data, metadata and models are standardized, they might use standards specific to a community of practice (CoP) instead of standards widely used. In order to keep the advantages of interoperability with this kind of CoPs, there is a need to make these specific heterogeneous CoP standards interoperable with more widely used standards so that they can also be discovered and accessed by clients, often implementing only the most used standards.

Two main architectural approaches to tackle this issue have been discussed (Nativi et al., 2009a; Nativi et al., 2012; Nativi et al., 2013a): the federation approach and the brokering approach, each having advantages and inconvenients. The federation approach pushes for common standards adoption whereas the brokering approach builds on existing data systems and federation systems by providing the necessary mediation and transformation functionalities. In other words, the federated approach requires the members of its ecosystem (e.g. a community of practice) to use the same standards, which might be a constraint in terms of time, funding or skills needed.

Alternatively, a brokering approach lowers the barriers for the CoPs as it is a middleware – a brokering layer with necessary components for ensuring interoperability between the various elements of the systems - that will be in charge of “translating” the heterogeneous standards into homogeneous ones. The best approach depends on the targeted systems to be interoperable. Nativi et al. (2012) state that federation approach works well in disciplinary infrastructures whereas brokering approach is better suited to build multidisciplinary and complex systems known as systems of systems. The second article of this chapter, called *“Enabling discovery of African geospatial resources”*, proposes the implementation of this brokering approach with Africa as a use case through the set up of an “Africa broker”. This use case proposes an innovative approach to gather in a same homogeneous geoportal useful African spatial data from diverse sources.

The third approach of this chapter focuses on metadata, that is a fundamental requirement for an efficient data discovery and should be given a high priority in SDI implementation (Masser, 2006) but remains unpopular and time-consuming. This leads to a situation where most data published in a SDI lack metadata (Trilles et al., 2014). One of the main reasons of this weakness is the decoupling between data and metadata; data being perceived as more important by users and producers, especially under time or financial constraints, metadata is often overshadowed in favor of data. Consequently, a solution needs to be found to either re-couple data with metadata, or to bypass the downside of decoupling through an semi-automatization of metadata creation/update.

This second solution is demonstrated in the third article of this chapter called *“Facilitating the production of ISO-compliant metadata of geospatial datasets”*, where OGC web services and open source geographic data server and metadata catalog are linked for obtaining automatic metadata. It is based on the possibilities offered by the ISO19115 metadata standard and its ISO19139 standard xml implementation, as well as the Catalogue Service for the Web (CSW) (Open Geospatial Consortium, 2007b) OGC web service, which is the key element of the proposed workflow. The innovative aspect of this approach gives another illustration of the benefits of interoperability and standardization for lifting technological barriers to data discovery.

4.2 SCOPED-W: SCalable Online Platform for extracting Environmental Data and Water-related model outputs (co-lead, submitted)

Yaniss Guigoz^{1,2*} / Pierre Lacroix^{1,2*}, Elham Rouholahnejad⁴, Nicolas Ray^{1,2}, Gregory Giuliani^{1,2,3}

* contributed equally to this work

¹University of Geneva, Institute for Environmental Sciences, EnviroSPACE Lab., 66 Bd Carl-Vogt, CH-1211 Geneva, Switzerland

²Global Resource Information Database (GRID) – Geneva, International Environment House, 11 chemin des Anémones, CH-1219 Châtelaine, Switzerland

³University of Geneva, Forel Institute, 66 Bd Carl-Vogt, CH-1211 Geneva, Switzerland

⁴Department of Environmental Systems Science, ETH Zurich, Universitaetstrasse 16, CH-8092 Zurich, Switzerland

4.2.1 Abstract

This paper presents SCOPED, an innovative approach for extracting environmental data using OGC services. In the field of water resource management, SCOPED-W ('W' for 'Water') is a method that was developed in the framework of EU/FP7 IASON and EOPOWER projects. This platform supports the collection of data required to build a Soil and Water Assessment Tool (SWAT) model and the uptake, spatialization and dissemination of raw data generated from the outputs of different SWAT models for the Black Sea region. Scientists are documenting the data served by the platform in ISO standardized metadata to support informed use. SCOPED-W primarily targets the community of SWAT users in the Black Sea region but it can easily be replicated in other geographical areas. Additionally, the SCOPED approach is based on data interoperability that makes it fully compatible with other domains of application as demonstrated here with three original use cases. The paper also highlights the benefits of the approach for the GEO community and discusses future improvements for supporting integration with other platforms such as UNEP Live.

4.2.2 Introduction

Environmental threats have become a major concern for humanity. At the economical level, it is estimated that losses from disasters such as earthquakes, tsunamis, cyclones, and flooding are now reaching an average of US\$250 billion to US\$300 billion each year, and future losses are estimated at US\$314 billion (UNISDR, 2015a). In order to better understand the threats and provide efficient solutions to reduce them, environmental analyses and modeling are necessary. This typically requires a huge amount of environmental data from diverse scientific disciplines. Additionally, given the broad spectrum of environment-related disciplines, communities of practice, and data formats, there is a strong

need for interoperability between technologies (e.g., operating systems, proprietary or open source GIS software), data, and standards (Nativi et al., 2012). Water resources in particular are a critical environmental component that already faces threats, both in terms of quantity and quality (IPCC, 2014a), and therefore is particularly deserving of attention.

Discovery of and access to existing environmental data and information is an essential pre-requisite to run environmental models or analyses. There is a growing interest among local and regional users to be able to replicate scale-specific models in their region (Stein et al., 2001).

However, discovery and access to ready-to-use data is often impeded by several technical factors such as the spatial and temporal scale, data format, geographic extent, and most of all interoperability of the data. This becomes even more problematic for users in areas with low bandwidth especially while dealing with large volumes of data (Mazzetti et al., 2014). Time and financial resources necessary for preparing the data at the appropriate scale and extent can also be a barrier especially for use in complex scientific models.

Much effort is currently being made to lower these barriers. Firstly, interoperability issues are being addressed by the Open Geospatial Consortium (OGC) (OGC, 2013b), the leading organization that develops open standards for geospatial data. The OGC standards comprise more than 35 standards for data visualization (Web Map Service – WMS) (Open Geospatial Consortium, 2006a), data access (Web Feature and Web Coverage Services – WFS, WCS) (Open Geospatial Consortium, 2005, 2006b), data processing (Web Processing Service – WPS) (Open Geospatial Consortium, 2007c), and data cataloguing (Catalog Service for the Web – CSW) (Open Geospatial Consortium, 2007b). Secondly, another essential and parallel effort consists in coordinating actions to raise awareness and promote implementation of spatial data infrastructures (SDIs) (Craglia et al., 2010; Giuliani et al., 2013c; Masser, 2005; Nebert, 2005). This effort is led by the Group on Earth Observations (GEO) (GEO, 2014c) and will support decision making through the provisioning of Earth Observation data. GEO coordinates a voluntary effort to build the Global Earth Observation System of Systems (GEOSS), a gateway between EO data and users (GEO secretariat, 2005a, 2008, 2011).

Standardization of data through OGC services has been discussed/implemented by many scientists (Castronova et al., 2013; Giuliani et al., 2013a; United Nations Economic and Social Council, 2013). A brokering approach for data discovery has been introduced by Nativi et al. (2012; 2013a), Giuliani et al. (2015) and others (Mazzetti et al., 2014; Pearlman et al., 2011) for integrating heterogeneous data coming from multidisciplinary domains. However, these methods redistribute thousands of existing resources (Giuliani et al., 2015), often as raw data, i.e. that does not fit the user's needs in terms of scale and format.

Consequently, an interoperable framework is needed for giving end-users the possibility to access customized data at a chosen geographical extent. Accessing data through a web framework is advantageous because it bypasses the need for desktop tools and expertise for processing the data. Such a framework would also reduce the time needed for downloading scientific data since data with a

reduced extent would be less voluminous, which reduces the technical barrier in low bandwidth areas. This would increase the interest of local/regional users to replicate scientific models in their region with their specific data and problems. In the field of hydrology, the “Soil and Water Assessment Tool” (SWAT) is widely used to predict the environmental impacts of land management practices, land use change, and climate change (Abbaspour et al., 2007; Gassman et al., 2007). Various efforts have been made to support the different phases of SWAT modeling activities. An attempt has been made by Giuliani et al. (2013c) to propose such an innovative, scalable, and interoperable framework, called OGC Web Services for SWAT (OWS4SWAT), that simplifies map and data production and facilitates exchange and integration of hydrological data with other sources. However, OWS4SWAT was limited by the types of SWAT output file that can be post-processed and the framework did not provide graphing tools.

Many other efforts to support the different phases of SWAT modeling activities have also been made. First, for data gathering and model preparation several software have been created to process SWAT inputs. Among others we can mention the Global Weather Data for SWAT, SWAT Weather Database, SWAT Precipitation Input Preprocessors (all available on the SWAT website: <http://swat.tamu.edu/software/links/>). These tools are mainly web interfaces or desktop applications that help connecting to a dedicated data provider and help users to download and handle data to get proper inputs for a SWAT model preparation. Compared to the OWS4SWAT framework (Giuliani et al., 2013c), these solutions have two limitations. First they only give access to a unique data source (e.g., NCEP CFSR for Global Weather Data), impeding users to access other data sources. Second, the access to data is not based on interoperable web services, which hampers efficient data sharing, makes data integration difficult, and limits the scalability of the software (e.g., use of other data sources or development of workflows). Consequently, implementing a framework entirely based on interoperable services could enhance data accessibility, reusability, and scalability.

The Hydrologic and Water Quality System (HAWQS⁴⁹) represents probably the most advanced effort in SWAT modeling. HAWQS is a web-based water quantity and quality developed by the US Environmental Protection Agency (US EPA) (White et al., 2015). It provides access to a set of databases, interfaces and models to evaluate the impacts of management alternatives, pollution control scenarios, and climate change scenarios of the quantity and quality of water from national to continental scales. This application helps users in preparing a SWAT model, generate SWAT input files, execute a SWAT model, analyze and download results. HAWQS can be considered as a specialized information management tool covering all major SWAT modeling aspects and enables users to explore results, scenarios, and alternatives. Though HAWQS is very well designed to execute SWAT models or scenarios exploration, it neither targets the same user audience nor has the same objectives as it does not provide extracting tools to facilitate the access of SWAT-related data sets to various types of users (e.g., SWAT specialists, environmental scientists, general public). Moreover, HAWQS does not have a module for publishing the results of SWAT models using interoperable

⁴⁹ <https://epahawqs.tamu.edu>

OGC services. Implementing such a module could enhance integration and facilitate linkages with other data sources or processing capabilities. It could help to incorporate SWAT results in data analysis workflows and could allow envisioning interactions with different scientific disciplines and coupling with other models.

Based on the identified gaps discussed above, we propose the SCOPED (Scalable Online Platform for extracting Environmental Data) approach that aims at providing a generic method to collect and extract environmental data and enables users to access data at specific scales and extents. SCOPED has been applied to the water domain in the framework of two EU FP7 projects (EOPOWER (EOPOWER, 2014) and IASON (IASON Consortium, 2015)). These projects led to the implementation of SCOPED-W ('W' for 'Water'), an innovative OGC compliant online platform for automatically preparing and serving water-related data with user-defined geographical extents.

More precisely, SCOPED-W has been developed to extract water-related data generated by the previous EU FP7 project enviroGRIDS (Lehmann et al., 2015). This project focused on the Black Sea watershed and produced numerous water-related data through the Soil and Water Assessment Tool (SWAT). Even though SWAT is a simulator of both hydrology and water quality, we focused on the hydrological aspects only due to water quality data limitations in some countries of the Black Sea Basin, which would have made the water quality modeling results unreliable. These data are available through OGC compliant web services (Giuliani et al., 2013d) and are a resource for the scientific community (<http://portal.envirogrids.net>). Although the platform currently focuses on the Black Sea catchment, other regions, thematic data, and existing products might also benefit from this automated, simple to use, and interoperable approach given the multidisciplinary nature inherent to environmental disciplines (George and Leon, 2007).

The five objectives of the SCOPED-W platform are the following:

- 1: address project specific requirements to create a hydrology-oriented module facilitating the assemblage of SWAT input data and optimizing the visibility and access to SWAT outputs and results;
- 2: give end-users the possibility to use these results at various geographic extents and/or open the way for replication in other regions than the Black Sea catchment;
- 3: perform automatic tailored geoprocessing functions through OGC standards while remaining simple to use;
- 4: be replicable for multiple domains in addition to water;
- 5: be easily integrated into other existing SDIs, platforms, and products.

This paper is structured as follows. Section 2 focuses on SWAT, its input and output data as well as its added value for specialists through the platform.

Section 3 gives details on the architecture, OGC services and data used in the platform. Section 4 suggests possible use cases while discussing the benefits, limitations and perspectives of the platform. In Section 5, conclusions are derived on the five objectives of the paper.

4.2.3 SWAT input and output

SWAT has been utilized in the framework of the FP7 enviroGRIDS project (Cau et al., 2013; Giuliani et al., 2013d; Gorgan et al., 2013a; Lehmann et al., 2014) to simulate the water quantity and water quality of Black Sea catchment with the area of 2 million km² over a 40-years period at a high spatial resolution and daily time step. The area of the catchment was divided into 12'982 sub-basins (Rouholahnejad et al., 2014). The sub-basins are further divided into 89'202 Hydrological Response Units (HRU). HRUs are the smallest calculation units in SWAT and consist of unique combination of land use, soil, and slope. The Black Sea SWAT model was calibrated and validated against a number of observation on stream flow and stream nitrate concentration at various sites. As there are often no data available on soil moisture, evapotranspiration, or aquifer recharge at such a large scale, crop yield was used as a surrogate to add confidence on the partitioning of water cycle components (e.g. infiltration and evapotranspiration) (Rouholahnejad, 2013). Further details on Black Sea SWAT model set up, development, and calibration can be found in Rouholahnejad et al. 2014.

SWAT requires a number of input data and generates a number of outputs, both are format specific. Minimal SWAT input data requirements include: digital elevation data, soil, land cover, river networks, and temperature and precipitation gauge data. NCDC climatic data (NOAA, 2015) and Climatic Research Unit (CRU) (Climatic Research Unit (CRU), 2008; Mitchell and Jones, 2005; National Center for Atmospheric Research Staff (Eds), 2014) are widely used as data sources for temperature and precipitation records. However, in order to include such core data in SWAT model, they need to be reformatted (e.g. in case of tabular data) and clipped on a specific geographic extent. As a consequence data often need some pre-processing work (e.g., clipping, reprojection, format transformation), given the fact that end users may be lacking the required GIS expertise/resources. Moreover, data discovery, preparation or extraction can be time consuming and users would usually largely benefit from pre-processed data. In addition, due to the fact that Black Sea catchment drains rivers of 23 European and Asian countries, and the few transboundary rivers in the basin, a dynamic routine for data and output extraction according to users' need would be of great interest.

SWAT-specific input and output variables such as precipitation, evapotranspiration, soil moisture, and groundwater recharge among many other variables are stored in several output files in tabular format at different temporal (at daily, monthly, and yearly time steps) and spatial resolutions (at sub-basin, HRU, or individual reach spatial scale). These tabular output files are very large and contain headers and special characters that require the following sequence of processing steps for being used in a geospatial context: (1) cleaning data by

removing header, special characters, and reformatting columns; (2) converting these large tables into separate tables (one table for each variable; one value for each sub-basin) in order to pick only the variables of interest; (3) joining these tables with geospatial data (e.g., a shapefile with sub-basins). An example of output.sub file is given in Figure 1.

```

1
SWAT May 8 2012    VER 2009/Rev 528                                0/ 0/ 0

General Input/Output section (file.cio):
15/11/2012 12:00:00 AM ARCGIS-SWAT interface AV

      SUB      GIS  MON   AREAkm2  PRECIPmm  SNOMELTmm    PETmm      ETmm      SWmm    PERCmm    SURQmm    GW_Qmm
TN03kg/ha
BIGSUB  1      0    1.16082E+03   12.200    0.135    3.417    3.312    211.680    0.000    0.000    10.065
0.000E+00
BIGSUB  2      0    1.15001E+03   12.200    0.135    3.457    2.992    211.814    0.000    0.000    8.778
0.000E+00
BIGSUB  3      0    1.11237E+03   12.200    0.135    3.471    2.687    211.883    0.000    0.000    7.625
0.000E+00
BIGSUB  4      0    1.64800E-01   12.200    0.000    3.474    2.116     0.000    0.000    0.000    0.000
0.000E+00
BIGSUB  5      0    1.16309E+03   12.200    0.135    3.488    2.896    211.813    0.000    0.000    9.273
0.000E+00
BIGSUB  6      0    1.37918E+03   14.400    0.000    2.646    1.603    145.309    0.000    0.000    0.000
0.000E+00
BIGSUB  7      0    1.19090E+03   14.400    0.000    2.669    1.614    118.846    0.000    0.000    0.000
0.000E+00
BIGSUB  8      0    1.57186E+02   14.400    0.000    2.674    1.628    143.611    0.000    0.000    0.000
0.000E+00

```

Figure 1: Example of a output.sub file generated by SWAT (first few lines and columns of the file)

Geographical visualization provides the means to explore both the information display and the data behind the information itself. In order to uptake and share the results of a SWAT model, the resulting layers can be published on a SDI. This provides the possibility for users to extract data on a given extent and enables them to do further detailed analysis on a desired geographic extent. These processing steps typically require working with many files and demand intricate processing capabilities from SWAT specialists (Yen et al., 2014). Consequently, automatization and tailorization of these processes, such as proposed in the SCOPED-W platform, represents an important added-value that makes it easier to re-use SWAT outputs at finer scales (Giuliani et al., 2013c).

SWAT results are not only of interest to SWAT experts, but also to various end-users: scientists of other disciplines (e.g., ecology, disaster management); decision-makers who need evidence-based data for environmental decisions; specific communities of users, such as the GEO communities of practice (GEO, 2010b). The metadata provided for each layer of the SCOPED-W platform through the enviroGRIDS catalogue (see section 3.2) is also of capital interest for decision-makers who need more information about the data layer and its relevance in a given context.

4.2.4 SCOPED-W: architecture and implementation

4.2.4.1 Architecture

The SCOPED-W platform is accessible through a web browser at <http://eopower.grid.unep.ch:8080/dataextractor/examples/swatappli.html>. It is

an expansion of the OWS4SWAT (Giuliani et al., 2013c) approach and contains five components. A Graphical User Interface (GUI: item 1 in Figure 2) gives the user three options for extracting SWAT-related data: by basin, by country, or by user-defined rectangle. The GUI also gives the possibility to access other useful SWAT-related data and tools. For each extraction, the user selects the layers of interest in the side menu and the desired extent. A geographic data server (item 2 in Figure 2) using the GeoServer software (Geoserver, 2010) allows to store and publish geographic layers through OGC web services (WMS, WFS and WCS). Data processing is based on PyWPS software (PyWPS, 2008) (see item 3 in Figure 2), an OGC WPS implementation written in Python language. Item 4 in Figure 2 consists of a series of scripts in Hyper Text Markup Language (HTML) and Javascript with the OpenLayers geospatial library. Data are retrieved in shapefile or geotiff formats on user's computer (item 5 in Figure 2). These five distinct components can easily communicate thanks to the interoperability provided by the OGC standards.

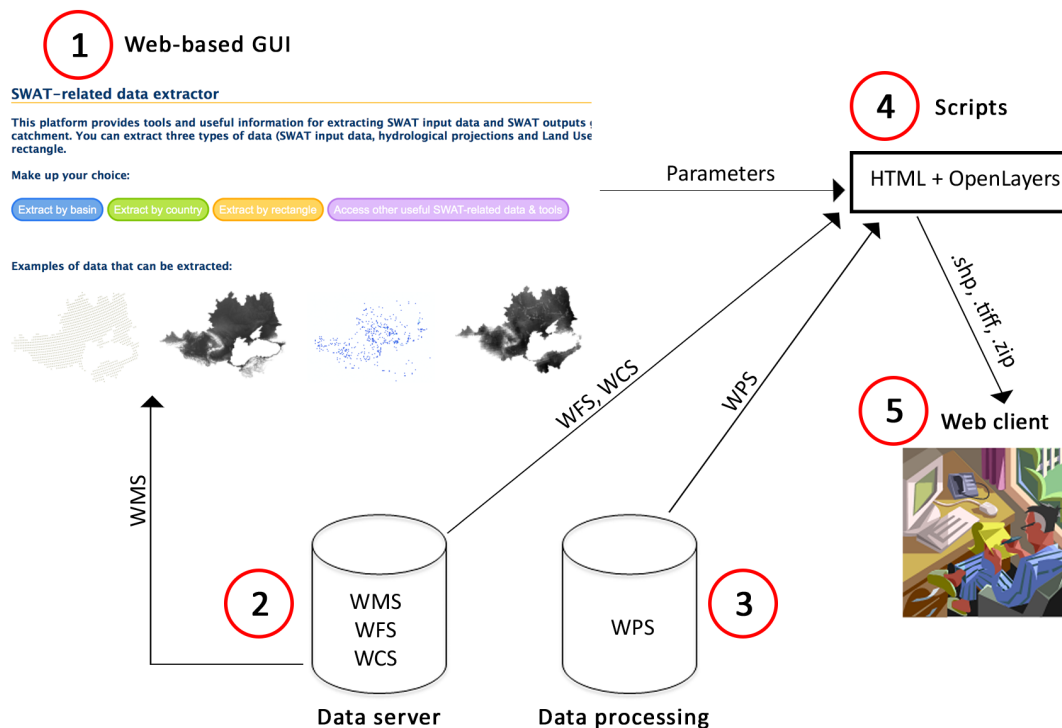


Figure 2: Architecture of SCOPED-W

All layers available in the platform for extraction are provided from the enviroGRIDS geospatial data publishing server. The advantage is that they are published as OGC web services (WMS, WFS and WCS), which are supported by geospatial libraries such as OpenLayers for further analysis and display in web applications. The OGC WMS, WFS, WCS and WPS web services are used as follows. The WMS is used for displaying the available basins and countries in the Graphical User Interface for choice of extraction. The WFS and WCS are used for exposing and downloading vector and raster geospatial layers, respectively. The WPS is used to apply the geoprocessing algorithms to the exposed WFS and WCS

layers. This geoprocessing part is performed on the server side in order to reduce resources use on the client side.

4.2.4.2 Data served by the platform

There are 40 data layers falling into four main categories.

(1) The first category entitled 'SWAT Input Data' contains six datasets that have been used as input parameters for the enviroGRIDS SWAT model (as recommended by Rouholahnejad, (2013): Shuttle Radar Topography Mission (SRTM) 90m Digital Elevation Data (Jarvis et al., 2008), Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type (Schneider et al., 2009, 2010), European Catchments and Rivers network System (ECRINS) rivers and lakes (European Environment Agency, 2011), and CRU precipitation and temperature datasets (Climatic Research Unit (CRU), 2008; Mitchell et al., 2005; National Center for Atmospheric Research Staff (Eds), 2014).

(2) The second category entitled 'Hydrological Projections' (13 layers of hydrological projections under future climate change and land changes scenarios) contains SWAT model outputs on green water flow, green water storage, blue water resources and water scarcity index map as well as a single layer for water resources with monthly and annual averages (1973-2006).

(3) The third category entitled 'Land Use Scenarios' contains 15 layers showing four scenarios of changes in land use based on interpretation (Mancosu et al., 2015) of IPCC's climate scenarios (IPCC, 2007) A1, A2, B1 and B2. The "HOT" scenario is based on IPCC's A1 climate scenario and expects the highest economic growth, with low population increase and high environmental pressure. The "ALONE" scenario is based on IPCC's A2 climate scenario and shows strong competition between agriculture and urban areas, regionally oriented economic growth and high environmental pressure. The "COOP" scenario is based on IPCC's B1 climate scenario and shows global cooperation, high economic growth, low population growth and low environmental pressure. The "COOL" scenario is based on IPCC's B2 climate scenario and shows an intermediate economic growth, a medium population growth and low environmental pressure.

Projections for years 2025 and 2050 are available for the four land use change scenarios, and can be extracted through the platform.

(4) Finally the fourth category 'SWAT output variables by subcatchment' includes six layers: annual nitrogen estimation (kg), annual flow estimation (m³/s), annual evapotranspiration estimation (mm), potential annual evapotranspiration estimation (mm), annual aquifer recharge estimation (mm), and annual precipitation (mm) (Rouholahnejad et al., 2014).

The layers under categories "hydrological projections", "land use scenarios" and "SWAT output variables by subcatchment" are results of the SWAT model execution and hence called SWAT outputs. This selection of layers was made

based on criteria linked to the platform requirements but the number or type of layers can be extended with additional work.

The extraction by rectangle is performed on the fly through a series of geoprocessing scripts that use the coordinates of user-defined extent. In the case of the extraction by basin or by country, layers of categories one ("SWAT Input data"), two ("Hydrological Projections") and three ("Land Use Scenarios") have been cropped on the desired extent and cached in order to minimize the delivery time. This is particularly useful for users in low bandwidth areas.

A different preparatory work was necessary for the layers part of the fourth category "SWAT output variables by subcatchment". It entailed a post-processing work on the SWAT model results that consists in two large tabular files of several millions records each with output.sub and output.rch extensions, so that these results can be included as geospatial layers into the platform. The following steps were applied:

- A sequence of chained Python and SQL scripts processes the output.sub and the output.rch files so that each output variable becomes available as an annual average instead of a monthly value. This processing resulted in 37 years (1970 to 2006) of data for each of the default 43 variables contained in output.sub and output.rch;
- Based on the model calibration (See Section 2) six variables were selected out of the 43 available ones: river nitrogen, river flow, actual evapotranspiration, potential evapotranspiration, aquifer recharge and precipitation per subcatchment;
- Each of these six variables was joined to a shapefile containing the delimitation of the Black Sea's subcatchments, based on a common field and saved as a new shapefile. Six shapefiles (1 per SWAT output variable) with 37 attributes (years 1970 to 2006) were then available at this stage;
- For quality check different persons compared randomly chosen resulting shapefiles with the original output.sub and output.rch files.

Each layer of the platform is linked to its metadata available in the enviroGRIDS metadata catalogue (enviroGRIDS, 2009) through its Universally Unique Identifier (UUID). This has two advantages: (1) to give access to standardized meta-information about the layer in a widely used metadata catalogue (GeoNetwork), allowing GIS users to discover the layer in a familiar environment; (2) to foster efficiency by avoiding redundancy and promoting use of already existing spatial data through discovery.

Finally, the platform homepage provides a link to a page with useful external SWAT-related data and tools. This page gives access to the following additional resources: the enviroGRIDS metadata catalogue to discover all the available enviroGRIDS data; an online tool provided by the National Climatic Data Center (NCDC) providing daily Climate Forecast System Reanalysis (CFSR) (National Centers for Environmental Prediction, 2010) data (precipitation, wind, relative

humidity, and solar) in a SWAT format for a given location and time period; a link to the SWAT-Plot and SWAT-Graph tools that generate graphs from SWAT output data; a link to the webpage of the Open Geospatial Consortium (OGC) standards; the endpoints of the enviroGRIDS web services.

4.2.4.3 Extraction workflow

The workflow for extracting data is streamlined, giving users the possibility to extract data in a few clicks on the desired area. The extraction workflow is presented in Figure 3 and detailed below:

1. Select extraction type: from the home page users can decide to extract data either by country (18 countries), by main river basin of the Black Sea catchment (26 basins) or by user-defined rectangle.
2. Select data to extract: after selection of the extraction type users are redirected to a page where they can choose from a dropdown list and select one or several layers to extract. They have the choice between 40 layers between the categories SWAT input data, hydrological projections, land use scenarios and SWAT output variables. Each layer can be dynamically previewed in the application when moused over.
3. View metadata: each layer has a link to a corresponding ISO19115/19139 metadata record in the enviroGRIDS catalogue.
4. Select extent for extracting: users select one basin or one country in the map, or drag a rectangle.
5. Extract: a request is sent to the WPS server. Python scripts are run and the data are cropped on the selected area. Results are delivered to the user as geotiff (for raster data) or shapefiles (for vector data) compressed in a zip file. This file contains one folder per layer. Layer names are prefixed with the basin name (e.g. INN_) or the ISO3 code of the country (e.g. BGR_).
6. Unzip: the zip file needs to be uncompressed for further use.
7. Use in desktop clients: the uncompressed layers can be displayed and further analyzed in desktop clients, e.g. QGIS, ArcGIS and uDig.

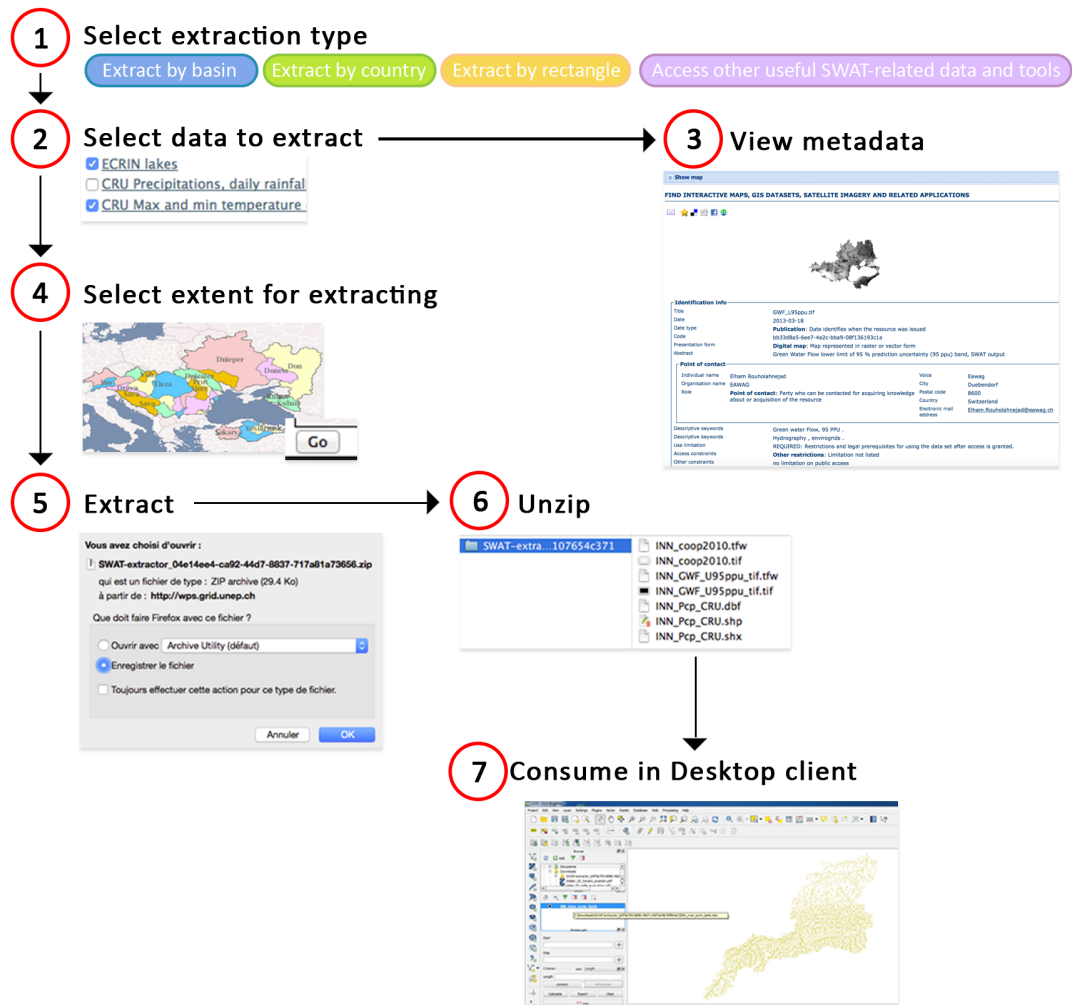


Figure 3: Extraction workflow

A summary of inputs and outputs of each step of the workflow is presented in Table 1.

Table 1: Inputs and outputs of the different processes

Process (as shown in Figure 3)	Inputs	Outputs
1. Select extraction type	Home page	New page with map and layer list
2. Select data to extract	Layer list	Layer(s) selected
3. View metadata	Layer	Metadata of the enviroGRIDS catalogue open in a new page
4. Select extent for extracting	Map	Basin or country selected, or

		rectangle dragged
5. Extract	‘Go’ button	Zip file with shapefiles and/or geotiff on the desired extent
6. Unzip	Zip file	Environmental data ready to be used in desktop clients
7. Consume in desktop client	Shapefiles and/or geotiff files on the desired extent	Further GIS analysis Integration of SCOPED outputs with other applications/projects Maps Decision-making

4.2.5 Discussion

To our knowledge, Cau et al. (2013), Cepek (2008) and Giuliani et al. (2013c) are the only authors to combine OGC services with SWAT inputs and outputs. The SCOPED approach has been applied to the water domain and implemented successfully as SCOPED-W. In this section we discuss the benefits, limitations, and possible improvements, as well as alternative applications of the approach outside the SWAT community.

4.2.5.1 Benefits of the approach

The SCOPED approach, and more specifically its SCOPED-W implementation for water-related applications, brings many benefits to the end-users and is in line with the objectives defined in the introduction of this paper.

The first benefit lies in its cross-platform compatibility, which makes it independent from the operating system of the end-user and therefore broadly accessible. On the client side no software is required except for a web browser, and tests were performed using the following web navigators: Internet Explorer 9.x or higher, Chrome 42.x or higher, Mozilla Firefox 37.x or higher and Safari 6.0 or higher.

Another advantage of the platform is that it is based entirely on open-source solutions and based on open standards. These are efficient solutions to share data and disseminate scientific results, both in transparent and interoperable ways. Users can reuse the code in other applications and projects or simply contribute to the improvement of the platform, e.g. by providing pieces of code for implementation of new functions. The possibility to download the data in widely used formats (e.g., shapefiles and geotiff) makes it useable in many GIS desktop application (e.g., ArcGIS, QGIS, uDig), which is a gain in term of

interoperability. Similarly, storage of the data in GeoServer exposes them through OGC standards, which enables communication between different software/server components and facilitates the discovery and access of data by different clients (both desktop and web-based). This might encourage the integration and exchange of hydrological data with other thematic data sources (e.g., climate, biodiversity). This is in line with the OGC vision and strategic goals (OGC, 2010).

The SCOPED approach can also be useful for the GEOSS to benefit from the potential of easy-to-use online (e.g., web and mobile) applications that demonstrate the value of standards-based access to environmental data and information. Currently, GEOSS is giving access to more than 65 million resources (GEO, 2014d). However, most of the resources available are raw data that are often difficult to use and/or interpret for scientists outside a specific scientific community. Therefore, GEO recognizes that geoprocessing tools are required to increase uptake of this data and provide understandable and useful information. The SCOPED approach is in line with these requirements and can offer a solution to address a specific user-driven problem using data served through interoperable services.

A particularly promising example of such interconnection of services is found in the environmental "big data" research stream. For example, the Australian Geoscience's 'Data Cube' (<http://www.datacube.org.au>) is a data infrastructure storing the exhaustive set of Landsat and MODIS satellite 25m-resolution images covering Australia for the past 30 years (totaling more than 1 Petabyte of data). The Data Cube can be thought as a high-end SDI for storing, accessing and performing analyses on EO data, as it runs on a large array of more than 10'000 computational cores from the Australian National Computing Infrastructure. The Data Cube is accessible by scientists worldwide upon request, and it has been used for example to get insights on high-resolution historical trends in country-wide flood events. As the Data Cube shared some of the same underlying software and standards (Geoserver, OGC web services) than SCOPED, one can foresee innovative links whereas SCOPED could be interfaced with the Data Cube (using the Data Cube API) to seamlessly interrogate its content, retrieve data for a given area of interest, and make it available as standardized web services to be input into environmental modeling tools.

The fact that the platform allows end-users to access geospatial data relevant to a specific scale (e.g., basin, region, country) is another important benefit and is in line with objectives 2 and 3. This consolidates grounds to make informed decision at a specific geographic level based on scale-specific data, which is usually more precise than national or supra-national (e.g., regional, global) data. More widely the platform is able – with additional data preparation – to integrate and serve data on other user-defined areas and delineations (e.g. provinces of a country, specific ecosystem area). This is interesting for scientists who want to replicate and run SWAT models in other regions, where it was never applied before, or obtain scale-specific SWAT output data for these regions.

A user-friendly interface has been developed to reduce complexity and to save time spent in searching and formatting data. The workflow for extracting the data is streamlined as was described in Section 3. This broadens the range of potential users by including also those with low GIS expertise who cannot analyze and use raw data generated by scientific models. More widely this can support scientists and in a second step decision-makers to make scale-specific and informed decision. In the case of SCOPED-W scientists from domains other than water resource management can integrate SWAT output data in their work and thus enlarge the scope of their research. This will be illustrated by three possible applications of SCOPED in the next section.

The SCOPED approach has also the potential to provide the users with an option to extract the SWAT input data (e.g., digital elevation model, soil, land use) to build region-specific SWAT models. The same approach had been used in Global Weather Data for SWAT (<http://globalweather.tamu.edu>) and the Climate Change Data for SWAT (<http://globalweather.tamu.edu/cmip>) where users are able to extract the desired data for the region of interest.

4.2.5.2 *Alternative applications of SCOPED outside of the SWAT community*

The SCOPED approach has primarily been developed for the SWAT community, meaning that it focuses on certain types of input and output data (e.g., temperatures, precipitations) relevant for particular societal benefit areas (SBA) such as water or soil. However, it can be extended to address the needs of other communities of practice with other types of data thanks to the use of OGC standards that ensure interoperability. Three possible use cases are presented.

(1) SCOPED can be extended with further developments to access other scientific data types such as sensor data that can be published using the OGC Sensor Web Enablement (SWE) standards (Lehmann et al., 2014; Simonis and Echterhoff, 2008). SWE standards enable developers to make all types of sensors, transducers and sensor data repositories discoverable, accessible and useable via the Web (OGC, 2015). In particular Sensor Observation Service (SOS) is a Web service specification defined by the OGC SWE group that aims at standardizing the way sensors and sensor data are discovered and accessed on the Web (Henson et al., 2009). 52°North has recently published an application programming interface (API) for advanced analysis and visualization of time series in maps and charts (52°North, 2015). The API integrates time series from many different sources, including the OGC SOS. This can be a good starting point for further extension of SCOPED to sensor data.

(2) In the Climate Science community there are important discussions about the concept of Climate Services. For the World Meteorological Organization (WMO) climate services are *“the dissemination of climate information to the public or a specific user”* (http://www.wmo.int/pages/themes/climate/climate_services.php). Having data readily available at appropriate scale and extent can allow timely communication of climate information, which is extremely important to prevent economical difficulties and humanitarian disasters that can result from changes in climatic conditions (GEO Secretariat, 2009; Geoff, 2011). The SCOPED

approach is therefore well suited to deliver efficient and effective climate services. Because it is based on the use of WMS and WCS standards it can easily integrate and expose climate data that are usually published in the NetCDF format (Domenico et al., 2006) using the Thematic Realtime Environmental Data Services (THREDDS) server (unidata, 2015).

(3) Another possible use case is when environmental monitoring or historical change detection is needed at a user-defined set of locations (points or polygons). The SCOPED approach could allow one to extract data from local, regional, or global data sets that are updated regularly and served through OGC services. Examples of such possibilities could be (i) deforestation analysis by offering change detection algorithm applied to temporal series of satellite imagery, (ii) historical analysis of landuse/landcover changes, (iii) live monitoring of small protected areas (e.g. wetlands, forest reserves) prone to adverse effects from changing local climatic conditions.

4.2.5.3 Limitations, possible improvements and perspectives

One of the primary objectives of SCOPED-W is to give users the easiest access possible to tailored geospatial SWAT-related data. However, some important limitations are currently imposed on the users.

