



Chapitre de livre

1997

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

## Requirements for a Hypermedia Electronic-Newspaper Environment Based on Agents: Working Paper

---

Morin, Jean-Henry

### How to cite

MORIN, Jean-Henry. Requirements for a Hypermedia Electronic-Newspaper Environment Based on Agents: Working Paper. In: Objects at large = Objets en liberté. Tschritzis, Dionysios (Ed.). Genève : Centre universitaire d'informatique, 1997. p. 177–193.

This publication URL: <https://archive-ouverte.unige.ch/unige:155512>

© The author(s). This work is licensed under a Creative Commons Attribution (CC BY)

<https://creativecommons.org/licenses/by/4.0>

# Requirements for a Hypermedia Electronic-Newspaper Environment Based on Agents<sup>1</sup>

*Working Paper*  
Jean-Henry Morin

## Abstract

The HyperNews project aims at developing a hypermedia electronic newspaper system, based on agent technology, that will help the readers reduce the time spent searching for relevant information. It will also enable information providers to commercialize hypermedia electronic news articles in terms similar to those of printed newspapers. Characteristic aspects of the system include enforcement of copyright control and payment on a "pay per use" basis, establishment of intra- and cross-newspaper hyperlinks according to the interests of the reader and the ability for information consumers to compose their own personalized hypermedia electronic newspaper through a Java-based World-Wide Web browser. In this paper, we identify and discuss the issues related to the development of an electronic newspaper environment based on agents and draw the corresponding requirements.

## 1 Introduction

The World-Wide Web [1] has paved the way to the global information age and what is about to become a major industry at the dawn of the third millennium: the electronic information industry. However, it is still in its infancy and information providers are reluctant to publish over the Internet for justifiable reasons. Namely, they require means to protect their copyright and intellectual property as well as to collect revenue from their electronic publishing activity. These are indispensable conditions for commercially viable electronic publishing services. From the point of view of information consumers such as managers and decision makers there is the wish to exploit the potential of this industry. They require means in order to be informed accurately with up to date information and presented in a way that best suits their needs [2].

Many newspapers and magazines are now on the Web [3][4]. For example, the Daily Telegraph went on-line November 1994 with the *Electronic Telegraph* [5]. It is a fully Web edited and based electronic information source. In Switzerland, *L'Hebdo* magazine went on-line September 1995 with the *Webdo* server [6] and has proven to be very successful according to access measurements. Such Web based services follow the "pull model of interaction" [7] where the user has to fetch the information by visiting all the desired Web sites. These type of Web based services suffer from major drawbacks when used by professionals. Namely they lack a unified, vendor independent and customized interface to electronic publishing of news articles. More-

---

1. This work was supported by the Swiss National Science Foundation with the SPP-ICS 1996-1998 project "HyperNews" (project number 5003-045333).

over, from the publishers point of view, issues such as copyright control and revenue collection are simply not addressed.

Recently, a new type of service emerged, offering a new model of interaction to their users: the push model [7] where the user is "fed" with information. PointCast [8] is an example of such services. Free information (supported by advertisement) are broadcast to the user's screen saver "while he is eating, on the phone or any time the computer is idle". Similarly, Netscape's Netcaster [9], formerly known as *Constellation*, is now part of Communicator (i.e., Netscape 4.0). The focus here is to automate off-line Web browsing, reception of information updates, automatic software upgrade and subscription to content channels as described above with PointCast. Netcaster relies on Marimba's Castanet [10] for distributing both applications and content to their users following a transmitter-channel-tuner metaphor. Such services offer the user limited personalization (i.e., through filters) and predefined channels. The focus is different from a commercially viable electronic newspaper system. It does not offer the information providers means either to protect intellectual property or to generate revenue from this electronic publishing activity. The idea is more on the side of a free (i.e., supported by advertisers) information news feed in a "one size fits all" way, hoping not to fall into a "shove" model.

The MIT Media Laboratory has developed an electronic newspaper system called Fishwrap [11][12]. It is used by the MIT community since 1993 and aims at combining individuals personal information needs (i.e., home town news) with the need to be informed on general issues. The Fishwrap system manages and takes into account the users profiles for personalizing their electronic newspaper. The profile can be updated through a feedback loop for fine tuning. The system however has a major limitation in that it is a centralized system to which information must be sent before dissemination and where the profiles are kept. Moreover, it does not consider the commercial and copyright needs of information providers.

The main idea of an electronic newspaper environment is that there is no central authority. In the real world newspaper publishing is characterized by the absence of a central authority. Information consumers and providers are inherently distributed and fully independent. Thus, new models and techniques need to be introduced in order to fulfil the needs of both information consumers and providers over an electronic information marketplace. The model of a hypermedia electronic newspaper system is rather straightforward. On the one hand, we have *information consumers* and on the other hand *information providers*. Between them, the *network* is used to achieve a commercial transaction in which published information units (i.e., *articles*) are offered by information providers against *payment* from the information consumers. Three major actors are identified: the information consumers, the information providers and the network. The first two are active, in the sense that they will initiate changes in the system, whereas the last one is best described as a passive actor upon which the system relies for communicating. We also identify the two agreed elements of the commercial transaction: the article and the payment as its counterpart.

