Semantic Trilogy '13

4th International Conference on Biomedical Ontology ICBO 2013
9th Data Integration in Life Science DILS 2013
4th Canadian Semantic Web Conference CSWS 2013

Definitions in Ontologies

Montreal, Quebec, Canada July 6th—12th 2013

UNIVERSITÉ
Concordia
UNIVERSITY

INTERNATIONAL WORKSHOP ON DEFINITIONS IN ONTOLOGIES (DO 2013)

Proceedings

edited by

Selja Seppälä
and
Alan Ruttenberg

DO 2013 is held in conjunction with the
4th International Conference on Biomedical Ontology (ICBO 2013)

Concordia University
Montreal, Quebec, Canada

July 7, 2013

Preface

Ontologies built using OBO Foundry principles are advised to include both formal (logical) definitions, as well as natural language definitions. Depending on the effort, one or the other can be underrepresented. Possible explanations to this bottleneck are the high cost of producing well-written definitions; an insufficient understanding of the nature of natural language definitions or of logic; the lack of an operational theory of definitions; the lack of studies that evaluate usability and effectiveness of definitions in ontologies; a paucity of tools to help with definition authoring and checking. Producing natural language definitions is time-consuming, costly and prone to all kinds of inconsistencies. Producing logical definitions that are effective, correct, and communicative is also difficult. It is therefore worth exploring different ways of assisting, with automation, creation and quality control of definitions.

This workshop gathers interested researchers and developers to reflect upon general themes as the selection and modeling of defining information; the relation between definitions in specific domains as opposed to domain-independent definitions; the theoretical underpinnings of definitions; tools that can facilitate relating logical and natural language definitions. In addition, we wanted to encourage participation by different communities using definitions so that their needs and solutions can be exposed. This interdisciplinary goal proved successful, as each of the three selected papers represents a different community: ontology, terminology, and natural language processing (NLP).

Topics

Topics of interest were split between foundational aspects, pragmatic issues and user perspectives. Below we list some possible topics.

Foundational aspect
- Theories of definition and their implications for the defining practice
- Realist versus conceptualist approaches in definition writing
- Definition modeling: what kinds of information are defining
- Domain-independent versus domain-specific definition models
- Formal versus natural language definitions

Pragmatic issues
- Quality control in definitions
- Ways of evaluating definitions
- Comparison and evaluation of different definition production techniques: handwritten, automatically generated from formal definitions, extracted from corpora or constructed from information retrieved from corpora
- Methods and tools to automate definition production and checking
- (Multilingual) definition generation

- Information retrieval for definition production
- Use of definition models to facilitate information retrieval
- Definition extraction from corpora
- Interactions between ontologies and lexical resources (WordNet, FrameNet)
- Consequences/Strategies of giving necessary versus necessary and sufficient definitions, or simply sufficient definitions
- Coordination of logical and textual definitions
- Alternatives to and variants of definitions: elucidations, explanations, glosses, figures

User perspectives
- Assessment of definitions used in current practice
- Balancing needs of within discipline use and wider use of definitions
- Use of specialized terminology versus general vocabulary
- Presentation of definitions to different user audiences
- Alternatives/Augmentations of textual definitions, such as figures and images for anatomy, where textual definitions may be harder to formulate

The articles published in these proceedings cover three distinct aspects of definitions in ontologies, and include a survey report summary on defining practices in the ontology community conducted in preparation of the workshop:

Summary of the survey results:
> Selja Seppälä and Alan Ruttenberg
> ***Survey on Defining Practices in Ontologies: Report Summary***

Automatically populating ontologies with multilingual definitions:
> Guoqian Jiang, Harold Solbrig and Christopher Chute
> ***A Semantic Web-Based Approach for Harvesting Multilingual Textual Definitions from Wikipedia to Support ICD-11 Revision***

Displaying user-oriented defining contents using domain-specific templates:
> Antonio San Martín and Pilar León Araúz
> ***Flexible Terminological Definitions and Conceptual Frames***

Enhancing term retrieval through the definition's contents:
> Gerardo Sierra-Martinez and Laura Elena Hernandez-Dominguez
> ***Automatic Construction of the Knowledge Base of an Onomasiological Dictionary***

In addition to these papers, the workshop's program includes three presentations aimed at widening the scope of the issues to be discussed:

Pragmatic aspects of defining diseases in the Disease Ontology:
    Lynn M. Schriml
    Definition of Disease Terms

Domain- and language-independent definition templates:
    Selja Seppälä
    Using BFO Categories for Creating Generic Definition Templates

Definition authoring tools:
    Alan Ruttenberg
    Tool Support for Definition Authorship and Management: Desiderata

Our invited speaker covers some foundational aspects of definitions in ontologies:
    Barry Smith
    Introduction to the Logics of Definitions

We expect to have a rich and insightful discussion with the participants and the attendees of the workshop that will yield a prioritized list of needs and useful recommendations regarding the defining practices in ontologies.

Acknowledgements

July 2013

Selja Seppälä
Alan Ruttenberg

Organizing Committee
Selja Seppälä (University at Buffalo, USA)
Alan Ruttenberg (University at Buffalo, USA)

Program Committee
César Aguilar (Pontificia Universidad Católica de Chile)
Nathalie Aussenac-Gilles (National Center for Scientific Research (CNRS), France)
Caroline Barrière (CRIM, Canada)
Thomas Bittner (University at Buffalo, USA)
Mélanie Courtot (British Columbia Cancer Research Centre, Canada)
Christiane Fellbaum (Princeton University, USA)
Natalia Grabar (Université de Lille 3, France)
Janna Hastings (European Bioinformatics Institute, Cambridge, UK)
Marie-Claude L'Homme (Université de Montréal, Canada)
James Malone (European Bioinformatics Institute, Cambridge, UK)
Alexis Nasr (Aix Marseille Université, France)
Fabian Neuhaus (National Institute of Standards and Technology (NIST), USA)
James Overton (Knocean, Toronto, Canada)
Richard Power (The Open University, UK)
Patrice Seyed (Tetherless World Constellation, Rensselaer Polytechnic Institute, USA)
Robert Stevens (The University of Manchester, UK)
Allan Third (The Open University, UK)
Sandra Williams (The Open University, UK)

DO 2013 PROGRAM

July 7, 2013

8:30–8:45   Welcome

8:45–9:00   Lynn M. Schriml
Definition of Disease Terms

9:00–9:15   Gerardo Sierra-Martinez and Laura Elena Hernandez-Dominguez
Automatic Construction of the Knowledge Base of an Onomasio-logical Dictionary

9:15–9:30   Guoqian Jiang, Harold Solbrig and Christopher Chute
A Semantic Web-Based Approach for Harvesting Multilingual Textual Definitions from Wikipedia to Support ICD-11 Revision

9:30–9:45   Antonio San Martín and Pilar León Araúz
Flexible Terminological Definitions and Conceptual Frames

9:45–10:15   **Discussion**

10:15–10:45   Coffee Break

10:45–11:15   **Invited Speaker**: Barry Smith
Introduction to the Logic of Definitions

11:15–11:30   Selja Seppälä
Using BFO Categories for Creating Generic Definition Templates

11:30–11:45   Alan Ruttenberg
Tool Support for Definition Authorship and Management: Desiderata

11:45–12:00   **Discussion**