The first limitation is that currently 40 SWAT layers considered as interesting by the authors are served by SCOPED-W. Enriching this list with new layers is possible but will require some time, for preparing the data and the metadata and publishing them. For example a contextualization of some of the available layers through normalization with population data would be an added value for end-users who could potentially access contextualized data such as SWAT outputs per capita. This could potentially highlight water issues in focused areas (e.g. inhabited areas) and hence provide critical additional information to decision-makers to reduce water stress. Such a development could take advantage of existing OGC services like the “Table Joining Service” (TJS) that enables linking statistical/attribute data with geospatial data.

Another limitation is that the extraction is currently limited to three types of extent (by country, by main river catchment, and by user-defined rectangle). It is also not possible to extract data at different geographic extents at once, for example for users who want to compare results of a model in different areas. Furthermore, SCOPED-W covers 18 countries and 26 sub-basins of the Black Sea catchment and any rectangular extent in this region. This means that the platform does not yet have the capacity to crop data on the fly using a free form user-defined area such as a sub-national administrative unit, a statistical unit, or an ecosystem area. To address these issues the platform would greatly benefit from publishing new cropping functionalities on the WPS server.

Similarly SCOPED-W could be integrated with other tools that support OGC services. A good example is the process brokering approach developed by the Institute for Atmospheric Pollution Research, National Research Council of Italy (Bigagli et al., 2015b). More specifically, a business process broker (BPB) is a distributed information system for creating, validating, editing, storing,

publishing, and executing geospatial-modeling workflows (Bigagli et al., 2015b; Colceriu et al., 2013). A BPB is a component that takes a formal description of a scientific business process, and translates it in an executable process which can be run on multiple and remote processing and workflow services. A BPB can be integrated with existing users workflows and it supports disaggregation of tasks into atomic processes. A BPB use case has recently been implemented by Mazzetti et al. (2015) where EO images are accessed by WFS (Process 1), mosaicked (Process 2), cropped on a user-defined area (Process 3), and published as OGC services (Process 4). The underlying technologies and standards are the same ones than for SCOPED-W (i.e., Python, GeoServer, PyWPS, WCS, WPS). In this specific BPB use case a processing service from the enviroGRIDS project for the Black Sea could be reused in the Mediterranean region by the IASON project. We see there a good opportunity to integrate the BPB approach with SCOPED-W, in particular for Process 3 and Process 4. The platform would serve more EO data and generate useful added value for various contextualized knowledge applications. For example UNEP-live (UNEP, 2014), which is the authoritative data platform of the United Nations Environment Programme (UNEP), could redistribute some of its regional and global data sets by hydrological basins.

Besides these possible improvements, automation of the links to the metadata records (currently static) would certainly improve and accelerate the publication of data from SCOPED-W. The platform could draw on the example of existing SDIs such as the Transboundary Waters Assessment Programme (TWAP) data portal (Transboundary Waters Assessment Programme (TWAP), 2015) that distributes hundreds of layers through OGC services with an automatic linkage to the metadata. The fact that the two SDIs are based on similar technologies and standards (GeoServer, GeoNetwork, OpenLayers; OGC services; ISO metadata) would ease the reproducing of this kind of functionality in SCOPED-W.

All these improvements, if implemented, would consolidate SCOPED-W, enhance data publication workflows, and enlarge the possible range of users. However, one of the challenges that hydrologists are facing is the increasing amount and complexity of input and output data. One way to tackle this is through distributed computing such as Grid computing that has found numerous applications in hydrology (Lecca et al., 2011). Distributed computing could also benefit the SCOPED approach if large and complex data sets must be processed. Distributed WPS services for example (see e.g. Gorgan et al. (2012c); Astsatryan et al. (2015)) could reduce the time needed for pre-processing and extraction of data.

4.2.6 Conclusions

The SCOPED approach provides a framework to collect and extract environmental data and enables users to access data at the most suitable spatial scale for their needs. More specifically the SCOPED-W platform gives end-users the possibility to visualize and access selected pre-processed SWAT data from the enviroGRIDS project, in accordance with expectations of objective 1.

The interactivity of the platform allows users to select SWAT raw data at a geographic extent of their choice. Additional layers of regions other than the Black Sea catchment can easily be added into the platform. This scalability and reproducibility fulfills objective 2.

The outcomes of objective 2 are made possible through interoperability proposed by the OGC standards used in the platform that link the various components of the platform, to ensure an automatic tailored geoprocessing workflow on the back-end while reducing complexity for the user as requested in objective 3.

Regarding objective 4, the reusability of the platform is not only possible for other regions by simple addition of new layers but also for other themes than water (e.g., climate services, land cover change).

Finally, the requirement of modularity and flexibility of objective 5 for integration in other products is also reached as individual components of the platform can be integrated in other existing products (e.g., GEOSS, UNEP-Live) with minor adjustments. This allows reaching broader communities of users (e.g., scientists, policy-makers) by giving them access to data already transformed into usable information.

Despite the defined objectives, the platform still needs some improvements for a complete automatization of the workflows (i.e., for having the pre-processing work completely integrated in a WPS chain). However, this platform opens new horizons to potential high-level technological or institutional developments, for the benefit of the EO global community.

4.3 Enabling Discovery of African Geospatial Resources (co-author, published)

Gregory Giuliani^{a,b,*}, Fabrizio Papeschi^d, Andiswa Mlisa^c, Pierre Lacroix^{a,b}, Mattia Santoro^d, André Nonguierma^e, Jan Cools^f, Yaniss Guigoz^{a,b}

^aUniversity of Geneva, Institute for Environmental Sciences, enviroSPACE Lab., Battelle – Building D, 7 route de Drize, CH-1227 Carouge, Switzerland.

^bUnited Nations Environment Programme, Division of Early Warning and Assessment, Global Resource Information Database – Geneva, International Environment House, 11 chemin des Anémones, CH-1219 Châtelaine, Switzerland.

^cGEO Secretariat, 7 bis avenue de la Paix, 1211 Geneva

^dCNR-IIA, Via Madonna Del Piano 10, 50019 Sesto Fiorentino, Firenze

^eUnited Nations Economic Commission for Africa, Information, Science & Technology Division, Menelik II Ave., Addis Abeba, Ethiopia

^fMilieu Ltd, Chaussée de Charleroi, 112, 1060 Brussels, Belgium

*Corresponding author: gregory.giuliani@unige.ch, +41223790709,
+41223790744

4.3.1 Abstract:

In Africa, natural resources are degrading, while being at the same time essential for maintaining or improving people's livelihood. The well being of African communities is highly correlated to changes in local ecosystem services. Their vulnerability to degradation of natural resources is extremely high and resilience against natural changes (e.g. climate variability) and socio-economic changes (e.g. fluctuations in food markets) is low.

Nowadays, it is widely accepted that reversing these trends and adapting to climate change require integrated responses tackling the underlying social, economic, political and institutional drivers of unsustainable natural resources use. Integrated approaches intrinsically ask for cooperation, exchange of information and communication to better understand complex interactions and assess environmental issues. Understanding these interactions requires collecting and integrating various data describing physical, chemical, biological and socio-economical conditions. However two common obstacles are currently preventing the implementation of such a integrated approaches: (1) difficulties to find data, and (2) difficulties to integrate data.

In response of these issues, this paper presents the *Africa Discovery Broker*, a web-based tool that enables users working in different domains to search through and access 32'442 heterogeneous African geospatial resources (e.g. remote sensing, geospatial data, socio-economical data) coming from 17 international, regional, national and research projects repositories.

4.3.2 Introduction:

Threats to the environment and natural resources, coupled with poor management, have serious implications for both poverty reduction and sustainable economic development (OECD, 2008). A proper natural resources management actually requires safeguarding food production, preserving livelihoods and socio-economic development. Various global environmental assessments (UNEP, 2005, 2006; UNEP/UNDP, 2009) have shown a continuous decline of natural resources, increasing the vulnerability of the poor as a result of ecosystem stress, competition for space, soaring food and energy prices and climate change.

Nowadays, it is widely accepted that reversing these trends and adapting to climate change require integrated responses tackling the underlying social, economic, political and institutional drivers of unsustainable land and water use. Many sectors and disciplines have developed integrated management frameworks such as Integrated Water Resources Management (IWRM),

Integrated Natural Resources Management (INRM), Integrated Coastal Zone Management (ICZM), Community-based forest management (CBFM), or Integrated Soil Fertility Management (ISFM) (GEO, 2014b; Rebelo et al.; Tripathi and Bhattacharya, 2004), to name but a few. Common elements in these frameworks are the integration of social and natural systems, the integration of different kinds of knowledge, the integration of different actors, stakeholders and institutions, and the integration across scales and sectors. Integrated responses require multiple instruments for their implementation and ask for fundamental shifts in governance institutions in terms of skills, knowledge capacity and organization (UNEP, 2005).

In Africa, natural resources are degrading, while being at the same time essential for maintaining or improving people's livelihood. The well being of African communities is highly correlated to changes in local ecosystem services. Their vulnerability to degradation of natural resources is extremely high and resilience against natural changes (e.g. climate variability) and socio-economic changes (e.g. fluctuations in food markets) is low.

Integrated approaches intrinsically ask for cooperation, exchange of information and communication to better understand complex interactions and assess environmental issues. Understanding these interactions requires collecting and integrating various data describing physical, chemical, biological and socio-economical conditions (e.g., population, ecosystems, biodiversity, vegetation, land cover, soils, water, wetlands, biomass) (Vicente-Serrano et al., 2012). These data have in common the description of a geographical location through a set of attributes and can be considered as geospatial data. Geospatial data & information have been recognized as essential for socio-economic planning and development (Ayanlade et al., 2008; Tripathi et al., 2004). Data describing the environment demonstrate their full potential when combined with other data sets allowing one to monitor and assess environmental status at different scales (e.g., global, regional, local), discover complex relationships between them, and to model future changes. This combination allows transforming data into information that can be used by decision-makers.

To facilitate geospatial data production, management, analysis and dissemination, Spatial Data Infrastructures (SDI) have been widely adopted (Craglia et al., 2012a). African countries have also embraced the concept of SDI (Lance, 2003; Rajabifard and Williamson, 2001b) but at a slower rate. Access to geospatial data of high quality is a pre-requisite for many stakeholders involved in various fields of activities. Therefore, it is a necessity to find, access, and integrate various types of data coming from different scientific or non-scientific sources. In other words, a multi-disciplinary geospatial framework is required to support efficiently and effectively integrated approaches like INRM or IWRM. However two common obstacles are currently preventing the implementation of such a framework: (1) difficulties to find data, and (2) difficulties to integrate data (Cooper and Gavin, 2005; Lance, 2003; Woldai, 2002). In scope with this paper, we only mention here the technical and data obstacles while a whole SDI also encompasses other aspects (e.g., laws, people and institutions). The amount of geospatial data is quickly growing but these data are not necessarily easy to access as they are often « siloed » in different locations (Gore, 1998). This leads

to useless duplication of efforts because users tend to re-create data that already exist. When accessible, these data might be very heterogeneous and hardly interoperable as they come from different disciplines and are based on different technologies, arrangements, protocols and formats.

The primary function of any SDI is data discovery, enabling users to search and evaluate data before accessing them (Nebert, 2005; Nogueras-Iso et al., 2005b). The fundamental requirement for an efficient and effective data discovery mechanism is that data is properly documented with metadata and stored in a catalog (Charvat et al., 2013a; Foresman, 2008). Otherwise without appropriate metadata a SDI will fail in its main objective of facilitating discovery and access to geospatial data (Masser, 2005). Unfortunately, most of the data produced are poorly documented or even worst are simply lacking metadata (Cooper et al., 2005; Guigoz et al., 2016; Woldai, 2002)). Potentially, there are a lot of useful data repositories both inside and outside Africa but unfortunately most of these data are hidden to users simply because they are difficult to discover. In order to support integrated frameworks and environmental assessments in Africa a solution for facilitating data discovery and access across various disciplines is fundamental.

Based on these considerations the aim of this paper is to present a proof of concept for an *Africa Discovery Broker*, a web-based tool for facilitating the discovery of heterogeneous geospatial resources in Africa (e.g. remote sensing, geospatial data, socio-economical data).

4.3.3 Methodology, implementation and preliminary results

Discovering existing geospatial resources in Africa supposes that these resources are structured in a way that allows reaching them, ideally through recognized standards, hence the importance of a SDI. If several frameworks exist to assess SDI status at national level (Eelderink et al., 2008b), such an assessment is not so obvious at continental level. Nevertheless, Europe is doing well with its INSPIRE State of Play (Vandenbroucke, 2010b). For Africa, an assessment methodology has also been proposed (Guigoz et al., 2016). The outcomes of this assessment show that Africa stands behind most other parts of the world in terms of SDI, even though regional differences exist. The SDI effort on the African continent is led by the UN Economic Commission for Africa (UNECA), that plays a key role in Africa (Schwabe and Govender, 2009b) and has been trying for years to promote the SDI concept across the continent, in particular through the Committee on Development Information, Science and Technology (CODIST) who hold regular meetings and comprise of SDI actors from African countries.

Despite this situation, some geospatial data repositories for Africa can be found on the Internet but are often linked to specific projects or themes. Some of these repositories contain metadata with links to the data custodians; others provide data for download while others are web services that allow for direct visualization of the data and metadata, which makes them more easily discoverable. One can mention the following data repositories focused on Africa: SERVIR-Africa (<https://www.servirglobal.net/Africa.aspx>), SERVIR Eastern & Southern Africa (<http://servirportal.rcmrd.org/>), ECOWREX

(<http://www.ecowrex.org/mapView/>), Africa Soil Information Service (<http://africasoils.net/services/data/>), Open Data for the Horn (<http://horn.rcmr.org/>), CREST (<http://ags.servirlabs.net/crestviewer/>), South Africa National Spatial Information Framework (<http://www.sasdi.net/>). SERVIR-Africa, SERVIR Eastern and Southern Africa as well as Open Data for the Horn are all hosted at the Regional Centre for Mapping of Resources for Development (RCMRD - http://rcmr.org/?page_id=4970) that is the regional center for Eastern Africa closely linked to UNECA. In terms of metadata, the most widely used standard is the ISO19115/19139. UNECA has contributed to the development of an African profile of this ISO standard for metadata. According to (Guigoz et al., 2016), only four African countries have adopted an official metadata standard: Botswana, Ethiopia, Nigeria, Senegal, and South Africa. But despite this official adoption, very scarce African metadata and data is available on the Internet, partly because very few African national geoportals such as the South African Earth Observation System of Systems (SAEOSS - <http://sageo.org.za/data-portals/saeoss/>), the Malawi Spatial Data Portal (MASDAP - <http://www.masdap.mw>) or the South African Environment Observation Network (SAEON - <http://data.saeon.ac.za/>) are currently accessible. RCMRD is outstanding through its metadata catalogue (<http://servir.rcmr.org/metacatalog/>).

Based on what precedes we can consider that SERVIR-Africa is a key resource to access/discover data across Africa. The SERVIR project builds upon a partnership involving eighteen African countries and provides through its metadata catalog access to 3947 resources (as of March 2015) in various areas such as biodiversity conservation, disaster management, agricultural development, and climate change adaptation. However, in a multi-disciplinary framework it would be better off being crossed with other data repositories to avoid users spend time searching in many different catalogs of data.

It is commonly known that approximately 50% of time is lost in searching data while doing environmental assessments (Craglia and Campagna, 2009a). Therefore facilitating data discovery across disciplines will certainly help to lower this percentage and will give more time to perform data analysis, a crucial step to better understand complex environmental issues and interactions.

Another issue with multi-disciplinary frameworks is that each discipline involved uses different technology, arrangements, protocols and formats to publish its resources. In order to make these various resources discoverable and interoperable it should not be requested to change or impose interoperability arrangements within a specific community but rather to lower entry barriers for both data users and providers. To tackle this issue, the Group on Earth Observation (GEO) (Nativi et al., 2012; Nativi et al., 2013a; Nativi et al., 2011; Vaccari et al., 2012) has adopted a brokering approach to implement multi-disciplinary interoperability within the Global Earth Observation System of Systems (GEOSS): *“Users and Data Providers are not asked to implement any specific interoperability technology but to continue using their tools and publishing their resources according their standards -as much as possible”*. As a System of Systems, GEOSS is composed of contributed Earth Observation systems. Through the GEOSS Common Infrastructure (GCI), GEOSS provides access to GEOSS

contributing systems – that operate independently within their own mandates. The GCI is a third-party layer that is in charge of transparently interconnecting GEOSS systems with GEOSS Societal Benefit Areas (SBAs) users (Nativi et al., 2013a). Due to the distributed and autonomous nature of GEOSS systems, an important requirement for the GCI was to be able to change dynamically the bindings – i.e. interconnections – between data providers and users. This was achieved by implementing the brokering approach; based on this approach, client applications – i.e. data users – and servers – i.e. data providers – are separated by a new intermediary component called *Broker*. When a client application needs a resource, it sends a query to the broker. The broker then forwards the query to connected servers, which process the request. Finally, the broker aggregates results and returns the result set back to the client application. The GCI builds on a broker for each main functionality: discovery (Nativi et al., 2009a), access (Boldrini et al., 2013), and semantic interoperability (Santoro et al., 2012). All the GCI brokering components are part of the GEO Discovery and Access Broker (DAB) framework. Currently, the GEO DAB brokers about 40 systems, providing access to about 14 million complex resources (e.g. time series) and more than 80 million individual resources (e.g. single satellite scenes) (Nativi et al., 2015).

To facilitate the discovery of heterogeneous resources across Africa, the *Africa Discovery Broker* (ADB) has adopted a brokering approach using the caching and mediation capabilities proposed by GI-cat (<http://essi-lab.eu/cgi-bin/twiki/view/GIcat/>) to broker heterogeneous resources (data catalog and access services) (Nativi et al., 2009c). GI-cat is an implementation of a broker catalog service developed by ESSI-Lab in the frame of the EU/FP7 EuroGEOSS project (<http://www.eurogeoss.eu>). It allows data providers to publish various catalog interfaces, enabling different clients to discover and evaluate geospatial resources over a set of heterogeneous data sources. A data provider can deploy his/her own GI-cat instance, grouping together disparate data sources, to accommodate his/her users' needs. GI-cat can access numerous catalog services, as well as inventory and access services to discover, and possibly access, heterogeneous resources. Specific components implement mediation services for interfacing heterogeneous service providers, which expose multiple standard specifications; they are called Accessors. These mediating components map the heterogeneous providers' metadata models into a uniform data model that extends the ISO 19115 Core profile. Accessors also implement the query protocol mapping; they translate the query requests expressed according to the interface protocols exposed by GI-cat, into the multiple query dialects spoken by the resource service providers. Currently, 45 discovery interfaces and 56 specifications are supported (Nativi et al., 2015), including OGC WCS (Open Geospatial Consortium, 2006b), OGC WMS (Open Geospatial Consortium, 2006a), OGC WFS (Open Geospatial Consortium, 2005), OGC WPS (Open Geospatial Consortium, 2007c), OGC SOS, OGC CSW (Open Geospatial Consortium, 2007b), THREDDS (Domenico et al., 2006), CDI, GBIF, GeoNetwork, Deegree, ESRI ArcGIS Geoportal, OpenSearch, OAI-PMH, NetCDF, NCML, ISO19115, GeoRSS, GDACS, DIF, File System, SITAD, INPE, HYDRO, and WaterML (Open Geospatial Consortium, 2007a). A complete list of supported sources and available catalog interfaces can be found at: [http://essi-](http://essi-lab.eu/cgi-bin/twiki/view/GIcat/)

lab.eu/do/view/GIcat/GIcatDocumentation.

A Discovery Broker can therefore transform query results to a uniform and consistent interface implementing metadata harmonization and protocol adaptation. Consequently, the *Africa Discovery Broker* enables users to search across various geospatial resources across Africa and to easily discover data that can fulfill their requirements (Fig.1).

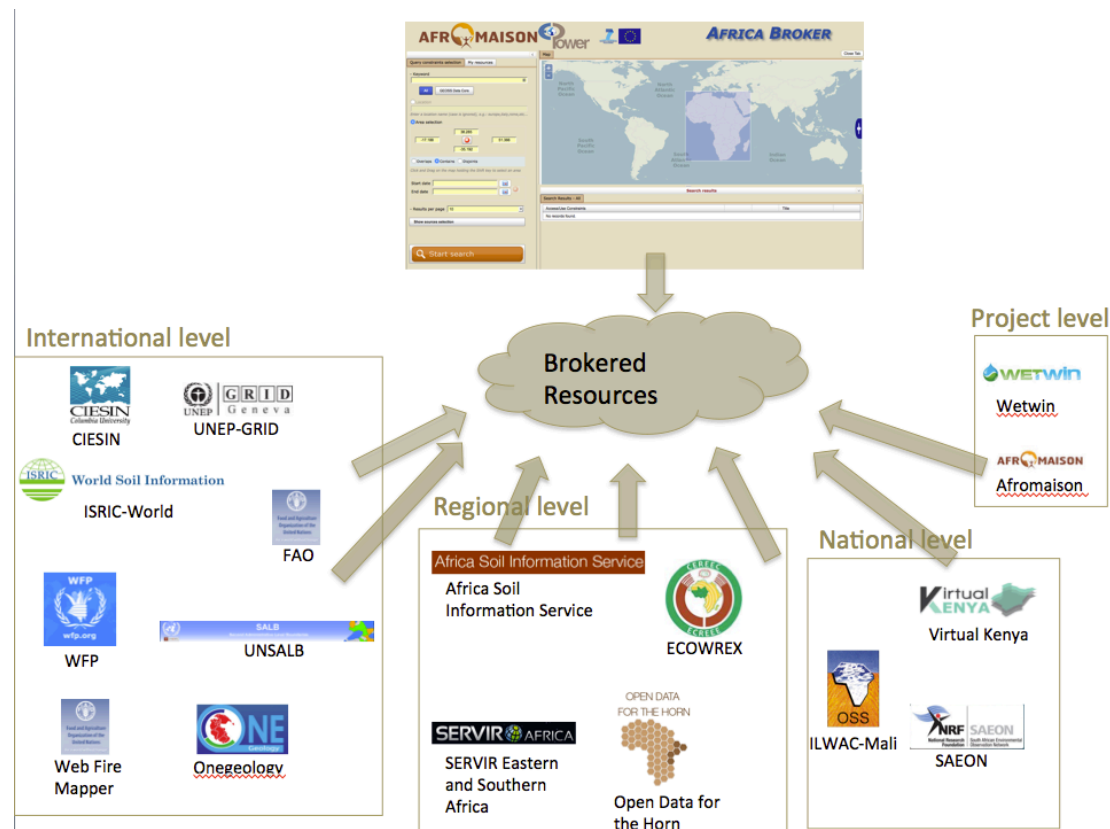


Figure 1: The concept of brokering African geospatial resources bottom: heterogeneous resources, middle: GI-cat, top: common interface)

Various heterogeneous resources have been registered (Tab.1) coming from various international (e.g., UNEP, FAO, UN WFP), regional (e.g., SERVIR-Africa, ECOWREX, Africa Soil Information Service), national (SAEON, ILWAC-Mali, Virtual Kenya) repositories and research projects (FP7 Afrimaison, FP7 WetWin). The *Africa Discovery Broker* will take care of the harmonization process and will expose results in a consistent way.

Table 1: *Africa Discovery Broker*, registered resources and links (as of March 2015)

Resource name	Endpoint	Type of service	Scale	Themes	Number of records
UNEP-GRID	http://metadata.grid.unep.ch:8080/geonetwork/srv/eng/csw?	CSW	International	Various	69
FAO	http://www.fao.org/geonetwork/srv/en/csw?	CSW	International	Various	6882
UNSALB	http://salbgeonetwork.grid.unep.ch/geonetwork/srv/en/csw?	CSW	International	Boundaries	131
OneGeology	http://onegeology-catalog.brgm.fr/geonetwork/srv/en/csw?	CSW	International	Geology	438
UN WFP	http://geonode.wfp.org/catalogue/csw?	CSW	International	Various	2204
Web Fire Mapper	http://geonetwork4.fao.org/wms/wms.php?	WMS	International	Fires	2
ISRIC – World Soil Information	http://meta.isric.org/srv/csw?	CSW	International	Soil	2156
Virtual Kenya	http://maps.virtualkenya.org/catalogue/csw?	CSW	National		626
South African Environmental Observation Network (SAEON)	http://app01.saeon.ac.za/PLATFORM_TEST/MAP/csw.asp?	CSW	National	Oceanography	14818
Integrated Land and Water Management for Adaptation to Climate Variability	http://ilwac.oss-online.org/ml-ilwac-gn2_10/srv/eng/csw?	CSW	National	Land, Water, Climate	365

y and Change (ILWAC) - Mali						
FP7 Afromais on	<a href="http://afromaison.grid.unep.ch:8080/geonetw
ork/srv/en/csw?">http://afromaison.grid. unep.ch:8080/geonetw ork/srv/en/csw?	CSW	Project	Various	531	
FP7 WetWin	<a href="http://sditest.unesco-
ihe.org:8080/geonetwo
rk/srv/en/csw?">http://sditest.unesco- ihe.org:8080/geonetwo rk/srv/en/csw?	CSW	Project	Various	18	
Open Data for the Horn	<a href="http://horn.rcmrd.org/
catalogue/csw?">http://horn.rcmrd.org/ catalogue/csw?	CSW	Regional	Various	123	
ECOWRE X	<a href="http://www.ecowrex.o
rg/geoserver/ows?">http://www.ecowrex.o rg/geoserver/ows?	WMS	Regional	Energy	24	
SERVIR - Eastern and Southern Africa	<a href="http://servirportal.rcm
rd.org/catalogue/csw?">http://servirportal.rcm rd.org/catalogue/csw?	CSW	Regional	Land cover	28	
Africa Soil Informati on Service	<a href="http://ciesin.columbia.
edu/geoserver/afsis/o
ws?">http://ciesin.columbia. edu/geoserver/afsis/o ws?	WMS	Regional	Soil	80	
SERVIR- Africa	<a href="http://servir.rcmrd.org
/metacatalog/csw/disc
overy?">http://servir.rcmrd.org /metacatalog/csw/disc overy?	CSW	Regional	Various	3947	

To ensure that data discovery is restricted to the African continent and enables search for data in repositories registered into the ADB, a tailored web-based application has been customized and is available at: <http://afromaison.grid.unep.ch/gi-cat/gi-portal>. It is centered on Africa by default, has customized colors close to the Afromaison project look and feel as this application was originally built during this project. Finally, it has specific logos making reference to the supporting projects (Fig.2).

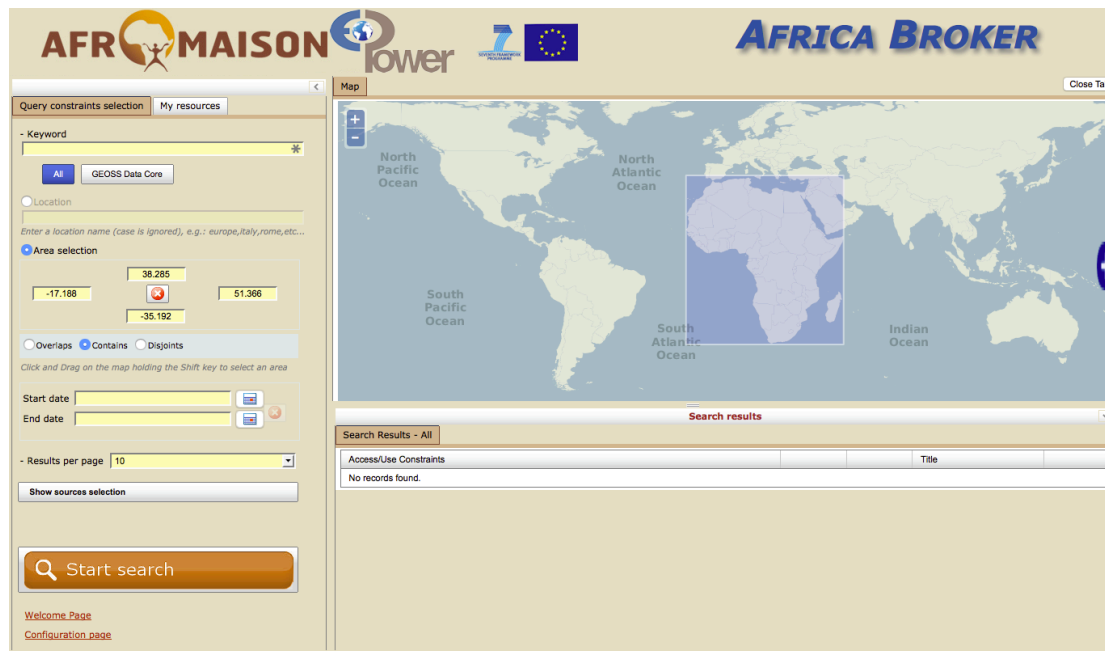


Figure 2: The *Africa Discovery Broker* customized portal

Queries can be formulated by composing the desired constraints corresponding to simple user needs (e.g. Where, What, When, Who). The results will be displayed on the map and listed in the bottom part of the screen.

- The *Where* constraint can be selected by using the mouse, by a selection directly on the map. The corresponding fields under selected area (W-N-S-E) will be updated. Alternatively the values can be entered directly from the keyboard.
- The *What* constraint can be inserted using the Keyword field in the top right part of the screen.
- The *When* constraint can be selected using the calendars inputs under the "Time" section at the right.
- The *Who* constraint can be selected from the left frame (hidden by default).

Once the desired constraints are selected (all are optional), the search/query can be run and results will be displayed in tabular list of matching resources (if any). Each row contains information on specific matching resources and buttons to perform further actions on it (e.g., view on map, view metadata).

Currently 17 repositories have been registered in the African Discovery Broker giving access to more than 32'442 resources. This allows to link resources published by data providers working in various disciplines and at different scales, adapting them to the tools commonly used by data users. The *Africa Discovery Broker* exposes several interfaces, including the OGC CSW/ISO, GI-CAT, REST, OpenSearch, OAI-PMH, CKAN, and ESRI-Geoportal. This enables various clients (e.g., QGIS: see Fig.3) to query directly the ADB without the need to use customized web applications.

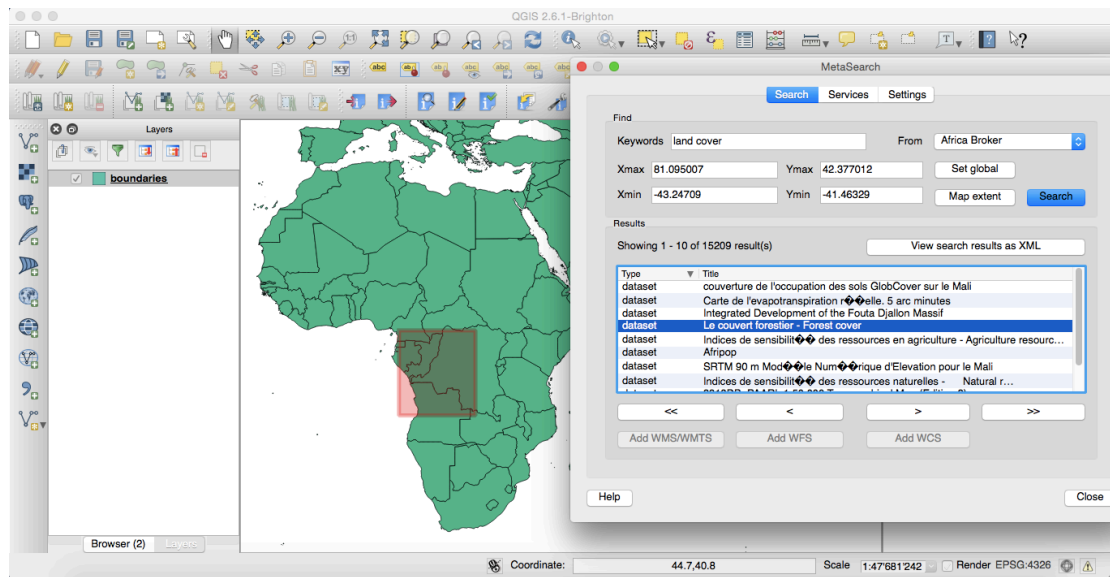


Figure 3: The *Africa Discovery Broker* directly queried in the QGIS application using the published CSW interface.

4.3.4 Discussion:

First tests/results show that the proposed proof-of-concept allows interconnecting heterogeneous data sources coming from various areas and at various geographical scales in a common, coherent and harmonized way. From both data providers and users perspectives, the proposed solution provides several benefits:

1. it keeps their existing capacities autonomous, meaning that they are not requested to implement or comply to a dedicated standard but can continue working with their own tools,
2. it supplements but not supersedes their system mandates: they first answer the needs/requirements of their own scientific community but at the same time contribute to a multi-disciplinary framework,
3. this brokering approach lowers the barriers for both resource providers and users,
4. it is flexible enough to integrate new systems/standards and consequently allows to build incrementally a system of systems by interconnecting additional resources, and finally
5. it provides other non-technical benefits such as identifying –and further mapping-- gaps (e.g., data, participation), offers a platform to coordinate information from various contributors/stakeholders, and can potentially foster a positive competition/emulation to make data discoverable.

Even if geospatial information can bring major benefits for the economy and development of African countries, most of them are still lacking timely access to proper geospatial data (Ayanlade et al., 2008). Several authors have identified the major barriers that are hindering efficient discovery and access to the vast amount of data existing across the African continent: (1) inadequate funding of geo-information services, (2) lack of people, skills, education in the field of geo-information, (3) lack of coordination at the continent scale, (4) lack of computing

and communication infrastructure (e.g., poor Internet connectivity), (5) amount of data still in analog format (e.g., paper), (6) political priorities & support, (7) lack of standardization, (8) social and cultural issues, and (9) lack of institutional policies, regulations, and guidelines (Ayanlade et al., 2008; Cooper et al., 2005; Woldai, 2002). These authors also emphasized the necessity of having tools to facilitate the management of digital metadata (e.g., production and maintenance) and the publication of this metadata on the Internet. The search of African geospatial resources to include into the AFB also showed that several geospatial data repositories or metadata catalog still exist but do not work (e.g., dead links), probably because the related infrastructure is not supported and/or maintained anymore.

To address these issues capacity building (at human, institutional, and infrastructure levels) appears a fundamental activity to be undertaken across the continent in order to develop skills about interoperability, standardization, metadata and data publication, data management, governance, fostering collaboration and cooperation (Donert, 2015a; Giuliani et al., 2014a; Giuliani et al., 2013e).

Another key enabler to succeed in leveraging geospatial resources in Africa is to build an efficient network of stakeholders across the continent and to develop an effective coordination mechanism and a robust governance structure. The African Earth Observations (EO) community is continuously growing and is establishing its presence in the region and in the global arena. This is supported by the development of the African Space Policy and Strategy led by the African Union Commission. This growing network takes also advantage of the national and regional programs and of the on-going cooperation initiatives with a great number of external partners. More specifically, the recently created AfriGEOSS initiative (<http://www.earthobservations.org/afrigeooss.php>), developed within the GEO framework, will strengthen the African EO network through establishing links between GEO activities and the existing capabilities and initiatives in Africa. AfriGEOSS provides the necessary framework for countries and organizations to access and leverage on-going bilateral and multilateral EO-based initiatives across Africa, thereby creating synergies and minimizing duplication for the benefit of the continent (GEO secretariat, 2012). This coordination initiative has been recognized essential to enhance Africa's capacity for producing, managing and using Earth observations, thus also enabling the Region's participation in the implementation of the Global Earth Observation System of Systems (GEOSS). The AFB provides a discovery and access to the GEOSS Data Core and is brokered by GEOSS.

The *Africa Discovery Broker* has been implemented with the vision of being hosted, managed, and promoted in Africa so that African stakeholders can be empowered to manage African resources. The UNECA a major actor in GIS/SDI/EO in Africa through its important network of partners/stakeholders represents a key enabler to support the adoption of interoperable solutions to share geospatial data and products, raise awareness about the benefits of increased access on geospatial data, and create commitments and active contributions in enabling and facilitating the discovery and access to geospatial data (EIS-Africa, 2002; UNECA, 2007a). In this regards, UNECA has established

the African caucus of the United Nations Global Geospatial Information Management (UN-GGIM:Africa - <http://ggim.un.org/knowledgebase/KnowledgebaseCategory19.aspx>) initiative, providing an overarching mechanism to coordinate geoinformation activities involving member States as the key players and putting in place a continental framework for common regional standards, standardization and compliance in line with international policy. Consequently, strengthening, extending and supporting UNECA with the proposed *Africa Discovery Broker* might (1) benefit UNECA's network of stakeholders, (2) bring new and relevant/significant African geospatial resources in the AFB, increasing their visibility and dissemination, and (3) be a major contribution to the AfriGEOSS and the UN-GGIM:Africa initiatives.

Finally, the *Africa Discovery Broker* is aiming to build bridges among the various African geoinformation communities and allows searching and discovering resources available from various heterogeneous repositories. In particular, on top of the AFB different catalogue interfaces have been published to support the development of tailored applications (e.g., desktop, web-based, mobile). This has the potential of valorizing African geospatial resources, helping African stakeholder empowerment, supporting integrated approaches and environmental assessments, and ultimately sustaining informed decision-making processes and supporting meaningful social and economic development. Ultimately, this can be seen as an implementation of a regional GEOSS portal supporting regional stakeholders to participate to the global mandate of GEO (Gorgan et al., 2013b).

4.3.5 Conclusions and perspectives:

Data discovery is a fundamental mechanism, enabling users to search and access data. In particular, for integrated approaches such as INRM, IWRM, ICZM, CBFM, ISFM or for environmental assessments, finding and accessing relevant data is a key requirement.

The *Africa Discovery Broker* facilitates the discovery of African Earth Observation resources by allowing searching across various heterogeneous repositories. It offers the possibility for users working in different domains to search through and access various metadata catalogues and data services in a common and customized frontend application. It includes 17 data sources giving access to 32'442 resources from international (e.g., UNEP, FAO, UN WFP), regional (e.g., SERVIR-Africa, ECOWREX, Africa Soil Information Service), national (SAEON, ILWAC-Mali, Virtual Kenya) repositories and research projects (FP7 Afromaison, FP7 WetWin). It is further connected to the GEOSS Data Core providing discovery and access to a growing number of data for Africa from the global community. Its flexibility enables users to register additional resources simply by entering new endpoints.

The proposed solution appears promising for creating synergies between various environmental related projects in Africa as well as fostering multidisciplinary collaboration/cooperation with environmental institutions (e.g., research, academic) in bringing various African stakeholders (e.g., decision-makers, scientists, local communities) relevant data on the environment.

However, this approach relies on the assumption that data repositories are well documented and may be accessible through the Internet. In order to improve data discoverability, accessibility, and integration, capacity building (at human, institutional, and infrastructure levels) is an essential prerequisite. This can be achieved for example through research collaborations coupled with workshops and trainings and this will help to raise awareness, gain commitments, convince and support African data holders to make available their data and metadata to a larger audience and to unlock the power of data, information, and services for the benefits of the African environment. The AfriGEOSS initiative and UN-GGIM:Africa are viewed as critical vehicles in raising awareness on the benefits of a solution such as the *Africa Discovery Broker*.

4.4 Facilitating the production of ISO-compliant metadata of geospatial datasets (co-author, published)

Gregory Giuliani^{1,2,*}, Yaniss Guigoz^{1,2}, Pierre Lacroix^{1,2}, Nicolas Ray^{1,2}, Anthony Lehmann¹

Addresses:

¹University of Geneva, Institute for Environmental Sciences, enviroSPACE Lab., Uni Carl-Vogt, 66 Boulevard Carl-Vogt, CH-1205 Genève, Switzerland.

²Global Resource Information Database – Geneva, International Environment House, 11 chemin des Anémones, CH-1219 Châtelaine, Switzerland.

*Corresponding author: gregory.giuliani@unige.ch, +41223790709, +41223790744

4.4.1 Highlights

- Facilitate the production of standardized metadata by embedding the generation of description in data production workflows.
- Link data with metadata. Metadata is permanently up-to-date and any changes in data will be automatically reflected thanks to the scheduled harvesting process.
- The ability to automatically generate standardized metadata from the content of a harvested data-publishing server significantly facilitates maintenance and management of the description of large volumes of data.
- The proposed approach is entirely based on an interoperable workflow using OGC standards and therefore is reusable.