The HyperNews (Hypermedia Electronic Newspaper) project, part of the MEDIA project [13], aims at developing an electronic newspaper environment based on agent technology. Such an environment should offer the publishers of newspapers, magazines and alike the means to commercialize electronically the information they hold under similar conditions as the printed

versions. Namely, by enforcing copyright control and revenue collection. Moreover, it should offer the information consumers the means to reduce the time spent in retrieving the information for which they express interest without being tied to a specific information provider. An electronic newspaper system should offer the information consumer an interface to an electronic newspaper system which is both *vendor independent* and customized according to his needs.

In this paper, we focus on the issues and requirements for a hypermedia electronic newspaper environment based on agents. In sections 2 and 3 we discuss the issues relevant to information consumers and information providers respectively and draw the requirements for an electronic newspaper environment. In section 4, we discuss the use of agent technology for the development of an electronic newspaper system. Finally, section 5 presents some concluding remarks, current achievements and future work.

## 2 Information Consumer's Requirements

The information consumer is the end user of the system, who needs a service providing him with high quality, up to date information. From this point of view, we discuss below the issues relevant to the design of an electronic newspaper system and identify the corresponding requirements.

### 2.1 User Interface

The World-Wide Web has become a very popular media and browsers are now available on most platforms. Moreover, the user interface of Web browsers is intuitive enough to be used even by children. Thus, it would be an advantage to enable the users to access in a seamless way both their electronic newspaper system and the World -Wide Web through the same interface. This leads to the following requirement:

- *information consumers should be able to use any Java enabled Web browser as an electronic newspaper reader.*

### 2.2 Free Choice of Information Providers

We assume that the information consumer is aware of the existing supply in a given information domain. The needed information sources, whether traditional or electronic, are chosen accurately, based on habits, experience and what is done in the company. Selecting information sources in an accurate way is not a random process but rather a conscious decision of the information consumer. This leads to the following requirement:

- *information consumers should be able to choose their information providers.*

### 2.3 Customized Information

In order to illustrate the notion of customized information, we consider how an information consumer accesses a printed newspaper or magazine. It is quite common to access the same information source for different reasons (e.g., finance, arts, sports) or to read only a specific section of a number of newspapers (e.g., finance, classifieds). Currently with paper based information sources the whole newspaper is bought although only a subset of it is used. The needed sections

are accessed almost directly and can then be browsed or scanned. Implicitly an information consumer has his mental representation of the structure of a given information source. This allows for direct access to the needed section of the information source depending on the context of work. For an electronic newspaper system, the *information profile* specifies an exhaustive set of information sources and information interests for each source. This leads to the following requirement:

- *information consumers should be able to specify their information interests for every information provider (information profile).*

## 2.4 Customizing Presentation

Finding and physically showing together related information from different information sources is time consuming and error prone. It can bring in-depth understanding of an information through different views. Cross information source *linking* is currently done “manually” by the reader. However, for an electronic newspaper system, this can be achieved by providing means for personal presentation of information. The goal is to offer both a consistent layout over time and physical proximity of related information independent from the source. For example, on the first page of an electronic newspaper, one could want the closing values of the Dow Jones, the Nikkei and the CAC40 from one provider and the weather forecast of the day from an other provider, etc. The *presentation profile* specifies how information is to be presented and can be thought of as the wire frame of the user’s electronic newspaper. It defines the general structure of an electronic newspaper as well as the content of the structure elements. We now have the following requirements:

- *information consumers should be able to specify the general structure of their electronic newspapers (presentation profile).*
- *information consumers should be able to specify the content of each structure element of their electronic newspapers (presentation profile).*

## 2.5 Context Dependent Electronic Newspapers

The situation in which information consumers hold concurrently different positions and responsibilities is rather common. It is also common to be involved in different projects, especially when the position in the hierarchy is high. Thus, depending on the role played by the information consumer, different information sources, sections, topics and presentations are needed. We call the elements defining the information needs and their presentation *information context*.

An information context is equivalent to an electronic newspaper. An information consumer can have multiple information contexts (i.e., electronic newspapers). An information context includes two elements: first, a list of information providers along with information interests for each source and second, the description of how the information should be presented to the consumer. The two parts comprising the information context are the *information profile* and *presentation profile*. For example, a user can have a work context, a private context, a context on a specific topic, etc. Thus we have the following requirement:

- *information consumers should be able to have multiple electronic newspapers (information contexts).*

In summary, the presentation profile together with the information profile are bound to an information context (i.e., an electronic newspaper). They provide the information consumer the way to specify where information is to come from, what information is of interest and how it is to be presented at retrieval time for each information context.

