



UNIVERSITÉ  
DE GENÈVE

Archive ouverte UNIGE

<https://archive-ouverte.unige.ch>

Article scientifique

Article

2008

Supplemental data

Open Access

This file is a(n) Supplemental data of:

Mapping proteins to disease terminologies: from UniProt to MeSH

Mottaz, Anais; Yip Sonderegger, Yum Lina; Ruch, Patrick; Veuthey, Anne-Lise

This publication URL:

<https://archive-ouverte.unige.ch/unige:2727>

Publication DOI:

[10.1186/1471-2105-9-S5-S3](https://doi.org/10.1186/1471-2105-9-S5-S3)

© This document is protected by copyright. Please refer to copyright holders for terms of use.

Regular Expressions used to extract the disease names from the Swiss-Prot  
disease comment lines

(1) Starter expressions	(2) Specific stop words	(3) Termination term
Cause(s) of /a involved in (can) contribute(s) to associated/association with correlated with responsible for contributor to result(s)/resulting in lead(s) to induce(s) defective in individual(s) with patient(s) with/suffering from reduce(s) influence(s) deleted in down-regulated in found in implicated in predispose(s) to favor antigen of antigen for thought to be an role in could impart mediate(s) candidate (gene)	susceptibility to development of genetic predisposition for developing pathogenesis of subset of various types of some form of increased risk of	also known as but which an due to in condition(s) such as . [MIM:

(1) Expressions used to extract the part of the string containing the disease name. (2) Terms removed from the string extracted. (3) Expressions indicating the end of the disease name.