4.4.2 Abstract

Metadata are recognized as an essential element to enable efficient and effective discovery of geospatial data published in Spatial Data Infrastructures (SDI). However, metadata production is still perceived as a complex, tedious and time-consuming task. This typically results in little metadata production and can seriously hinder the objective of facilitating data discovery.

In response to this issue, this paper presents a proof of concept based on an interoperable workflow between a data publication server and a metadata catalog to automatically generate ISO-compliant metadata.

The proposed approach facilitates metadata creation by embedding this task in daily data management workflows; ensures that data and metadata are permanently up-to-date; significantly reduces the obstacles of metadata production; and potentially facilitates contributions to initiatives like the Global Earth Observation System of Systems (GEOSS) by making geospatial resources discoverable.

4.4.3 Introduction

Spatial Data Infrastructures (SDI) are recognized as an effective environment for digital geospatial data production, management, analysis and diffusion (Craglia et al., 2012b). The primary function of any SDI is data discovery, enabling users to search and evaluate data before downloading them (Nebert, 2005; Nogueras-Iso et al., 2005a). The fundamental requirement for an efficient and effective data discovery mechanism is that data is properly documented with metadata and stored in a catalog (Foresman, 2008; Ma, 2006). Without appropriate metadata, a SDI does not facilitate the discovery of, and access to geospatial data (Masser, 2005).

The primary role of metadata and catalogs for data discovery is recognized in data sharing initiatives such as the Infrastructure for Spatial Information in the European Community (INSPIRE) (European Commission, 2007) and the Global Earth Observation System of Systems (GEOSS) (GEO secretariat, 2005a). This important role is also reinforced with the increasing momentum gained by Open Data access policies (Wessels et al., 2014). These policies highlight the importance of using standards to enable interoperability for both metadata description (e.g., ISO19115-1:2014, FGDC, Dublin Core) (Díaz et al., 2012b) and searching (e.g. OGC Catalog Service for the Web) (Nogueras-Iso et al., 2005a). Having interoperable metadata allows various systems to exchange metadata ensuring that metadata records can be discovered, accurately interpreted, and subsequently used or integrated into other platforms or applications (Nativi et al., 2013a).

Despite the importance of having metadata and associated catalogs, most data that are currently published via SDI are lacking metadata (Batcheller, 2008; Batcheller et al., 2009; Trilles et al., 2014). There are several reasons for this: lack of funding (i.e., financial costs), time commitment, no perceived added value, complexity of standards, and tedious process for creating metadata (Kalantari et al., 2010; Lehmann et al., 2014; Myroshnychenko et al., 2015; Trilles et al., 2014). Moreover, data and their description (metadata) are often published and produced with different software, leading to the duplication of efforts to enter relevant information (e.g., title, abstract, keywords), and consequently cause data and metadata to be disconnected (Giuliani et al., 2013d; Kalantari et al., 2010). This can be an important issue because when a dataset is updated the changes must be also reflected in the related metadata. Another issue related to data-metadata disconnection is that data providers are often confused in the choice of their publication workflows. Some of them publish data first and then create metadata while others do the opposite. This confusion contributes to fragmentation, disconnection and lack of good and reliable data documentation (Díaz et al., 2012a).

Means to facilitate the production of standardized metadata and ensuring that data and metadata remain linked should be beneficial for both data providers and users (Ellul et al., 2013; Olfat et al., 2012). A lot of research has been conducted to overcome some of these issues and various solutions have been proposed: (1) automatic generation of standardized metadata from Earth

Observation products (Yue et al., 2010), (2) automatic inventories while scanning data folders (Moura, 2012; Prunayre et al., 2013), (3) using new file format (e.g., NetCDF) where data and metadata are stored in the same file (Lehmann et al., 2014), and (4) innovative workflows to extract information based on web services, semantic enablement or tagging (Florczyk et al., 2012; Kalantari et al., 2010; Manso Callejo et al., 2010; Yue et al., 2012). These authors recognize the necessity to embed metadata production in data creation, automating the generation of metadata where possible. Unfortunately, most of these implementations require a high level of SDI expertise to develop tailored and often complex solutions. Therefore, convincing data providers to produce metadata can remain a major barrier.

Based on these considerations the aim of this paper is to present a proof of concept using an interoperable workflow between a data publication server and a metadata catalog to: (1) automatically generate standardized descriptions of geospatial data, (2) establish a permanent link between data and metadata (e.g., changes in data are automatically reflected in corresponding metadata), and (3) facilitate data-metadata publication workflows through a single entry point.

4.4.4 Methodology

The proposed approach is designed to meet the following requirements:

1. The use of a classical workflow: data providers usually store data on a server, publish them as services, then generate the proper documentation and store them in a metadata catalog (figure 1). From a data provider point of view, this workflow is easier than first creating the metadata (e.g., requires additional work, time-consuming, monotonous, complex) and then publishing the data;
2. The introduction by data providers of relevant metadata (e.g., title, abstract, keywords) into only one place (e.g., data publication server when uploading data), avoiding duplication of efforts and offering a single entry point; and
3. The use of interoperable services based on Open Geospatial Consortium (OGC) standards. This will allow using the proposed approach with various software solutions and therefore enhance reusability of the method.

The general scenario of the workflow assumes that a data server can be harvested by a metadata catalog, and automatically generates basic and standardized metadata.

This scenario is composed of three steps (figure 1 – right): (1) a data provider manually publishes a range of different datasets (e.g., satellite images, vector, raster) on a data publication server, (2) while he publishes data, he needs to fill basic descriptive elements for each layers (e.g., name, title, abstract, keywords, projection, bounding boxes, point of contact) required by the data publication server, (3) the metadata server harvests the data publication server, generates ISO-compliant metadata using both the CSW interface for harvested descriptive elements and a mapping file for fixed elements, stores the generated metadata, and publishes them in the catalog.

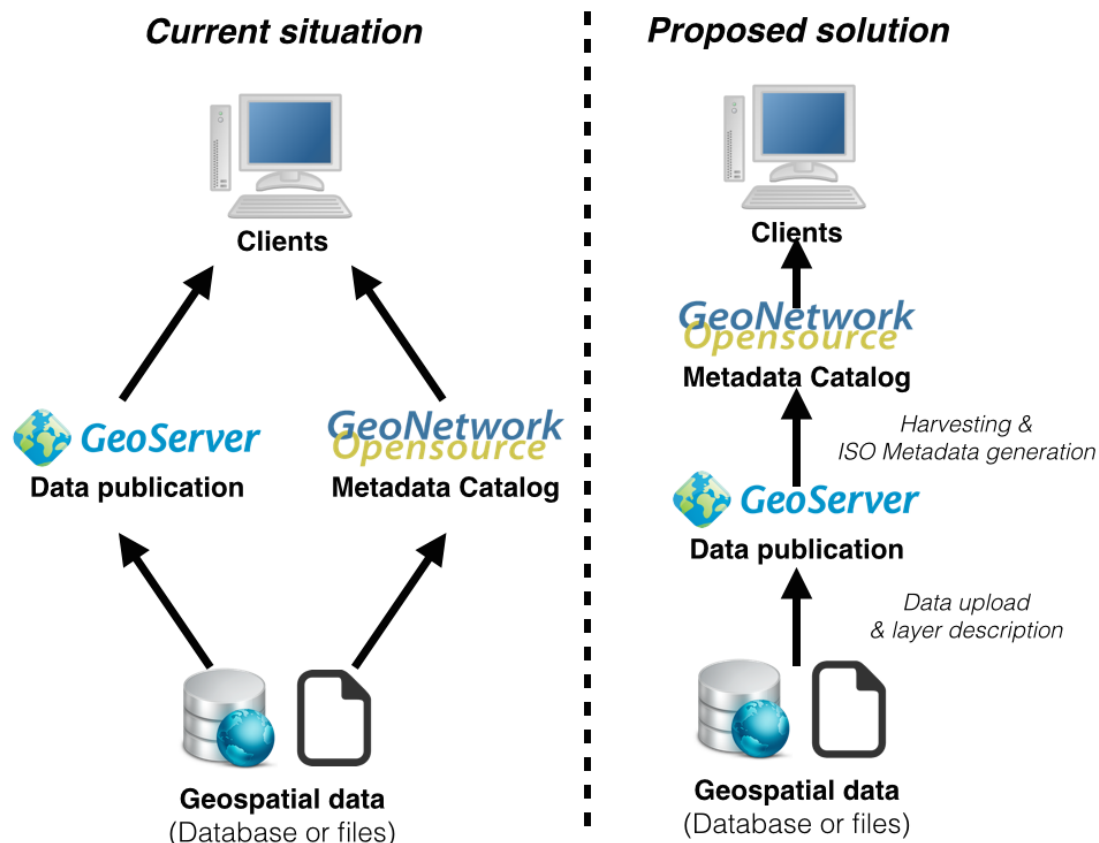


Figure 1: Workflow schema of the proposed solution for producing ISO-compliant metadata compared to the current situation. On the left (current situation): data and metadata are published and managed separately leading to a disconnection of data and metadata. On the right (proposed solution): data provider only publishes its data on the data server and describes the basic elements (layer name, title, abstract, keywords, projection, bounding boxes, point of contact). Then the Metadata catalog harvests the data publication server and generates ISO-compliant metadata using both the CSW interface for dynamic elements and the mapping file for fixed elements. Metadata are linked to the data and are permanently up-to-date.

4.4.5 Implementation

The implementation used to validate the method is based on two components:

- A data publishing server (e.g., GeoServer⁵⁰) together with its CSW extension;
Geoserver is an open source web server designed to publish data from different major sources (e.g. shapefile, geotiff, PostGIS) using OGC

⁵⁰ <http://www.geoserver.org>

standards (e.g., WMS, WFS, WCS) and allowing the users to share their data in an interoperable and standardized way.

- A metadata catalog (e.g., GeoNetwork⁵¹) that is able to harvest the content of the data publishing server through a CSW request. GeoNetwork is also an open source project that acts as a web-based metadata catalog to manage geospatial resources. It implements OGC CSW specification allowing users to search, query, discover, publish, and manage metadata on different data layers.

The workflow uses the OGC Catalogue Service for the Web (CSW) standard (Open Geospatial Consortium, 2007b) as the interface between the various components involved. The CSW defines protocol to facilitate management, communication, discovery of data (through standardized mechanisms) and development of metadata-driven user interfaces in a machine-readable format using an open standard. An interesting feature of CSW is that through a *Harvest* request, metadata can be created, updated, or deleted on the server that performs the request.

Metadata is queried and constructed directly from GeoServer's internal catalog. Currently two metadata schemes are supported by GeoServer's CSW plugin: Dublin Core and ISO Metadata Profile. All elements from both profiles can be customized through mapping files using OGC Common Query Language (CQL) expression against GeoServer catalog properties. This enables a data provider to customize/create a profile according to a metadata standard and facilitates the production of metadata when harvesting the data-publishing server through the CSW interface. In the proposed scenario, descriptive elements of layers published in GeoServer are not sufficient to generate metadata that comply with ISO19115-1:2014 Core Elements standard (ISO, 2014) (ISO 2014). Consequently, a mapping template has been developed using the ISO19115-1:2014 Core Elements that are shown in Table 1. Bold elements are those that can are coming from GeoServer descriptions and italic elements are those that are coming from the mapping file. By applying a mapped template while harvesting a node, ISO-compliant metadata can be constructed using the various elements that are either dynamically harvested (e.g., descriptive elements from GeoServer like title, bounding boxes, abstract) or fixed (e.g., recurrent elements like metadata language, metadata standard name) that can be directly defined in the mapping file.

Field Name	XML Path for Mapping
Dataset title	MD_Metadata/MD_DataIdentification.citation/CI_Citation.title
Dataset reference date	MD_Metadata//MD_DataIdentification.citation/CI_Citation.date
Abstract	MD_Metadata/MD_DataIdentification.abstract
<i>Dataset language</i>	MD_Metadata/MD_DataIdentification.language
Metadata point of contact	MD_Metadata.contact/CI_ResponsibleParty

⁵¹ <http://geonetwork-opensource.org>

Metadata date stamp	(MD_Metadata.dateStamp
<i>Dataset topic category</i>	MD_Metadata/MD_DataIdentification.topicCategory
Geographic location of the dataset	MD_Metadata/MD_DataIdentification.extent/EX_Extent/EX_GeographicExtent/EX_GeographicBoundingBox or EX_GeographicDescription
<i>Dataset character set</i>	MD_Metadata/MD_DataIdentification.characterSet
<i>Metadata language</i>	MD_Metadata.language
<i>Metadata character set</i>	MD_Metadata.characterSet
<i>Additional extent information for the dataset</i>	MD_Metadata/MD_DataIdentification.extent/EX_Extent/EX_TemporalExtent or EX_VerticalExtent
<i>Lineage</i>	MD_Metadata/DQ_DataQuality.lineage/LI_Lineage
<i>Spatial representation type</i>	MD_Metadata/MD_DataIdentification.spatialRepresentationType
Distribution format	MD_Metadata/MD_Distribution/MD_Format.name and MD_Format.version
<i>Metadata standard version</i>	MD_Metadata.metadataStandardVersion
Spatial resolution of the dataset	MD_Metadata/MD_DataIdentification.spatialResolution/MD_Resolution.equivalentScale or MD_Resolution.distance
Dataset responsible party	MD_Metadata/MD_DataIdentification.pointOfContact/CI_ResponsibleParty
On-line resource	MD_Metadata/MD_Distribution/MD_DigitalTransferOption.onLine/CI_OnlineResource
Metadata file identifier	MD_Metadata.fileIdentifier
<i>Metadata standard name</i>	MD_Metadata.metadataStandardName
Reference system	MD_Metadata/MD_ReferenceSystem

Table 1: ISO19115-1:2014 Core Elements (in bold, elements that can be generated automatically during harvesting, in italic elements that are generated from the mapping file).

To test the workflow, three different data files have been published in GeoServer: vector (i.e., country borders), raster (i.e., land cover map), imagery (i.e., a single Landsat 8 scene). In all these three cases, accurate ISO19115-1:2014 Core Elements descriptions have been generated with a CSW harvesting request from GeoNetwork and metadata were stored in the catalog. The metadata generation workflow has been experimented with two other popular metadata catalogs who support CSW harvesting (e.g., Glocat⁵², ESRI Geoportal Server⁵³) and in both cases they also generate and store standardized data descriptions. Consequently, this

⁵² <http://essi-lab.eu/do/view/Glocat>

⁵³ <http://www.esri.com/software/arcgis/geoportal>

workflow is flexible (e.g., can be used with various software solutions), reusable, and enables automatic creation and storage of standardized and harmonized ISO19115-1:2014 compliant metadata.

4.4.6 Discussion

Results show that the proposed solution is simple to implement, facilitates the automatic production of ISO-compliant metadata, embeds the generation of metadata in data provider's workflows, and links data and metadata together. Because the workflow generates ISO19115-1:2014 Core Elements metadata, the proposed approach is sufficient for the purpose of data discovery through general description of vector, raster, and satellite imagery data. However, it cannot answer complex description requirements like hierarchies (e.g., parent/child metadata for data sets collection) that need more detailed information. We argue that proposing a methodology to automatically generate adequate description of geospatial data in general terms is essential to raise awareness and reach commitments from data providers to give at the very least a minimal, harmonized and coherent documentation of their resources. This may help to increase the number of good quality and standardized metadata, and can have a positive impact on data discovery in initiatives like GEOSS or INSPIRE. It will in turn help to demonstrate the added-value of good and reliable metadata to facilitate the discovery of resources to a wider audience and to different communities of users.

Another important benefit of this approach is that it links data and metadata. Each time data is modified, the related metadata is automatically updated at the next harvesting run. This means that data and metadata are permanently up-to-date. Moreover, metadata production is embedded in the workflows of data providers. When publishing their data, these providers only have to take care of filling the mandatory description fields. The remaining tasks are executed automatically through the harvesting mechanism. Consequently, this process is neither time-consuming nor repetitive. It helps to lower the barrier of associated costs of metadata production, to avoid duplication of efforts in entering the same information at several places, and to facilitate the management of metadata. However, even if the proposed approach associates data and metadata through a synchronization mechanism this does not resolve the issue of integrating data and metadata in one single file (e.g., NetCDF). This is an interesting solution and further investigations are required to understand how the presented workflow can interact with data and metadata stored in a same file.

Other advantages of the proposed solution are that: (1) additional metadata standards profiles can be implemented (e.g., FGDC, INSPIRE), (2) once metadata is generated, the data description can be further edited in the metadata catalog and manually completed with any useful/missing information.

Compared to other approaches that are more complex to implement, this workflow is completely based on standardized and interoperable services. This facilitates the communication between relevant software components and eases the production of harmonized metadata. In particular, this allows implementing this workflow with different metadata catalogs, because users are not restricted

to a dedicated software solution. However, to our knowledge, not all data publishing software implement a CSW interface. The use of GeoServer is currently a necessity.

This approach can be seen as similar to harvesting an OGC Web Map Service (WMS) endpoint and generating metadata from each layer from the WMS instance. However, WMS harvesting differs on the following points: (1) no mapping templates are applied during harvesting and thus various metadata elements can not be efficiently handled, (2) under the *distribution information* section it will only manage the creation of WMS links (and not WFS/WCS links like in the proposed approach), and (3) it can generate a thumbnail of the data which is not the case with the presented workflow.

These first results are encouraging and prove the feasibility of using an interoperable and scalable workflow between a data publication server and a metadata catalog following OGC standards. However, further research is required in order to tackle issues like semantics interoperability (Vaccari et al., 2008; Yue et al., 2012), support of various metadata schemas (e.g., INSPIRE, FGDC), interoperability across disciplines (e.g., multi-disciplinarity), data quality description (Díaz et al., 2012b) and Open Data policies (Wessels et al., 2014).

4.4.7 Conclusions & perspectives:

Recognizing both the importance of metadata to enable efficient and effective data discovery and the fact that data providers are serving increasingly large volumes of data, managing and maintaining a metadata catalog can be challenging.

The proposed approach:

- (1) facilitates the production of standardized metadata by embedding the generation of description in data production workflows.
- (2) links data with metadata. Through the proposed approach metadata is permanently up-to-date and any changes in data will be automatically reflected thanks to the scheduled harvesting process.
- (3) produces standardized metadata following the ISO standard. The core elements (e.g., title, abstract, extent) are generally sufficient for most users to efficiently discover data.

The ability to automatically generate standardized metadata from the content of a harvested data-publishing server significantly facilitates the maintenance and management of the description of large volumes of data.

The proposed approach appears to be a valid solution to reduce the barriers of metadata production (e.g., duplication of efforts, cost, time-consuming, monotonous process, complexity of standards), can potentially convince data providers to generate metadata, facilitate their contribution to data sharing initiatives like GEOSS, and may help to demonstrate the added value of having properly documented data to facilitate data discovery and access to the largest possible audience.

4.5 Chapter key outcomes

- Getting ready-to-use data might be challenging because of many technical issues, which might affect interest and efficiency of data users.
- A solution is needed, that would improve data readiness for a better user experience.
- The customization approach is of interest for local/regional communities that have the possibility to easily and efficiently access relevant geographical extent data.
- Interoperability and OGC standards are well suited to set up such a customization approach.
- SCOPED is a generic implementation of the customization approach that addresses interoperability and standardization requirements. It allows to automatically extract specific extent spatial data and automatically links to the corresponding metadata record.
- SCOPED has been developed using the water domain use case (SCOPED-W) but can be used for other environmental domains.
- The SCOPED approach lowers complexity for users by asking them to only interact with a web browser and reporting the complex tasks on automatic workflows allowed by OGC Web Processing Services (WPS).
- Such an approach contributes to data valorization by giving it an added value and more visibility for a second life.
- Interactivity offered by extractors might trigger a sense of community membership and participate to growing use and potential sharing of scale specific datasets.
- Environmental multidisciplinary approaches require integration of diverse geospatial resources, often heterogeneous in terms of standard or format.
- Geospatial resources might exist for a given area (e.g. Africa) but in specific standards or formats, making their discovery and access not efficient.
- The brokering approach allows gathering heterogeneous resources in a homogeneous interface, giving hence a chance to a particular geographic or thematic community to discover and access geospatial resources that might otherwise be disparate and maybe unknown.

- The brokering approach is based on the possibilities offered by interoperability implemented through various types of standards. It keeps the existing capacities as autonomous as possible by interconnecting them.
- This approach benefits geospatial resources producers by promoting resources that might otherwise not be discoverable or accessible; it also benefits users through increased geospatial resources availability.
- This approach lowers complexity for both users and producers, shifting it to an intermediate level managed by IT specialists. It supplements but not supplants systems mandates and governance arrangements.
- Despite the benefits brought by this approach, raising awareness to understand its potential and building capacity to set it up is essential. To this end, key local stakeholders having understood and accepted its principles are necessary.
- Metadata is fundamental for an efficient data discovery since “without appropriate metadata services which help them to find this information it is unlikely that SDIs will achieve their overarching objective of promoting greater use of geographic information” (Masser, 2006).
- Making metadata interoperable is essential as it allows various systems to exchange metadata, expanding its reach and related data discovery.
- Many factors affect systematic metadata availability for spatial data, requiring new approaches.
- One of the issues is the decoupling of data and metadata, which requires double work from the data producer, often to the detriment of metadata. This affects data production and updates. A sustainable solution integrating both data and metadata while minimizing metadata production and update efforts is needed.
- Interoperability offers a solution to obtain an semi-automatized and standardized metadata production and update workflow combining a geospatial data server with a metadata catalog. This is possible through the Catalog Service for the Web (CSW) standard.

5 ASSESSMENT

5.1 Introduction to SDI assessment

The previous chapter demonstrated that technological solutions exist to lower the complexity barriers hindering a wider adoption of SDI principles and tools. Building capacity and addressing complexity are major actions to support SDIs implementation worldwide. However, performing these actions requires resources, for which justifications are needed. Hence policy makers, government representatives and the public are increasingly interested in rational assessment studies measuring the benefits of SDIs and the level of realization of their objectives (Grus et al., 2011). Assessment can also be perceived as a kind of surveillance, that lies somewhere between care and control, which can easily be politically motivated as “the power to see is also the power to influence” (Taylor and Broeders, 2015) and requires caution. Assessment is anyway an essential component of an SDI implementation process and permits to identify successes or weaknesses, justify the realization of intended goals, evaluate the real impact of the actions, address the identified weaknesses and formulate new funding requests. For these reasons it should always be considered right from the beginning of an SDI implementation process.

As discussed in chapter 2, several approaches exist to assess SDIs, which depend on the purpose of the assessment. But no ready-to-use approach exists (Grus et al., 2011) to measure an entire SDI, only a framework (the multi-view assessment framework) is proposed to guide the choice to a best comparable methodology, that will need to be adapted to use cases. An assessment of the African use case has been performed in the second chapter (stocktaking) with goal to take a rapid snapshot of the whole SDI in Africa at a given time. Such a macro or systemic assessment can be replicated at a later stage to re-evaluate the evolution of the whole system. Some smaller aspects of an SDI, such as a particular component (e.g. the infrastructure component, or even one aspect of this component such as a geoportal) are more convenient to assess, which can be done more deeply than in a macro assessment.

Besides the SDI itself or its components, there is also a need to assess impact. We can differentiate the impact that a SDI has on its target (e.g. a company, an institution, a society) from the impact that actions taken have on the SDI implementation or improvement. GEO abounds in this direction by stipulating that it is important to define indicators for measuring progress in capacity building for Earth Observations (GEO secretariat, 2006). Two main types of assessment can then be formulated: (1) SDI assessment, which can consist in assessment of the whole SDI system (stocktaking) status or assessment of its components; and (2) assessment of the impact, which can be further differentiated between impact that actions performed (e.g. capacity building) have on the SDI or impact that the SDI has on its target (Figure 10). Both assessment types are complementary and respond to particular aspects that give a complete overview of the SDI status and an answer to expectations. In each case, measurements need to be performed, which requires specific assessment indicators that should be comparable through time for regular interval assessments. The ease to determine these indicators depends on the precision of

the SDI goals definition (Grus et al., 2011), which highlights the importance of the vision for an SDI.

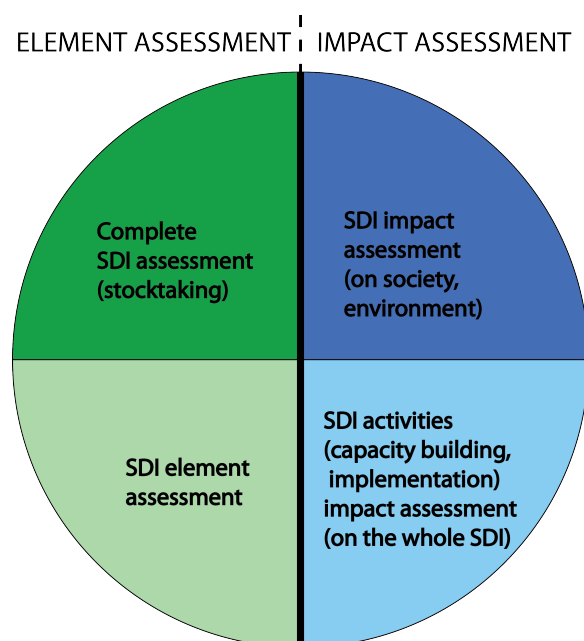


Figure 10: elements of SDI assessment

The first paper of this chapter, “*EGAL: a methodology for Environmental Geoportals Assessment and Label*”, presents a methodology for assessing one of the key elements of SDI: the geoportals, that are themselves part of the “infrastructure” SDI component. Geoportals are the visible parts of SDI and gateways to geographic content and capabilities (De Longueville, 2010; Maguire et al., 2005; Nushi et al., 2015). Being able to visualize geographic data is of particular importance for the users who can interact indirectly with environmental observations and models in a reduced complexity environment (Karpouzoglou et al., 2016). Geoportals can also play the role of an enabling platform to support the chaining of services across participating organizations, which is required in a SDI (Williamson et al., 2006).

In order to be efficient, geoportals should meet certain characteristics in terms of visibility as well as facilitating data discoverability and access. To this end, they should take advantage of the latest technological developments (e.g. web services) and principles (e.g. open data) in the geospatial domain. Being able to measure through particular indicators if these requirements are met is important for a better geospatial data use and sharing. Beyond the measurement of these indicators, a simple and illustrative labeling of the geoportals could indirectly participate to a global capacity building effort to improve this important tool linking users to data.

The second part of this chapter gathers some work performed in the frame of the FP7 European project “EOPower” that aims at creating conditions for sustainable economic development through the increased use of earth observation products and services for environmental applications (Noort, 2015a). This work is of particular importance for the present chapter as it proposes a methodology to assess the impact of Earth Observation solutions: the

“EOPOWER impact assessment framework”, that not only captures economic aspects but also benefits that are currently not captured in economic calculations, such as those relating to sustainable management of natural resources and climate change. Such an impact assessment has been performed on the capacity building activities presented in the “Capacity Building” chapter of this thesis: the “Bringing GEOSS Services into Practice” workshop as well as Earth Observation activities promotion performed in Armenia.

A last aspect of assessment would still need to be performed, which is the assessment of the impact that has SDI on a society. Some elements of such an impact will be discussed but measuring the effects of an SDI in a country requires a full assessment that can only be performed after several years, when the effects become measurable and visible.

5.2 EGAL

EGAL: a methodology for Environmental Geoportals Assessment and Label

Yaniss Guigoz^{1,2}, Pierre Lacroix^{1,2}, Nicolas Ray^{1,2}, Anthony Lehmann^{2,3}, Hy Dao^{1,4},
Martin Lacayo², Gregory Giuliani^{1,2}

Addresses:

¹University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, Switzerland.

²University of Geneva, Institute for Environmental Sciences, EnviroSPACE Lab., Bd Carl-Vogt 66, CH-1211 Geneva, Switzerland.

³University of Geneva, Dpt. F.-A. Forel for Environmental and Aquatic Sciences, Uni Carl-Vogt, CH-1211 Geneva 4, Switzerland.

⁴University of Geneva, Institute for Environmental Governance and Territorial Development, Uni Carl-Vogt, CH-1211 Geneva 4, Switzerland.

5.2.1 Abstract

Geoportals are the gateway to access relevant geospatial data, which is of particular importance in environmental sciences. The ease of finding geoportals for specific themes or regions is especially needed from a community of practice perspective; in any case, the ease of discovering and providing access to relevant useable data is key in an efficient geospatial data quest. Efficient data discovery and access is facilitated by data interoperability supported by specific standards, often put in place in a formal Spatial Data Infrastructure (SDI) framework. It is then important to have a methodology to easily determine if a geoportal is compliant with established efficiency characteristics for geospatial data discovery, access, and use.

This paper provides the “Environmental Geoportals Assessment and Label” (EGAL) methodology to easily assess geoportals discovery and their ability to propose the discovery and access of useable data. Several geoportals are selected, categorized in a typology based on certain characteristics, and assessed through this methodology. Then a visual “EGAL” label is proposed, efficiently conveying the elements of success determined for a geoportal. We believe that such an efficient visual representation contributes to a useful spatial data infrastructure capacity building and can help community geoportals developers or stakeholders to improve these gateways to geospatial data.

5.2.2 Introduction

Environmental threats linked mostly to global warming are bigger than ever (IPCC, 2014;Rahmstorf et al., 2011;Secretariat of the Convention on Biological Diversity, 2014;UNEP, 2012;UNEP, 2014), requiring political decisions at the global level to be implemented rapidly, before reaching a critical tipping point (Fang et al., 2015;Steffen et al., 2015) that could affect the planet for centuries⁵⁴. Simultaneously, the exponential technological progress drives a “data revolution” (Independent Expert Advisory Group on a Data Revolution for Sustainable Development, 2014) by making it possible to collect and store big data and to generate unparalleled information crucial for addressing environmental issues.

In order to address environmental challenges, the political decisions should be based on scientific evidences resulting from environmental data and models. Despite the abundance of such data, their discovery and access is not always optimal and is hindered by several barriers such as price, no conformance to standards, non accessibility and license concerns (Giuliani, 2011). Addressing these issues requires a specific framework taking into account all the necessary aspects leading to an efficient data flow. Such a framework exists and is called “Spatial Data Infrastructures” (SDIs). It has been developed since the early 1990s by many authors and organizations (Grus et al., 2011) and can be defined as “*the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data*” (Nebert, 2005). SDIs are meant to avoid fragmentation, gaps in availability of geographic information, duplication of data collection, and problems of identifying, accessing, or using the

⁵⁴ https://www.ipcc.ch/publications_and_data/ar4/wg1/en/spmsspm-projections-of.html

available data (Van Orshoven, 2003). Several authors (Mansourian et al., 2006; Rajabifard, 2002) distinguish five main SDI components: data (geospatial data), people (human resources), access network (networking technology), policy (institutional framework), and standards (technical standards).

A successful SDI is essential for data discovery and access, which are the building blocks supporting informed decision-making. The definition criteria of a successful SDI depend on the perspective of the SDI: for example success can be present from the institutional perspective (i.e., there is a good (inter)national coordination for data exchange), but weak from the technological perspective due to poor geoportals or bad internet connectivity. Chan et al. (2001) group SDI definitions from a four-perspective classification: identificational, technological, organizational, and productional. The identificational perspective describes the uniqueness of SDI that distinguishes it from other systems; this might be useful for example to get specific funding. The technological perspective describes the technical aspects of SDI (e.g. software). The organizational perspective describes the SDI components and their relationships. The productional perspective defines the way SDIs are used by an organization to generate products and services. The authors argue that from a technological perspective, an ideal SDI is a hierarchy of geospatial datasets that users at different levels can access to meet their needs. They also state that framework data (seven themes of geospatial data used by most GIS applications⁵⁵, also called baseline data), standards, and the delivery mechanism of SDI, although only part of the whole SDI, are key elements. Similarly, Maguire et al. (2005) highlight that geoportals -- gateways to geographic content and capabilities -- are a key element of SDI. Nushi et al. (2015) stress the importance of geoportals by stating that anyone wishing to access datasets must make use of the technological components.

In this paper we focus on the technological perspective of SDI and in particular on the key SDI elements defined by Chan et al: data, standards, and geoportals as delivery mechanisms of SDI. Even though these elements are only the visible parts of a whole SDI (De Longueville, 2010), they are essential in meeting the needs of users looking for data. Kok et al. (2005) state that an SDI develops gradually through different stages, addressing in priority the most pressing issues such as collecting and sharing data before including the political aspect (i.e. defining data access and use policy and getting national commitment), highlighting the importance of geoportals. Despite the fact that the technological side of SDI is only one of the several SDI components, a web portal serving geospatial data or services is often called a "SDI" in the common language. Despite this misnomer, such geoportals share however SDI objectives and concepts as gateways facilitating discovery of and access to geospatial data (Georis-Creuseveau et al., 2015) and their management costs alone were estimated at around 120 million euros worldwide in 2006 (Crompvoets, 2006).

Maguire (2005) states that a portal is a web environment interface that allows an organization or a community of information users and providers to aggregate

⁵⁵ <https://www.fgdc.gov/framework/handbook/appendixA>

and share content and create consensus. De Longueville (2010) suggests that community-based geoportals can be a good platform to enable sharing functionalities as exchange of information is expected to be more intense inside communities. Besides, Manso Callejo and Castelein (2010) say that *“portals can be a good source to collect statistical data of availability of data and services and their use”*. We use the term “geoportal” in this paper to designate a web portal serving geospatial data. Environmental geoportals are then essential interfaces for bridging the environmental data needs from the scientific and political worlds with environmental data producers. Nativi et al. (2015) stress the importance of the “community geoportals”, a concept introduced by GEOSS to provide different communities of practice (Nativi et al., 2012) – user-led communities of stakeholders⁵⁶ – with specialized functionalities serving their needs. This is all the more important since *“Web 2.0 allows communities that share common centers of interest to exchange information from peer-to-peer, collectively discussing the relevance and/or quality of any piece of information, as well as commenting on each other’s contribution to a community’s collective knowledge.”* (De Longueville, 2010). As the communities of practice are the main targets of geoportals, it is important to understand their characteristics to find out if these are well embodied in the geoportal. A geoportal’s usefulness depends on its customization (Tait, 2005) to the targeted community’s needs (e.g. thematic information served, complexity of the user interface, language) and on its ability to provide easy discovery and access to data. We argue that for these two last elements, geoportals must be as compliant as possible with SDI cutting edge standards and should for example implement standardized Open Geospatial Consortium (OGC)⁵⁷ web services that ensure data interoperability and hence an optimal data discovery and access.

Considering the importance of the geoportals for the data discovery and access in the SDI framework, we propose in this paper a methodology named “Environmental Geoportals Assessment and Label” (EGAL) to assess a selection of geoportals with the objective of evaluating if they meet expectations in light of relevant SDI criteria. We also want to explore if geoportals have certain characteristics depending on targeted communities of users and see if certain communities are more SDI compliant than others. This might help to derive some community-specific recommendations to better address technical SDI issues. The objectives of this paper are then to: (1) establish a **SDI typology** based on selected fundamental criteria: geographical scale, theme, and technology used for the geoportal; (2) classify a **selection of geoportals** in this typology covering different geographic scales and themes; (3) define some **basic indicators** to make a rapid assessment of environmental geoportals; (4) **assess** the indicators defined for each geoportal selected and give a simple and visual message summarizing the assessment; and (5) determine some **characteristics** of geoportals based on their typology and use them to issue general best practices and recommendations for certain communities or practice, as well as

⁵⁶

https://www.earthobservations.org/documents/committees/uic/200905_11thUIC/07%20us0901b_cop.pdf

⁵⁷ <http://www.opengeospatial.org/>

general recommendations based on the assessment.

5.2.3 Methodology

In order to address the various objectives of this paper, the methodology that we propose depends on several requirements. The first one is to remain focused on the technological aspect of the SDIs for the reasons explained in the introduction, even though several methodologies exist to assess a general SDI status (Al Shamsi et al., 2011;Guigoz et al., 2016).

The second requirement consists of proposing a simple methodology for a rapid assessment of the technical SDIs – the geoportals – that can be applied by anyone with an interest in the SDI domain. For example, such a rapid assessment methodology can greatly benefit people who want to develop a geoportal or improve an existing one by taking into account explicit criteria for immediately improving their geoportal's quality in terms of SDI conformity. This also contributes to a capacity building effort, making for example a community rapidly aware of geographical-related best practices.

The last requirement, in line with pedagogic aspects of a rapid assessment, is to remain highly illustrative in the assessment. Some successful examples exist, in the SDI domains and beyond: EIONET⁵⁸ scoring criteria, GEO-label⁵⁹ – a label “to recognize the scientific relevance, quality, acceptance and societal needs for activities in support of GEOSS” (Science & Technology Committee, 2010, p.2), the linked open data 5-star badges,⁶⁰ and the Creative Commons licenses⁶¹. All of these examples are highly illustrative and convey a simple visual message that has the advantage of immediately making the scoring appreciable through an intuitive scheme. Among these examples, the GEO-label corresponds to the simple and illustrative message we would like to transmit. However, it targets the quality of geospatial datasets for promoting trust in GEO labeled datasets, whereas we want to label the quality of geoportals through several of their components. We therefore need to develop a new specific label for a multidimensional geoportal assessment.

Based on these requirements, the methodology below describes the necessary chronological steps for a rapid assessment of geoportals (Figure 1): i) establishing some typology indicators, ii) classification in this typology of a selection of geoportals, iii) definition of variables for assessing the ease to discover the selected geoportals and the data they serve, as well as accessing and using this data; and iv) finally, this assessment must be translated into a simple and visual message resulting from the scoring of the assessed geoportals and indicating their SDI best practices compliance.

⁵⁸ <https://www.eionet.europa.eu/dataflows/pdf2016/criteria>

⁵⁹ <http://www.geolabel.info/Index.htm>

⁶⁰ <http://5stardata.info/en/>

⁶¹ <https://creativecommons.org/choose/>

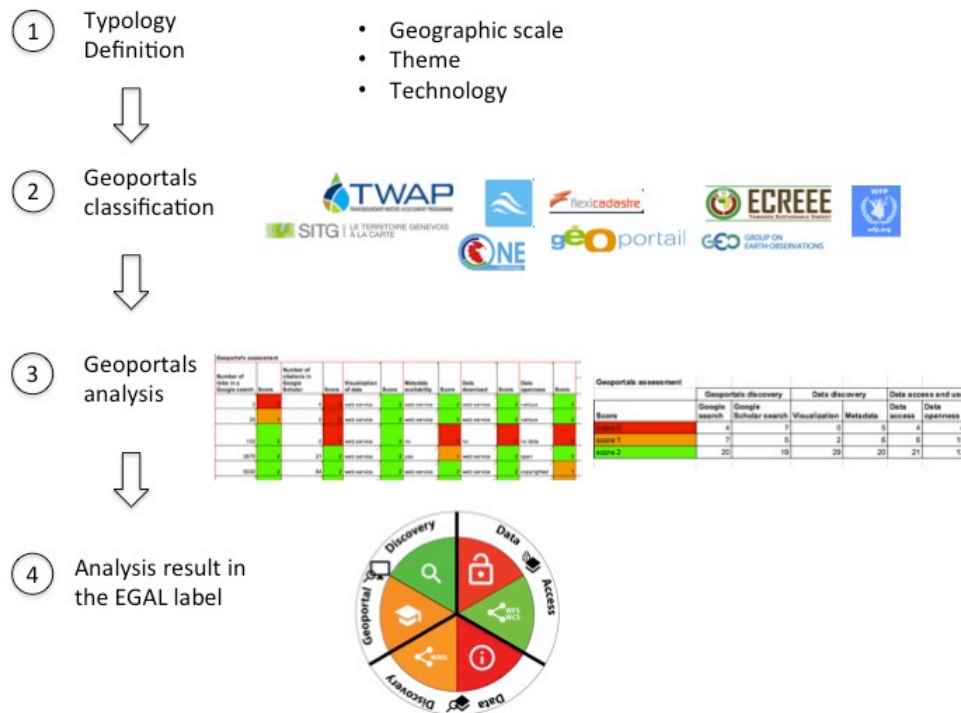


Figure 1: the "EGAL" workflow

5.2.3.1 Typology

De Longueville (2010) argues that SDI communities are typically based on themes, technology, or geographic area of interest. For establishing the typology of geoportals, we therefore consider these three main categories. The "themes" and "geographic area of interest" are obvious characteristics of the data served by the geoportal that go in line with the EU's flagship SDI monitoring process called "State of Play" (SoP) (Vandenbroucke, 2011). It states that SDIs can be developed for different spatial extents and for general or thematic user communities (Van Orshoven, 2003). For the geographic aspect we consider the hierarchical scale as it matches with politico-administrative levels, and has importance for governance. The "technology" aspect is relevant to find out if some technologies are used more in certain communities than others. This technological choice is a driver that has long-term consequences on the sustainability policy of the platform, mainly due to the proprietary licenses cost, and reflects an institutional choice that might be different depending on the communities.