## 2.6 Active Information or Notification of Update Availability

Currently, newspapers, magazines and alike are published on a regular predefined time interval basis. The information contained in a given edition is considered static in the sense that it was valid when printed. Nothing however guarantees that it has not evolved by the time it is read. Any evolution of the information will only be traced in the next edition. For immediate information updates, other means and media have to be used, for example on-line information feeds such as Reuters or Bloomberg. In an electronic newspaper system, means are needed to offer the information consumer the possibility of being notified upon availability of information updates (i.e., active information). This follows a push model of interaction where the information consumer can request to be “fed” with information updates on issues for which he has expressed interest. This leads to the following requirement:

- *information consumers should be able to be notified when information updates are available on issues of interest (active information).*

## 2.7 Information Access and Information Evolution

There are situations where consumers are either not able or not willing to read their news for many reasons including vacation, illness, sudden high priority events, etc. In such cases, newspapers will be piled up for later reading. In order to catch up, the information consumer might jump to the most recent edition or just start with the first unread edition. In any case, for tracing either backward or forward the evolution of a subject the reader will have to browse through all the editions building his own information evolution with risks of missing something important. Such a task is time consuming, error prone and cumbersome. An electronic newspaper system can very easily provide links between each step of the life-cycle of an information (events happen, evolve and terminate) through hypertext links since information are related one to another. For example, when writing an article, a columnist can easily include a “follow up on” link to relate it to a previous article. Moreover, hyperlinks can also be used to refer to archives, other encyclopedic information or related material both on an internal (i.e., bound to the same information provider) and external basis (i.e., referring to material of a different information provider). This leads to the following requirements:

- *information consumers should be able to access easily any point of the historical evolution of an article,*
- *information consumers should be able to access directly any referenced material both inside and outside the domain of the information provider.*

## 2.8 Anonymity

In the traditional newspaper publishing industry, the information consumer has the choice of revealing or hiding his identity to the publisher. By revealing his identity, through, for example, a subscription, the reader can enjoy several benefits such as, lower price, frequent buyer bonus, personalized information, special issues, etc. However, by buying the edition from a newspaper seller the information consumer retains his full anonymity. Accessing information in an anonymous way is very important since information providers can easily build a reader's profile based on what is read, often leading to unwanted solicitations and targeted advertisement. Besides, the piece of information read is itself information which the reader might not want to divulge. For example, knowing that a well known businessman is suddenly interested in some small venture company could lead to unpredictable speculations or even have consequences on the stocks of that particular company. Furthermore, anonymity and privacy can be crucial when considering new job opportunities or even desirable when reading "sensitive" articles.

In an electronic newspaper system the information consumer should retain the possibility to reveal or hide his identity. The electronic information consumer should be able to pay for information without having to reveal in any direct or indirect way his identity to the information provider. The problem however is that unlike real cash, the provider can easily find out who the reader is and what is read from the information included in the electronic payment [14]. A possible solution is to either introduce an intermediary acting as a trusted third party between the consumer and the provider, or to use an electronic cash system based on anonymous money. This leads to the following requirement:

- *information consumers should be able to retain full anonymity when desired.*

## 2.9 Information Payment, Usage and Access

In the traditional newspaper publishing industry, the smallest information unit that can be put on the market is the *issue* and its price is fixed accordingly. The price does not vary if an issue holds more or less information than a previous one. It is also not possible to either negotiate a discount when claiming that only a sub section of the issue is relevant to the reader, nor is it possible to ask a refund if the issue has not been read at all.

In an electronic newspaper system however, the granularity of the marketable information unit can be brought down to the level of the article. This allows for a more flexible pricing policy where each article has a price and can be sold independently from an other. Combined with a security scheme, enforcing payment upon access to the content of an article, the information consumer is provided with a *pay per use* system. This offers the major advantage of allowing information consumers to freely hold and distribute articles without infringing any copyright or intellectual property laws.

Many electronic commerce schemes and protocols exist. They can be classified according to how they cope with issues such as support for micro-transactions, security, anonymity, off-line payment, dispute handling, pricing policies etc [15]. In order to offer maximum flexibility to the information consumers, various payment methods should be accommodated by the system by providing a common abstraction layer to electronic commerce. Work has been done in this

direction by SunSoft with the Java Electronic Commerce Framework (JECF)[16][17][18]. Thus, we have the following requirements:

- *information consumers should be able to pay for articles on a usage basis,*
- *information consumers should be able to hold and distribute freely articles,*
- *information consumers should be able to access subsequently articles they have already paid for,*
- *electronic articles should be self contained,*
- *an electronic newspaper system should accommodated existing electronic commerce systems.*

## 2.10 Off-line Activity and Information Consumer Mobility

Situations in which people require high geographical flexibility and mobility are common nowadays. Consequently, they often have access to many different hosts like for example a workstation at the office, a desktop at home, a laptop when traveling and a palmtop on vacation. This raises two issues. First, off-line activity when a network connection is not available or possible. Second, synchronizing or transferring environments between different hosts. Concerning off-line activity, reading electronic newspaper articles should be possible without any network connection. Of course, in order to acquire the data one must connect to a network but after that, the system should offer means for reading the downloaded information off-line, for example while going to work, on vacation or even from a CD-ROM. This in turn raises the problem of off-line payment systems. Thus, mechanisms for off-line payment must be considered for an electronic newspaper system. For example, smart card technology could offer a clean solution to such off-line problems. This aspect is enforced by the recent release of Java Cards [19] which provides an API for programming smart cards. This leads to the following requirement:

- *information consumers should be able to access articles off-line.*

Concerning the issue of environment mobility, such situations require means to transfer easily a part or the whole environment between the different hosts. Therefore, we have the following requirement:

- *information consumers should be able to move their electronic newspaper environment between hosts easily.*

## 2.11 The Information Consumer as an Information Provider

This last issue introduces in some way the next section discussing the issues and requirements of the information provider. In fact, we consider here the information consumer as an information provider. To illustrate this issue, we consider the following two examples where a user receives an article: (i) the user decides to forward it to a friend together with some free comments. (ii) the user decides to forward it to a client together with some paying comments. In the first case, the comments are free. However, in the second case, the user has added value to the original article and will want to collect revenue for this. Thus in both cases a new article is created, which contains the original article and the comments with its corresponding price.

