

Data integration through data elements: Mapping data elements to terminological resources

Fleur Mouglin

EA 3888, IFR 140, Faculté
de Médecine, Université de
Rennes I, France

fleur.mouglin@univ-
rennes1.fr

Anita Burgun

EA 3888, IFR 140, Faculté
de Médecine, Université de
Rennes I, France

anita.burgun@univ-
rennes1.fr

Olivier Bodenreider

National Library of Medicine,
Bethesda, Maryland
USA

olivier@nlm.nih.gov

Abstract

Data integration is a crucial task in the biomedical domain. Data elements (DEs) play an important role in data integration and we propose to map DEs to terminological resources as an approach to data integration. We extracted DEs from eleven disparate biomedical sources. We compared these DEs to concepts and/or terms in biomedical controlled vocabularies and to reference DEs. We also exploited DE values to disambiguate under-specified DEs. Results suggest that data integration can be achieved automatically with limited precision and largely facilitated by mapping DEs to terminological resources. Finally, the use of general lexical resources and of more powerful techniques to exploit DE values would improve our method.