Regarding the first category (themes), there are a variety of thematic geoportals (e.g. coastal geoportals (Georis-Creuseveau et al., 2015; Longhorn et al., 2005; Wright, 2009), geologic geoportals (OneGeology, 2012)). Making an exhaustive list of thematic geoportal is extremely difficult. However, as we want to focus on environmental geoportals, we use a classification intended to cover most of the relevant environmental themes: the GEOSS Societal Benefit Areas (SBAs), that are stated as *"the domains in which Earth observations are translated into support for decision-making"* (GEO Secretariat, 2015). These consist in: (1)

food security and sustainable agriculture; (2) biodiversity and ecosystem sustainability; (3) disaster resilience; (4) infrastructure and transportation management; (5) energy and mineral resources management; (6) public health surveillance; (7) water resources management; (8) sustainable urban development; as well as climate, that has an impact across all SBAs. We also add two other themes that might be individually relevant for specific environmental communities: (10) cadastre and (11) topography. The classification we propose is non-exclusive, in order to take into account the possible multi-thematic dimension of geoportals. For example, the Swisstopo⁶² and the Environmental Data Explorer (EDE)⁶³ geoportals cover all SBAs.

For the second category (geographic data of interest), we consider the geographic scale targeted in the geoportal. Rajabifard (Rajabifard et al., 1999) defines five SDI levels (scales) that require different levels of data details: local, state, national, regional and global levels. Van Orshoven (2003) in the EU State of Play, also defines five SDI territory extents, with a slightly different nomenclature regarding sub- and super-national levels: local, regional, national, multi-national and global levels. We also choose to use a five-tier geographic classification, slightly modified from the two previous classifications in order to take into account the continental level that is relevant for geoportals such as INSPIRE⁶⁴ and Africa soil⁶⁵. Moreover, we do not consider it relevant to differentiate several sub-national levels. Therefore we define “local” as anything below the national level. Based on these considerations, we propose the following nomenclature: local, national, regional, continental, and global.

For the third category (technology), we differentiate the geoportals developed using proprietary technologies (e.g., ESRI, that is the commercial leader in GIS) or open source technologies (e.g., Geoserver or Mapserver) to find out if some spatial or thematic communities are more committed to a technology than other communities. In order to find the technology used by the geoportal, the endpoint structure of the web services served by the geoportal is analyzed as most of the time proprietary geoportals are based on the RESTful service by opposition to geoportals developed on open source technologies that use OGC web services. We distinguish three main sources of geospatial servers that have their own web services syntax and allow us to determine the technology, as will be described in the next section:

- Geoserver: contains “/geoserver” in the endpoint
- Mapserver: contains “/mapserv?” in the endpoint⁶⁶
- ArcGIS server: contains “rest” in the URL of the webservices⁶⁷

⁶² <http://map.geo.admin.ch>

⁶³ <http://ede.grid.unep.ch>

⁶⁴ <http://inspire-geoportal.ec.europa.eu>

⁶⁵ <http://africasoils.net/services/data>

⁶⁶ http://mapserver.org/ogc/wms_server.html

⁶⁷ <http://www.iowaview.org/wp-content/uploads/2014/07/A-Brief-Explanation-of-Basic-Web-Services.pdf>

5.2.3.2 Classification

Based on the typology defined in the previous section, we have created an online table⁶⁸ (Google document) that allows classifying numerous geoportals covering different geographic scales and themes. The choice of these geoportals represents a wide sample covering a variety of scales and themes. All the information collected for assessing the various geoportals is reported in this online table that can publicly be viewed, including the formula used.

5.2.3.3 Assessment

In order to measure the completeness of the identified and selected environmental geoportals, we have selected some simple indicators for a rapid assessment. The goal of these indicators is to quickly evaluate if a geoportal fulfills certain criteria that make it fit-for-purpose as *“information accessibility is a key factor in allowing a virtual community to reach its goals”* (De Longueville, 2010).

In order to assess the facility to discover and access geospatial data through selected geoportals, we have grouped the indicators into three logical and sequential categories: (1) geoportal discovery; (2) data discovery in the geoportal; (3) data access and use in the geoportal. We did not take into consideration the design and usability of the geoportals as this requires a deeper analysis that cannot be performed in a rapid assessment on several dozens of geoportals.

We have defined two indicators for each of the three categories. This choice remains in line with the simplicity required by a rapid assessment. Each indicator is assessed through three possible values. The scoring system for each value is discussed in section 5.2.3.4.

5.2.3.3.1 Geoportal discovery

The facility to discover a geoportal is important in a community perspective as its participants need to quickly find a gateway to relevant geospatial data, saving them the time to look for each data set available on the Internet. Despite the existence of initiatives proposing unique entry points to Earth Observations data such as the Global Earth Observation System of Systems (GEOSS)⁶⁹ geoportal, the first reflex for many end-users remains to perform Internet searches, probably using the dominant Google search engine. Users might define their search using the word “spatial data” associated with the geographic place and/or theme of interest (e.g. “spatial data”, “West Africa”). They might also go one step further and look for an existing geoportal by using the appropriate syntax in the Internet search. As described and justified in Annex 1, we considered the number of results of the exact geoportal’s URL as the relevant information to determine the degree of facility for discovering a geoportal. However, a small regional

⁶⁸ <https://docs.google.com/spreadsheets/d/1d354bTEnAwpYx14YsZb-cNLBZQKR6iDIYpkwDQKsiho/edit?usp=sharing>

⁶⁹ <http://www.geoportal.org>

geoportal based on a specific environmental theme cannot claim to have the same number of Internet references as global multi-thematic geoportal. Therefore, the raw results must be weighted by the number of themes and the geographic scale targeted, as detailed in the section 5.4.1.

In addition to an Internet search, we also considered as relevant the number of scientific articles or more general literature categories such as reports referring to a given geoportal since environmental sciences require scientific data available through geoportals. To this end, we performed a series of tests (Annex 2) to determine the best syntax to use as well as the best literature search engine. We chose to use Google Scholar, that *“is an online, freely accessible search engine that lets users look for both physical and digital copies of articles”*⁷⁰, and the geoportal’s URL in quotes for performing this search. Compared to other literature search engines (e.g. Web of Science), Google Scholar has the advantage of referencing wider types of articles (e.g. conference proceedings, book chapters, technical reports). As with the geoportal’s web search, we need to weight differently global geoportals from local ones regarding their frequency of appearance in literature. Also, results of a geoportal’s search through literature search engines should be considered with caution given the time lag between the launch of a geoportal and its wider use and citations in scientific literature.

The two indicators selected for assessing geoportal discovery facility were then:

- The number of links pointing to the geoportal in a Google engine search
- The number of citations of the geoportal in Google Scholar

5.2.3.3.2 Data discovery

Once a geoportal is found, the second step consists in finding the desired data through this geoportal as the primary function of any SDI should be data discovery, enabling users to search, and evaluate data before downloading them (Nebert, 2005;Nogueras-Iso et al., 2005). We distinguish two ways of discovering data in a geoportal: (1) visual discovery: displaying various layers in a graphical interface, which allows one to determine if this corresponds to the data needed; (2) textual discovery: a description of the available data (metadata) is also essential as it gives the end-user all the details about the data (e.g. the method of creation, the quality).

The importance of data visualization in the discovery process is illustrated in Nativi’s ranking algorithms for the GEO Discovery and Access Broker (GEO DAB) where resources provided with a preview are ranked first (Nativi et al., 2015). The importance of metadata is highlighted by Maguire et al. (2005) who state that a catalog of metadata is a key component of any SDI. Moreover, web services, which can be defined as an *“applications running on a computer connecting to a remote web service via a URL allowing access to distributed data and services”* (Giuliani, 2011), are shaping today’s distributed architectures for geographic information (De Longueville, 2010) and in a service-oriented framework, discovery services are commonly provided by catalog and registry

⁷⁰ http://www.library.illinois.edu/ugl/howdoi/use_google_scholar.html

components (Nativi et al., 2009). Many applications show the importance of metadata catalogs (Giuliani et al., 2016; Giuliani et al., 2015; Nativi et al., 2015; Nogueras-Iso et al., 2005), standards, and web services for visualizing and accessing data and metadata. The web service typically used for visualizing data is the Web Mapping Service (WMS) (Open Geospatial Consortium, 2006) that renders an image (e.g. jpg, png, etc.), while the web service for querying metadata catalogs is the Catalogue Services for the Web (CSW) (Open Geospatial Consortium, 2007).

We have then selected the two following indicators for data discovery in the geoportal:

- Possibility to visualize the data available in the geoportal
- Availability of metadata through the geoportal

We have not considered the total number of layers available in the geoportal as an appropriate criteria of success because a local or thematic geoportal might only have a few relevant layers to share but may still completely fulfill their goal toward their particular user audience; and conversely, global geoportals may have many layers but be inadequate for the intended purpose.

5.2.3.3.3 Data access and use

After data discovery, the end-user also needs to be able to access the data itself, either through a simple download or more elegantly with web services conveying the data. The most commonly used geospatial web services for accessing data are the Web Feature Service (WFS) (Open Geospatial Consortium, 2005) for vector data and Web Coverage Service (WCS) (Open Geospatial Consortium, 2006) for raster data. The advantage of accessing the service instead of simply downloading the data is its accessibility from both conventional desktop GIS applications, as well as from thin client browser (Maguire et al., 2005). Accessing data through a web service also requires less manipulation than downloaded data, and ensures that data is always up-to-date since directly connected to the source.

Even if data can be accessed, its use and/or re-publishing might have some restrictions or might be totally free, in the sense of open data principles⁷¹ that state that data should be *“available without restrictions on use as part of the public domain”*. An illustration of the importance of and the efforts made for providing access to open data can be found in many projects, such as the open data inception (OpenDataSoft, 2015) or the Open Data for Africa (African Development Bank, 2014) projects.

We considered the two following indicators for evaluating data accessibility and use:

- Data can be downloaded or accessed through web services without a login
- Data is available as open data

⁷¹ <http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

5.2.3.4 Assessment and labeling

For each selected environmental geoportal, we assessed the various criteria defined above (Figure 2). The goal of this evaluation is to be able to put in place an easy and rapid visual appreciation of the geoportals to inform the developers of the geoportals and ultimately to extract some trends linked to community geoportals.

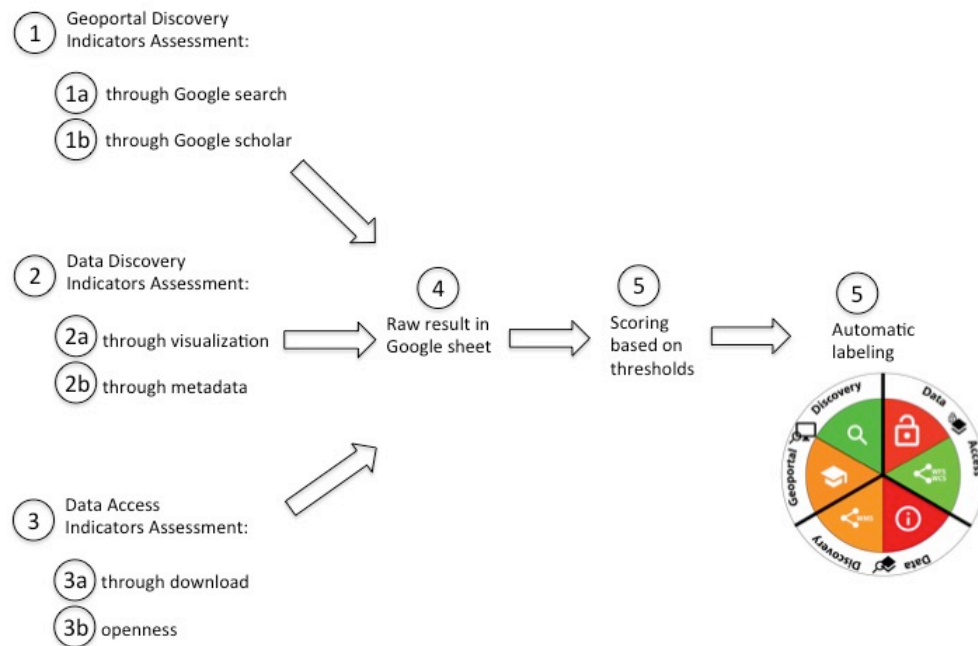


Figure 2: Geoportals assessment workflow

The assessment scoring is defined as follows and summarized in Table 4.

5.2.3.4.1 Geoportal discovery indicator

For the indicator showing the number of links pointing to the geoportal in a Google search (excluding the links from the domain name of the geoportal itself), we defined the following scoring scale:

Table 1: scoring of geoportal discoverability through a Google Search based on the number of results

Score	Number of results for a global scale geoportal	Number of results for a non-global scale geoportal
0	0-9	0-4
1	10-999	5-99
2	>1000	>100

This differentiation between global and non-global geoportals is made necessary by the need to take into account the typical broader audience of global geoportals. The thresholds chosen are static and a first attempt to differentiate

three categories of scores. The minimum and maximum limits of the class scoring 1 are particularly subjective and based on the authors best estimate but might need further refinement with more experience and the assessment of additional geoportals.

For the indicator showing the number of citations of the geoportal in Google scholar, we also differentiated the scoring of global geoportals with a higher citation expectation than the local ones by using the following scoring scale, also based on authors best estimate:

Table 2: scoring of geoportal's discoverability in Google scholar

Score	Number of results in Google scholar (non global geoportal)	Number of results in Google scholar (global geoportal)
0	0	0
1	1-3	1-9
2	>3	>9

The classification is automatically performed in the Google document, and for most indicators, except the ones linked to geoportal discovery, the users are limited to only a few possibilities of information capture, which has the advantage of standardizing the answers and facilitating automatization of the scoring through formulas. For all indicators, the possibilities of adding comments in the cells have been used by giving details on the search (e.g. syntax, endpoints found), which allows to reproduce the search if necessary.

5.2.3.4.2 Data discovery indicator

For the data discovery category, the indicator “possibility to visualize data available in the geoportal” can consist in the following answers:

- “no” in case no visualization is possible, which scores 0;
- “yes” in case visualization is possible but not as a web service (WMS or ESRI REST) or in case it is not possible to determine if a web service is used, which scores 1;
- “web service” in case data can be visualized through a web service (WMS or ESRI REST), which scores 2.

In order to determine this scoring the very first step consists in visually checking the geoportal and find if the layers can be visualized. If it is not the case, a score of 0 can be directly assigned. If layers can be visualized, it means that the next step will consist in determining if this visualization is made possible through web service or not. To this end, the first step is to read if this information is written in the geoportal; if it is not the case, three methodologies are proposed in annex 3: (1) to use a dedicated website interface, (2) to use a particular syntax for a search in Google, and (3) to use add-ons in web browsers. With these different methodologies, it is possible to determine the presence of web services

for most geoportals. We also made use of the possibility to add comments in the cells of the Google document to indicate which methodology has been used for each geoportal, which allows reproducing the search if necessary.

For the data indicator “availability of metadata in the geoportal”, the possible answers can be:

- “no” in case no metadata is available, which is scored as 0;
- “yes” in case metadata is available but not as a web service (CSW), which scores 1;
- “web service” in case the metadata is available through the CSW web service, which scores 2. The Catalogue Service for the Web (CSW) is the preferred and most widely used OGC web service for exchanging metadata over the Internet.

In order to determine if metadata was available through CSW from a geoportal, we started by visually checking if a link to a layer’s metadata or to a general metadata catalog was available. If it was the case, we checked the metadata catalog and tried to directly determine if it serves metadata through CSW. This is typically the case for Geonetwork instances. If this information could not be determined after a visual check, we performed a Google search with syntax similar to the one used for WMS. However, it is frequent that the CSW service corresponding to the metadata is not hosted on the same domain name as the geoportal for various reasons. This requires to perform a careful check of the geoportal and sometimes combine it with Google searches with different syntaxes (e.g. geoportal URL + inurl:csw; geoportal url+csw; geoportal url+geonetwork) to obtain the information. We systematically included in the comment of the Google sheet’s corresponding cell the syntax of the search performed for getting the information in the interests of reproducibility.

5.2.3.4.3 assessment of data access indicators

The scoring of the data download ability indicator is performed as follows:

- “no” in case data cannot be downloaded or accessed via web services (WFS or WCS), which scores 0;
- “direct” in case data can be accessed through direct download but not through a web service, which scores 1;
- “web service” in case data can be accessed via WFS or WCS web services, which scores 2.

The first step for this indicator’s assessment consists in visually checking in the geoportal if there is any indication for a data download or a service availability. The next step consists in determining if a WFS/WCS web service exists, which can be done through the same methods as the ones explained above for the WMS service (e.g., Spatineo search, Google search, use of web browser add-ons). Additionally, if a WMS service is found during the assessment of the visual discovery possibilities, a simple change from WMS to WFS and WCS in the endpoint syntax in the search might directly give the information about the

availability of such web services.

There might also be some specific cases where data served through the geoportal is linked to several data sources with different data policies, where some layers are available without restriction for download or via web services. In such cases, the interest of this assessment being simply to check if the geoportal includes web services, we scored it “Web service” if at least a few layers are available via web service.

For scoring the last indicator “data openness”, several cases might occur:

- no data is accessible as determined in the previous indicator, which means that it is assessed “no data” and scores 0;
- data might be accessible but only through a login for registered users (with or without an associated fee); it is also assessed “no data” and scores 0 as it is not openly accessible to the public;
- data might be accessible with a restriction regarding its use; this is assessed “restricted” and scored 1 as it cannot be freely used. This also applies to geoportals serving various data, including a mix of open and restricted data with a clear mention of use restriction for some of the data served;
- data is freely accessible and no particular restriction about its use/republication is indicated on the geoportal or in the metadata. We assumed that if the data had some use restrictions, this would be indicated either in the geoportal or in the accompanying metadata; we assessed such data as “open” and it scored 2. It often happens that commercial use of the data is prohibited; in such a case, we still consider the data as open since the commercial use of data is another debate that should not hide the open characteristic of data for noncommercial use;
- various data sources are served through the geoportal and metadata is provided for some or each of them (e.g. in the case of a metadata catalog). Given the difficulty of individually checking each metadata record, we considered such a case as “various” and it is scored as 2 since we assumed that at least part of the data could be used.

These different cases and their scoring are summarized in Table 3 below:

Table 3: data openness scoring

Case	Standardized attribution	Score
No data accessible	No data	0
Data accessible through login	No data	0
Data accessible but with restricted use	Copyrighted	1
Data accessible and no particular indication on data restriction on the geoportal	Open	2

Various data sources served through the geoportal and potential diverse data policies	Various	2
---	---------	---

To determine data openness, the first step consisted of checking the previous indicator of the ability for data to be downloaded. If it was not the case, it was directly assessed as “no data”. If it was downloadable, we checked in the geoportal or in the metadata if there was any data-related legal use indicated, even though this might be bypassed when using a web service as discussed in Section 5.2.5.

5.2.3.4.4 Labeling

The scores for each indicator are not cumulative because geoportal discovery, data discovery, and data accessibility are independent aspects of the assessment of geoportals, and these indicators cannot be combined into an overall score. Instead, obtaining an evaluation of each defined indicator gives a finer assessment of the completeness of the various geoportal components. Besides, the translation of each indicator’s score into a self-speaking qualitative representation to deliver a simple and strong visual message should further help promoting SDI best practices among SDI practitioners.

To this end, we proposed the “EGAL” label wheel (Figure 3) that is made of six distinct parts, visualizing the status each of the six indicators for a given geoportal. It has the advantage of giving each indicator an individual visibility symbolized by a corresponding logo and an assessment value with an associated color. The six parts of the label are distributed into three equal parts, represented by thicker lines, representing the three main assessment categories (geoportal discovery, data discovery, data access). Moreover, each category background can be of three different colors: red in case the assessment’s result is 0, orange in case the assessment’s result is 1, and green in case the assessment’s result is 2. These colors are self-speaking as most people have an immediate understanding of the positive aspect of green, the negative aspect of red and the intermediate aspect of orange (e.g. daily visualization of traffic lights using these colors).

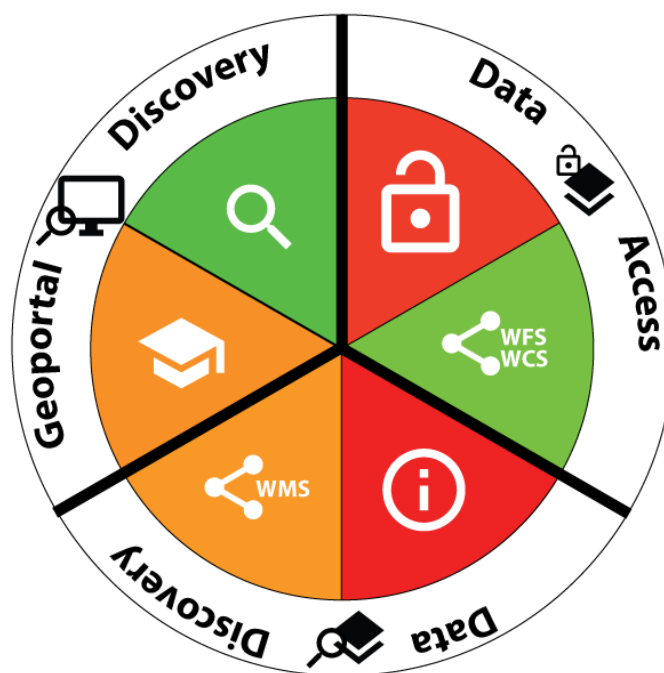


Figure 3: the "EGAL" label

The Table 4 below summarizes the various categories, indicators, scoring criteria, and associated symbols:

Table 4: Geoportals assessment and scoring criteria

Category	Indicator	Assessment criteria	Score	Color
Geoportal discovery 	Number of links pointing to the geoportal in a Google engine search 	• 0-9 (if global scale) • 0-4 (if non-global scale)	0	red
		• 10-999 (if global scale) • 5-99 (if non-global scale)	1	orange
		• > 1000 (if global scale) • > 100 (if non-global scale)	2	green
	Number of citations of the geoportal in Google scholar 	0 citations	0	red
		• 1-9 citations (if global scale) • 1-3 citations (if non-global scale)	1	orange
		• > 9 (if global scale)	2	green

		scale) • > 3 (if non-global scale)		
Data discovery 	Possibility to visualize the data available in the geoportal 	No visualization of the data possible	0	red
		Data can be visualized but not through a web service	1	orange
		Data can be visualised through a web service	2	green
	Availability of metadata through the geoportal 	No metadata available	0	red
		metadata available but not through a web service	1	orange
		Metadata available through a web service	2	green
Data access and use 	Data can be downloaded without login or accessed through web services 	Data cannot be downloaded or accessed through WFS/WCS	0	red
		Data can be downloaded but not through a web service	1	orange
		Data can be downloaded through a web service	2	green
	Data is usable 	No data accessible Data accessible through login	0	red
		Data available but copyrighted	1	orange
		Data available and no data use indication Diverse data policies	2	green

5.2.4 Results

5.2.4.1 Typology

At the time of writing this paper, 31 geoportals were assessed using the proposed methodology. The results can be visualized in a dedicated website⁷² and the details in a publicly visible Google sheet⁷³.

Table 5: Geoportals assessment results by typology category

Category	Attributes	Raw number	Unique number
Geographic scale	Local	7	1
	National	19	4
	Regional	18	3
	Continental	15	1
	Global	6	4
	Multi-scale	18	-
	unknown	0	-
Theme	Food security and sustainable agriculture	17	0
	Biodiversity and ecosystem sustainability	20	0
	Cadaastre	12	1
	Climate	19	2
	Disaster resilience	15	2
	Infrastructure and transportation management	13	0
	Energy and mineral resources management	12	0
	Public health surveillance	12	0
	Topography	15	3
	Water resources management	21	1
	Sustainable urban development	11	0
	Multi-theme	22	-
	unknown	0	-
	Proprietary	2	2
Technology	Open Source	22	22
	Multi-technology	0	-
	unknown	7	-

Out of these 31 geoportals, the vast majority (18) has multiple spatial scales, meaning that it does not address a unique scale. Among the geoportals addressing a unique scale, four target the national and global scale, three target the regional scale, and one addresses both the local and the continental scales.

Similarly to the scale, the majority (22) of the geoportals assessed are multi-thematic. Among the ones targeting a single theme, there are three geoportals for the “topography” theme only; two geoportals for the “climate” and “disaster resilience” themes only; and one geoportal for “cadaastre” or “water resources management” themes only.

The vast majority of the geoportals assessed (22) use open source technology (e.g., Geoserver or Mapserver) to serve the geospatial data through the geoportal, while only three geoportals use proprietary technology. One geoportal uses no map server of any kind, and it was not possible to determine the technology used for six geoportals.

⁷² <http://gala.unige.ch/EGAL>

⁷³ <https://docs.google.com/spreadsheets/d/1d354bTEAwPYx14YsZb-cNLBZQKR6iDIYpkwDQKsiho/edit?usp=sharing>

5.2.4.2 Scoring

Regarding the indicators assessed to evaluate the geoportals, the majority of them scored well regarding the scoring criteria defined in the methodology as illustrated in Table 6, and further discussed in the next section.

Table 6: results of the indicators assessment

	Geoportals discovery		Data discovery		Data access and use	
Score	Google search	Google Scholar search	Visualization	Metadata	Data access	Data openness
score 0	4	7	0	5	4	4
score 1	7	5	2	6	6	3
score 2	20	19	29	20	21	24

The majority of the assessed geoportals (20) obtained a good scoring in a Google or Google scholar search, whereas a smaller number could not be found or only had a few references. The trend is similar in both types of searches.

The data visualization assessment is the indicator where the geoportals perform the best as we did not find a single geoportal without data visualization. Moreover, most of them (29) give access to this visualization through a web service (WMS). Data discovery through metadata performed also well with the majority of geoportals (20) listing a web service to access metadata. Five geoportals do not provide metadata at all, while six of them do not provide metadata through a web service.

In the same trend as data discovery, data access through web services was found for the majority (21) of the geoportals with only a few (four) not providing access to data, and six of them providing access to data not through web services.

Finally, the majority (24) of the data served through the geoportals linked to terms of use that allow a free data re-use, which we considered as open.

5.2.5 Discussion

5.2.5.1 Main findings

The classification of the geoportals in the proposed typology revealed that geoportals serving multi-scale and multi-thematic data are much more frequent than the ones serving monoscale or monothematic data. The environmental field being multi-disciplinary by nature, finding various related thematic data in

environmental geoportals is not surprising. For example, a geoportal on drought is expected to serve at the same time data on climate, water, weather, and possibly other themes. The reasoning is similar for scale as depending on the scale data might come as an aggregate of smaller scales. For example, a European geoportal might serve at the same time data at local level, national level and at regional (European) level. A geoportal proposing such a choice can even be considered more complete than a geoportal serving only monothematic and/or monoscale data as users will not only find data they were originally looking for but also related data considered as relevant by geoportals developers.

If we consider the geoportals covering at least 7 themes out of 11 (which is the case for 12 geoportals), it is interesting to note that in most cases they are not global geoportals but smaller scale (local/national/regional) ones. This might be explained by a general better availability of data at local/national levels than at global level. It might also signify that local/regional/national geoportals focus more on the geographic extent specificity of the data served than on the theme itself, whereas global geoportals usually target a particular thematic. This typology should be further explored by the analysis of more geoportals to confirm the supposed trend.

Regarding the technology used in the geoportals analyzed, the proposed methodological approach of endpoints analysis was relevant as it allowed to determine the technology used in 24 out of the 31 geoportals. It showed that the vast majority of the geoportals assessed (22) use open source technology, either Geoserver or Mapserver, while only two geoportals use a proprietary technology. Given the small number of geoportals using proprietary technology, it is not possible to determine a community-specific technological trend. We can simply note that open source technology seems to have the preference for environmental geoportals developers. The open source technology advantages such as performance, possibilities of customization, low price, favorable national or institutional policies might particularly suit research communities.

The geoportals assessment through the six indicators revealed a general positive scoring for most indicators, which might be due to subjective and too generous scoring thresholds that might need further discussion and refinement.

The Google and Google Scholar search methodologies allowed geoportals discovery in most cases but this implies one knows the geoportal's URL for the reasons explained in Annex 1, which might be a barrier to an optimal discovery. This shows the importance of having a tool other than Google to search for such information, which is exactly what the Global Earth Observation System of Systems (GEOSS) wants to achieve through the GEOSS geoportal⁷⁴. Unfortunately at the time of writing some back-end technical weaknesses, mainly due to poor semantic results not providing a satisfactory user experience, still prevent the expected role of the GEOSS platform to be fulfilled. However, GEOSS has the potential to become the ideal discovery tool for geoportals and data if these associated issues are solved. The concept of Linked (Open) Data (Knibbe, 2014;Kuhn et al., 2014), aiming at allowing navigation and search directly

⁷⁴ <http://www.geoportal.org>

through data on the web as it is done currently with webpages, is another promising development, which still needs further implementation and diffusion.

Data visual or textual discovery was also quite satisfactory in the geoportals analyzed as most of them allow for data visualization or provide metadata. However, the difficulty in finding the web services endpoints for data visualization (WMS) or metadata (CSW) is one of the most obvious weaknesses of most geoportals assessed. Finding this information was the most challenging and time consuming operation whereas web services are supposed to facilitate data discovery and access. Most of the time it was necessary to either use a web browser add-on (e.g. Firebug) or to perform a Google Search with a particular syntax described in the methodology to find this information. In several cases the web service is not hosted on the same domain name as the geoportal, making the search even more difficult.

The same discussion applies to data access as the scoring was quite good since it was possible most of the time to access data, including by web services (WFS or WCS), but finding these web services was the biggest challenge. An additional difficulty is that a geoportal is often just a gateway serving layers from various sources (e.g. portal.geology.org). The data access service is then dependent on the layers providers, making it difficult to determine the geoportal's data accessibility unless clearly stated. An interesting workaround to this issue is the possibility to save the "Web Map Context" of a cartographic composition made on the geoportal. The resulting file provides the various web service endpoints used for the map composition.

In addition to the web services discoverability and access issues just mentioned, the complex overall architecture of some geoportals also made it very difficult to discover and access data. These findings show that even though technology exists (e.g., geoportals, web services), the main issue remains a proper coordination of the geoportal's components that should implement the vision of the geoportal's owner. Simple solutions exist to make the web services endpoints more easily discoverable. One can make more systematic use of geospatial content management systems (e.g. Geonode⁷⁵) that optimize geoportal basic functionalities and already contain a dedicated section grouping the various endpoints. Alternatively, a geoportal not using a geospatial content management system could simply contain a dedicated section listing the web services endpoints, taking hence advantage of the usefulness of web services.

For the data openness indicator, the majority of geoportals (24) scored 2 as in most cases a general data policy indicates the possibility to re-use the data without restriction (except for commercial use) unless otherwise stated. A difficulty lies in the geoportals serving data from various sources, which might all have a different data policy. In such cases it is difficult to know if there are restrictions on the data use, making it crucial to have a metadata catalog that allows defining the data policy independently for each layer. Another alternative consists in accepting some agreement before accessing data (e.g. by a simple click on a "legal" box). But these solutions still do not prevent someone knowing

⁷⁵ <http://geonode.org/>

the web service access endpoints (e.g., WFS or WCS) to bypass these restrictions, which advocates for a broader use of open data that is not concerned by these restrictions and thus easier spread and used.

5.2.5.2 Limitations

Firstly, the number of geoportals assessed allows testing the methodology but is too small to derive solid general characteristics of the geoportals in link with their typology, but we think that this situation will evolve as more geoportals will be assessed.

Secondly, there is some subjectivity in several parts of the methodology:

- when categorizing the geoportals in the typology based on the number of layers served, it might be difficult to check each of them to determine with certitude all scales and themes proposed, which requires to take into account a margin of error; and
- thresholds between the categories of some indicators are potentially subject to refinements as more geoportals are analyzed. This is particularly true for the indicators linked to geoportal discovery (Google and Google scholar searches). One solution to overcome the subjectivity inherent to the thresholds of these indicators would consist in making them dynamic: these thresholds would be permanently automatically calculated in function of raw results in order to always separate three equal categories. Assuming that regular assessments of all geoportals are made, this could mean that a given geoportal might change of category depending on iteration, which highlights the need for geoportals to be referenced and cited for constantly being more discoverable. Apart from the thresholds, the difference between the categories 1 (in orange) and 2 (green) is sometimes fuzzy and subject to interpretation and possible biases. This is for example the case when it is necessary to determine if a web service can really be considered as relevant and justifying that the geoportal is labeled “web services”. As an illustration, in the case where a web service endpoint exists but is not working after several attempts, we decide to score it 0. This raises the need to have additional sub-indicators for several of the defined indicators. For example, the facility to discover web services might be a useful sub-indicator for data discovery and data access. Regarding data copyright, a further distinction might be made by differentiating copyrights just asking for source acknowledgment from other one that are more restricting.

Thirdly, the use of Google scholar to determine if a geoportal is known through citations does not take into account the time lag between the online launch of a geoportal and the time at which articles mentioning it are published and referenced by Google Scholar.

5.2.5.3 Recommendations and perspectives

Based on the outcomes of the assessment and the discussion, we think that certain elements should be taken into account in geoportals development or updates for the benefit of the environmental community:

- A geoportal should always contain a dedicated section where users can easily find the webservices endpoints (URLs). In case of a geoportal serving data from several sources, such a dedicated section should also contain the webservices endpoints of the various sources;
- The data policy should clearly be stated in the geoportal. In case of various sources, a proper link should be added to each individual data policy;
- Unless there are reasons to copyright the data, open data should be privileged as it really unlocks the power of data for societal benefits;
- A small description of data types and themes available in the geoportal on its homepage would be an asset allowing users to directly know if the geoportal suits their needs; and
- Coordination being the major success factor for an SDI (SoP 2003), this also applies to geoportals that are the visible part of an SDI and should reflect the vision and coordination of the developers, for example through a proper architecture, data policy, and web services presentation.

In addition to the elements presented so far, this paper highlights the needs for a better management of geoportals. Some solutions such as live monitoring dashboards for geoportals similar to the one existing for metadata in Europe⁷⁶ but using criteria defined in this article might help to ensure a constant quality of web services and might be part of a set of best practices that would need promotion at international level, for example through the Group on Earth Observations (GEO) level.

We foresee as future steps the need to refine the indicators used and add sub-indicators to this end. Coupling this with a live monitoring dashboard and an appropriate communication strategy on the existence and usefulness of our geoportal assessment could help constitute useful capacity building material. A deeper analysis using a larger sample of geoportals is also needed to better reveal trends in geoportals typology depending on communities.

The current website⁷⁷ developed for visualizing geoportals assessment results could evolve towards a more interactive solution allowing geoportals developers to submit their geoportal's url for assessment through the EGAL procedure. This could favor an increase of the quality and usefulness of geoportals over the long run. Such a website would still require human intervention for some parts (e.g. for check the level of data openness) but some of the assessment procedures could be automatized upon uploading of a geoportal url, for example:

- an automatic check of the server at regular intervals to determine if it is up or down;
- an automatic check of the availability of the various web services (e.g. CSW, WMS, WFS, WCS)

A further research analyzing the other SDI components (people, policy) of some

⁷⁶ <http://www.europeandataportal.eu/mqa-service/en/dashboard.html>

⁷⁷ <http://gala.unige.ch/EGAL>

geoportals, especially the ones that obtained a good or a weak score would complement the geoportal typology for best practices recommendations at a given geographical scale or for a given thematic community. The methodology and label presented in this paper are discussed from the environmental perspective, using for example the environment-related GEOSS SBAs themes for classifying the geoportals. But this is not limited to environmental geoportals and could very easily be extended to other scientific or even non-scientific domains using geospatial data, mainly by replacing the geoportals themes for classification.

A more general question should also be raised about the complexity of data discovery, access and representation through the traditional division between data, metadata, and portrayal. It would be interesting to determine to what extent the assessed criteria would improve with data, metadata and legend integration in a common format such as netCDF⁷⁸. This would reduce the complexity of data discovery, access and portrayal⁷⁹, the latter being of particular importance for Communities of Practice (e.g. geological maps) or institutional representations (e.g. national topographic maps).

5.2.6 Conclusions

The overarching goal of this paper was to determine if environmental geoportals play their role of gateways to useable geospatial data for addressing environmental challenges, and if some characteristics can be extracted for thematic/geographic communities of practice based on a typology of geoportals. To address this question, specific objectives have been set up and implemented.

1) A general SDI typology was successfully fulfilled. This typology is based on fundamental criteria on which communities of practice are built: geographic scale, theme and technology used. For each of these categories, relevant attributes have been determined based on literature review and knowledge of the SDI domain.

2) Geoportals on the environment were identified, selected and classified according to this typology. This was successfully achieved, even though subjectivity issues were discussed in the previous section and might require some consolidation. The main outcomes of this classification revealed that most geoportals assessed are multiscale, multithematic, and use open source technology.

3) Basic indicators to perform a rapid assessment of the selected geoportals were developed and tested to assess geoportals discovery on the Internet, data discovery, access, and use through these geoportals. This methodology highlighted the importance of the syntax in Google searches and the advantages of standardized web services.

⁷⁸ <http://www.unidata.ucar.edu/software/netcdf/>

⁷⁹ <http://www.geo-solutions.it/blog/netcdf-grib-support-geoserver>

4) The proposed methodology was applied to obtain a score for each of the six indicators tested for each geoportal. It also consisted in providing a simple and visual message for which we proposed the “EGAL” label. The importance of weighing certain results according to the type of geoportals for the final scoring was demonstrated and the application of the methodology revealed important weaknesses in the discovery of web services, which affects data discovery and access. The importance of geoportals architecture, for example regarding data policy, was also discussed in light of the findings of the assessment and some solutions were proposed to easily overcome some weaknesses.

5) Characteristics of geoportals could be deduced from the typology and the assessment. This could only partly be addressed as bigger samples of geoportals need to be classified and assessed. Another part of the objective was to issue some recommendations, which could be achieved even if these recommendations remain general to all geoportals, independently from targeted communities.

Finally, this assessment through the methodology developed could give a first overview of the geoportals quality and revealed common weaknesses preventing them to optimally play their role as gateways to environmental data. Improvements and future steps have been proposed to further refine the assessment methodology and we are confident that some of the quick wins proposed, along with recommendations, and visual labeling can help to improve geoportals quality to support more efficient technical SDI implementation.

5.2.7 Annexes

Annex1: Google search syntax results for known geoportals

The Google search engine allows performing searches based on proprietary algorithms. In order to determine the best way to discover a geoportal, various combinations of keywords in a Google search have been performed on 07.02.2016. For this test we considered two geoportals, well known by the authors for having participated to their development: the PREVIEW geoportal⁸⁰ giving access to global data linked to risk from natural hazards and the ECOWREX geoportal⁸¹ showing West African data on Renewable Energy.