The idea here is that an information consumer can become an information provider of his own added value and a reseller of other information provider's material without infringing any copyright or intellectual property law. This leads to a recursive definition of an article where an article can hold its own added value and other articles provided that each article is able to protect and release its content as discussed in the next section. Thus we have the following requirement:

- *information consumers should be able to publish new articles embedding articles of other providers together with their own added value.*

### 3 Information Provider's Requirements

The information provider wants to commercialize electronically the information that he holds in a way that will secure his intellectual property rights. From this point of view, we discuss a number of issues related to the transition from the printed to the electronic media and identify the corresponding requirements.

#### 3.1 Copyright Management and Revenue Collection

Copyright management is a crucial issue for publishers [20][21][22]. A legal base exists for printed material in order to protect intellectual property and rights of author. These laws enforce correct usage of copyrighted material. However, very little has been done yet for electronic documents and data. The problems raised by copyright and rights of author in the electronic world represent a blocking factor for the publishing industry. Information providers need to be assured that electronic information items will be paid for, used as they should be (i.e. view only, authorized copying, distribution, etc.), retain copyright information both as a whole or part, etc.

As far as articles are concerned, a scheme is needed to protect their content from being hacked and tampered with while enforcing payment to the copyright holder. This can be achieved by encryption and digital signature techniques. The scheme used does not need to be 100% secure but secure enough to prevent generic attacks or failures, for example, the breaking of one article key leading to the access of all the articles. Moreover, due to the limited life time and the low cost of the information content, breaking a key should involve sufficient effort to be either too costly or simply too long. This is inherent to the document's type which is a newspaper or a magazine article (i.e., such information almost become public domain after a day or so). The access to the article content (i.e., decryption) should not occur before payment has been made. This means that the client has been debited and the copyright holder credited with the right amount of money. Subsequent access to an article that has already been paid must be possible by providing a valid proof of purchase. However, only the user that has paid for the article should be granted subsequent access right through the proof of purchase.

It should be noted that all these security aspects should not prevent copying, holding and free distribution of the articles since they embody the necessary mechanisms for protecting, paying and releasing their content. Thus, when one finds that an article could be of interest to a friend, it can be directly copied and forwarded to him since any access to the content will generate a new payment to the copyright holder. This forwarding action must not involve the information provider.

Accessing the article content off-line without being tied to a network is an important issue. However it requires the use of tamper proof external devices such as secure smart card readers for off-line electronic commerce in order to undertake the role of a credit institution. This raises however the problem of correctly dispatching the revenue collected off-line to the right information providers. This issue will need careful attention when considering off-line usage.

Currently Revenue collection for electronic material does not support the pay per use scheme. If one is interested in only one section of a newspaper, it will not be possible to pay only for that specific section but one will have to buy the whole issue or pay a full subscription price in order to be granted access rights. Whereas in the electronic world having every piece of information attached with a price associated to authorized actions results in a highly flexible *pay per use* scheme.

It is noteworthy to mention that issues considering whether or not information is to be subsidized by advertisement is a policy decision of the information provider. Free information can be considered as paying information with price zero. Free does not mean loosing copyright and intellectual property. Thus these issues need careful attention even in a “free information” digital world trend.

We identify the following requirements:

- *information providers should be able to collect revenue from their electronic publishing activity,*
- *information providers should be assured that their copyrights and intellectual property are protected,*
- *information providers should be assured that article content access will generate payment to the copyright holder.*
- *information providers should be assured that article content will be protected and secured against illegal access and use.*

### 3.2 Information Classification

Each publication has its own way of structuring and classifying its content. For example, a newspaper has a front page with headlines and is composed of columns and sections (e.g., politics, finance, sports, arts, classifieds, etc.), sub-sections (e.g., international, national, local, etc.). Thus, the classification scheme used by an information provider for a given publication can be easily derived from the structure of that publication. From there on, and in the scope on an electronic newspaper system, an electronic article belonging to a given *section* can be assigned a classification identifier by the publisher. This is done at the time the article is published and is similar to inserting the article in a given section of a hard copy newspaper. Moreover, due to the electronic nature of the article and the possibility to dynamically process them, it is possible to assign the article alternate classification identifiers together with “weights” reflecting the level of relevance of that particular article to the alternate classifications. For example, an article can be classified as national politics and also be relevant to social issues, etc. Publications are different and their structure depends on many parameters such as field, theme, periodicity, target audience, etc. It is important to maintain this diversity of information sources and not to attempt

to unify information classification globally. A global unified classification scheme of information would result in a difficult to manage structure, which would in any case end up being an unsatisfactory common denominator. Thus, it is capital that each electronic publication keeps its own classification scheme from which the users will be able to make accurate and fine grained decisions about their interests. There is a problem however when the classification scheme changes due to reorganization of the publication. In the printed media, changes are simply reflected in the new editions together with comments. In the electronic world, careful attention needs to be given to this issue in order to ensure synchronization of the classification scheme used by the information consumer with the *official* one of the information provider. This leads to the following requirement:

- *information providers should be able to have their own classification schemes.*

### 3.3 Transition from Print and Web Publishing to Electronic Publishing

Today, most if not all publishers use computer assistance in their publishing process. It offers two major advantages. First, as opposed to the printed media, the marginal cost of publishing electronic documents once the initial copy exists, is close to nothing. Additional copies can be produced in a cost effective way. The second advantage is that almost all the material is already in some sort of digital form. Thus, when a new electronic publishing standard appears, it only takes the time to build the corresponding converters for using it. For example, when the Web appeared, it was relatively easy for the publishing industry to generate html files from the different electronic formats they already used. This leads to the following requirement:

- *information providers should be able to easily transform existing electronic material into self contained electronic newspaper articles.*

### 3.4 Hypermedia Electronic Publishing

Hypermedia electronic publishing is the result of combining hypertext techniques for information navigation with multimedia components. Hypertext systems offer a navigation mechanism allowing users to reach a target object through the referencing link also called an *anchor*. Thus, an electronic newspaper system must offer the information provider means to relate (i.e., link) articles to one another for historical evolution and access to archives as well as access to other resources. This issue is closely related to of the information consumer's requirements on information access and evolution discussed in section 2.7. There are two different types of links. First, the internal link for which the target is bound to the same information provider. Second, the external link for which the target can be anything outside the domain of the information provider (e.g., a World-Wide Web address). This leads to the following requirements:

- *information providers should be able to establish historical evolution links between articles,*
- *information providers should be able to insert links in articles to reference other material inside and outside their own domain.*

### 3.5 Using Standards for Information Composition and Rendering

In the present context, the use of de facto standards such as the html formatting language [23] and the http protocol [24] offer many advantages. Both of them have become widely available standards. Html allows the use of any html editor for formatting electronic newspaper articles. This prevents from building yet another proprietary standard for composing and rendering information. Moreover, combining a commercially viable electronic newspaper system and the Internet protocols in a transparent way will benefit to the users. Such an approach is a key factor in user acceptance. This leads to the following requirement:

- *information providers should rely on a standard widely accepted formatting language for electronic newspaper article composition and rendering.*

### 3.6 Editorial Process

Electronic publishing raises a fundamental question concerning the editorial process. Namely, what is an *edition* or an *issue* in the electronic context? In the traditional newspaper industry when the chief editor gives his final consent he triggers the publishing process of an edition. This is done on a day by day basis for a daily newspaper. Thus, the editorial process for the printed media is a discreet notion based on periodical issues. Whereas in an electronic newspaper infrastructure, one could say that an edition is triggered whenever a new information is made available (i.e., any time). Thus, the editorial process becomes a continuous notion. Although this issue is in essence a policy decision of the information provider, we need to satisfy the following requirement:

- *information providers should be able to reflect the editorial process in the electronic newspaper articles.*

### 3.7 Policy and Marketing Issues

An electronic newspaper system must offer full flexibility to the information provider to accommodate any policy issue the information provider chooses. Such issues cover but are not limited to brandware, pricing, payment methods, layout, frequent buyer programs, subscription, advertisement, direct marketing etc. For example, the industry is very keen on logos, presentation styles, layouts and special services they offer. They are often protected by copyright and reflect in an unambiguous way a brand, a product or even a life-style. From a marketing point of view, this is a very important issue. Thus we have the following requirement:

- *information providers should be able to accommodate and customize policy dependent issues according to their needs.*

### 3.8 Service Availability

High availability of the information providers is also of prime importance. The expectations facing the electronic publishing industry require a twenty four hours per day availability of the system due to the inherent global nature of the used media. Bad servicing over time will lead to a loss of trust from the clients which will consider other information sources offering better quality of service. This leads to the following requirement:

- *information providers should strive to offer high service availability to information consumers.*

## 4 An Agent Based Electronic Newspaper System

The usage of agent technology for an electronic newspaper system is a choice. However, there are a number of issues relevant to the electronic publishing field, for which agent technology, also called mobile object systems [25] [26], can provide solutions.

### 4.1 Portability and Architecture Independence

From a network point of view, an electronic newspaper system is likely to span over a great number of interconnected heterogeneous hosts for both information providers and consumers. Besides, it is highly probable that these hosts will not be running the same operating system. To this respect, agent technology provides a very clear abstraction of distributed heterogeneous systems by running on each participating node of the network a common agent execution environment also called an agent execution platform. Thus, all the collaborating agent execution platforms together form the agent system as a whole and share common characteristics.

From the point of view of the underlying technology over which an agent system relies, issues such as architecture independence (i.e., portability) and network awareness are of prime importance. The Java [27][28] language has become a de facto standard for network aware software development. Moreover, the generated Java byte code can be interpreted on any Java Virtual Machine [29] and is thus fully portable due to its wide availability. The Java language per se does not support mobile computations in the sense of an agent system. However, it may be extended to support the agent paradigm as in [30][31][32][33][34]. Thus, we make the basic assumption that a Java virtual machine exists, or can be easily downloaded and installed. This assumption is also supported by the recent release of the JavaOS [35] operating environment, the JavaStation [36] and the picoJava [37] microprocessor.