The search criteria took into account the language (searches in English and French, relevant in both cases), the possibilities offered by Google search operators⁸² to refine the search: quotes (“), dash for exclusion (-) and site (site:) for a whole website’s exclusion in the results. In each case we analyzed the first two pages of results (corresponding to 20 results) to determine if the result returned by Google was relevant. Our “relevance” criteria were the following:

⁸⁰ <http://preview.grid.unep.ch>

⁸¹ <http://www.ecowrex.org/mapView>

⁸² <https://support.google.com/websearch/answer/2466433>

- the result should mention the exact geoportal's URL or at maximum be one click away from the geoportal itself
- we did not take into account results that would link to a layer served by a cartographic server linked to the geoportal (ex: WMS service), but only the geoportal itself
- we accepted any result's format (html page, pdf document, powerpoint, ...)

The results when searching by keywords were disappointing and only a few links to the geoportal could be found through this combination of keywords. Besides, these results were not consistent as for example a strict search "ECOWREX geoportal" would return zero result, whereas a similar strict search in French or a search without quotes in English and in French would return two results. Consequently, we cannot consider this methodology as relevant to be used for determining if a geoportal is easily discoverable or not as the results are not robust enough.

On the other hand, we have considered relevant a Google search with the exact geoportal's URL from which we exclude the results of the geoportal's hosting domain name. This exclusion is necessary as the results would be biased by the number of internal links pointing to the geoportal. Such a Google search is relevant as it returns an information on the geoportal's presence through the number of links pointing to the geoportal, which is an indicator of other resources available through Internet that judged relevant to refer to the geoportal, and hence an indicator of success.

Basing the search syntax on the geoportal's URL with exclusion of the domain name, we still tried a few combinations (with or without quotes for the URL and with or without the "http://" in front of the URL). Here again the number of results varies drastically. We decided then to:

- keep the quotes for the URL to make sure only the path to the geoportal is considered and not only part of it
- remove the "http://" in front of the URL as: (1) it is frequent that URLs are indicated without the http protocol in front and (2) the rest of the URL is anyway encompassed in results containing the http.

It means that we have performed searches on the following model:

"preview.grid.unep.ch" -site:.grid.unep.ch

Table 7: Google test geoportals search results based on various syntaxes

Syntax	Google number of relevant results
ECOWREX geoportal	2
"ECOWREX geoportal"	0

ECOWREX geoportail	2
"ECOWREX geoportail"	2
ECOWREX portal	3
"ECOWREX portal"	0
ECOWREX portail	0
"ECOWREX portail"	0
http://www.ecowrex.org/mapView - site:http://www.ecowrex.org	75
"http://www.ecowrex.org/mapView" -site:http://www.ecowrex.org	9
www.ecowrex.org/mapView - site:http://www.ecowrex.org	76
"www.ecowrex.org/mapView" site:http://www.ecowrex.org	37
PREVIEW geoportail	3
"PREVIEW geoportail"	2
PREVIEW geoportail	1
"PREVIEW geoportail"	0
PREVIEW portal	0
"PREVIEW portal"	0
PREVIEW portail	0
"PREVIEW portail"	0
http://preview.grid.unep.ch -site:.grid.unep.ch	316'000
"http://preview.grid.unep.ch" -site:.grid.unep.ch	1'130
preview.grid.unep.ch -site:.grid.unep.ch	71'200
"preview.grid.unep.ch" -site:.grid.unep.ch	5'000

Annex2: Scientific articles search websites comparison

Based on the findings of annex1 where it appeared that a free text search for a geoportal is too random to be used, we decided to also use the geoportal's URL to make a search in the scientific literature. We assume that if a scientific article refers to a geoportal, its URL is mentioned in the article in most cases.

Besides deciding that we would use the geoportal's URL as the search string in scientific articles, we also needed to decide the exact syntax of the URL as well as the search engine for performing the searches as a growing number of online tools allow to search for scientific literature. Among the most famous tools, we selected: Google Scholar⁸³, Web of Science⁸⁴, Sciencedirect⁸⁵ and Researchgate⁸⁶. Haddaway (2015) differentiates academic citation databases (e.g. Web of

⁸³ <https://scholar.google.com/>

⁸⁴ <https://apps.webofknowledge.com>

⁸⁵ <http://www.sciencedirect.com/>

⁸⁶ <https://www.researchgate.net/>

Science) from academic citation search engines (e.g. Google Scholar, Microsoft Academic Search) and concludes that Google Scholar is a useful platform for searching for environmental science grey literature, which is an asset in our purpose of analysing geoportals presence in scientific literature. In order to find the most appropriate syntax and tool, we compared several geoportal search strings using several different tools as illustrated in Table 8 on 08.02.2016.

Table 8: comparison of search strings and tools

Syntax	Google Scholar relevant results	Microsoft Academic Search	Science direct relevant results	Researchgate relevant results	Web of Science relevant results
http://www.ecowrex.org/mapView	0	0	0	0	0
" http://www.ecowrex.org/mapView "	1	0	0	0	0
http://www.ecowrex.org/mapView - site:http://www.ecowrex.org	1	0	0	0	0
" http://www.ecowrex.org/mapView " - site:http://www.ecowrex.org	1	0	0	0	0
www.ecowrex.org/mapView - site:http://www.ecowrex.org	1	0	0	0	0
"www.ecowrex.org/mapView" - site:http://www.ecowrex.org	1	0	0	0	0
"www.ecowrex.org/mapView"	1				
http://preview.grid.unep.ch	94	1	0	0	0
" http://preview.grid.unep.ch "	81	0	8	0	0
http://preview.grid.unep.ch -site:.grid.unep.ch	81	1	0	0	0
" http://preview.grid.unep.ch " -site:.grid.unep.ch	81	0	0	0	0
preview.grid.unep.ch -site:.grid.unep.ch	86	0	0	0	0
" preview.grid.unep.ch "	86	0	0	0	0

-site:.grid.unep.ch					
"preview.grid.unep.ch"	94	0	0	0	0

These results show the undeniable better performance of Google Scholar for providing relevant results of the appearance of a given string search. We decided then to use Google Scholar as the tool and the geoportal's URL between quotes as the syntax, to make sure that the exact URL is being searched for in both peer-reviewed and grey literature.

Just like for the Google Search engine described in Annex1, we did not use the "http://" before the URL, and we used quotes for the URL. On the other hand, we did not exclude the hosting website from the search as it might contain useful literature about the geoportal, which does not bias the results. The final string of the searches in Google scholar is similar to the following example:

"preview.grid.unep.ch"

Annex3: Visualization web services search

Once it is established that a geoportal provides a preview of geographic layers, it is still needed to determine objectively if the layers displayed are served through a web service, such as the Web Mapping Service (WMS) or the ESRI REST. To this end, we propose three methodologies, in order of increasing difficulty:

The first method after a visual check of the geoportal consists in using a dedicated interface called “Spatineo”⁸⁷ that is a directory service allowing performing searches to retrieve Web services from URLs. It allows also limiting the results by WMS, ESRI REST and/or other types of web services. The only action needed is to enter the geoportal’s URL in the search interface and it will return matching web services.

In case Spatineo does not return results for the searched URL, two other methods are still possible to try. The first one uses the Google Search engine and consists in determining if it exists an URL containing the web service. This is made possible by the fact that web services follow a particular syntax, containing:

- an endpoint (e.g. <http://preview.grid.unep.ch:8080/geoserver/ows?>)
- the service name after the endpoint (e.g. service=WMS)

Alternatively, the endpoint can directly be a wms endpoint (e.g. <http://preview.grid.unep.ch:8080/geoserver/wms?>)

In both cases, the word “wms” appears in the URL. Kliment et al. (2013) propose a syntactic method to discover available OGC services endpoints, which consists in using the “inurl:wms” word in a Google search containing the geoportal’s name. For example, if we take the “PREVIEW” geoportal (<http://preview.grid.unep.ch/>) and we want to know if it has a WMS service, we could perform the following Google search: <http://preview.grid.unep.ch/> inurl:service=WMS. This returns several results showing that links exist with the WMS service available in this geoportal. We did this same type of search on three geoportals (results are available in Table 9 below) with various syntaxes and except for the Preview geoportal, the results are not always satisfactory with this methodology. We can nonetheless make the following recommendations with this type of search:

- use the general domain name (e.g. <http://www.ecowrex.org>) of the geoserver and not the precise geoportal’s address (e.g. <http://www.ecowrex.org/mapView>) as the geoserver instance serving the web service is most of the time located in a different folder of the domain name (e.g. www.ecowrex.org/geoserver/ecreee/wms)
- use the “http://” syntax in front of the general domain name and the quotes for the URL as this narrows down the search

A Google search with this methodology should hence have the syntax:

“<http://www.ecowrex.org>” inurl:wms

But as shown in the table below, this methodology might not work in some cases, reason why we suggest a last option, which is to use web browser add-ons (e.g.

⁸⁷ <http://directory.spatineo.com/>

Firebug or Console in Firefox, Console in Chrome, ...). Such add-ons allow to monitor live behavior of a website by displaying a large panel of information returned. Using the “network” tab in Firebug allows to see which requests are performed when displaying a map. As shown in Figure 4, this clearly shows that WMS services are called. The advantage of this method is that it can also show if ESRI REST web service is used.

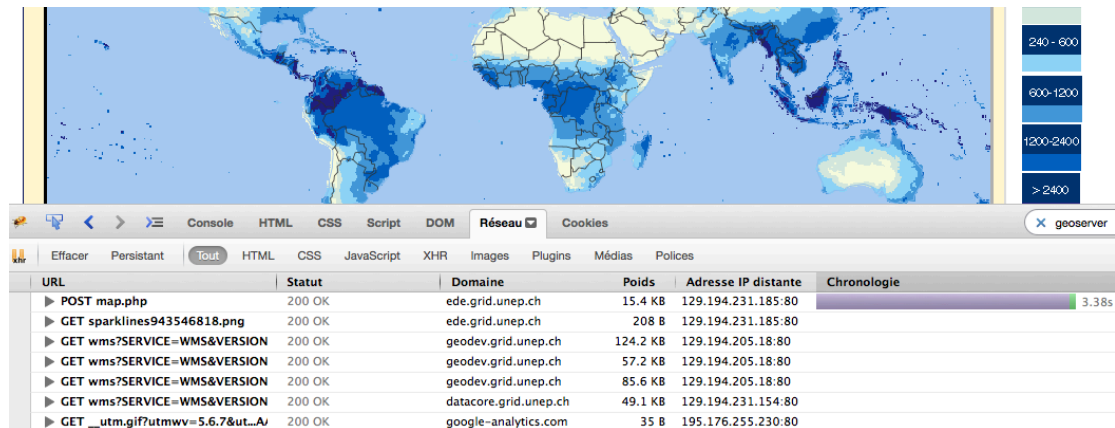


Figure 4: Firebug result for a WMS search on the <http://ede.grid.unep.ch> geoportal

Table 9: Google search to determine if the word "wms" is contained in a URL

Syntax	Number of relevant results returned
http://www.ecowrex.org/mapView inurl:wms	0
" http://www.ecowrex.org/mapView " inurl:wms	0
www.ecowrex.org inurl:wms	2 (out of 4)
" www.ecowrex.org " inurl:wms	2 (out of 4)
http://www.ecowrex.org inurl:wms	2
" http://www.ecowrex.org " inurl:wms	2
http://preview.grid.unep.ch inurl:wms	5 (out of 6)
" http://preview.grid.unep.ch " inurl:wms	5 (out of 6)
preview.grid.unep.ch inurl:wms	6 (out of 7)
" preview.grid.unep.ch " inurl:wms	6 (out of 7)
http://ede.grid.unep.ch inurl:wms	0
" http://ede.grid.unep.ch " inurl:wms	0
ede.grid.unep.ch inurl:wms	0
" ede.grid.unep.ch " inurl:wms	0

5.3 The EOPOWER impact assessment framework

The EOPOWER impact assessment framework has been specifically developed to assess the impact of the EOPOWER project activities and of selected Earth Observation solutions (Noort, 2014, 2015b). The originality of this framework lies in the possibility it offers to capture not only economic benefits of technological innovations such as Earth Observation solutions, but also other aspects currently not captured in economic calculations, such as benefits related to sustainable management of natural resources. This framework is necessary in a context where usefulness of Earth Observation application is not self-evident and its benefits need to be demonstrated.

The framework consists in three major parts: (1) a general reference framework; (2) a set of indicators; and (3) a rating of the business environment. Depending on the Earth Observation solution assessed, the paradigm shift or the business environment parts might not be relevant and hence be skipped.

The general reference framework part aims at determining if the Earth Observation solution assessed provides an added value compared to the current situation. To this end, a series of questions are asked and next steps are suggested (Figure 11):

- Does the new solution cause a paradigm (i.e. model) shift? As an illustration, a technical innovation such as very high resolution (VHR) satellite imagery could create a shift of the paradigm from landscape level to farm scale in the sense that information previously available at landscape level could now be available at a farm's level given the additional accuracy.
- Is the current process (business process or organization process) improved with the Earth Observation application?
- Does the application provide economic value that can be quantified?
- Is a clear measurable goal defined to which the Earth Observation application contributes?
- Is a future payment scheme or other economic mechanism foreseen in which the Earth Observation application fits?

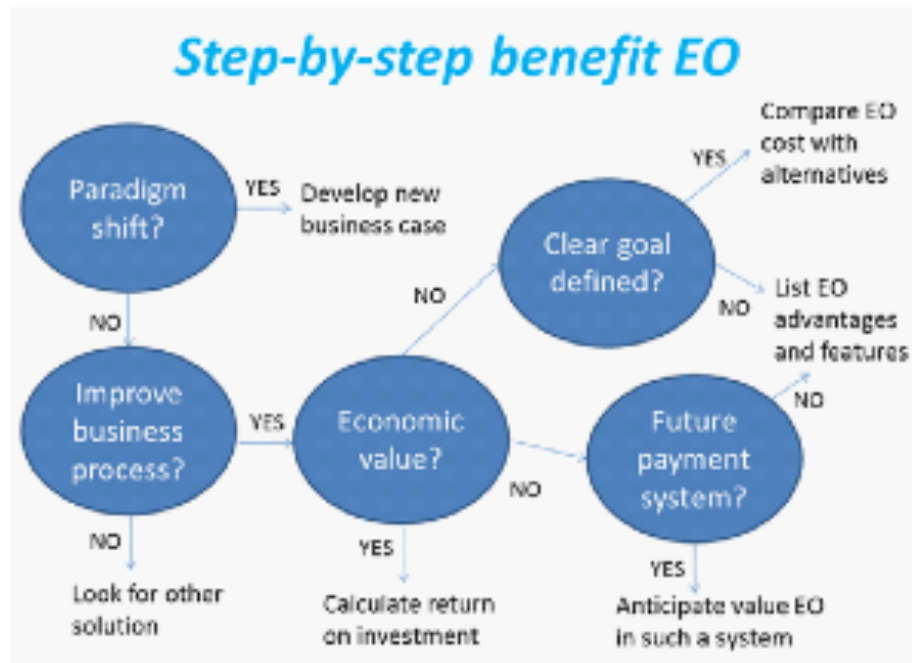


Figure 11: step-by-step benefit of Earth Observation application (Noort, 2013)

If the Earth Observation application assessed presents an added value, the next step consists in examining the various aspects of the application through a series of indicators. These indicators allow taking into consideration not only the technical aspects of the application but also their social and economic aspects. There are eleven indicators to assess for a given Earth Observation application:

1. **Fit-for-purpose:** does the product or service do what it is supposed to do to solve a certain problem?
2. **Comparative advantage:** what does the product or service do significantly better than other solutions to the same problem?
3. **Complexity/ease-of-use:** are the users able to work with the product or service?
4. **Elegance:** does the product or service create a sense of community belonging once it is mastered?
5. **Cost-benefit:** is the cost-benefit of the product or service attractive even in the long term?
6. **Sustainability:** can the product or service be delivered when needed, taking into account its sustainability (long term availability, availability of funding to sustain it in the long term, long-term institutional support, long-term user interest)?
7. **Resilience:** in case of breakdown in one of its elements, can the product or service still be delivered at an acceptable level? Are alternatives

available?

8. **Reproduction** capacity/flexibility: can the product or service be easily applied or adapted for use in another region or another situation?
9. **Acceptance**: do the users intuitively understand what the product or service is about and are they interested in it as a solution?
10. **Level of knowledge transfer required**: are the training requirements and associated costs and efforts clear and acceptable?
11. **Ethics, transparency, public accountability, objectivity and impartiality**: does the product or service increase the level of objectivity and impartiality in decision-making? Does it improve transparency and public accountability? Does it raise ethical issues?

In all cases the indicators can be assessed qualitatively using a 1 (poor) to 5 (excellent) scale. When possible, they are also assessed quantitatively. The rating of these various indicators allows assessing strengths and weaknesses of the Earth Observation application, which gives an indication of the suitability of the application to the issue it wants to address.

The last part of the assessment framework consists in assessing the broader business environment by differentiating four categories:

- The willingness (and capacity) of the Earth Observation application's beneficiaries to pay for it
- The possibility to embed the product or service in general organizational or business processes
- The Openness of the government's data policy and free access to public data
- The institutional environment ability to accept and support new solutions

An assessment of two Earth Observation related applications, both in capacity building, has been performed. The first one concerns the workshop "Bringing GEOSS Services into Practice" and the second one concerns all the recent Earth Observation promotion activities performed in Armenia in the frame of several projects in collaboration with the University of Geneva.

5.3.1 Impact assessment of the "Bringing GEOSS Services into Practice" workshop

5.3.1.1 General background

5.3.1.1.1 Presentation of the case study

This case study focuses on the Black Sea and South Caucasus regions, in which the former FP7 enviroGRIDS⁸⁸ project has widely promoted the GEO/GEOSS data sharing principles. During this project, and in line with the GEO⁸⁹ data sharing principles, the workshop “Bringing GEOSS Services into Practice” (BGSIP) was developed and taught to various audiences in the regions. The enviroGRIDS project ended in year 2013 but in parallel two other EU FP7 projects, EOPOWER⁹⁰ and IASON⁹¹, took up the BGSIP workshop by upgrading and translating it.

The "Bringing GEOSS services into practice" workshop aims at teaching how to configure, use and deploy a set of open source software to build up a spatial data infrastructure (SDI). Trainees learn how to publish and share data and metadata using OGC and ISO standards and how to register services into GEOSS. The workshop has been developed in the train-the-trainers format and consists in a tutorial, a virtual machine and a PowerPoint presentation in seven languages. All this material can be downloaded for free from a dedicated website⁹².

A version for beginners (based on the use of GeoNode) is in preparation while the program of the current version (workshop for advanced users) is the following:

- Concepts on spatial data infrastructures
- How to store geospatial data? (PostGIS and flat rasters)
- How to publish geospatial data? (GeoServer, WMS, WFS, WCS, KML, SLD)
- How to document and search geospatial data? (GeoNetwork, CSW, ISO metadata)
- How to process geospatial data? (Python, WPS, PyWPS)
- How to view geospatial data? (WMS, OpenLayers, QGIS, KML)
- How to download geospatial data? (WFS, WCS, QGIS)
- How to analyze geospatial data? (WPS local/remote)
- How to share geospatial data? (GEOSS, Discovery and Access Broker)

Since 2010 the workshop has been given to about 480 trainees and it has reached a large audience that may potentially be very active in Earth Observation. Therefore the workshop is the product / service that will be used to assess the impact of Earth Observation in the BS and SC regions in the frame of EOPOWER D11.01. The objective of the present case study is to measure the impact of the workshop itself. Unlike other products developed in EOPOWER

⁸⁸ <http://www.envirogrids.net/>

⁸⁹ <https://www.earthobservations.org>

⁹⁰ <http://www.eopower.eu/>

⁹¹ <http://www.iason-fp7.eu>

⁹² <http://www.geossintopractice.org>

based on enviroGRIDS outcomes⁹³, the workshop exists for long enough to measure its impact. Moreover, it is widely used in other projects (especially IASON and ClimVar⁹⁴).

5.3.1.1.2 Methodology used to assess the impact

The methodology used in this case study is based on:

- 1) A list of predefined indicators (See below section IMPACT INDICATORS)
- 2) An online questionnaire (cf. Annex 5.3.1.4) designed to get some feedback from the past attendees of the workshop, especially in the Black Sea and South Caucasus regions. The expected feedback relates to how they have put into practice the knowledge acquired with the workshop in their own institutions. 409 trainees have been surveyed in Bucharest (May 2010), Tbilissi (Nov. 2010), Delft (April 2011), Istanbul (Sept. 2011), Novi Sad (Sept. 2013), Geneva (course on SDI, University of Geneva, Feb. 2014), Istanbul (Oct. 2014) and Batumi (Oct. 2014). Through this questionnaire, we want to extract quantifiable information such as: how many past attendees have published OGC web services in their institution, how many have built a SDI, and how many have trained others in order to improve the sharing of environmental data in their country / region.
- 3) A graph showing the number of downloads of the workshop material from the Open Archive of University of Geneva⁹⁵ in function of time (period covered: March 2014 – December 2014).
- 4) A particular focus on two success stories in Armenia and Georgia. In terms of geospatial data sharing these two countries started from scratch a few years ago and a few organizations dealing with environmental studies have followed the workshop since that time. We think that this has contributed, among other elements, to help the two countries to overtake EO tools, to gain in GIS and SDI capacity, and to pave the way towards their GEO membership.

Note that the methodology used in this case study has some limitations. Firstly, the conclusions that we draw do not only apply to the Black Sea (BS)/South Caucasus (SC) regions because we surveyed former attendees of the workshop from the BS, SC, Mediterranean and African regions. Secondly only 46 out of the 409 people surveyed completed the questionnaire. Thirdly, the workshop has been regularly improved since the beginning of EOPOWER⁹⁶ therefore we have

⁹³ e.g. the EOPOWER online platform of environmental base data and water-related model output available at:

<http://eopower.grid.unep.ch:8080/dataextractor/examples/swatappli.html>

⁹⁴ <http://climvar.grid.unep.ch>

⁹⁵ <https://archive-ouverte.unige.ch/>

⁹⁶ 1st June 2013

measured the impact of the product at a given time while it is continuously evolving.

Similarly it is difficult to quantify the economic value per se but it definitely has a positive impact as will be shown in the IMPACT INDICATORS section.

5.3.1.2 Assessment

5.3.1.2.1 Paradigm shift

The workshop is planned to remain free and hence no future payment scheme or economic mechanism other than project funding or university teaching is foreseen at the present time. In the frame of IASON and EOPOWER a new version of the workshop is being produced, with latest software version. We think that with this new version the workshop material will remain up-to-date for one year after the end of EOPOWER (May 2015). Afterwards the PowerPoint presentations and the dedicated website will be maintained by UNIGE in the frame of the Certificate of Geomatics but the maintenance of the whole material will require integrating the workshop in projects with official and dedicated funding. In that sense, the fact that UNIGE will continue disseminating the workshop (e.g. Thessaloniki, May 2015; FOSS4G, Como, July 2015; Prague, Summer 2016⁹⁷) is a smart way of provoking potential funding opportunities. More generally the fact that this workshop is now a flagship product at UNIGE in terms of capacity building, as well as the continuous investment of UNIGE in projects with open source based capacity building will probably support partially the maintenance of the workshop.

The workshop does not create a paradigm shift per se but it contributes to it given this “open” nature. It allows indeed building capacity in Earth Observation from data collection and storage to data publication and sharing in a transparent and non-commercial way through open source software and open data. This complete open data workflow is something quite innovative and important that can contribute to change people’s view and approach to environmental issues, improve the organization process and reduce complexity.

5.3.1.2.2 IMPACT INDICATORS

Each indicator is assessed from the workshop “Bringing GEOSS Services into Practice” perspective. In some cases, we could corroborate the assessment with the results of the online survey.

No.	Indicator	Quantitative assessment	Qualitative assessment (to be indicated on a scale of 1 (= poor) to 5 (= excellent))
1	Fit-for-purpose	According to a survey sent to 409 former attendees of the workshop,	5: The BGSIP product gives end users the necessary knowledge and tools to

⁹⁷ in discussion with a IASON partner

		65% of the respondents said they taught or are planning to teach in their organization what they learnt during the workshop.	take advantage of Earth Observation benefits, both conceptually and technically.
2	Comparative advantage	Not applicable	4: Comparative advantage of BGSIP based on open source solution vs commercial solutions: Pro: - all data & software are free and/or open source - products can sometimes be tailored to users needs (for ex. in a project ⁹⁸). Con: Customer's service sometimes less professional (e.g. mailing lists, forums for developers) than with commercial solutions (e.g. dedicated hotline).
3	Complexity (to user) / ease-of-use	Not applicable	3: even though this workshop tries to simplify as much as possible the complexity of geospatial data processing, some parts remain technically complicated for newcomers. For this reason it is planned to develop an even lighter part of the workshop, in order to reduce the complexity as much as possible. The results of the online survey show that technical difficulties do not prevent attendees to transmit the knowledge acquired in their organization but that there are rather institutional, financial barriers, or time constraint.
4	Elegance	Not applicable	5: The solution provided is very smart: a global community of users, often very dynamic, reactive and good willing, supports the solutions presented in the workshop. This means that end users of the workshop not only have access to simple cutting-edge solutions for managing geospatial data but also to almost real-time answers to their questions from the internet community. Moreover, all the material available can be downloaded for free, including the software to be used. This gives a full freedom to users who can in turn become members of the community, supporting others. This is a smart community spirit.
5	Cost-benefit	The workshop needs to be	3: the workshop maintenance is

⁹⁸ the "Afromaison" broker was customized for specific users needs (<http://afromaison.grid.unep.ch:8080/gi-cat/gi-portal/>)

		maintained, and regularly presented. This requires funding, e.g. through projects. A rough estimate is the salary of 1 person/month over a year only for the maintenance (plus extra costs such as travels). If the workshop can be framed in a project or in 'official' courses in an academic institution, the cost-benefit will be increased; otherwise, it might be insufficient.	totally dependent on funding available, as mentioned in the quantitative column. But the fact of providing it for free through internet in a train-the-trainer spirit allows local trainers to present the workshop themselves. Hence no other cost than the ones of maintenance is required.
6	Sustainability	Not applicable	5: the workshop is being continuously maintained and upgraded, mostly through projects or university commitments (teaching at university). This situation is likely to be long-term; it can hence be said that it is sustainable.
7	Resilience	Not applicable	5: the workshop material can be downloaded from the internet and the creators can be reached by email. In consequence, it is very unlikely that there is a breakdown in the value chain.
8	Reproduction capacity / flexibility	As mentioned for indicator 1, 65% of the respondents of the survey of the workshop's attendees said they taught in their organization what they learnt at the workshop. Even if this does not mean they have reproduced the exact workshop, it shows that the workshop's principles can be reproduced.	5: From a technical perspective the workshop is reproducible and flexible. It is based on universal internet technologies. Apart from South Caucasus and Black Sea it has been presented to people from Northern Africa, Europe and Mediterranean Sea countries and can be potentially presented in any region of the world. Language is not a barrier as the workshop material is available in Arabic, Croatian, English, French, Russian, Serbian, Spanish and soon in Czech. The only barriers against reproducibility might be financial, institutional or political.
9	Acceptance	Not applicable	4: Generally the workshop is well accepted but the topics might be unknown and difficult to understand depending on the audience. This is the reason why a big effort on dissemination and simplification is needed. In any case, after a possible initial reluctance of non-technical audience, the workshop is well accepted and used: the workshop material has been downloaded 800 times over the 8 past months, with a constant frequency (see Figure 12).
10	Level of knowledge transfer required	According to the survey, only 32% of the respondents created an SDI in their organization. This is low but promising because in the meantime	3-4: The workshop targets people with some knowledge in IT and if possible GIS/SDI concepts and practice. In reality, the audience's

		76% of them say they are in a position to teach this workshop in their organization and 65% say they have already done it. Similarly 65% say they have / are planning to share data through the GEOSS	level is very heterogeneous and the level of knowledge might sometimes be an issue. For this reason a lighter version of the workshop is under development, in order to satisfy all levels of knowledge. Moreover the workshop can be delivered in various formats depending on the audience: (1) general presentations (1-3 hours) for conferences targeting decision-makers, (2) hands-on workshops (1-2 days) for scientists and students, or (3) a combination of the two former.
11	Ethics, transparency, public accountability, objectivity and impartiality	Not applicable	5: All the material and content of the workshop is transparently put at disposal in a collaborative and open spirit in view of improving the global geospatial data workflow and sharing. The workshop material is licensed under a Commons Creative license.

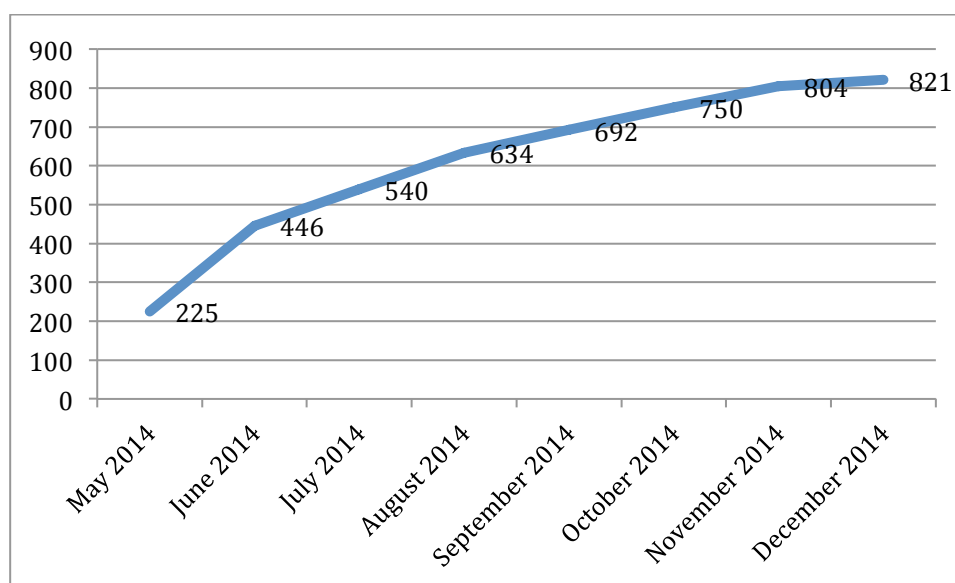


Figure 12: Cumulative number of downloads of the workshop material in function of time (March to December 2014) from the open archive of Geneva University

5.3.1.2.3 BUSINESS ENVIRONMENT

The goal of the workshop is to remain free and be disseminated freely as much as possible for the benefit of the environment. However, to be able to maintain it still has a cost (salary) that is generally supported by projects or university teaching commitments. The willingness to pay is hence not of interest for this product.

The workshop can easily be embedded into organizational processes, either on an institutional basis (university or Group on Earth Observations) or on a projects basis.

In terms of openness, the product is transparent as all material is freely published and available using open source tools and data. As previously mentioned, the goal is not to do business with it or to access markets, but to participate to a global effort of data sharing. This being said, data sharing can also be profitable to countries and help them spare time and money in the long run. This requires of course a strong institutional involvement, in order to convert individual organizational advantages into national benefits.

In that vein the example of two countries of the South Caucasus region, Armenia and Georgia is significant. These two countries have become new members of the Group on Earth Observations (GEO) in 2014 whereas they had started GIS and SDI activity very recently. These two success stories, published in the frame of EOPOWER WP13⁹⁹, are the result of the global EO capacity building put in place in the two countries through projects and collaboration, as well as a strong personal commitment and involvement of the country's representatives in these projects. The fact that the workshop was delivered several times in Armenia and Georgia plus the strong links that were maintained further between the workshop's organizers and the national focal points, are two elements that helped the two countries to raise awareness on data sharing principles, to improve their institutional workflows, to better manage their environmental data and further to be accepted as GEO members. The Earth Observation promotion activities that took place in Armenia and led to their successful uptake by Armenian institutions will be the focus of another impact assessment specific to Armenia.

5.3.1.3 Conclusion and perspectives

It is difficult to scientifically measure the impact of the workshop for several reasons explained above. In particular the frontier between the effects of the workshop and other promotion activities is not always clear. Similarly, most of the presentations of the workshop have been done in the frame of various projects (enviroGRIDS, IASON, EOPOWER, ClimVar) coupled with other activities in relation with GEO data sharing principles.

However it is certain that the workshop has played a significant role in the acceptance of two countries as GEO members: Armenia and Georgia. Also, a high percentage of past attendees of the workshop have trained (or are planning to train) others; one third of the respondents have built up a SDI based on the knowledge acquired; and 65% have / are planning to share data through the GEOSS. It is therefore possible to assert that the workshop had impacts at various levels: technical, institutional and from the 'capacity building' perspective. The concept of 'selected persons' (national focal points or trainees in position to train others in turn) has proved to work and should be renewed.

⁹⁹ http://www.eopower.eu/success_stories/Armenia.pdf and http://www.eopower.eu/success_stories/Georgia.pdf

Moreover the indicators show that there is a strong potential for increasing the impact of the workshop in the future (hence the measurement of it). This will go through maintaining it and upgrading its material regularly with new software versions that would come out. To this end, advantage will be taken of this workshop being also taught at the University of Geneva. This will give the possibility to the team to dedicate some time to this maintenance. Additionally, if the team is part of other capacity building projects, time in these projects could also be dedicated to such upgrading tasks in order to always propose latest software versions for this workshop that must remain cutting-edge. Besides, it is foreseen to enrich the course with practical exercises based on thematic data, e.g. in fields such as climate change, water resource and mineral exploration.

More importance should also be given to the audience, as it is very often heterogeneous. The workshop might become counter-productive if not fitting the audience: too simple or too complicated. That is why it is planned to develop a lighter version of the workshop in order to lift even more the technological barriers that might cause some reluctance among certain end-users or decision-makers.

Moreover, decision-makers should be systematically targeted and invited in the audience besides technical people in order to orientate the workshop to a “train-the-decision-makers” aspect besides the “train-the-trainers” one as this might increase the impact.

Finally, the workshop should ideally be combined with presentation and/or practical use of a product that is familiar to the attendees. This can happen if the workshop is presented in the frame of an on going project. The experience of the “ClimVar” project is significant: the workshop BGSIP was given to end-users and was immediately followed by a half-day presentation/use of the SDI platform supporting the project. The theoretical knowledge acquired during the workshop was thus immediately put into practice through practical exercises of data publication into the SDI platform. With this combination, the workshop became more comprehensive to the attendees and more in line with their specific needs and expectations.

These conclusions point out that there is still some room for improving the workshop and its impact. A secured long term funding of the workshop is an issue that deserves a particular attention. Its integration in projects with official and dedicated funding is a first answer; besides, the continuous dissemination activities of the workshop at various events should also present additional potential funding opportunities. But its “institutionalization” through UNIGE teaching is a supplemental guarantee that should avoid any breakdown in the value chain.

5.3.1.4 *Annex: online questionnaire to people who attended the BGSIP workshop since 2010*

* First name
* Last name
* Email
* Country Choose one of the following answers
Please choose...
* Company/Organisation name
* Type of organization Check any that apply
☐ Private ☐ Government ☐ Education/Academic ☐ Not for profit ☐ Personal use ☐ Other:

*** Position/Job role**
Check any that apply

- ☐ Technical lead
☐ GIS professional
☐ Decision maker
☐ Manager
☐ Other:

*** Have you created a Spatial Data Infrastructure in the frame of your institution, including webservices?**
In case of "yes", please indicate the urls, endpoints, ... of your webservices

Choose one of the following answers

- ☐ Yes
☐ No

Please enter your comment here:

*** Have you shared or are you planning to share data of your institution through GEO/GEOSS (http://www.geoportal.org/web/guest/geo_home_stp)?**
In case of "no", please indicate if this is due to:

- institutional barriers,
- financial barriers,
- technical barriers,
- time constraint,
- other

Choose one of the following answers

- ☐ Yes
☐ No

Please enter your comment here:

*** In your institution, are you in a position to train others with material acquired during this workshop?**

☐ Yes ☐ No

Have you or are you planning to teach what you learnt during the workshop in your organization in order to improve data sharing?
In case of "no", please indicate if this is due to:

- institutional barriers,
- financial barriers,
- technical barriers,
- time constraint,
- other

Choose one of the following answers

☐ Yes ☐ No

Please enter your comment here:

Do you have any other comments or wishes for future editions of the workshop?

5.3.2 Earth Observation promotion activities and their impact in Armenia

5.3.2.1 General background

Armenia is part of the former Soviet Union and was one of the most industrialized Soviet republics with large-scale industrial activities such as mining, chemical and electrical industry, machinery etc. After the collapse of USSR, this industrial production collapsed but showed some recovery in the mid-1990s, mainly due to the activation of mining companies.

The economic policy shifted towards a strong support to industrial development greatly ignoring ecological interests. As a result, mining-related industries were permitted to operate without environmental regulations. Consequently, the unfavorable ecological situation inherited from the Soviet industrial era substantially worsened, becoming a national and regional issue to trans-boundary aspects ecological aspects (water for example).

In order to address this ecological situation, international cooperation was vital for the country's future and for the greater area. Hence, several projects using

Earth Observation¹⁰⁰ have taken place in the country since 2009. Given this completely new Earth Observation approach in the country, a lot of capacity building and promotion activities needed to take place for an efficient knowledge transfer. Most of these Earth Observation activities were proposed by the University of Geneva (UNIGE) to two Armenian institutions: the Center for Ecological-Noosphere Studies (CENS) and the Institute for Informatics and Automation Problems (IIAP). CENS is the main Armenian institution responsible for fundamental and applied studies in the area of ecology and environmental protection, while IIAP is the leading ICT research and technology development institute. These two institutions can now be considered as the two leading Armenia Institutions regarding the management, processing, and sharing of geospatial data in the environmental domain.

Through the projects mentioned, various Earth Observation (EO) and Spatial Data Infrastructure (SDI) principles and software were introduced to key persons of these institutions. It allowed then to perform the same activities to a broader audience (students, national academy of sciences, other national key persons from other institutions) with their support during various events. These promotion activities could be further illustrated and enriched by the growing work done in the national context by CENS and IIAP, such as an Interoperable Cloud-based scientific gateway for NDVI time series analysis¹⁰¹. All these EO activities contributed to the writing of scientific articles that became in turn a promotion activity.

These EO activities continued in parallel to the EOPOWER project and synergies were found whenever possible to further promote EO activities in Armenia. Armenia was used as a pilot in EOPOWER to test the EGIDA methodology. This activity allowed to further raise awareness on the advantages of EO activities and the Group on Earth Observation (GEO) principles benefits. The consecration of all these efforts was the GEO membership of Armenia (that became official on 13th of November 2014) during the EcoARM2ERA and EOPOWER projects, getting hence a national and international visibility.

Finally, the success stories in EO (including one in Armenia) developed and disseminated in the EOPOWER project is another promotion activity that might have an impact in the Black Sea region. This type of promotion activities highlights local/regional EO successes that can potentially be replicated by other regional partners. This not only shows regional stakeholders that there are potential EO solutions to develop, but also fosters regional collaboration to this end.

¹⁰⁰ enviroGRIDS (<http://www.envirogrids.net/>), ARPEGEO (<http://www.arpegeo.sci.am/>), EcoARM2ERa (<http://www.ecoarm2era.eu/>)

¹⁰¹ Astsatryan H., Hayrapetyan A., Narsisian W., Asmaryan S., Saghatelyan A., Muradyan V., Guigoz Y., Giuliani G., Ray N. An Interoperable Cloud-based scientific gateway for NDVI time series analysis. Computer Standards & Interfaces, submitted.

5.3.2.2 Assessment

The nature of the elements assessed (promotion activities) is different from a usual Earth Observation application, as it is made of set of heterogeneous actions instead of a single product. As it is obvious that these heterogeneous promotion activities provide an added value for the country, the impact assessment focuses on the indicators part instead of the general reference framework or the business environment. For the indicators part, we differentiate input indicators (elements that have been used for the promotion activities) from output indicators (elements resulting from these promotion activities) as elements for assessing the impact indicators.

5.3.2.2.1 Input indicators

The following elements can be considered as input indicators that are relevant for the promotion of EO activities in Armenia: success stories, EGIDA pilot project, promotion at events, flyers, personal meetings.

Success stories in general and the success story on Armenia in particular¹⁰² were presented at several events in which Armenian environmental/EO stakeholders were present. This allowed to better promote the expected EO benefits for the country.

The participation of Armenia as a pilot for the EGIDA methodology in the frame of the EOPOWER project gave CENS a boost to establish a stakeholder network, to organize several events, to develop EO/SDI promotion material and to raise awareness about GEO/GEOSS in the country.

Several events were organized during the various projects. This allowed raising awareness for various audiences, including the state academy of science, students, environmental/geographic national stakeholders, ... as well as to establish network with neighboring countries (e.g.: Georgia, Russia) or international networks (e.g. GEO or OGC)

Dissemination material such as flyers or posters was produced either for the promotion events, project's meetings, EGIDA pilot, ... These were done both in English and Armenian in order to reach as many stakeholders as possible as language might be a barrier.

And finally many personal meetings were organized between different key persons. This was first of all the case between UNIGE, CENS and IIAP, but later between CENS and the National Academy of Sciences, CENS and GEO principals, CENS and national environmental actors.

5.3.2.2.2 Output indicators

¹⁰² Available from the GEOCAB platform:
<http://www.geocab.org/#/results/success%20story>

Interest in EO products has definitely increased in Armenia as a result of various promotion activities. Some of these promotion activities were performed in the frame of EOPOWER but the general effort in this sense was done back in 2010 with the participation of Armenia to the enviroGRIDS project; only the network of environmental stakeholders put in place through the EGIDA methodology can be considered as a proper EOPOWER output. We should then consider the various projects already mentioned for the promotion activities. In terms of economic benefits, it is still too early to measure them as this whole process has just been put in place.

The following output indicators can be considered: the GEO membership of Armenia, the SDI teaching just put in place at CENS, the portal of parallel computing for environmental indices¹⁰³ (ex: NDVI), the number of EO-related scientific articles, the network of environmental stakeholders establishment, the number of country's stakeholders part of this network

5.3.2.2.3 Impact indicators

Each indicator is assessed from the perspective of promotion activities in Armenia.

No.	Indicator	Quantitative assessment	Qualitative assessment (to be indicated on a scale of 1 (= poor) to 5 (= excellent))
1	Fit-for-purpose	Some output indicators such as some of the ones described in section 5.3.2.2.2 can be used to assess the impact and fit-for-purpose status of the promotion activities chosen.	5: Promotion activities in Armenia definitely contributed to raise Earth Observation and its benefits awareness in the country. All the output indicators show that the impact is consequent (e.g. the Armenian membership of GEO), which proves that the activities are fit-for-purpose.
2	Comparative advantage	Not applicable	5: The positive impact of EO activities in the country makes them indispensable for valuation of the country's resources. EO provides a lot of added value at a very

¹⁰³ arpegeo.sci.am/gisservices/

			reasonable cost, especially through the GEO membership.
3	Complexity (to user) / ease-of-use	Not applicable	3: The concepts linked to Earth Observation primarily target the scientific community directly working with Earth Observation products. Besides, these concepts also need some political and legislative support (e.g. for laws facilitating data sharing). This results in a brand new conceptual workflow among various institutions, that needs to be well understood and applied.
4	Elegance	No quantitative data are available but the number of people in Armenia aware of Earth Observation activities is increasing.	5: The Armenian stakeholders are convinced that EO promotion activities are important to share the successes encountered so far in the country either internally (e.g. new teaching at university level, web analytic portal focused on the country) or internationally (GEO membership). These successes represent elegant and positive means of promoting EO activities.
5	Cost-benefit	The cost of the various activities (events, teaching, travels linked to GEO membership, ...) could be calculated but is not available to us.	5: the promotion activities are cost-efficient as they do not require many financial resources compared to the enormous advantages that Earth Observation could bring to the country in terms of efficiency.
6	Sustainability	The number of teachers, speakers, ... necessary for a sustainable EO activities promotion can be	4: Through the network of Armenian environmental data providers built in the frame of the EGIDA methodology used during EOPOWER, as well as specific teaching at university level, there is a good hope that

		calculated.	promotion activities in the country will become sustainable. This is reinforced by the strong commitment of EO focal points at CENS and IIAP, and further reinforced by the GEO membership and related commitment.
7	Resilience	Not applicable	4: as for sustainability, the commitment of EO focal points at CENS and IIAP, as well as the network being created in the country and the GEO membership makes it very unlikely to have a breakdown in the promotion of EO activities chain.
8	Reproduction capacity/flexibility	The cost of EO activities promotion is very low compared to the added value for the country.	5: the EO promotion activities can easily be reproduced in other places given their universal nature (e.g. flyers, presentations, teaching,...) . This is also the case for example with the workshop “Bringing GEOSS Services into Practice” (BGSIP), that aims at teaching how to configure and deploy a set of open source software to set up a spatial data infrastructure(SDI). This workshop can easily be reproduced in other parts of the world and in other languages.
9	Acceptance	Not applicable	5: The promotion activities have generally been well received as an important knowledge transfer for the benefit of the country.
10	Level of knowledge transfer required	Not applicable	4: These promotion activities target primarily the scientific community as well as decision-makers. As a good governance needs to accommodate the scientific workflow by supporting laws

			and institutional commitments, the audience is broad (politicians, decision-makers, scientists) and must be quite high level. This is the case in Armenia, where the two high-level institutions (CENS and IIAP) have the expected level and will further redistribute the acquired knowledge in the country.
11	Ethics, transparency, public accountability, objectivity and impartiality	Not applicable	5: The promotion activities are done transparently in view of improving country's GIS and SDI capability to conform to international recognized standards. These promotion activities are included as much as possible in existing projects or activities in order to optimize use of resources. Besides, they are based on free and open source technologies.

5.3.2.3 Conclusion

Earth Observation activities have definitely much increased in Armenia compared to a few years ago. A substantive part of this increase is due to promotion activities and knowledge transfer that could take place in international projects in the country. These projects are indeed the drivers of Earth Observation promotion, discovery and adoption through funding. Another part is the strong commitment of local key actors to SDI establishment in the country. This led to a full success and positive impact with Armenia's GEO membership, the creation of an environmental SDI, several related scientific articles in international journals, and the instauration of GIS teaching.

It is however difficult to distinguish between the particular impact of EOPOWER promotion activities compared to the other projects. But the clear input of EOPOWER is the testing of the EGIDA methodology in the country, that allowed a clear progress in environmental network of stakeholders, dissemination means such as flyers, ...

One may wonder why these EO promotion activities worked better in Armenia than in other countries that were also targeted by EO promotion activities in the frame of projects. Here are a few tentative explanations:

- The national focal points in Armenia were very committed to EO in general as a tool to help their country. Posters, scientific articles, EGIDA methodology were means taken seriously to make the country progress in EO activities.
- Armenia was part of several consecutive funded projects (EU, Swiss, ...) that allowed to perform EO activities and capacity building

Armenia has now taken over the EO in the country and can play an important regional role, politics allowing, as an EO leader in environmental issues, addressing the transboundary issues inherent to environmental problems.

5.4 Chapter key outcomes

- Assessment is an essential component of an SDI implementation process that allows targeting specific actions, justifying the actions taken and giving accountability for mobilizing resources.
- Several types of assessments can be performed: on a whole SDI, on one or some components of the SDI, on the impact of the actions performed for the SDI or on the impact of the SDI on a society. Each assessment's type requires a customized methodology.
- The resources to mobilize strongly depend on the assessment's type as well as the scale of the SDI or its impact to assess.
- A geoportal is a key element of an SDI that allows discovering, visualizing and accessing geospatial data.
- Given their importance, a specific framework for geoportals assessment is needed, that can deliver a simple and illustrative message about the geoportals capacity to provide discovery, visualization and access to geospatial data.
- Geoportals are especially important for providing community-specific information. Besides complying with best practices common to any geoportal (e.g. use of web services, use of open data) community-specific best practices (e.g. use of proprietary/non proprietary technology, language, provision of mono/multi-scale or -thematic information) would be an asset. This requires defining community characteristics.
- Geospatial data discovery, visualization and access through web services is quite satisfactory in most geoportals assessed that use web services as a major asset allowing standardized exchange and interoperability.
- An impact assessment framework exists to capture the impact of innovative Earth Observation applications and activities: the EOPOWER impact assessment framework. It captures a whole range of impact aspects: economic, social, political and technical.

- Such an assessment framework allows an objective justification of the suitability of the product/service/action assessed through the broad range of aspects examined. This justification is often required for giving legitimacy to the product and allowing its potential expansion.
- Given the importance of the impact that actions taken should have on their target, it is essential to plan from the beginning of the whole process the impact's measurement indicators because of the possible time lag of certain indicators. This is for example the case if one of the indicators is the number of visits on a project's capacity building material web page. Therefore, assessment tools (e.g. questionnaires, websites statistics,) should be planned and set up right from the planning of an SDI solution deployment (e.g. capacity building workshop) in order to rapidly be operational.
- The impact assessment of promotion activities showed that impact might sometimes be difficult to attribute to a single product or set of products, but is the result of a whole process, for example through several projects.
- Capacity building activities are key for empowering a society with SDI. However, capacity building activities results become only visible after some time. There is hence a shift between the CB activities and the moment when it becomes possible to measure their impact on the SDI. This issue, that might temporarily prevent justification of the actions performed, should be clearly stated at the beginning of the process.

6 CONCLUSIONS

The accelerating rhythm of environmental issues requires urgent answers that are usually taken by politicians and decision-makers in the public and private sectors. In order to support sound decisions, the scientific world is responsible for providing the necessary knowledge. This knowledge is the result of complex analyses that require data and information, and must remain simple and clear in order to be understood and used. To this end, visualization of information, for example on maps, is a powerful support. However, data discovery, visualization, access, availability and sharing is still often difficult because of several issues (e.g. lack of interoperability, legal restrictions) that SDIs should help to address.

Despite an increasing number and an improved acceptance of SDIs, efforts remain to be made for properly establishing SDIs in several parts of the world where data access remains problematic and negatively influences the possibilities to tackle environmental issues. For establishing or improving SDIs, it is necessary to know where the gaps are and what actions should be performed in priority, for example on the technological or on the institutional side. This requires a good understanding of the complex nature of the SDIs and an appropriate methodology to assess them.

Based on these considerations, we argued that Capacity Building plays a key role in all the SDI components and should be the central element of any SDI implementation or improvement from the beginning of the process. The central role of Capacity Building is intrinsically linked to the human factor it targets, that has an influence on all the components of the SDI and is able to remove the different barriers to data access. For this reason it is essential to find in priority the key individuals who will be the target of the capacity building, who will then become a necessary support for wider capacity building activities, for example as first pieces of a stakeholders network. Consequently, the aim of this research was to:

Examine the role of capacity building in successful SDI implementation by lifting tangible and intangible barriers.

Three associated research questions have then been formulated to support this research and were explored through four chapters.

6.1 Answers to research questions

1. What are the existing methodologies to evaluate SDI implementation?

Chapter 2 reviewed and discussed various existing frameworks aiming at measuring the status of SDIs implementation, which is defined as stocktaking or inventory of the situation. Such an assessment is a snapshot of the various components of an SDI and is the first necessary step for getting a global overview of the situation, in order to organize the next actions. It is as necessary in a

situation where nothing exists than in a situation where an SDI is already well advanced and requires adjustments. If we take an extreme situation of a country where nothing exists in terms of SDI, it still remains necessary to identify the potential basic maps existing, the available GIS or IT skills among the population, the existing institutions and their relationships, that will constitute an embryo on top of which the next actions will be built. At the opposite, in a society with an advanced SDI level, it still remains necessary to assess the SDI status in a perspective of maintaining this advanced situation to ensure long-term sustainability. In both cases, regular assessments are necessary to measure the progress, which is for example the aim of the European INSPIRE State of Play.

The chapter revealed the difficulty of this exercise given the complex nature of the SDIs that are dynamic and multifaceted. Their dynamic nature makes it difficult to take snapshots and requires to constantly re-evaluating the situation of their various components. Their multifaceted aspect supposes to assess them from several different angles (e.g. technological aspect, organizational aspect, legal aspect). Both these constraints require huge resources and knowledge to be performed thoroughly, which makes the human (knowledge) and financial aspects of stocktaking essential and potentially unaffordable. The geographic scale of the SDI is another constraint that has an influence on the affordability of the SDI status assessment, as a small-scale SDI (e.g. the one of a city or a commune) will require much less resources than a bigger scale one (e.g. a country or a continent). These different constraints (human or financial resources, geographic scale, time available) will have a fundamental influence on the possible level of details of the assessment (Figure 13), making it necessary to find a compromised balance between all these elements while remaining credible with an acceptable accuracy of the assessment, keeping in mind that none of the proposed methods can fully meet the requirement.

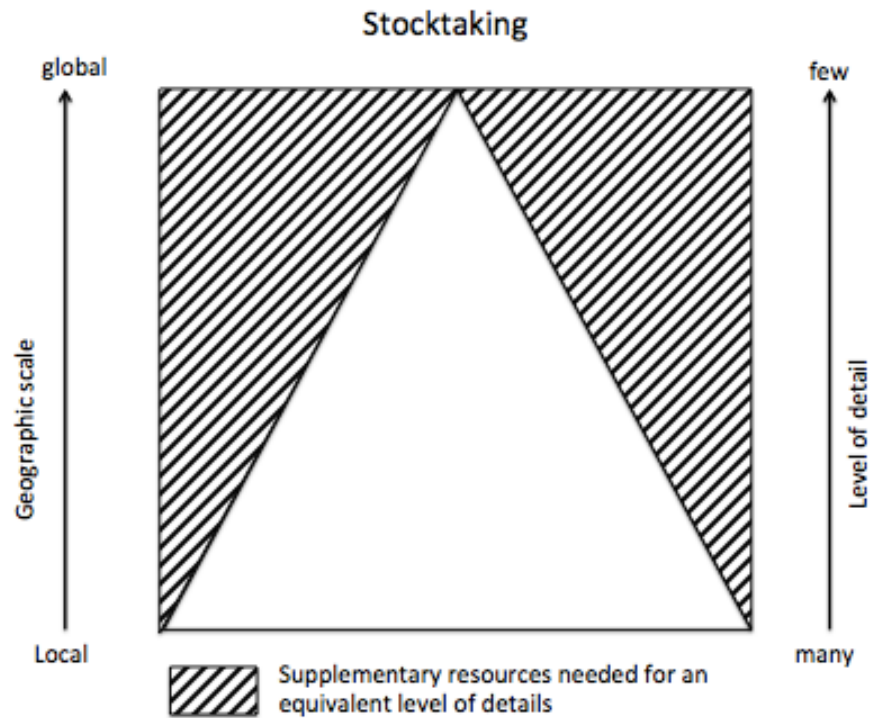


Figure 13: Resources needed for stocktaking in relation to the level of details obtained

Besides the difference of scale that impedes to have a uniform assessment methodology, there are also differences for assessing entities of a same scale (e.g. national scale) because of organizational, technological and financial differences (Delgado Fernandez et al., 2005).

Various authors have developed several SDI assessment methodologies with specificities making them focus on different parts of the SDI (e.g. technical, institutional). Recognizing the multifaceted reality of SDIs reflected in these different approaches, the Multi-View Assessment Framework has been developed by Grus et al. (2007) and discussed in chapter 2. It aims at guiding users towards the most suitable existing approaches and is open to new approaches. This framework is innovative in the sense that it groups most of the existing methodologies into an overarching classification primarily based on the purpose of the assessment. Being able to answer this first question – the reason why an assessment is needed – supposes to have defined a clear goal resulting from a vision and a set of sequential steps for establishing or improving a SDI. For each methodology, several indicators have been identified to assess the focused components, and some of these indicators can be used in different approaches.

Many of the SDI approaches examined use indicators well suited to the national scale. But the need to assess SDIs at supra-national level (e.g. continental level) requires another approach given the constraints described before that would require too much resources depending on the level of detail. With such a challenge, even if the assessment indicators defined in the other methodologies can be used, the assessment methods need to be adapted to a large-scale use case with limited resources. This is where innovative solutions are needed that allow

to bypass the supplemental resources needed for an acceptable level of detail given the scale.

The solution proposed in the article “Spatial Data Infrastructures in Africa: a gap analysis” makes use of information available through the Internet and also strongly depends on a local key SDI player (UNECA), able to gather information that would otherwise be difficult to obtain in the African context. The combination of these sources of information allowed assessing proxy variables to measure the defined indicators. This might sometimes be subjective but still gives a first overview of the situation, that can later be refined with other methodologies, for example at the country level. From the assessor’s side, it also presents the advantage to be low demanding on necessary resources.

The weaknesses detected in Africa during the assessment as well as the implication of a local key SDI player can also be seen as an opportunity to seize for boosting SDI implementation in Africa. Some mechanisms could easily be put in place, such as more systematic and standardized use of the Internet for exchanging information useful for SDI reporting. This simple measure would necessitate to formally establishing a network of focal points (e.g. one per country and one for Africa) that would agree on some basic indicators to report on at regular intervals. An embryo of such a network already exists, that consists in the countries delegates who gather every two years for the CODIST meeting at UNECA headquarters. Besides, some of these countries already report on some of the indicators (in a Word document), through a survey sent from UNECA that has been used for gathering information of the article.

An impulsion would be needed to formalize an Internet mechanism and related architecture, by setting up national/regional focal points and reporting websites. This impulsion could come by demonstrating to countries the benefits they would get in playing the game. The benefits could be the following: (1) better vision of their country’s situation through recognized indicators; (2) better institutional cooperation in the country (e.g. between the Statistics Office and the National Mapping Agency); (3) participation in a regional effort leading to a better regional integration; (4) potential international collaboration. Demonstrating these potential benefits requires awareness raising, which is part of the Capacity Building process. This shows that Capacity Building can already be necessary for stocktaking, which is the first stage of the SDI implementation process. If successfully implemented in a dashboard, such an automatic mechanism can fulfill one of the fundamental requirements of SDI assessment that consists in constantly re-evaluating the situation of the various components to address the dynamic nature of SDIs.

Based on all these elements, we can conclude by saying that even if no single assessment methodology exists for SDI because of their complexity, it is still possible to perform assessments at all scales, with reasonable resources, if a compromise on the level of details is accepted. This chapter also illustrated some specific elements of importance when performing SDI stocktaking at supranational level:

- Leadership is essential in the region to centralize the SDI efforts. This includes not only leadership at the regional level (UNECA in the case of Africa) but also at other scales (e.g. in countries).
- Leadership often requires a clear mandate that has to come from the political side.
- Based on this leadership, a network can be constituted, that is essential to distribute the work and resources needed.
- Commitment of key players (leaders) is necessary to trigger the necessary actions (e.g. better use of Internet in information exchange) for improving SDI assessment indicators.
- Commitment might require to understand the potential benefits of an improved situation, which can be provided through Capacity Building activities

2. What are the innovative solutions to lift the barriers of SDI implementation?

Acknowledging SDI weaknesses in Africa found through the assessment performed in chapter 2, the chapter 4 discussed innovative solutions to lift some barriers to SDI implementation. The technological angle has been chosen in this chapter as one kind of barriers to lift, as technological progress in the geospatial domain has great potential to support better SDI implementation. This concerns primarily the infrastructure SDI component, which is faster to address than the institutional or individual components as it usually involves less actors and can be set up in a few weeks or months with no structural changes needed. Innovative technological solutions have been demonstrated through three different approaches: (1) assistance to data access through customization of the download; (2) homogeneous discovery of heterogeneous African data; (3) semi-automatization of metadata generation.

The three approaches have in common to reduce the technological complexity for the data user and/or producer, by shifting it to another level managed by IT specialists, such as programmers. In the first case (data customization), the simplification lies in the possibility for the user to automatically obtain extent-specific data through a web browser, leaving the scripts performing the processing work. In the second approach (data homogenization), the user is facilitated in data of interest (data about Africa) discovery by the discoverability of heterogeneous African data in a single web search interface. Additionally, on the data producer side, it is simply required to register a standardized web service in a third-party software (GI-cat) that manages the complexity by homogenizing data discovery. In the third approach, the user does not need to do anything and benefits from systematic improved information through semi-

automatic metadata production. In this approach, the data producer needs to do small extra effort during data publication by adding some systematic metadata information, leaving the fastidious task of metadata publication in a separate metadata catalog to the automatic workflow between Geonetwork and Geoserver.

The possibility to externalize the technological complexity successfully demonstrated through these three examples relies on some essential elements: (1) the interoperability between the software (e.g. Geoserver, Geonetwork) and the languages/libraries (e.g. Javascript, OpenLayers, Python) that are able to communicate; (2) the existence of commonly recognized standards (e.g. WMS, WFS, WCS, WPS, CSW) that are key enablers of the interoperability; (3) the free availability at no cost of the different elements (software, languages, libraries, standards) that allows development, testing and deployment of multi instances of these solutions without restrictions. In all cases (SCOPED-W, GI-Cat, Geonetwork), the user benefits from these improvements through a simple interface of a web browser presenting the added-value resources, which stimulates data use and production (Figure 14). With additional knowledge, users can also take advantage of interoperability and standards by directly connecting to web services through dedicated GIS software (e.g. QGIS).

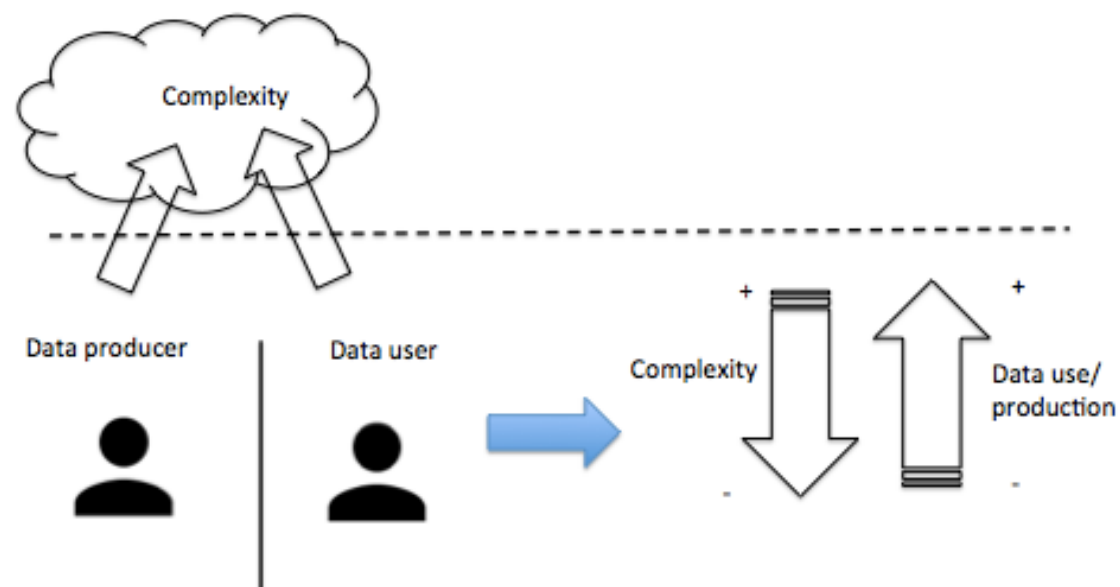


Figure 14: complexity reduction impact on data use and production

In addition to the advantages of interoperability and standards, these examples demonstrate the benefits of the possibilities offered by the recent “open” developments, such as open source software, open standards and open data. In terms of software, the openness of the source code allows a whole dedicated community to participate to its development for improving it and fixing the bugs, which boosts a global coordinated effort to improvement of geospatial solutions. In addition to being open source, all the software used in these developments are also free in the sense that a license stipulates that they can be freely used, copied,

modified, improved and re-distributed¹⁰⁴. Even if this definition of “free” does not explicitly requires that the software have no cost, all the ones used in the developments presented (Geoserver, Geonetwork, SWAT, GI-cat, PyWPS) are available at no cost. This means that the benefits demonstrated by these technological solutions can also be implemented in developing countries, often relying on small or no budget for geospatial activities.

The term “open standard” is defined by ITU¹⁰⁵ as a standard made available to the general public and developed and maintained in a collaborative and consensual process aiming at facilitating interoperability and data exchange for widespread adoption. Such open standards contribute to prevent lock-in from proprietary solutions that could become barriers to interoperability. While the Open Geospatial Consortium (OGC) standards used in the technological developments presented in this chapter are open standards, the ISO standards used for metadata (ISO19115 and ISO19139) cannot be considered as open standards^{106, 107} as they are not available for the general public because of their price. However, this restriction is not a barrier to their use in our developments since their implementation (e.g. in Geonetwork) is freely accessible to the users.

All data used in the developments of this chapter can be considered as open data, in conformity with the definition in introduction of this thesis as it can be used, reused and redistributed without restriction. In addition to being served openly, data used are often valorized and given a second life (e.g. in the case of the SWAT data through the SCOPED-W application, or in the case of project specific data served through the Africa broker), an added-value (extent-specific data obtained through automatic processing has an added-value compared to a whole dataset), additional visibility and discoverability (semi-automatic metadata generation).

Altogether, these several open concepts constitute a meaningful and enabling environment that can be of added value by removing institutional and geographical barriers associated with information flows (Karpouzoglou et al., 2016), which would be more difficult to put in place with proprietary technology or standards and restricted data. It is an illustration of the possibilities offered by the openness concept, presenting advantages particularly well suited in the environmental community where research is often led by public agencies (e.g. universities) or public-private partnerships (e.g. project consortia). For this type of actors, the goal is not to make immediate financial gain out of the activities performed, but to address environmental and societal challenges that will have a broad and long-term positive societal impact. This is illustrated by the famous example of the Landsat data policy change¹⁰⁸ that occurred in 2009. The archives have been freely released to the public since that date, boosting the number of satellite images downloads and use for the benefit of the society, generating an benefit of \$1.7 billion to the US (Ryan, 2016). A recent study by the McKinsey

¹⁰⁴ <https://www.gnu.org/philosophy/free-sw.html>

¹⁰⁵ <http://www.itu.int/en/ITU-T/ipr/Pages/open.aspx>

¹⁰⁶ <http://www.computerweekly.com/blogs/public-sector/2012/03/whats-an-open-standard-says-is.html>

¹⁰⁷ <http://www.robweir.com/blog/2010/09/recipe-for-open-standards.html>

¹⁰⁸ <http://earthobservatory.nasa.gov/IOTD/view.php?id=85703>

Global Institute (Manyika, 2013) shows the numerous benefits of open data, that can potentially generate up to US\$3 trillion of annual value in the following seven domains: education, transportation, consumer products, electricity, oil and gas, health care and consumer finance.

This chapter mainly addresses the infrastructure SDI component through innovative technological approaches aiming at reducing the complexity for the data users and producers by shifting it to another level managed by IT specialists. This reduced technological barrier has a direct influence on the individual SDI component as more people become able to discover and access customized geospatial data since a simple web browser is needed. The required skills and necessary time for finding and/or sharing geospatial data is reduced, to the benefit of other tasks or re-enforcement of other skills.

In all the approaches taken, very small efforts are requested from the data producers and only consist in exposing their data through standardized formats (e.g. through a geospatial data server) over the Internet so that they can eventually become accessible by data users. This is particularly true in the brokering approach that promotes a decentralized architecture where disparate nodes of a network corresponding to different spatial or thematic communities are bridged to participate in a global effort. A parallel can be drawn between this architecture and the network of focal points described as a solution for SDI assessment in the previous question of this conclusion. Connecting the various nodes of a network allows re-enforcing the whole structure by combining the individual possibilities for the benefit of the system. A central node might also be necessary to coordinate the system, which is concretized by a leader institution in the case of assessment and by the broker implementation (e.g. GI-cat) in the brokering approach.

We can conclude by saying that innovative solutions definitely exist to lift the barriers of SDI implementation. This chapter examined some innovative technological possibilities that highlighted a few key points for supporting better SDI implementation:

- Simplicity and reduction of complexity are necessary for an increased data production, sharing and use that are beneficial to the global community.
- Customization of data and information improves the user experience and participates to increased data use.
- The system of systems approach, implemented through brokering frameworks, is particularly adapted to multi-disciplinary environments as it keeps existing capacities as autonomous as possible while interconnecting them for the benefit of the whole system.
- Interoperability between data and components is essential in a system; interoperability is enabled by standards.

- The “open” principles are well suited for an increased data production, exchange and use in the environmental domain. Additionally, the financial affordability of software promoted through the open principles is particularly suitable to scientific research and developing countries, often relying on low budgets
- Reduction of technological complexity has an influence on the individual SDI component (less skills required), which can be exploited to build capacity of other technological aspects

3. How to measure the impact of capacity building activities on Spatial Data Infrastructures?

Williamson (2003) states that establishment of long-term commitment to education and research in university of a host country is the only way to address human resources development in support of SDI projects. This statement highlights the importance of individual skills in SDI development and the necessary capacity building actions to support them. Based on this ascertainment, chapter 3 examined a set of capacity building material and activities performed in several projects to discuss the benefits and limitations encountered and the resulting best practices determined. This chapter confirmed the essential role played by capacity building in supporting SDI implementation, which was extremely successful in certain cases such as Armenia and Georgia. Such successes not only empower specific countries but also give them a potential regional role (e.g. in environmental coordination efforts), which might pave the way to better political integration with neighboring countries.

This chapter also allowed to detail the various components of capacity building activities and highlighted the essential role of awareness raising that must be performed at the very beginning of the capacity building activities, with primary role to create understanding and commitment among the targeted audience. The importance of commitment has been described several times in the previous discussions as a fundamental element having a crucial influence on the further steps of SDI-related activities. Commitment of key stakeholders of an SDI implementation process guarantees a successful development but lack of commitment is likely to cause a failure of the whole process. The key stakeholders can be found at several levels: politicians responsible for geospatial data of their country or region; donors with specific targets in a given country/region; high level national scientists leading projects requiring geospatial data; IT specialists or GIS technicians part of specific projects. They must all be sensibilized to the benefits of SDI through capacity building activities.

Such activities shall of course be tailored to the audience as the topics discussed for raising awareness and creating commitment are different depending on the educational background. The formula of combined events was mentioned as a successful option to address heterogeneous backgrounds. In the case of politicians and donors, the approach will almost consist in giving an overview by describing the issues and benefits, showcasing success stories. For more

specialized people this awareness raising will also be important, but a second phase of specific capacity building activities (e.g. learning to collect, manage or share geospatial data) will be necessary and of equal importance. Innovative concepts presented in an SDI approach, such as data sharing and open principles might face resistances in societies with a restrictive approach on data sharing. But potential benefits, success stories, including successful financial examples such as the one of the Landsat archives, and a global trend towards the same direction should allow to overcome the resistances in the long term.

Apart from the “ideological” resistances just mentioned, there are also more “passive” resistances occurring when stakeholders do not see the necessity for a change and won’t provide any support in that direction. This corresponds to the “stand-alone” stage, which is the first out of four stages of NSDI development described by Kok (2005). This stage is characterized by lack of willingness and commitment to change, no vision, no leadership, internal communication only. Dessers (2015) mentions the complexity and ambiguity of inter-organizational structures as a factor of resistance to inter-organizational data sharing, that indeed corresponds to this “stand-alone” stage. It is followed by an “exchange and standardization” stage, where stakeholders start to recognize some bottlenecks, start prioritizing problems after acknowledging they can not all be solved at once, start to have a vision, envisage leadership and use communication for data exchange. It is followed by an “intermediary” stage where support for the change becomes high and implies an increasing awareness for the need of cooperation among stakeholders and an intense activity with development of standards and policies. The “network” stage is the fourth and last stage; only few organizational bottlenecks remain, there is a broad support for the goals of changes, a vision, cooperation, open communication and proactive community innovative solutions for societal problems.

Kok defines four critical organizational components of the SDI to move from one stage to the other for reaching the ideal fourth stage: (1) leadership; (2) a vision; (3) communication channels; (4) ability of the spatial information community for self-organization. Capacity Building activities are able to address these different components: (1) SDI leadership can be formed through capacity building activities, which is the aim of the BGSIP workshop described in chapter 3 under its train-the-trainer format; (2) a common SDI vision in a community naturally arises when the various issues are understood after sensibilization and awareness raising; (3) based on understanding of SDI basic concepts such as non redundancy and use of web services, external communication evolves from purely internal to an external one with other stakeholders. Education to these SDI principles plays hence a central role in the communication process; (4) ability of a community to solve problems evolves through the different stages from the delegation of these problems to others, to the ability to self-solving them through innovative solutions based on a committed approach. This self-organization is the result of Capacity building work. In summary, we can say that Capacity Building supports the various aspects of transition from no or embryonary SDI to mature SDI corresponding to the “network” stage.

During this long term capacity building process accompanying the SDI development, capacity building activities need to be tailored for addressing not

only the various audiences, but also the three levels defined in the GEO capacity building strategy (individual, institutional and technical) as well as the three hierarchical levels (individual, entity and societal levels) defined by Williamson. To our opinion, the individuals are the common denominators of all these capacity levels as the upper hierarchical levels (institutions or society levels) are made of groups of individuals. All the same, the institutions or infrastructures of a country need individual skills for their re-enforcement. This shows the specific necessity to build individual capacities, whose knowledge will influence the technical and political spheres for the benefit of the whole society. The various experiences reported in chapter 3 confirmed the importance of individuals, who are key in a SDI implementation process.

Because of the importance of capacity building activities that can be seen as the backbone of solutions to SDI implementation, chapter 5 discussed the possibilities to assess the real impact of the different activities part of the processes of building a SDI, focusing in particular on impact of the capacity building activities. Such assessments are necessary to justify the actions taken and their continuity, the funds spent, and allow re-evaluating the actions taken in case they are not satisfactory. Giff (2006) states that good performance indicators should be specific, relevant and timely. He adds that although qualitative performance indicators are usually used to measure the socio-political impact of a program, they often need to be quantified, which is a complex task, as funding agencies need quantitative information for taking cognitive decisions and performing comparative analyses. The essential role of indicators for measurement was highlighted, along with the challenge of defining the necessary measurable variables.

Chapter 5 also discussed the specific assessment of geoportals, which can be seen as an exercise of capacity building aiming at giving guidelines for improving this essential part of an SDI that allows discovery, view and access to data. Despite the critical human aspect that is at the heart of SDIs, the technological aspect should not be neglected, as it is indispensable to link the various concerned stakeholders through the Internet. Besides, it often reflects the institutional view of its funders and promoters: a geoportal not fulfilling its role might show that capacity building is needed both at the technical level for improvement and at institutional level for awareness raising on the concepts and benefits of SDIs. The same remark goes to web services, which are an indicator of geoportals assessment: very often they exist but are so difficult to find that they become useless, whereas they should be a fundamental technical element of SDI. Small efforts to make them more accessible would make the difference. This demonstrates the importance of both capacity building necessary to promote principles of coherent SDI architecture, and also importance of assessments that allow highlighting such issues. This example illustrates the importance of a clear vision for an SDI, that should clarify what the SDI aims for (Grus et al., 2006) and that requires coordination, concertation and awareness, which can be acquired through capacity building.

Specific assessments measuring impact of activities or specific SDI components are logically discussed in the fifth and final chapter of this thesis, whereas a global SDI assessment – the stocktaking – was discussed in the 1st chapter after

the introduction. This is because assessment under different forms is always needed at regular intervals in the SDI processes: for the whole SDI (stocktaking) and for all the various actions performed during the implementation. This is not only the case for SDI implementation but also for SDI maintenance and improvement. We can in fact see SDI as a cyclic process made of many interlinked sub-cyclic processes, that all need assessment for improvement or maintenance as illustrated in Figure 15 showing the SDI cycle. The main SDI cycle starts with (1) a first assessment of the whole SDI (stocktaking), based on which a vision is drawn, goals are set up and next actions defined; then (2) the capacity building phase starts with the definition, implementation and impact assessment of capacity building actions (e.g. awareness raising, technical workshops), followed by (3) an SDI implementation phase with definition, implementation and impact assessment of specific actions for implementing/improving the SDI (e.g. modification of data policies, development of geoportals). At this stage, the impact of the whole SDI on the society should ideally be assessed but given the shift between the time when the SDI is built and produces effects, and the time when these effects become measurable it is not possible to perform this assessment in the first years. Instead, it is possible to assess again the whole SDI and determine if the goals defined during the last stocktaking are met. Then new goals can be set up and the cycle goes on at a regular rhythm in view of constantly improving the SDI.

In parallel, similar sub-cycles of assessment, actions definition and impact assessment of these actions can be taken at several levels, for example for the capacity building actions or for the implementation actions (e.g. geoportals assessment). These cycles can even occur at lower scale. As an illustration, the EGIDA methodology, which is part of capacity building activities already includes for itself such an auto-assessment activity. We can see these cycles as a system of smaller interconnected systems that need each other for always improving their performance that will benefit the whole system.

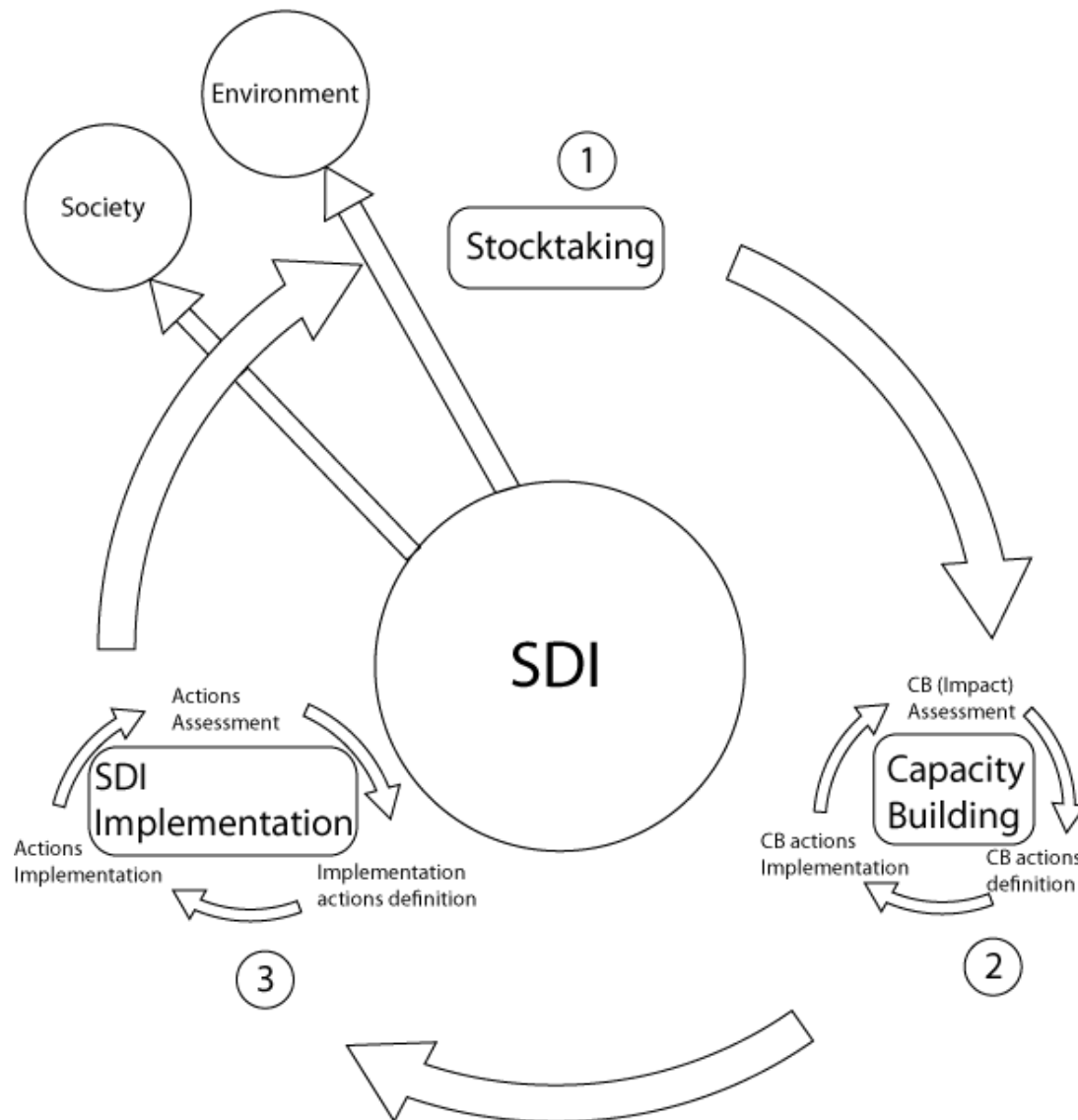


Figure 15: SDI cycles

Based on the outcomes of chapters 3 and 5, we can retain the following key messages about capacity building and measurement of its impact:

- Different stages characterize the implementation of a SDI. During these phases, the human component towards innovation should ideally evolve from resistance or passivity to commitment. Given the critical aspect of the human factors that will determine success or failure of the SDI implementation, tailored capacity building is needed all along the process.
- Among the various capacity building activities, awareness raising is of utmost importance as only awareness from the various stakeholders can create the necessary commitment. Commitment is necessary to lift the various barriers for next steps of SDI activities.