One of the most important requirements for an agent based electronic newspaper system concerns the agent platform. That is, in order to participate in the agent based commercial exchange of hypernews articles against its counterpart (i.e., electronic payment) both the information providers and the information consumers must agree on a common underlying agent execution platform. Portability issues must be considered in the choice of this technology in order to avoid as much as possible high hardware architecture dependencies. The Java Virtual Machine could be considered to this respect. Issues surrounding interoperation of heterogeneous agent systems fall out of the scope of this paper. It represents an active research area [38] and could be addressed in the future separately in order to achieve true interoperation of agents and thus to drop this requirement about the same execution environment.

### 4.2 Security

Security is a crucial issue in the scope of an electronic newspaper system since the target is to distribute commercially the electronic information and thus to protect as much as possible intellectual property, copyright of content and revenue collection. From the point of view of the agent

execution platform, special attention must be given to three aspects. First, authentication of the communicating parties. This means that when an agent migrates between two platforms there must be a way to ensure that the agent really comes from the announced source. Second, the platform must be assured that the received agent has not been tampered with while migrating. Third, a minimal level of security is required to prevent malicious agents to enter or at least to execute on the agent platform. This involves some way of detecting such agents and defining what malicious actions cover and what to do with them.

Most of the issues regarding security and mobile object systems are described in detail in [40]. They are classified in five categories: transfer security, authentication and authorization, host system security, computational environment security and mobile object system security. However, from the point of view of a user of agent technology (i.e., agent programmer), this is not sufficient. Granting access and execution rights to a foreign agent requires more than knowing that the agent is authenticated and not malicious. For example, in an electronic newspaper system, it is important for the user not to be bothered by unwanted solicitations even if they are not malicious and come from identified sources. Thus, the platform must provide a *hook* to a second level of security which can be customized according to the needs of the application (e.g., an access control system). Such a hook appears to be fundamental in the scope of electronic newspaper system in order to prevent flooding of unwanted information, advertisement agents and other solicitations. Similarly, after considering incoming agents, attention should be given to outgoing agents. Since the agent environment will be hosting foreign agents, it is capital to constrain any migration or communication request to be granted clearance for leaving or authorization to communicate with the outer world. Migration and communication should not be possible without the information consumer being explicitly aware of it.

The raw material of an electronic newspaper system is the article. By encapsulating both the article content and the necessary primitives to protect, pay and release this content inside an autonomous entity (i.e., an article agent) is an advantage that agent technology provides. Thus, the article agents are fully responsible for their own security. They control access to their content, generate payment and enable them to be freely copied and distributed. The free copying and distribution of hypernews articles (i.e., agents) is a capital requirement for the targeted electronic newspaper system. Since payment must occur prior to the access of the content, there is no problem of misuse or copyright infringement.

### 4.3 Persistency

Persistency is an important issue in the field of mobile agents [39]. In the scope of an electronic newspaper system, we consider this issue from two viewpoints: the agent platform and the agents themselves. From the platform point of view, it should be possible to freeze the agent environment as a whole any time (this also involves the freezing of the currently executing threads) and to restore it either on the same machine or after migration on a new host, resuming execution at the exact previously frozen state. This could be found useful in case of *desktop to laptop or palmtop* transfer for mobile users.

Concerning the agent instances, it should be possible to save agents to secondary memory and reload them subsequently. Since the agents are executing, two cases arise. In the first case,

the currently executing threads of the agents must be frozen immediately and the agent saved. In the second case, the agent is instructed to complete current execution of critical sections prior to the freezing operation. This type of operation requires a way to mark the critical sections of an agent. The freezing operation on an agent should also be possible without saving it to secondary storage. This is, putting it in idle state until some condition is verified. This could be very useful for scheduling purposes and to save system resources.

#### 4.4 Migration

An electronic newspaper system requires that the participating actors exchange payment against the provided service (i.e., the article). It is also most likely that such an environment will evolve over time, requiring to update software. Mobility and flexibility of persons have become common constraints of every day life. Therefore moving around at such a fast pace require means for freeing computer environments from being tied to specific hosts or geographical locations by allowing them to be easily transferred from host to host without losing any state while migrating. The issues raised here cover software upgrade, mobility of both the running software and environment while preserving code and state<sup>1</sup> during migration. Agent technology should accounts for these issues together with the inherent security issues.

The agent system should offer means for sending agents (i.e., migration) on either a one to one, or a one to many basis. In both cases it should be possible to specify whether a response is expected or not. When sending agents on a one to many basis, it should also be possible to specify that the agent is to be issued (i.e., copied) as many times as the number of destinations or that it should have a unique copy travelling from destination to destination until they have all been reached. When an agent migrates, many things can happen. The issue here is to provide the mechanism to instruct an agent what to do in situations such as unreachable destination, denied access, normal or abnormal termination, security violation, etc. For example, to instruct the agent what to do upon completion when it has migrated, a final action could be self destruct, return to original location, stay and sleep, execute, save, etc.