- Among the different levels of capacity building or the SDI components, the capacity building of individuals is key as it can then influence the company, institution, whole society regarding all the aspects of SDI, including infrastructural and organizational.
- Assessment of the various SDI implementation activities are necessary to ensure they have a real impact and are on track regarding the goals set up at the beginning of the process. This is particularly true for capacity building activities that are key to the whole SDI process.
- In an assessment process, one of the main challenges is to determine and agree on the necessary measurement variables corresponding to the defined indicators.
- SDIs can be seen as a cyclic system of systems that all have the same sequential cycles of stocktaking – actions – impact assessment – stocktaking for constant re-evaluation and improvement of each sub-system and eventually the whole SDI.

6.2 Contributions, innovations, and their relevance to scientific and practical advancements

The overarching goal of this research is to contribute to improving data discovery and access, which are essential to address environmental issues but still lacunar worldwide. As SDIs are the recognized framework to best manage data sharing issues, this thesis addresses the potential of the three different pillars of SDI: infrastructure, people and institutions. For each of them, it proposes an innovative contribution through specific methodologies or applications, while keeping in mind the necessity to remain simple to improve understanding and up taking by the wider user range.

The second chapter (SDI stocktaking), through the “Spatial Data Infrastructures in Africa: a gap analysis” article, describes a pioneering methodology for a rapid large-scale assessment of SDI implementation. It proposes an alternative to traditional methods that are resource demanding and most of the time performed at national level or smaller scales. It allows getting a broad regional overview, which can be further refined at smaller scales with usual assessment methods that require more resources. It is based on SDI fundamental indicators defined by experts. Besides giving coherent results when applied on Africa, it allows determining some strengths and weaknesses, and gives the opportunity to step back and propose solutions for most SDI components.

The third chapter (capacity building) proposes a new capacity building methodology called “Bringing GEOSS Services into Practice” that contributes to the development of individual geospatial skills. It is innovative in the sense that it provides a complete training for the whole chain of geospatial data concepts and management, from awareness raising to data sharing, promoting open source software and open data. It allows individual trainees to have sufficient knowledge for laying the foundations of an SDI in their institution.

The other contribution of chapter three, through the article “Leading the way towards an environmental National Spatial Data Infrastructure in Armenia”, consists in a capacity building success story at national level that led to the participation of Armenia’s to become a full GEO member. This article produced a set of best practices and recommendations that might be useful to other countries in the process of establishing a Spatial Data Infrastructure. Furthermore, the EGIDA institutional capacity building methodology applied in Armenia as a pilot represents an innovative methodological set of practices and guidelines for countries contribution to GEOSS and European INSPIRE initiatives. Its description and application in Armenia are clearly a new contribution to institutional capacity building activities.

The fourth chapter (implementation) proposes innovative solutions for facilitating data discovery and access through automation, brokering or metadata simplification. They all have in common to reduce the technological complexity that might hinder a broader data use and production. All these solutions rely on interoperability, standards, open source, open data and open standards. They contribute to a global effort towards open data principles and data sharing for SDI implementation. The “SCOPED-W” article proposes an advanced way to access SWAT customized data through a data extractor using OGC web services in a dedicated geoportal. The “Enabling discovery of African geospatial resources” article contributes to African geospatial data valorization through the innovative brokering technology. The “Facilitating the production of ISO-compliant metadata of geospatial datasets” article is a contribution to tackle the metadata production issue that remains a weakness in the geospatial data production workflow. The innovative aspect of this methodology lies in the semi-automatization of the metadata creation and update made possible by linkages established between a geographic data server and a metadata catalog through OGC web services.

The fifth chapter (assessment) is a contribution to the improvement of SDI specific elements or activities related to its general assessment. The “EGAL” article proposes a methodology for assessing geoportals based on their ability to be discovered and to facilitate data discovery and therefore open access. Such a methodology shall encourage geoportals developers to take into consideration best practices. The criteria and methodology proposed are innovative considering our literature review, as well as the capacity building aspect of the method through a simple, illustrative and incentive label.

The “EOPOWER impact assessment framework” is an innovative methodology for assessing impact of Earth Observation solutions or activities. Its application in two use cases is described in the second part of chapter five. This part contributes significantly to SDI capacity building activities, their performance and justification through the measurement of their impacts.

As a synthesis of the various chapters, the conclusion of the thesis contributes to a better understanding of the cyclic stages of SDI implementation at several levels. It also highlights the key role of individuals for a successful SDI and the necessity to take this cultural aspect into account in any SDI implementation. The necessity to reduce complexity as much as possible, which was considered in the

innovative methodologies proposed, is also a contribution to this “human” aspect that is the key to successful SDI solutions.

6.3 Recommendations and perspectives

The previous discussions and key messages highlighted strengths and weaknesses of the SDI approaches and implementations, and also suggested solutions to overcome the most obvious issues. This section summarizes some of the main recommendations and opens some perspectives based on the findings of this study, keeping simplicity as a leitmotiv whenever possible to address the complexity inherent to SDIs :

- **Privilege quick wins:** some simple quick wins are often possible to implement for rapidly obtaining small gains that added together can make a remarkable difference. They often concern the technical aspect of SDI (e.g. centralized database for African SDI monitoring indicators, better discoverability of web services). Given the long-term timeline of the SDI general mechanisms set up and effects visibility, such quick wins implementations could keep a SDI building process alive and should systematically be proposed to their respondent when feasible.
- **Initiate a dialogue with relevant stakeholders towards a regional SDI architecture:** based on the findings of the article on SDI assessment in Africa, one of the solutions proposed is to establish a regional architecture of key players and national focal points to improve SDI reporting and implementation. It would be important to establish dialogue with relevant stakeholders and to set up an agenda for implementing a pilot project with a few interested countries as a start. This would allow applying the quick wins just mentioned, contextualizing the SDI architectural suggestions and fostering long-term SDI relationships and commitments among the participants. Such a project requires long term funding and commitment that necessitate an attracting packaging. To initiate the dialogue with the stakeholders, it is possible to use previous collaborations (e.g. in the frame of the SALB project¹⁰⁹) or previous technological development such as the Africa broker
- **Promote the “open” principles:** the advantages of the open source software, open data and open standards have largely been demonstrated and discussed in this thesis. Keeping in mind the occasional need of complementary with proprietary solutions for software (e.g. in the case of large corporations), the open principles and solutions largely address the needs of SDIs in the environmental domain and represent a sustainable solution (e.g. long-term cost). This guarantees less dependence to funding issues, which is of particular importance in developing countries or in periods of budget restrictions. Promoting these principles and tools during capacity building processes ensures an integration of the trainees in a sustainable and supportive community sharing these principles.

¹⁰⁹ http://ggim.un.org/3rd%20Prep%20Meeting/SALB_GGIM-3.pdf

- **Increase the reach of SDI Capacity Building:** recognizing the fundamental aspect of Capacity Building for promoting the SDI principles and facilitating their wider use worldwide, there is a constant need to reach remote audiences and raise awareness of the geospatial concepts. The increasing number of people connected to the Internet¹¹⁰ (3,3 billion people at end 2015¹¹¹) and technological progress (e.g. the Google “LOON” project¹¹²) should have a multiplier effect. The Internet becomes then an incomparable mean to reach more people and capacity building material adapted for the Internet should be privileged. This is for example the case of the Massive Open Online Courses (MOOCs) that allow education to overcome the distance barrier. SDI concepts conveyed in this thesis could for example be packaged and spread in a larger MOOC.
- **Privilege solutions useable through a web browser for exchanging data:** as discussed in the chapter on implementation, reducing the complexity for the data user and producer by sending it to another level is key, and all the technical solutions proposed go to this direction. A web browser represents one of the simplest and familiar interface for a data user or producer, is free and available by default on a desktop, tablet or mobile device. It is then perfectly suited to be used as the interface between data users and producers, that meets the criteria linked to data exchange such as data discovery, visualization, access and processing.
- **Explore further solutions that address metadata issues and necessity together:** as discussed in chapter 4, metadata is a fundamental element in a SDI that allows data discovery and consequent benefits. But metadata production remains a burden for the data producers as it is fastidious to complete and requires precious time that could be saved for more interesting tasks. A solution has been proposed in this thesis through the article “facilitating the production of ISO-compliant metadata of geospatial datasets” that automatizes basic metadata production. Other solutions have been suggested that couple data and metadata through a specific format such as NetCDF. Given the recent development of the automatized methodology proposed in the article presented, further tests and implementation in projects workflows need to be evaluated. A potential integration of this methodology with coupled data/metadata formats such as NetCDF also need to be explored. The current automatized methodology requires to fill some basic metadata fields (e.g. name, title, abstract, keywords, projection, bounding box, point of contact) in the data publication server (e.g. Geoserver). Giving the possibility to the data producer of filling this same information directly in the GIS software (e.g. QGIS) would present the advantage of completely preparing the layer, including its metadata, in the GIS client. Combined with plugins such as OpenGeo Explorer that allow to

¹¹⁰ <http://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>

¹¹¹ <http://www.internetworldstats.com/stats.htm>

¹¹² <https://www.google.com/loon/>

directly publish a layer to the data production server from the GIS client, this could even further simplify the data production workflow but would require some development from the plugin developers.

- **Explore further participative solutions to data production:** the technological evolution of mobile devices (e.g. smartphones) has opened the way to new and increased possibilities of collecting data. Each mobile device can potentially be turned into a sensor able to transmit data collected to an SDI. This represents a huge potential of data contribution and citizen science involvement that necessitates mechanisms of integration and control. A success story of crowd-sourced data such as OpenStreetMap¹¹³ could potentially be replicated to environmental domains (e.g. the UNEP air quality monitoring system¹¹⁴). Moreover, the OGC proposes the “Sensor Web Enablement”¹¹⁵, which is a suite of standard encodings and web services that enable discovery of sensors, processes, and observations. Given the potential added-value of such sensor data in SDIs, a capacity building procedure for collecting, managing and sharing sensor data would represent an important contribution to the environmental community.
- **Promote technological best practices:** unlike certain types of best practices that cannot easily be replicated to other environments because of cultural issues, the technological best practices are universal and can then be promoted independently from the cultural environment. A label, such as the one proposed in the “EGAL” article of chapter 5, is a simple, illustrative and elegant way to promote best practices that can also be used in a capacity building approach. Such a label could be refined to more accurately reflect best practices linked to geoportals and related data discovery and access, and could also be extended to other SDI technological aspects for promoting best-recognized practices. An example of best practices promotion is the “GEOSS best practices wiki”¹¹⁶.
- **Develop a methodology to assess the impact of SDI on the society and on the environment:** there is a time lag between the moment when a SDI is developed and the time when impact on the society or the environment can be measured. Such measurement is necessary as discussed in chapter 5, but needs the definition of many measurable indicators. Such a methodology, if successful, would be a major asset for future SDI funding, which is of importance in an increasing competitive environment driven by budget cuts in many areas.

¹¹³ <https://www.openstreetmap.org>

¹¹⁴

http://uneplive.unep.org/media/docs/news_ticker/Air_Quality_Leaflet_Letter_size.pdf

¹¹⁵ <http://www.ogcnetwork.net/SWE>

¹¹⁶ <http://wiki.ieee-earth.org/>

- **Build guidelines including all the different aspects of the SDI cycle presented in this thesis:** even if it is not possible to get a one-fits-all solution for SDI implementation, some general best practices highlighted along the various chapters of this thesis could be summarized and packaged to give a global overview of the SDI cycle and raise awareness on its different aspects. This would constitute a general framework, with some parts where best practices can be clearly mentioned (mainly the technological ones that are universal) and other parts more demonstrating success stories in specific regions. Such guidelines would constitute additional capacity building material that would allow to have a global overview of SDI implementation elements.

6.4 Final conclusion

This research started by the ascertainment that geospatial data are essential to support sound political decisions necessary to address increasing environmental issues. Spatial Data Infrastructures are recognized as a framework able to provide the necessary structure and tools to efficiently discover, visualize and access geospatial data. Despite this promising possibility, it was not possible to find out the degree of SDIs implementation worldwide or in specific areas. It was not possible either to find a straightforward methodology for performing myself this assessment. Besides, in spite of the existence of SDIs for decades, finding useful and usable environmental data on the Internet remains challenging depending on time and location factors. From these statements, a deeper analysis was necessary to better understand the issues and barriers causing such difficulties in deployment of SDIs worldwide.

As technology already exists to efficiently discover, visualize and access geospatial data, we wanted to focus on other possible causes. For that reason the human factor was selected, which is at the core of each aspect of SDIs, as a possible cause of some SDIs failures. As capacity building represents the most suitable way to address human aspects, we decided to examine the role it plays in SDI components and processes, and how this can be optimized to positively influence SDIs. This was done through a deepening of three related aspects: (1) the way to assess the status of SDI implementation, as it is necessary to determine where capacity building is needed; (2) the type of innovative actions that can be taken to improve SDI implementation and (3) the way to assess if the actions taken are effective, appropriate and have a positive impact.

The three research questions associated to these related aspects gave the opportunity to take the measure of their complexity. Their multi-faceted aspect and dynamic nature that makes them so complex mirrors the human nature that has similar features. The complex and diverse human cultures must invariably be integrated in SDIs since the institutional and individual components are direct products of particular cultures. Integrating diversity in the equation prevents a one-fits-all solution or a simple transposition of best practices for SDIs but requires case-by-case solutions based on a common framework in which local specificities can be included. A parallel can be drawn with globalization of the

economy, which accentuates the gap between the developing and developed world, as models working in a specific context might not be directly transposable to other cultures. One of the problems in both economic and SDI cases is that the technical side (e.g. infrastructure SDI component and economic models) requires a uniform and standardized approach to work. These models leave aside the cultural aspects that are a barrier to uniformity and might cause a failure of the whole system if not integrated.

A balance needs then to be found between the necessary technical uniformity and the cultural plurality inherent to mankind. Masser (2006) distinguishes a top down vision that emphasizes the need for standardization and uniformity from a bottom up vision that stresses the importance of diversity and heterogeneity. A general solution might consist in decoupling as much as possible the technical components from the cultural ones. A given community should be able to use its own technology and standards while still being connected to the system. This is made possible in a system of systems approach and concretely illustrated in the article “enabling discovery of African geospatial resources” of chapter 3 where a tool (the broker) bridges the gap between different scientific communities (Nativi et al., 2012). In this kind of approach, very little effort is requested from the target community. Even though this technical solution goes into the right direction by finding a compromise between the technical necessities and community-specific singularities, it was not understood and uptaken as expected in Africa. This shows the need to work on several levels as a single solution will not make the difference, but many smaller improvements on the several SDI components might help the whole system implementation to progress. This also shows the necessity to permanently assess the impact of the actions taken, to evaluate if they meet the goals defined, which was largely discussed in chapter 5 and in the answer to the third research question.

Capacity building tackles the issue from another angle as it builds knowledge in other cultures about the technical uniformity principles required. In this case, an effort of adaptation is expected in the long run from the audience. This is intended to reduce the gap between cultural specificities and internationally recognized technical solutions. This is typically the role of awareness raising and requires a good knowledge of the cultural aspects of the targeted audience. This approach is a global one as it can address all the SDI components whereas the system of system approach targets the technical level. However, as previously mentioned several solutions are needed and not exclusive; good technical solutions are completely compatible with capacity building actions, and can in turn be promoted as capacity building material.

Besides the positive aspects played by all the solutions discussed in this thesis for an improvement of data discovery and access, one should also keep in mind the necessity of quality data for sound political decisions. Data quality and uncertainty inherent to data and models need to be taken into account. This is being addressed through standards such as UncertML¹¹⁷ and QualityML¹¹⁸ that

¹¹⁷ <http://www.uncertml.org/>

¹¹⁸ <http://qualityml.geoviqua.org/>

enrich metadata with a more comprehensive suite of terms that describe a broader range of quality measures.

The various discussions along the chapters also highlighted the crucial importance of the networks that are indispensable in SDIs whose fundamental role is exchange. Networks can be digital ones for exchanging data but can also be human ones in charge of building an SDI. In both cases networks are made of interconnected nodes. This connexion is possible through understanding from each node of the information received and the next action necessary in the network for support of the whole system. In the case of digital networks, this understanding is called interoperability and is made possible by the standards. In the case of human networks, the understanding is called cooperation and is made possible through capacity building. In both cases a framing normative institution (e.g. OGC for technical standards and GEO for coordination at global levels) is necessary to ensure coherence of the networks, which might trigger a balance of power.

In conclusion, given the prominent human factor that is central in SDIs, capacity building is definitely one of the most critical aspects for SDI improvement and is a concept of choice for bridging science and policy at all levels. Individual smaller capacity building actions such as freely and openly sharing capacity building material (e.g. online tutorials on YouTube or openclassrooms¹¹⁹) have a place of choice as they can intensively contribute to individual capacity building. But capacity building can only make a SDI agenda reach its goal along with other solutions such as technical ones that need to be as tailored as possible to the cultural specificities of a given community. One should also keep in mind that SDIs are not an end in itself but are a framework for data efficiency in support of actions necessary for addressing environment challenges. Among these actions, building capacity of everyone, especially the younger generations, to understand the challenges and provide the tools to address them is essential for a hopeful future.

¹¹⁹ <https://openclassrooms.com/>

ANNEXES

List of figures

<i>Figure 1: the Decision-Information-Knowledge-Wisdom (DIKW) pyramid (based on Ackoff, 2010).</i>	<i>19</i>
<i>Figure 2: Nature and relations between SDI components (source: Rajabifard, 2002)</i>	<i>26</i>
<i>Figure 3: SDI components</i>	<i>26</i>
<i>Figure 4: SDI hierarchy (based on Rajabifard, 1999)</i>	<i>28</i>
<i>Figure 5: Capacity Building levels (based on Williamson, 2003 and the GEO Capacity Building strategy, 2006)</i>	<i>32</i>
<i>Figure 6: SDI components and GEO Capacity Building levels</i>	<i>33</i>
<i>Figure 7: Synthesis of projects and articles contribution to the chapters of the thesis</i>	<i>40</i>
<i>Figure 8: the Multi-View Assessment Framework (Grus et al., 2007)</i>	<i>42</i>
<i>Figure 9: the EGIDA methodology activities (Mazzetti et al., 2013)</i>	<i>91</i>
<i>Figure 10: elements of SDI assessment</i>	<i>167</i>
<i>Figure 11: step-by-step benefit of Earth Observation application (Noort, 2013)</i>	<i>200</i>
<i>Figure 12: Cumulative number of downloads of the workshop material in function of time (March to December 2014) from the open archive of Geneva University</i>	<i>207</i>
<i>Figure 13: Resources needed for stocktaking in relation to the level of details obtained</i>	<i>223</i>
<i>Figure 14: complexity reduction impact on data use and production</i>	<i>226</i>
<i>Figure 15: SDI cycles</i>	<i>233</i>

List of acronyms

ADB	African Discovery Broker
AFREF	African Geodetic Reference Frame
BGSIP	Bringing GEOSS Services into Practice
BPB	Business Process Broker
CB	Capacity Building
CENS	Center for Ecological-Noosphere Studies
CODIST	Committee on Development Information, Science and Technology
CoP	Conference of Parties
CSW	Catalog Service for the Web
DAB	Discovery and Access Broker
DIKW	Decision-Information-Knowledge-Wisdom
EBV	Essential Biodiversity Variable
ECOWAS	Economic Community of West African States
ECV	Essential Climate Variable
EO	Earth Observation
EOV	Essential Ocean Variable
ERA	European Research Area
ESA	European Space Agency
EU	European Union
EV	Essential Variable
FOSS	Free and Open Source Software
GBIF	Global Biodiversity Information Facility
GCI	GEOSS Common Infrastructure
GEO	Group on Earth Observations
GEOCAB	Global Earth Observation CApacity Building
GEOSS	Global Earth Observation System of Systems
GIS	Geographic Information System
GSDI	Global Spatial Data Infrastructure

GUI	Graphical User Interface
ICT	Information and Communication Technology
IIAP	Institute for Informatics and Automation Problems
ITU	International Telecommunications Union
INSPIRE	Infrastructure for Spatial Information in the European Community
IPCC	Intergovernmental Panel on Climate Change
ISO	International Organization for Standardization
IWRM	Integrated Water Resources Management
MOOC	Massive Open Online Course
NMA	National Mapping Agency
NSDI	National Spatial Data Infrastructure
OGC	Open Geospatial Consortium
OSGeo	Open Source Geospatial Foundation
PNF	Permanent Networking Facility
SBA	Societal Benefit Area
SDI	Spatial Data Infrastructure
SoP	State of Play
SOS	Sensor Observation Service
SWAT	Soil and Water Assessment Tool
TWAP	Transboundary Waters Assessment Programme
UN	United Nations
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information Management
UNCED	United Nations Conference on Environment and Development
UNECA	United Nations Economic Commission for Africa
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
UNSDI	United Nations Spatial Data Infrastructure

USGS	United States Geological Survey
WCS	Web Coverage Service
WFS	Web Feature Service
WHO	World Health Organization
WMO	World Meteorological Organization
WMS	Web Mapping Service
WPS	Web Processing Service

REFERENCES

- 52°North (2015) "Advanced time series analysis on the web with R."
<http://blog.52north.org/2015/04/22/advanced-time-series-analysis-on-the-web-with-r/>
- Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R. (2007). "Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT." *Journal of hydrology* **333**(2): 413-430.
- Ackoff, R. L. (2010). "From data to wisdom." *Journal of applied systems analysis* **16**: 3-9.
- AEGOS (2011) AEGOS project. WWW document, <http://www.aegos-project.org>
- African Development Bank (2014) Open data for Africa. WWW document, <http://www.afdb.org/en/knowledge/statistics/open-data-for-africa/>
- Afromaison (2011) The Afromaison project. WWW document, <http://www.afromaison.net>
- Afromaison (2013) Africa broker. WWW document, <http://afromaison.grid.unep.ch:8080/gi-cat/gi-portal>
- Aleksandryan, A. (2006). "Contamination of water basins of Armenia by persistent organic pollutants. ." *NATO Security through Science Series*.
- Annoni, A., Craglia, M., Smit, P. (2002). *Comparative analysis of NSDI*. Proceedings of the 8th EC-GI&GIS Workshop. Dublin, Ireland.
- Arshad, N. H. and Hanifah, F. (2010). *Issues and challenges in NSDI implementation*. Proceedings of the 9th WSEAS international conference on System science and simulation in engineering.
- Asmaryan, S., Saghatelyan, A., Astsatryan, H., Bigagli, L., Mazzetti, P., Nativi, S., Guigoz, Y., Lacroix, P. M. A., Giuliani, G., Ray, N. (2014). "Leading the way toward an environmental National Spatial Data Infrastructure in Armenia." *South-Eastern European Journal Issue of Earth Observation and Geomatics* **3**: 53-62.

Astsatryan, H., Hayrapetyan, A., Narsisian, W., Saribekyan, A., Asmaryan, S., Saghatelyan, A., Muradyan, V., Guigoz, Y., Giuliani, G., Ray, N. (2014). "An interoperable web portal for parallel geoprocessing of satellite image vegetation indices." Earth Science Informatics: 1-8.

Astsatryan, H., Narsisian, W., Asmaryan, S., Saghatelyan, A., Muradyan, V., Giuliani, G., Guigoz, Y., Ray, N. (2015). "An interoperable cloud-based scientific GATEWAY for NDVI time series analysis." Computer Standards & Interfaces **41**: 79-84.

Astsatryan, H., Narsisian, W., Ghazaryan, V., Saribekyan, A., Asmaryan, S., Muradyan, V., Guigoz, Y., Giuliani, G., Ray, N. (2012). "Toward the Development of an Integrated Spatial Data Infrastructure in Armenia."

Ayanlade, A., Orimoogunje, I. O. O., Borisade, P. B. (2008). "Geospatial data infrastructure for sustainable development in sub-Saharan countries." International Journal of Digital Earth **1**(3): 247-258.

Bacu, V., Mihon, D., Stefanut, T., Rodila, D., Abbaspour, K., Rouholahnejad, E., Gorgan, D. (2013). "Calibration of SWAT hydrological models in a distributed environment using the gSWAT application." International Journal of Advanced Computer Science and Applications: 66-74.

Bastin, L., Cornford, D., Jones, R., Heuvelink, G. B. M., Pebesma, E., Stasch, C., Nativi, S., Mazzetti, P., Williams, M. (2013). "Managing uncertainty in integrated environmental modelling: The UncertWeb framework." Environmental Modelling & Software **39**: 116-134.

Batcheller, J. K. (2008). "Automating geospatial metadata generation—An integrated data management and documentation approach." Computers & Geosciences **34**(4): 387-398.

Batcheller, J. K., Gittings, B. M., Dunfey, R. I. (2009). "A method for automating geospatial dataset metadata." Future Internet **1**(1): 28-46.

Bektas, F. B., Mihon, D., Colceriu, V., Allenbach, K., Goksel, C., Dogru, A. O., Giuliani, G., Gorgan, D. (2013). "Remotely Sensed Data Processing on Grids by Using GreenLand Web Based Platform." International Journal of Advanced Computer Science and Applications **3**(3): 58-65.

Bell-Pasht, K. and Krechowicz, D. (2015). Why does access to good climate data matter? WMO Bulletin. **64** (2).

Beniston, M., Stoffel, M., Harding, R., Kernan, M., Ludwig, R., Moors, E., Samuels, P., Tockner, K. (2012). "Obstacles to data access for research related to climate and water: implications for science and EU policy-making." environmental science & policy **17**: 41-48.

Berners-Lee, T. (2006) "Linked Data."
<https://www.w3.org/DesignIssues/LinkedData.html>.

Bigagli, L. and Lipiarski, D. (2015a). Science Valorisation Methodology and Use-Cases.

Bigagli, L., Santoro, M., Mazzetti, P., Nativi, S. (2015b). "Architecture of a Process Broker for Interoperable Geospatial Modeling on the Web." ISPRS International Journal of Geo-Information **4**(2): 647-660.

Bizer, C., Heath, T., Berners-Lee, T. (2009). "Linked data-the story so far." Semantic Services, Interoperability and Web Applications: Emerging Concepts: 205-227.

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., Zemp, M. (2014). "The concept of Essential Climate Variables in support of climate research, applications, and policy." Bulletin of the American Meteorological Society **95**(9): 1431-1443.

Boldrini, E., Santoro, M., Papeschi, F., Nativi, S. (2013). Data access services interoperability in the geosciences by means of the GI-axe Brokering Framework. EGU General Assembly Conference Abstracts.

Castronova, A. M., Goodall, J. L., Elag, M. M. (2013). "Models as web services using the open geospatial consortium (ogc) web processing service (wps) standard." Environmental Modelling & Software **41**: 72-83.

Cau, P., Manca, S., Soru, C., Muroli, D., Gorgan, D., Bacu, V., Lehmann, A., Ray, N., Giuliani, G. (2013). "An interoperable, GIS-oriented, information and support system for water resources management." International Journal of Advanced Computer Science and Applications **3**(3): 75-82.

Cepek, A. (2008). Proceedings of the Workshop Geoinformatics. FCE CTU.

Chan, T. O., Feeney, M.-E., Rajabifard, A., Williamson, I. P. (2001). "The dynamic nature of spatial data infrastructures: a method of descriptive classification." **55**: 65-72.

Charvat, K., Vohnout, P., Sredl, M., Kafka, S., Mildorf, T., De Bono, A., Giuliani, G. (2013a). "Enabling Efficient Discovery of and Access to Spatial Data Services." International Journal of Advanced Computer Science and Applications **EnviroGRIDS Special Issue on "Building a Regional Observation System in the Black Sea Catchment"**: 28-31.

Charvat, K., Vohnout, P., Sredl, M., Kafka, S., Mildorf, T., De Bono, A., Giuliani, G. (2013b). "Enabling efficient discovery of and access to spatial data services." International Journal of Advanced Computer Science and Applications **3**(3): 28-31.

Clarke, D., Holland, P., Nairn, A., Swift, A., and Robertson, D. (1999) "The contribution of the Global Map to a GSDI. ." <http://www.gsdi-docs.org/docs1999/globalmap.htm>.

Climatic Research Unit (CRU) (2008) CRU Time Series (TS) high resolution gridded dataset. WWW document, http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_dataent_1256223773328276

Clinton, W. J. (1994). "Coordinating geographic data acquisition and access: the National Spatial Data Infrastructure." Executive Order **12906**(17671.17674).

CODATA (2015). The Value of Open Data Sharing. GEO XII. Geneva, Switzerland.

Coetzee, S. (2011). "Results from a normative dependency analysis of geographic information standards." Computer Standards & Interfaces **33**(5): 485-493.

Coetzee, S. and Eksteen, S. (2012). "Tertiary education institutions in Africa: Cloudy with a chance of GISc education in some countries." South African Journal of Geomatics **1**(2): 13.

Colceriu, V., Mihon, D., Minculescu, A., Bacu, V., Rodila, D., Gorgan, D. (2013). "Workflow based description and distributed processing of satellite images." International Journal of Advanced Computer Science and Applications: 50-57.

Cooper, A. K. and Gavin, E. (2005). Spatial metadata in Africa and the Middle East. World Spatial Metadata Standards.

Cooper, A. K., Moellering, H., Hjelmager, J., Rapant, P., Delgado, T., Laurent, D., Coetzee, S., Danko, D. M., Düren, U., Iwaniak, A. (2013). "A spatial data infrastructure model from the computational viewpoint." International Journal of Geographical Information Science **27**(6): 1133-1151.

Craglia, M. (2010). "Building INSPIRE: The Spatial Data Infrastructure for Europe." ArcNews **32 No1**(Spring 2010): 5-7.

Craglia, M. and Campagna, M. (2009a). "Advanced Regional SDIs in Europe: comparative cost-benefit evaluation and impact assessment perspectives." International Journal of Spatial Data Infrastructures Research.

Craglia, M. and Campagna, M. (2009b). "Advanced regional spatial data infrastructures in Europe." Luxembourg: Office for official publications of the European Communities.

Craglia, M., de Bie, K., Jackson, D., Pesaresi, M., Remetey, F., pp, G., bor, Wang, C., Annoni, A., Bian, L., Campbell, F., Ehlers, M., van Genderen, J., Goodchild, M., Guo, H., Lewis, A., Simpson, R., Skidmore, A., Woodgate, P. (2012a). "Digital Earth 2020: towards the vision for the next decade." International Journal of Digital Earth **5**(1): 4-21.

Craglia, M., de Bie, K., Jackson, D., Pesaresi, M., Remetey-Fülöpp, G., Wang, C., Annoni, A., Bian, L., Campbell, F., Ehlers, M. (2012b). "Digital Earth 2020: towards the vision for the next decade." International Journal of Digital Earth **5**(1): 4-21.

Craglia, M., Pavanello, L., Smith, R. (2010). "The use of spatial data for the preparation of environmental reports in Europe."

Crompvoets, J. (2006). National spatial data clearinghouses: worldwide development and impact, WUR Wageningen UR.

Crompvoets, J., Rajabifard, A., Bregt, A., Williamson, I. P. (2004). Assessing the worldwide developments of national spatial data clearinghouses. **18**: 665-689.

Crompvoets, J., Rajabifard, A., Van Loenen, B., Fernández, T. D. (2008). A multi-view framework to assess SDIs, Wageningen University, RGI.

De Longueville, B. (2010). "Community-based geoportals: The next generation? Concepts and methods for the geospatial Web 2.0." Computers, Environment and Urban Systems **34**(4): 299-308.

Delgado Fernandez, T., Lance, K., Buck, M., Onsrud, H. J. (2005). Assessing an SDI Readiness Index. FIG Working Week 2005 and GSDI-8. Cairo, Egypt: 10.

Dessers, E., Vancauwenberghe, G., Vandenbroucke, D., Cromptvoets, J., Van Hootegem, G. (2015). "Analysing spatial data performance in inter-organisational processes." International Journal of Digital Earth **8**(5): 403-420.

Díaz, L., Granell, C., Huerta, J., Gould, M. (2012a). "Web 2.0 Broker: A standards-based service for spatio-temporal search of crowd-sourced information." Applied Geography **35**(1): 448-459.

Díaz, P., Masó, J., Sevillano, E., Ninyerola, M., Zabala, A., Serral, I., Pons, X. (2012b). "Analysis of Quality Metadata in the GEOSS Clearinghouse." Int'l J. Spatial Data Infrastructures Research **7**: 352-377.

Díaz Sánchez, L., Remke, A., Kauppinen, T., Degbelo, A., Foerster, T., Stasch, C., Rieke, M., Schaeffer, B., Baranski, B., Broering, A. (2012). "Future SDI-impulses from geoinformatics research and IT trends." International Journal of Spatial Data Infrastructures Research **7**: 378-410.

Dietz, S., Bowen, A., Dixon, C., Gradwell, P. (2016). "Climate value at risk/' of global financial assets." Nature Clim. Change **advance online publication**.

Domenico, B., Nativi, S., Caron, J., Bigagli, L., Davis, E. (2006). A standards-based, web services gateway to netCDF datasets. Proc. of AMS-22nd IIPS Conference, Atlanta, Georgia, abstr.

Donert, K. (2015a). Digital Earth – Digital World: strategies for geospatial technologies in 21st century education. Geospatial Technologies and Geography Education in a Changing World. Springer.

Donert, K. (2015b). Digital Earth–Digital World: Strategies for Geospatial Technologies in Twenty-First Century Education. Geospatial Technologies and Geography Education in a Changing World.

Eade, D. (1997). Capacity-building: an approach to people-centred development, Oxfam.

Econstats (2013) Econstats statistics. WWW document, http://www.econstats.com/wdi/wdiv_533.htm

ECOWREX (2014) ECOWREX map viewer. WWW document, <http://www.ecowrex.org/?q=node/906>

Eelderink, L. (2006). Toward key variables to assess National Spatial Data Infrastructures (NSDIs) in developing countries. MSc, Geographical Information Management and Application (GIMA).

Eelderink, L., Cromptvoets, J., de Man, E. (2008a). Chapter 15: Towards key variables to assess National Spatial Data Infrastructures (NSDIs) in developing countries. A Multi-View Framework to Assess SDIs. R. A. Cromptvoets Joep, van Loenen Bastiaan & Delgado Fernández Tatiana.

Eelderink, L., Cromptvoets, J., Erik De Man, W. H. (2008b). Towards key variables to assess National Spatial Data Infrastructures (NSDIs) in developing countries. A Multi-View Framework to Assess Spatial Data Infrastructures.

EIONET (2012) EIONET dataflow. WWW document, <http://www.eionet.europa.eu/dataflows/pdf2012/>

EIS-Africa (2002). Geo-Information Supports Decision-Making in Africa: an EIS-Africa Position Paper.

EIS-Africa (2014) EIS-Africa overview. WWW document, <http://www.eis-africa.org/>

El-Askary, H., Park, S., Ahn, M., Prasad, A., Kafatos, M. (2015). "On the Detection and Monitoring of the Transport of an Asian Dust Storm Using Multi-sensor Satellite Remote Sensing." Journal of Environmental Informatics.

Ellul, C., Foord, J., Mooney, J. (2013). "Making metadata usable in a multi-national research setting." Applied ergonomics **44**(6): 909-918.

enviroGRIDS (2009) enviroGRIDS metadata catalog. WWW document, <http://envirogrids.grid.unep.ch:8080/geonetwork>

EOPOWER (2014) The EOPOWER project. WWW document, <http://eopower.eu>

Erskine, M. A., Gregg, D. G., Karimi, J., Scott, J. E. (2013). "Business decision-making using geospatial data: A research framework and literature review." Axioms **3**(1): 10-30.

ESA (2013) Monitoring and managing Africa's water resources. WWW document, http://www.esa.int/Our_Activities/Observing_the_Earth/Monitoring_and_managing_Africa_s_water_resources

ESA (2014) ESA Global Land Cover Map. WWW document, http://www.esa.int/Our_Activities/Observing_the_Earth/Space_for_our_climate/ESA_global_land_cover_map_available_online

ESRI (2010). GIS best practices - Spatial Data Infrastructures (SDI).

ESRI (2012) ESRI Offices worldwide. WWW document, <http://www.esri.com/about-esri/offices>

ESRI (2016) GIS Dictionary. WWW document, <http://support.esri.com/en/knowledgebase/Gisdictionary/browse>

ESSILab (2014) GI-Cat tool. WWW document, <http://www.essi-lab.eu/do/view/GIcat/WebHome>

European Commission (2007). "Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)." Official Journal of the European Union.

European Commission (2012). "Communication from the commission to the European Parliament, the Council, the European economic and social committee and the committee of the regions: a reinforced European Research Area Partnership for Excellence and Growth."

European Commission (2013a) INSPIRE geoportal. WWW document, <http://inspire-geoportal.ec.europa.eu/>

European Commission (2013b) INSPIRE Member State Contact Points. WWW document, <http://inspire.ec.europa.eu/cms/contactpoints.cfm>

European Commission (2014a). Capacity Building for GEOSS: European contributions to capacity building activities in GEO.

European Commission (2014b) INSPIRE spatial data themes. WWW document, <http://inspire.ec.europa.eu/index.cfm/pageid/2/list/7>

European Environment Agency (2011) ECRINS. WWW document, <http://projects.eionet.europa.eu/ecrins>

EyeonEarth (2012) EyeonEarth geoportal. WWW document, <http://www.eyeonearth.org>

FAO (2013) Africover. WWW document, http://www.glcnet.org/activities/africover_en.jsp

FGDC (2013). Toward A New Strategic Vision for the National Spatial Data Infrastructure: A Report to the Federal Geographic Data Committee.

FGDC (2014) National Spatial Data Infrastructure WWW document, <http://www.fgdc.gov/nsdi/nsdi.html>

Florczyk, A. J., López-Pellicer, F. J., Nogueras-Iso, J., Zarazaga-Soria, F. J. (2012). "Automatic generation of geospatial metadata for web resources." International Journal of Spatial Data Infrastructures Research **7**: 151-172.

Forbes (2016) "How Satellite And Imaging Technologies Are Changing The World As We Know It (Part 2)."
<http://www.forbes.com/sites/truebridge/2016/05/06/how-imaging-technologies-are-changing-the-world-part-2/-4096dcb41a93>.

Foresman, T. W. (2008). "Evolution and implementation of the Digital Earth vision, technology and society." International Journal of Digital Earth **1**(1): 4-16.

Gallagher, S. (2012) "The future of bandwidth: how Africa is embracing "the cloud" on its own terms." Ars Technica.
<http://arstechnica.com/business/2012/06/why-africa-embraces-cloud-computing/>.

Garcia Almirall, P., Moix Bergadà, M., Queraltò Ros, P., Craglia, M. (2008). "The socio-economic impact of the spatial data infrastructure of Catalonia." JRC Scientific and technical reports.

Gassman, P. W., Reyes, M. R., Green, C. H., Arnold, J. G. (2007). "The soil and water assessment tool: historical development, applications, and future research directions."

GBIF (2012) Global biodiversity information facility. WWW document, <http://www.gbif.org/>

Gemeda, D. O. (2012). Assessing the development of Ethiopian National Spatial Data Infrastructure, Wageningen University.

GEO (2009). Implementation Guidelines for the GEOSS Data Sharing Principles.

GEO (2010a). Crafting geoinformation: the art and science of Earth observation.

GEO (2010b). GEO Communities of Practice. GEO-VII Plenary and 2010 Ministerial Summit, Beijing, China.

GEO (2012) GEOSS portal. WWW document, <http://www.geoportal.org>

GEO (2013a) Geonetcast. WWW document, <http://www.earthobservations.org/geonetcast.shtml>

GEO (2013b). The GEOSS Water Strategy: From Observations to Decisions. Executive Summary.

GEO (2014a) AfriGEOSS. WWW document, http://www.earthobservations.org/art_019_002.shtml

GEO (2014b). The GEOSS Water Strategy: From Observations to Decisions.

GEO (2014c) The Group on Earth Observations overview. WWW document, <http://www.earthobservations.org/index.php>

GEO (2014d). Report on Progress 2011-2013.

GEO secretariat (2005a). GEOSS 10-Year Implementation Plan: Table of Work Plan Targets: 1-22.

GEO Secretariat (2005b). The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan.

GEO secretariat (2006). GEO Capacity building strategy: 13.

GEO secretariat (2008). Building a Global Earth Observation System of Systems: 7.

GEO Secretariat (2009). "Delivering climate services through GEOSS."

GEO Secretariat (2011). "How GEOSS supports decision making in nine Societal Benefit Areas (SBAs)."

GEO secretariat (2012). AfriGEOSS: An Initiative to Reinforce GEOSS in Africa.

GEO Secretariat (2015). The value of open data sharing.

Geoff, H. (2011) "The Global Framework for Climate Services ".
<http://www.earthzine.org/2011/05/12/the-global-framework-for-climate-services/>.

George, C. and Leon, L. F. (2007). "WaterBase: SWAT in an open source GIS." The Open Hydrology Journal 1: 19-24.

Geoserver (2010) GeoServer User Manual. WWW document,
<http://docs.geoserver.org/stable/en/user/index.html>

Geraci, A. (1991). IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries, IEEE Press.

Giff, G. (2006). The value of performance indicators to Spatial Data Infrastructure development. GSDI-9, Santiago, Chile.

Giff, G. A. and Cromptvoets, J. (2008). "Performance indicators a tool to support spatial data infrastructure assessment." Computers, environment and urban systems **32**(5): 365-376.

Giuliani, G. (2011). Spatial data infrastructures for environmental sciences.

Giuliani, G., Dubois, A., Lacroix, P. M. A. (2013a). "Testing OGC Web Feature and Coverage Service performance: Towards efficient delivery of geospatial data." Journal of Spatial Information Science **7**: 1-23.

Giuliani, G. and Gorgan, D. (2013b). "Editorial: EnviroGRIDS Special Issue on "Building a Regional Observation System in the Black Sea catchment"." International Journal of Advanced Computer Science and Applications **3**(3): 4-8.

Giuliani, G., Guigoz, Y., Lacroix, P., Ray, N., Lehmann, A. (2016). "Facilitating the production of ISO-compliant metadata of geospatial datasets." International Journal of Applied Earth Observation and Geoinformation **44**: 239-243.

Giuliani, G., Lacroix, P., Guigoz, Y., Bigagli, L., Ray, N., Lehmann, A. (2014a). Bringing GEOSS services into practice.

Giuliani, G., Lacroix, P., Guigoz, Y., Bigagli, L., Ray, N., Lehmann, A. (2014b). Bringing GEOSS Services into Practice. GIS Open Source Workshop Material. . University of Geneva, United Nations Environment Programme, National Research Council of Italy.: 189.

Giuliani, G., Papeschi, F., Mlisa, A., Lacroix, P. M. A., Santoro, M., Nonguierma, A., Cools, J., Guigoz, Y. (2015). "Enabling Discovery of African Geospatial Resources." South-Eastern European Journal Issue of Earth Observation and Geomatics **4**(1S): 1-16.

Giuliani, G. and Peduzzi, P. (2011a). "The PREVIEW Global Risk Data Platform: a geoportal to serve and share global data on risk to natural hazards." Nat. Hazards Earth Syst. Sci. **11**(1): 53-66.

Giuliani, G., Rahman, K., Ray, N., Lehmann, A. (2013c). "OWS4SWAT: Publishing and Sharing SWAT Outputs with OGC standards." International Journal of Advanced Computer Science and Applications **3**(3): 90-98.

Giuliani, G., Ray, N., Lehmann, A. (2013d). "Building Regional Capacities for GEOSS and INSPIRE: a journey in the Black Sea Catchment." International Journal

of Advanced Computer Science and Applications (IJACSA)(EnviroGRIDS Special Issue on "Building a Regional Observation System in the Black Sea Catchment").

Giuliani, G., Ray, N., Lehmann, A. (2013e). "Building Regional Capacities for GEOSS and INSPIRE: a Journey in the Black Sea Catchment." International Journal of Advanced Computer Science and Applications **EnviroGRIDS Special Issue on "Building a Regional Observation System in the Black Sea Catchment"**: 19-27.

Giuliani, G., Ray, N., Schwarzer, S., De Bono, A., Peduzzi, P., Dao, H., Van Woerden, J., Witt, R., Beniston, M., Lehmann, A. (2011b). "Sharing environmental data through GEOSS."

Golledge, R. G., Marsh, M., Battersby, S. (2008). "Matching geospatial concepts with geographic educational needs." Geographical Research **46**(1): 85-98.

Gore, A. (1998). "The Digital Earth: Understanding our planet in the 21st Century." Australian surveyor **43**(2): 89-91.

Gorgan, D., Abbaspour, K., Cau, P., Bacu, V., Mihon, D., Giuliani, G., Ray, N., Lehmann, A. (2011). Grid based data processing tools and applications for black sea catchment basin. Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2011 IEEE 6th International Conference on, IEEE.

Gorgan, D., Bacu, V., Mihon, D., Rodila, D., Abbaspour, K., Rouholahnejad, E. (2012a). "Grid based calibration of SWAT hydrological models." Natural Hazards and Earth System Science **12**(7): 2411-2423.

Gorgan, D., Bacu, V., Mihon, D., Rodila, D., Stefanut, T., Abbaspour, K., Cau, P., Giuliani, G., Ray, N., Lehmann, A. (2012b). "Spatial data processing tools and applications for Black Sea catchment region." International Journal of Computing **11**(4): 327-335.

Gorgan, D., Bacu, V., Mihon, D., Stefanut, T., Rodila, D., Cau, P., Abbaspour, K., Giuliani, G., Ray, N., Lehmann, A. (2012c). "Software Platform Interoperability Throughout EnviroGRIDS Portal." Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of **5**(6): 1617-1627.

Gorgan, D., Giuliani, G., Ray, N., Lehmann, A., Cau, P., Abbaspour, K., Charvat, K., Jonoski, A. (2013a). "Black sea catchment observation system as a portal for

GEOSS community." International Journal of Advanced Computer Science and Applications **3**(3): 9-18.

Gorgan, D., Giuliani, G., Ray, N., Lehmann, A., Cau, P., Abbaspour, K., Charvat, K., Jonoski, A. (2013b). "Black Sea Catchment Observation System as a Portal for GEOSS Community." International Journal of Advanced Computer Science and Applications **EnviroGRIDS Special Issue on "Building a Regional Observation System in the Black Sea Catchment"**: 9-18.

Grus, L., Bregt, A., J., C. (2006). Workshop exploring Spatial Data Infrastructures report, Wageningen, the Netherlands.

Grus, L., Castelein, W., Crompvoets, J., Overduin, T., van Loenen, B., van Groenestijn, A., Rajabifard, A., Bregt, A. K. (2011). "An assessment view to evaluate whether Spatial Data Infrastructures meet their goals." Computers, Environment and Urban Systems **35**(3): 217-229.

Grus, L., Crompvoets, J., Bregt, A. K. (2007). "Multi-view SDI assessment framework." International Journal of Spatial Data Infrastructures Research **2**: 33-53.

GSDI (2011) The SDI Africa Discussion List. WWW document, <http://lists.gsdi.org/mailman/listinfo/sdi-africa>

GSDI (2012) SDI-Africa Newsletter. WWW document, <http://www.gsdi.org/newsletters - Africa>

GSDI (2013) Spatial Data Infrastructure links. WWW document, <http://www.gsdi.org/SDILinks>

Guigoz, Y., Giuliani, G., Nonguierma, A., Lehmann, A., Mlisa, A., Ray, N. (2016). "Spatial Data Infrastructures in Africa: a gap analysis." Journal of Environmental Informatics.

Henricksen, B. (2007). "UNSDI Compendium: A UNSDI Vision, Implementation Strategy, and Reference Architecture." Proceedings of United Nations Geographical Information Working Group (UNGIWG). <http://www.ungiwg.org/unsdi.htm>.(accessed October 15, 2010).

Henson, C., Pschorr, J. K., Sheth, A. P., Thirunarayan, K. (2009). SemSOS: Semantic sensor observation service. Collaborative Technologies and Systems, 2009. CTS'09. International Symposium on, IEEE.

Hewitt, C., Silva, V. B. S., Golding, N., Gao, R., Coelho, C. A. S., Duell, R., Pollock, J., Onogi, K., WMO Secretariat (2015). Managing risk with climate prediction products and services. WMO Bulletin. **64** (2).

IASON Consortium (2015) IASON Project Home Page. WWW document, <http://www.iason-fp7.eu>

Idrees, M. O., Saeidi, V., Yusuf, Y. A., Rashid, A. (2012). "Comparing Approaches and Strategies for NSDI Implementation Between the Developed and Developing World." Article under Review for the International Journal of Spatial Data Infrastructures Research.

IPCC (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A. (eds.)]. Geneva, Switzerland, IPCC: 104 pp.

IPCC (2013). "IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis.

Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker,

T.F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)].".

IPCC (2014a). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Geneva, Switzerland, IPCC: 151 pp.

IPCC (2014b). Summary for policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)].

IPCC (2014c). Summary for Policymakers. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

ISO (2014). ISO 19115-1:2014: Geographic information -- Metadata -- Part 1: Fundamentals.

ITU (2011). Measuring the Information Society, International Telecommunication Union.

ITU (2014) ITU statistics. WWW document, <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>

Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E. (2008). "Hole-filled SRTM for the globe Version 4." available from the CGIAR-CSI SRTM 90m Database (<http://srtm.csi.cgiar.org>).

Kakarekaa, S., Gromovb, S., Kukharchyka, T. (2004). "Estimation of heavy metal emission fluxes on the territory of the NIS." Atmospheric Environment **38**: 7101-7109.

Kalantari, M., Olfat, H., Rajabifard, A. (2010). "Automatic spatial metadata enrichment: reducing metadata creation burden through spatial folksonomies."

Karpouzoglou, T., Zulkafli, Z., Grainger, S., Dewulf, A., Buytaert, W., Hannah, D. M. (2016). "Environmental Virtual Observatories (EVOs): prospects for knowledge co-creation and resilience in the Information Age." Current Opinion in Environmental Sustainability **18**: 40-48.

Koch, H., Liersch, S., Hattermann, F. (2013). "Integrating water resources management in eco-hydrological modelling." Integrated Water Resources Management in a Changing World: Lessons Learnt and Innovative Perspectives: 13.

Kok, B. and van Loenen, B. (2005). "How to assess the success of National Spatial Data Infrastructures?" Computers, Environment and Urban Systems **29**(6): 699-717.

Kozma, R. and Wagner, D. (2005). Core indicators for monitoring and evaluation studies in ICTS for education. Monitoring and Evaluation of ICT in Education Projects.

Kurkjian, R. (2000). "Metal contamination in the Republic of Armenia." Environmental management **25**(5): 477-483.

Kurkjian, R., Dunlap, C., Flegal, A. R. (2002). "Lead isotope tracking of atmospheric response to post-industrial conditions in Yerevan, Armenia." Atmospheric Environment **36**(8): 1421-1429.

Lance, K. (2003). "Spatial Data Infrastructure in Africa: Spotting the elephant behind trees." GIS@development **7**(7): 35-41.

Laney, D. (2001). "3D data management: Controlling data volume, velocity and variety." META Group Research Note **6**: 70.

Lautenbacher, C. C. (2006). "The global earth observation system of systems: science serving society." Space Policy **22**(1): 8-11.

Lecca, G., Petitdidier, M., Hluchy, L., Ivanovic, M., Kussul, N., Ray, N., Thieron, V. (2011). "Grid computing technology for hydrological applications." Journal of Hydrology **403**(1): 186-199.

Lehmann, A., Giuliani, G., Mancosu, E., Abbaspour, K., Sözen, S., Gorgan, D., Beel, A., Ray, N. (2015). "Filling the gap between Earth observation and policy making in the Black Sea catchment with enviroGRIDS." Environmental Science & Policy **46**: 1-12.

Lehmann, A., Giuliani, G., Ray, N., Rahman, K., Abbaspour, K. C., Nativi, S., Craglia, M., Cripe, D., Quevauviller, P., Beniston, M. (2014). "Reviewing innovative Earth observation solutions for filling science-policy gaps in hydrology." Journal of Hydrology **518**: 267-277.

Longhorn, R. A. and Blakemore, M. (2007). Geographic information: value, pricing, production, and consumption, CRC Press.

Ma, J. (2006). "Managing metadata for digital projects." Library Collections, Acquisitions, and Technical Services **30**(1-2): 3-17.

Maguire, D. J. and Longley, P. A. (2005). "The emergence of geoportals and their role in spatial data infrastructures." Computers, Environment and Urban Systems **29**(1): 3-14.

Makanga, P. and Smit, J. (2010). "A Review of the status of Spatial Data Infrastructure Implementation in Africa." SACJ **45**: 18-25.

Mancosu, E., Gago Da Silva, A., Barbosa, A., De Bono, A., Ivanov, E., Lehmann, A., Fons, J. (2015). "Future land-use change scenarios for the Black Sea catchment." Environmental Science & Policy **46**: 26-36.

Manso Callejo, M.-Á. and Castelein, W. (2010). "Monitoring of Spatial Data Infrastructures."

Manyika, J. (2013). Open data: Unlocking innovation and performance with liquid information, McKinsey.

Masser, I. (1998). The first generation of national geographic information strategies. Proceedings of Selected Conference Papers of the Third Global Spatial Data Infrastructure Conference, Canberra, Australia.

Masser, I. (2005). GIS Worlds: Creating Spatial Data Infrastructures, ESRI Press.

Masser, I. (2006). "Multi-level Implementation of-SDIs: Emerging Trends and Key Strategic Issues." GIM INTERNATIONAL **20**(2): 31.

Masser, I., Rajabifard, A., Williamson, I. (2008). "Spatially enabling governments through SDI implementation." International Journal of Geographical Information Science **22**(1): 5-20.

Maué, P., Stasch, C., Athanasopoulos, G., Gerharz, L. (2011). "Geospatial standards for web-enabled environmental models." International Journal of Spatial Data Infrastructures Research (IJSDIR) **6**(2011): 145-167.

Mazzetti, P., Bonora, N., Cinnirella, S., Nativi, S. (2013). The EGIDA Methodology. Towards a sustainable GEOSS.

Mazzetti, P., Nativi, S., Santoro, M., Boldrini, E. (2014). Big Data challenges and solutions in building the Global Earth Observation System of Systems (GEOSS). EGU General Assembly 2014. **16**.

Mazzetti, P., Roncella, R., Nativi, S., Lacroix, P., Guigoz, Y., Ray, N., Giuliani, G., Gorgan, D., Bacu, V., Danut, M. (2015). Integration of data and computing infrastructures for Earth Science: an image mosaicking use-case. Earth Science Informatics.

Mihon, D., Colceriu, V., Bacu, V., Gorgan, D. (2013a). "Grid based processing of satellite images in Greenland platform." International Journal of Advanced Computer Science and Applications: 41-49.

Mihon, D., Colceriu, V., Bacu, V., Rodila, D., Gorgan, D., Allenbach, K., Giuliani, G. (2013b). "Ogc compliant services for remote sensing processing over the grid infrastructure." International Journal of Advanced Computer Science and Applications **3**(3): 32-40.

Mitchell, T. D. and Jones, P. D. (2005). "An improved method of constructing a database of monthly climate observations and associated high-resolution grids." International journal of climatology **25**(6): 693-712.

Molloy, J. C. (2011) "The open knowledge foundation: open data means better science." PLoS Biol **9**, e1001195.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3232214/pdf/pbio.1001195.pdf>
<http://www.plosbiology.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pbio.1001195&representation=PDF>
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3232214/>.

Monett, H. Á. and McLeod, P. (2013). Spatial Data Infrastructure (SDI) Manual for the Americas. Tenth United Nations Regional Cartographic Conference for the Americas, New-York, USA.

Moura, J. (2012) "Metadator - Création automatisée de métadonnées."
<http://www.portailsig.org/content/metadator-creation-automatisee-de-metadonnees>.

Myroshnychenko, V. (2010). "Gap Analysis Report Phase I and Phase II."

Myroshnychenko, V., Ray, N., Lehmann, A., Giuliani, G., Kideys, A., Weller, P., Teodor, D. (2015). "Environmental data gaps in Black Sea catchment countries: INSPIRE and GEOSS State of Play." Environmental Science & Policy **46**: 13-25.

Najar, C., Rajabifard, A., Williamson, I. P., Giger, C. (2007a). A framework for comparing Spatial Data Infrastructures on the basis of web services and metadata management: an Australian-Swiss case study.

Najar, C., Rajabifard, A., Williamson, I. P., Giger, C. (2007b). A framework for comparing Spatial Data Infrastructures on the basis of web services and metadata management: an Australian-Swiss case study.

NASA (2013) Safari2000. WWW document,

National Center for Atmospheric Research Staff (Eds) (2014) The Climate Data Guide: CRU TS3.21 Gridded precipitation and other meteorological variables since 1901. WWW document, <https://climatedataguide.ucar.edu/climate-data/cru-ts321-gridded-precipitation-and-other-meteorological-variables-1901>

National Centers for Environmental Prediction (2010) NCEP Climate Forecast System Reanalysis (CFSR). WWW document, <http://rda.ucar.edu/pub/cfsr.html>

Nativi, S. and Bigagli, L. (2009a). "Discovery, mediation, and access services for earth observation data." Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of **2**(4): 233-240.

Nativi, S., Bigagli, L., Mazzetti, P., Boldrini, E., Papeschi, F. (2009b). GI-cat: a mediation solution for building a clearinghouse catalog service. 2009 International Conference on Advanced Geographic Information Systems & Web Services.

Nativi, S., Bigagli, L., Mazzetti, P., Boldrini, E., Papeschi, F. (2009c). GI-Cat: A Mediation Solution for Building a Clearinghouse Catalog Service. Advanced Geographic Information Systems & Web Services, 2009. GEOWS '09. International Conference on.

Nativi, S., Craglia, M., Pearlman, J. (2012). "The brokering approach for multidisciplinary interoperability: a position paper." International Journal of Spatial Data Infrastructures Research **7**: 1-15.

Nativi, S., Craglia, M., Pearlman, J. (2013a). "Earth Science Infrastructures Interoperability: The Brokering Approach." Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of **6**(3): 1118-1129.

Nativi, S., Khalsa, S., Domenico, B., Craglia, M., Peralman, J., Mazzetti, P., Russ, R. (2011). The Brokering approach for earth science cyberinfrastructure.

Nativi, S., Mazzetti, P., Craglia, M., Pirrone, N. (2013b). "The GEOSS solution for enabling data interoperability and integrative research."

Nativi, S., Mazzetti, P., Santoro, M., Papeschi, F., Craglia, M., Ochiai, O. (2015). "Big Data challenges in building the Global Earth Observation System of Systems." Environmental Modelling & Software **68**: 1-26.

Nativi, S., Plag, H.-P., Mazzetti, P. (2013c). Towards a sustainable GEOSS - Some Results of the Egida Project.

Nebert, D. D. (2005). Developing Spatial Data Infrastructure: The SDI Cookbook.

Nebert, D. D. (2008) "Developing Spatial Data Infrastructures: The SDI Cookbook." http://www.gsdi docs.org/GSDIWiki/index.php/Main_Page.

Nedovic-Budic, Z., Feeney, M.-E. F., Rajabifard, A., Williamson, I. (2004). "Are SDIs serving the needs of local planning? Case study of Victoria, Australia and Illinois, USA." Computers, Environment and Urban Systems **28**(4): 329-351.

Neteler, M., Bowman, M. H., Landa, M., Metz, M. (2012). "GRASS GIS: A multi-purpose open source GIS." Environmental Modelling & Software **31**: 124-130.

NOAA (2015) NCDC data. WWW document,
<http://gis.ncdc.noaa.gov/map/viewer/>

Nogueras-Iso, J., Zarazaga-Soria, F. J., Muro-Medrano, P. R. (2005a). "Geographic information metadata for spatial data infrastructures." Resources, Interoperability and Information Retrieval.

Nogueras-Iso, J., Zarazaga-Soria, F. J., Muro-Medrano, P. R. (2005b). Geographic Information Metadata for Spatial Data Infrastructures: Resources, Interoperability and Information Retrieval, Springer.

Noort, M. (2011) "Marketing Earth Observation products and services, part 1." <http://www.eopower.eu/docs/marketing1.pdf>.

Noort, M. (2012). The GEONetCab approach to Capacity Building.

Noort, M. (2013) "Marketing Earth Observation products and services part2."
<http://www.eopower.eu/docs/marketing2.pdf>.

Noort, M. (2014). Methodological framework for impact assessment of earth observation application.

Noort, M. (2015a) "Capitalising on the Power of Earth Observation for Economic Development."
http://www.eopower.eu/downloads/EOPOWER_popular_final_report.pdf.

Noort, M. (2015b). EOPOWER impact assessment and EGIDA methodology connect.

Norris, J. (2015). Future trends in geospatial information management: the five to ten year vision (second edition).

Nushi, B., Van Loenen, B., Besemer, J., Crompvoets, J. (2012). "Multi-View SDI Assessment of Kosovo (2007-2010) - Developing a solid base to support SDI strategy development."

Nushi, B., Van Loenen, B., Crompvoets, J. (2015). "The STIG: A new SDI assessment method." International Journal of Spatial Data Infrastructures Research 10, 2015.

OECD (2008). Natural resources and pro-poor growth: the economics and politics. DAC guidelines and reference series.

OGC (2010) OGC vision. WWW document,
<http://www.opengeospatial.org/ogc/vision>

OGC (2013a) OGC Membership by Level & Region. WWW document,
<http://www.opengeospatial.org/ogc/members/report?sortBy=%27gov%27#Government-National>

OGC (2013b) The Open Geospatial Consortium. WWW document,
<http://www.opengeospatial.org/>

OGC (2015) Sensor Web Enablement (SWE) WWW document,
<http://www.opengeospatial.org/ogc/markets-technologies/swe>

Olfat, H., Kalantari, M., Rajabifard, A., Senot, H., Williamson, I. P. (2012). "Spatial metadata automation: A key to spatially enabling platform." International Journal of Spatial Data Infrastructures Research 7: 173-195.

OneGeology (2012) One Geology geoportal. WWW document,
<http://portal.onegeology.org/>

Open Geospatial Consortium (2004). Geospatial Portal Reference Architecture: 23.

Open Geospatial Consortium (2005). Web Feature Service Implementation Specification: 131.

Open Geospatial Consortium (2006a). OpenGIS Web Map Server Implementation Specification: 85.

Open Geospatial Consortium (2006b). Web Coverage Service (WCS) Implementation Specification, 143.

Open Geospatial Consortium (2007a). CUAHSI WaterML: 76.

Open Geospatial Consortium (2007b). OpenGIS Catalogue Services Specification: 218.

Open Geospatial Consortium (2007c). OpenGIS Web Processing Service: 87.

Open Geospatial Consortium (2007d). Sensor observation service.
Implementation Standard OGC.

Open Geospatial Consortium (2007e). Styled layer descriptor profile of the web map service implementation specification. Open Geospatial Consortium Inc. OGC. 1.

Open Geospatial Consortium (2009). "OGC KML." Open Geospatial Consortium (OGC), Standard OGC.

Open Geospatial Consortium (2015) "OGC E-learning tutorial module."
<http://cite.opengeospatial.org/pub/cite/files/edu/index.html>.

OSGeo (2015) OSGeo. WWW document, <http://www.osgeo.org/about>

Pearlman, J., Craglia, M., Bertrand, F., Nativi, S., Gaigalas, G., Dubois, G., Niemeyer, S., Fritz, S. (2011). EuroGEOSS: an Interdisciplinary Approach to Research and Applications for Forestry, Biodiversity and Drought, International Center for Remote Sensing of Environment (ICRSE).

Pereira, H. M., Ferrier, S., Walters, M., Geller, G., Jongman, R., Scholes, R., Bruford, M. W., Brummitt, N., Butchart, S., Cardoso, A. (2013). "Essential biodiversity variables." Science **339**(6117): 277-278.

Petrosyan, V., Orlova, A., Dunlap, C. E., Babayan, E., Farfel, M., von Braun, M. (2004). "Lead in residential soil and dust in a mining and smelting district in northern Armenia: a pilot study." Environmental research **94**(3): 297-308.

Plag, H.-P., Jules-Plag, S., Callaghan, C., McCallum, I. (2013). Linking science and technology communities to GEOSS. Towards a sustainable GEOSS.

Prunayre, F., Coudert, M., Lafage, B. (2013) "Keep your data and metadata synchronized with Talend Spatial."

PyWPS (2008) PyWPS software. WWW document,
<http://pywps.wald.intevation.org/>

Rajabifard, A. (2002). Diffusion for Regional Spatial Data Infrastructures: particular reference to Asia and the Pacific, University of Melbourne.

Rajabifard, A. (2016). "Integrating Geospatial information into sustainable development goals." Geospatial World.

Rajabifard, A., Chan, T. O., Williamson, I. P. (1999). The Nature of Regional Spatial Data Infrastructures.

Rajabifard, A. and WILLIAMSON, I. (2004). SDI development and capacity building.

Rajabifard, A. and Williamson, I. P. (2001a). Spatial data infrastructures: concept, SDI hierarchy and future directions.

Rajabifard, A. and Williamson, I. P. (2001b). Spatial Data Infrastructures: Concept, SDI Hierarchy and Future directions. Geomatics'80, Tehran, Iran.

Rajabifard, A., Williamson, I. P., Holland, P., Johnstone, G. (2000). From Local to Global SDI initiatives: a pyramid of building blocks. 4th Global Spatial Data Infrastructure Conference. Cape Town, South Africa.

RCMRD (2014) African Geodetic Reference Frame. WWW document, http://rcmrd.org/?page_id=4974

Rebelo, L.-M., Johnston, R., Hein, T., Weigelhofer, G., Dhaeyer, T., Kone, B., Cools, J. "Challenges to the integration of wetlands into IWRM: The case of the Inner Niger Delta (Mali) and the Lobau Floodplain (Austria)." Environmental Science & Policy(0).

Riecken, J., Bernard, L., Portele, C., Remke, A. (2003). North-rhine westphalia: Building a regional sdi in a cross-border environment/ad-hoc integration of sdis: Lessons learnt. 9th EC-GI\ & GIS Workshop ESDI.

Robinson, M. (2008). A History of Spatial Data Coordination.

Rockström, J., Steffen, W. L., Noone, K., Persson, Å., Chapin III, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J. (2009). "Planetary boundaries: exploring the safe operating space for humanity."

Rouholahnejad, E. (2013). Modeling the hydrology of the Black Sea Basin and assessing the impacts of climate change and land use change on water resources, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 21603, 2013.

Rouholahnejad, E., Abbaspour, K. C., Srinivasan, R., Bacu, V., Lehmann, A. (2014). "Water resources of the Black Sea Basin at high spatial and temporal resolution." Water Resources Research **50**(7): 5866-5885.

Rowley, J. E. (2007). "The wisdom hierarchy: representations of the DIKW hierarchy." Journal of information science.

Ryabinin, V. (2015). Interview: Vladimir Ryabinin, Executive Secretary of the Intergovernmental Oceanographic Commissions (IOC) of UNESCO. WMO Bulletin. **64** (2).

Ryan, B. (2016). "The benefits from open data are immense." Geospatial World.

Saghatelyan, A. (2007). Geochemical characteristic of heavy metal contents for Armenia's rivers. Geophysical Research Abstracts.

Saghatelyan, A. and Sahakyan, L. (2007). Heavy metals in system" soil-farm produce-organism" within the area of environmental impact of ore-mining production. Geophysical Research Abstracts.

Saghatelyan, A., Sahakyan, L., Belyaeva, O. (2013). "FOOD SAFETY ISSUES OF THE MINING IMPACT TERRITORIES." 13th SGEM GeoConference on Ecology, Economics, Education And Legislation 1(International Multidisciplinary Scientific GeoConference SGEM2013): 489-496 pp.

Santoro, M., Mazzetti, P., Nativi, S., Fugazza, C., Granell, C., Díaz, L. (2012). "Methodologies for augmented discovery of geospatial resources." Geographic Information Systems: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications **305**.

Schneider, A., Friedl, M. A., Potere, D. (2009). "A new map of global urban extent from MODIS satellite data." Environmental Research Letters **4**(4): 044003.

Schneider, A., Friedl, M. A., Potere, D. (2010). "Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'." Remote Sensing of Environment **114**(8): 1733-1746.

Schwabe, C. and Govender, S. (2009a). Getting geoinformation and spatial data infrastructure (SDI) to work for Africa: the rationale behind AESI-ALIGN. Proceedings of GSDI 11 world conference, 15-19 June, Rotterdam.

Schwabe, C. and Govender, S. (2009b). Getting geoinformation and spatial data infrastructure (SDI) to work for Africa: the rationale behind AESI-ALIGN. GSDI-11. Rotterdam, The Netherlands.

Sebake, M. D. and Coetzee, S. (2013). "Address Data Sharing: Organizational Motivators and Barriers and their Implications for the South African Spatial Data

Infrastructure." International Journal of Spatial Data Infrastructures Research **8**: 1-20.

SERVIR-Africa (2014) SERVIR-Africa overview. WWW document, <https://www.servirglobal.net/Africa.aspx>

Simonis, I. and Echterhoff, J. (2008). GEOSS and the Sensor Web. GEOSS Sensor Web Workshop Report.

Spiegel, E. (2015) "Why Does Big Data Need to Become Smart Data?" **2015**. <http://www.bloomberg.com/sponsor/siemens/digitalization-nation/?mvi=a4ff1c5ce0f24a049c9a26c4d64ec91a>.

Stein, A., Riley, J., Halberg, N. (2001). "Issues of scale for environmental indicators." Agriculture, Ecosystems & Environment **87**(2): 215-232.

Steudler, D., Rajabifard, A., Williamson, I. (2008). Evaluation and Performance Indicators to Assess Spatial Data Infrastructure Initiatives. A Multi-View Framework to Assess SDIs.

Su, L., Christensen, P., Liu, J. (2013). "Comparative study of water resource management and policies for ecosystems in China and Denmark." Journal of Environmental Informatics **21**(1): 72-83.

Taylor, L. and Broeders, D. (2015). "In the name of Development: Power, profit and the datafication of the global South." Geoforum **64**: 229-237.

The Economist (2009) Political Instability Index. WWW document, http://viewswire.eiu.com/site_info.asp?info_name=social_unrest_table&page=noads&rf=0

Transboundary Waters Assessment Programme (TWAP) (2015) Transboundary Waters Assessment Programme (TWAP) data portal. WWW document, <http://twapgeoportal.grid.unep.ch/app/>

Trilles, S., Díaz, L., Gil, J., Huerta, J. (2014). "Assisted Generation and Publication of Geospatial data and metadata." International Journal of Spatial Data Infrastructures Research **9**.

Tripathi, N. and Bhattacharya, S. (2004). "Integrating Indigenous Knowledge and GIS for Participatory Natural Resource Management: State of the Practice." The Electronic Journal of Information Systems in Developing Countries **17**(3): 1-13.

UNDP (2009). Capacity Development: a UNDP primer.

UNECA (2007a). Accelerating Africa's Development to meet the MDGs: Challenges and the Way Forward for Southern Africa. Lusaka, Zambia.

UNECA (2007b). Determination of Fundamental Datasets for Africa: Geoinformation in Socio-Economic Development. Addis Ababa, Ethiopia, United Nations Economic Commission for Africa: 51.

UNECA (2013) Committee on Development Information, Science and Technology WWW document, <http://www.uneca.org/codist>

UNEP (2005). Millennium Ecosystem Assessment. Living beyond our means - Natural Assets and Human Well-being.: 28p.

UNEP (2006). Africa Environmental Outlook 2: Our environment, our wealth.

UNEP (2012). Global Environment Outlook (GEO-5), summary for policy makers.

UNEP (2014) UNEP-live. WWW document, <http://uneplive.unep.org/>

UNEP/UNDP (2009). Mainstraming poverty-environment linkages into development planning: a handbook for practitioners.

unidata (2015) THREDDS Data Server (TDS). WWW document, <http://www.unidata.ucar.edu/software/thredds/current/tds/>

UNIGIS (2014) UNIGIS overview. WWW document, <http://www.unigis.net/>

UNISDR (2015a). Making Development Sustainable: the future of Disaster Risk Management, Global Assessment Report on Disaster Risk Reduction. Geneva, Switzerland.

UNISDR (2015b). The Pocket GAR 2015. Making Development Sustainable: The Future of Disaster Risk Management.

UNITAR (2014) Haiti Earthquake 2010: Remote Sensing based Building Damage Assessment Data. WWW document, <http://www.unitar.org/unosat/haiti-earthquake-2010-remote-sensing-based-building-damage-assessment-data>

United Nations Economic and Social Council (2013). Establishment and implementation of standards for the global geospatial information community. Committee of experts on global geospatial information management. Cambridge, UK.

Universities UK (2013) "Massive open online courses Higher education's digital moment?".
<http://www.universitiesuk.ac.uk/highereducation/Documents/2013/MassiveOpenOnlineCourses.pdf>.

University of Illinois (2014) Daily state of the cryosphere. WWW document, <http://arctic.atmos.uiuc.edu/cryosphere/>

USGS (2010) Shuttle Radar Topography Mission. WWW document, <http://srtm.usgs.gov/>

USGS and USAID (2013) FEWS NET Africa Data Portal. WWW document, <http://earlywarning.usgs.gov/fews/africa/index.php>

Vaccari, L., Craglia, M., Fugazza, C., Nativi, S., Santoro, M. (2012). "Integrative Research: The EuroGEOSS Experience." IEEE Journal of Selected Topics In Applied Earth Observations And Remote Sensing **5**(6): 1603-1611.

Vaccari, L., Shvaiko, P., Marchese, M. (2008). "An emergent semantics approach to semantic integration of geo-services and geo-metadata in Spatial Data Infrastructures." International Journal of Spatial Data Infrastructure Research.

van Loenen, B. (2006). Developing geographic information infrastructures: the role of information policies, IOS Press.

Van Loenen, B. and Kok, B. (2007). Survey on organisational SDI aspects.

van Loenen, B. and van Rij, E. (2008). Assessments of Spatial Data Infrastructures from an organisational perspective. A Multi-View Framework to Assess SDIs.

Van Orshoven, J. (2003). Spatial Data Infrastructures in Europe: State of Play Spring 2003.

Vandenbroucke, D. (2009). Inspire & NSDI State of Play: Methodology. D 1.1 - Report on Methodology.

Vandenbroucke, D. (2010a). INSPIRE & NSDI SoP. D4.1 - Summary report regarding the results of the European Assessment of 34 NSDI (first year) - September 2010.

Vandenbroucke, D. (2010b). Spatial Data Infrastructures in Europe: State of play spring 2010: 1-72.

Vandenbroucke, D. (2011). INSPIRE & NSDI SoP. D4.2 - Summary report regarding the results of the European Assessment of 34 NSDI (first year) - September 2011.

Vandenbroucke, D., Cromptvoets, J., Vancauwenberghe, G., Dessers, E., Van Orshoven, J. (2009). "A Network Perspective on Spatial Data Infrastructures: Application to the Sub-national SDI of Flanders (Belgium)." Transactions in GIS **13**(s1): 105-122.

Vicente-Serrano, S. M., Begueria, S., Gimeno, L., Eklundh, L., Giuliani, G., Weston, D., El Kenawy, A., López-Moreno, J. I., Nieto, R., Ayenew, T., Konte, D., Ardö, J., Pegram, G. G. S. (2012). "Challenges for drought mitigation in Africa: The potential use of geospatial data and drought information systems." Applied Geography **34**(0): 471-486.

Volfson, I. F., Farrakhov, E. G., Pronin, A. P., Beiseyev, O. B., Beiseyev, A. O., Bogdasarov, M. A., Oderova, A. V., Pechenkin, I. G., Khitrov, A. E., Pikhur, O. L. (2010). Medical Geology in Russia and the NIS, Springer.

Weinberger, D. (2010) "The Problem with the Data-Information-Knowledge-Wisdom Hierarchy." <https://hbr.org/2010/02/data-is-to-info-as-info-is-not>.

Wessels, B., Finn, R. L., Linde, P., Mazzetti, P., Nativi, S., Riley, S., Smallwood, R., Taylor, M. J., Tsoukala, V., Wadhwa, K. (2014). "Issues in the development of open access to research data." Prometheus **32**(1): 49-66.

White, M., Gambone, M., Yen, H., Daggupati, P., Bieger, K., Deb, D., Arnold, J. (2015). "Development of a Cropland Management Dataset to Support US Swat Assessments." JAWRA Journal of the American Water Resources Association.

WHO (2013) Global Health Observatory Data Repository. WWW document, <http://apps.who.int/gho/data/node.main>

Wikipedia (2015) Open data. WWW document, https://en.wikipedia.org/wiki/Open_data

Williamson, I. P., Rajabifard, A., Binns, A. (2006). "Challenges and issues for SDI development."

Williamson, I. P., Rajabifard, A., Enemark, S. (2003). Capacity building for SDIs.

Woldai, T. (2002). Geospatial Data Infrastructure: The Problem of Developing Metadata for Geoinformation in Africa. International conference of the African Association of Remote Sensing of the Environment AARSE : Geoinformation for sustainable development in Africa. Abuja, Nigeria.

World Bank (2011) "World Bank SDI report."
http://lgosmgb2.nottingham.ac.uk/elopeowiki/index.php/World_Bank_SDI_Report.

World Bank (2013) World Bank data. WWW document, <http://data.worldbank.org/topic/economy-and-growth>

World Bank (2014) The Worldwide Governance Indicators (WGI) projet. WWW document, <http://info.worldbank.org/governance/wgi/index.aspx#home>

Xia, X., Wu, Q., Mou, X., Lai, Y. (2014). "Potential impacts of climate change on the water quality of different water bodies." Journal of Environmental Informatics **201400263**.

Yang, W. and Yang, Z. (2014). "Evaluation of Sustainable Environmental Flows Based on the Valuation of Ecosystem Services: a Case Study for the Baiyangdian Wetland, China." Journal of Environmental Informatics **24(2)**: 90-100.

Yen, H., Ahmadi, M., White, M. J., Wang, X., Arnold, J. G. (2014). "C-SWAT: The soil and water assessment tool with consolidated input files in alleviating computational burden of recursive simulations." Computers & Geosciences **72**: 221-232.

Yue, P., Gong, J., Di, L. (2010). "Augmenting geospatial data provenance through metadata tracking in geospatial service chaining." Computers & Geosciences **36**(3): 270-281.

Yue, P., Gong, J., Di, L., He, L. (2012). "Automatic geospatial metadata generation for earth science virtual data products." Geoinformatica **16**(1): 1-29.