#### 4.5 Practical Issues for an Agent System

Ideally, the agent system should be totally transparent to the user especially in the scope of an electronic newspaper system which is aimed at non computer experts. The platform should be able to install itself almost automatically from scratch and remove itself likewise. Thus, a new user could download the environment through a plain web browser and simply launch it to use the electronic newspaper system without even noticing the existence of an agent environment.

Concerning the platform configuration at start up time, it should be able, upon request, to automatically configure, in a kind of *plug 'n play* way, most of the system resources like printers, databases, etc. Different levels of automatic configuration could be introduced depending on the level of the user (e.g., fully automatic, semi-automatic, personalized for experts). The reason is that no matter where a system is, such system resources are most likely to exist and to be needed

1. This includes preserving the state of the currently executing threads.

by the user and the system. This way, when a platform is launched at a given location, those common resources can be automatically found and configured to meet the users needs.

There is an important issue concerning agent embedding and duplication. For example, one receives an electronic newspaper article agent and decides to make comments about its content. The issue is how to achieve this from an agent point of view. The ideal situation would be to create a new container agent holding the comments and embedding the initial untouched article agent. Conceptually, this is a very clean approach in the sense that it reflects the different levels of added value. Thus, for example, an information provider offering a news digest service can easily compose an agent on a given topic, holding article agents from other providers related to the given topic. The added value of the digest service is reflected by the containing agent which can have a price while preserving the embedded articles of other providers. This approach raises however the issue on how an agent is to be duplicated. In the first example of an article and its comment, there are three ways one might want to forward the information: with the comment, without the comment or only the comment with a reference to the article.

Finally, there are a number of "housekeeping" issues. Having an environment hosting and executing software agents requires means to monitor them. For example, retrieving the state of a foreign agent which is faulty requires to remove it. An agent system must offer to the runtime environment the means to create dynamically new agents and to tailor them according to the needs of a running application. For example an electronic newspaper system must be able to create an article agent on the fly with specific properties. Finally from the point of view of the programmer of an agent based system a clear interface (API) to the agent system is essential.

## 5 Conclusion

In this paper we have presented the requirements for a hypermedia electronic newspaper system. We have considered and discussed a series of issues relevant to information consumers and providers allowing to draw a set of requirements for an electronic newspaper system. We have then considered and discussed the requirements related to the use of agent technology for implementing an electronic newspaper system.

Based on the requirements described in this paper, we have designed a scheme for the commercial distribution of electronic documents enforcing copyright control and payment at the time the article content is accessed for reading [41]. Furthermore, we have designed an architecture for an electronic newspaper system based on agents called HyperNews [42]. Currently, we are working on the implementation of the HyperNews prototype which will be installed by the end of 1997 at our industrial partner's site "L'Hebdo" for evaluation in a commercial environment.

## References

- [1] T. Berners-Lee, R. Cailliau, A. Luotonen, H. Frystyk Nielsen and A. Secret, "The World Wide Web", *Communications of ACM*, Vol. 37, No 8, August 1994, pp. 76-82

- [2] Jean-Henry Morin and Dimitri Konstantas, "Towards Hypermedia Electronic Publishing", Proceedings of second IASTED/ISMM International Conference on Distributed Multimedia Systems and Applications, Stanford, California, August 7-9 1995.
- [3] Webdo, Les Journeaux sur le Web, [http://www.webdo.ch/webactu/webactu\\_presse.html](http://www.webdo.ch/webactu/webactu_presse.html)
- [4] University of Waterloo, Canada, *Directory of Electronic Journals and Magazines*, <http://www.lib.uwaterloo.ca/journals/>
- [5] The Electronic Telegraph, Web site of the Daily Telegraph, <http://www.telegraph.co.uk/>
- [6] Webdo, Web site of L'Hebdo, <http://www.webdo.ch/>
- [7] The Economist Newspaper Limited, "Pushmepullyou", November 16, 1996
- [8] PointCast Network, <http://pioneer.pointcast.com/>
- [9] Netscape Communications Corporation, *Netscape NetCaster, Dynamic Information-delivery Tool for Intranet and Internet Users*, [http://home.netscape.com/comprod/products/communicator/netcaster\\_frameset.html](http://home.netscape.com/comprod/products/communicator/netcaster_frameset.html)
- [10] Marimba Inc, "Castanet", <http://www.marimba.com/datsheets/castanet-ds.html>
- [11] Pascal R. Chesnais, Matthew J. Mueklo, and Jonathan A. Sheena, "The Fishwrap Personalized News System", *Proceedings of the 1995 2nd International Workshop on Community Networking*, pp 275-282, Princeton, NJ, June, 1995.
- [12] Fishwrap Documentation center, <http://fishwrap-docs.www.media.mit.edu/docs/>
- [13] Dimitri Konstantas, Jean-Henry Morin and Jan Vitek, "MEDIA : A Platform for the Commercialization of Electronic Documents", in *Object Applications*, Ed. Denis Tschritzis, CUI, University of Geneva, 1996.
- [14] Ralf Hauser, Gene Tsudik, "On Shopping Incognito", in *Proceedings of The Second USENIX Workshop on Electronic Commerce*, pp 251-257, Oakland, California, November 18-21 1996.
- [15] P. Janson, M. Waidner, "Electronic Payment Systems". Final SEMPER Activity Paper 211ZR017, Vers. 7, SEMPER / IBM Zürich Research Lab, January 1996.
- [16] Java Soft, "Java Commerce Home Page", <http://www.javasoft.com/products/commerce/>
- [17] Java Soft, "White Paper: The Java Electronic Commerce Framework (JECF)", [http://www.javasoft.com/products/commerce/doc.white\\_paper.html](http://www.javasoft.com/products/commerce/doc.white_paper.html)
- [18] Theodore Goldstein, "The Gateway Security Model in the Java Electronic Commerce Framework", Sun Microsystems Inc, 1996.
- [19] Sun Microsystems Inc, "Java Card: API Specification", Version 1.0, October 1996.
- [20] J. Ebersole, "Protecting intellectual property rights on the information superhighways", *International Publishers Association Bulletin*, Volume X, No. 3, 1994, pp. 3-4.
- [21] The Cyberspace Law Institute, <http://www.cli.org/>
- [22] Pamela Samuelson, "The Copyright Grab", *Wired Magazine*, Issue 4.01, January 1996
- [23] The National Center for Supercomputing Applications, "A Beginner's Guide to HTML", <http://www.nc-sa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>
- [24] T. Berners-Lee, R. Fielding and H. Frystyk, "Hypertext Transfer Protocol-HTTP/1.0", RFC 1945, May 1996.
- [25] J. Vitek and C. Tschudin (Eds.), "Mobile Object Systems: Towards the Programmable Internet", Springer-Verlag, LNCS 1222, 1997.
- [26] K. Rothermel and R. Popescu-Zeletin (Eds.), "Mobile Agents", Springer-Verlag, LNCS 1219, 1997.
- [27] James Gosling, Bill Joy and Guy Steel, "The Java Language Specification", The Java Series, Addison-Wesley, 1997.
- [28] Ken Arnold and James Gosling, "The Java Programming Language", The Java Series, Addison-Wesley, 1996.
- [29] Tim Lindholm and Frank Yellin, "The Java Virtual Machine Specification", The Java Series, Addison-Wesley, 1997.
- [30] J. White, "Mobile Agents", General Magic White Paper, General Magic, Inc. 1996.
- [31] General Magic, "Agent Technology: Odyssey", General Magic, Inc., <http://www.genmagic.com/agents/>

- [32] IBM Research, "IBM Aglets Workbench - Home Page: Programming Mobile Agents in Java", <http://www.trl.ibm.co.jp/aglets/>
- [33] M. Ranganathan, A. Acharya, S. Sharma and J. Saltz, "Network-Aware Mobile Programs", Technical Report, University of Maryland, CS-TR-3659, June 1996.
- [34] Markus Strasser, Joachim Baumann and Fritz Hohl, "Mole - A Java Based Mobile Agent System", *Second ECOOP Workshop on Mobile Object Systems*, University of Linz, July 8-9, 1996.
- [35] P. Madany, "JavaOS: A Standalone Java Environment", White Paper, Sun Microsystems, Inc., 1997, <http://java.sun.com/products/javaos/javaos.white.html>
- [36] Sun Microsystems, "JavaStation - An Overview", White Paper, Sun Microsystems, Inc., 1996, [http://www.sun.com/javastation/whitepapers/javastation/javast\\_ch1.html](http://www.sun.com/javastation/whitepapers/javastation/javast_ch1.html)
- [37] Sun Microsystems, "picojava I Microprocessor Core Architecture", White Paper, Sun Microsystems, Inc., 1996, <http://www.sun.com/sparc/whitepapers/wpr-0014-01/>
- [38] P. Dömel, A. Lingnau, O. Drobnik, "Mobile Agent Interaction in Heterogeneous Environments", in *Mobile Agents, Proceedings of First International Workshop, MA'97*, K. Rothermel and R. Popescu-Zeletin (Eds.), Springer-Verlag, LNCS 1219, 1997.
- [39] M. Mira da Silva and A. Rodrigues da Silva, "Insisting on Persistent Mobile Agent Systems", in *Mobile Agents, Proceedings of First International Workshop, MA'97*, K. Rothermel and R. Popescu-Zeletin (Eds.), Springer-Verlag, LNCS 1219, 1997.
- [40] J. Vitek, M. Serrano and D. Thanos, "Security and Communication in Mobile Object Systems", in *Mobile Object Systems: Towards the programmable Internet*, pp 177-199, Springer-Verlag, LNCS 1222, 1997.
- [41] Vassilis Prevelakis, Dimitri Konstantas and Jean-Henry Morin, "Issues for the Commercial Distribution of Electronic Documents", *Proceedings of CMS'97, Communications and Multimedia Security Joint Working Conference IFIP TC-6 and IFIP TC-11*, Athens, Greece, September 22-23 1997 (A shorter version is published in the *Proceedings of ERSADS'97, Second European Research Seminar on Advances in Distributed Systems*, Zinal, Switzerland, March 17-21 1997).
- [42] Jean-Henry Morin, "HyperNews, a Hypermedia Electronic-Newspaper Environment based on Agents", Technical Report, Centre Universitaire d'Informatique, University of Geneva 1997.