LREC-COLING 2024

ISA-20: The 20th Joint ACL - ISO Workshop on Interoperable Semantic Annotation @LREC-COLING-2024

Workshop Proceedings

Editor
Harry Bunt

20 May, 2024
Torino, Italia

Proceedings of the 20th Joint ACL - ISO Workshop on Interoperable Semantic Annotation @LREC-COLING-2024

Copyright ELRA Language Resources Association (ELRA), 2024
These proceedings are licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ISBN 978-2-493814-32-6
ISSN 2951-2093 (COLING); 2522-2686 (LREC)

Jointly organized by the ELRA Language Resources Association
and the International Committee on Computational Linguistics

Preface

Twenty years after the first ISA workshop (Tilburg, 2004), this is the year of the 20th edition of the series. The ISA workshops were inspired by the decision of the International Organisation for Standards ISO to start developing annotation standards for language data, including lexical information, typed feature structures, morphological and syntactic information, and the Semantic Annotation Framework (SemAF), a multi-part standard for annotating aspects of meaning. As the development of such standards is the work of small groups of experts nominated by ISO member countries, the ISA workshops were set up with the intention to (a) promote the involvement of all interested scholars in these processes, and (b) to inform scholars in language studies and developers of language resources and linguistic applications of the ISO activities and the standards under development. To support this two-way interaction, the ISA workshops were organised as a joint initiative of the ACL Special Interest Group in Semantics (SIGSEM) and of ISO Working Group TC 37/SC 4/WG 2, Semantic annotation. apart from ISA-3 (Marina del Rey, 2008) and ISA-8 (Pisa, 2012), all ISA workshops have been organised as satellite events of large conferences such as LREC, ACL, IWCS and COLING.

This year's workshop at LREC-COLING 2024 has seen a higher number of submissions, full papers as well as short papers, than any previous edition of the ISA series. Since most of the submissions were of excellent quality, we have been forced to follow the LREC policy of having not only the accepted short papers but also some of the accepted full papers presented as a poster plus a flash presentation. True to the original intention of the ISA workshop series, the ISA-20 program features a mix of papers, presented in these proceedings, on developing new ISO standards such as VoxML (visual information) and QuantML (quantification), on ways to use multiple standards defined by SemAF parts, and on topics not directly related to ISO activities but to semantic annotation as such.

Many thanks are due to the Program Committee members for their diligent and fast review work.

The organisers,

Harry Bunt, Nancy Ide, Kiyong Lee, Volha Petukhova, James Pustejovsky, and Laurent Romary.

Organizing Committee

Harry Bunt
Nancy Ide
Kiyong Lee
Volha Petukhova
James Pustejovsky
Laurent Romary

Program Committee

Jan Alexandersson
Maxime Amblard
Johan Bos
Harry Bunt
Jae-Woong Choe
Stergios Chatzykiriakidis
Robin Cooper
Rodolfo Delmonte
David DeVault
Simon Dobnik
Jens Edlund
Alex Fang
Robert Gaizauskas
Koiti Hasida
Nancy Ide
Elisabetta Jezek
Kiyong Lee
Philippe Muller
Rainer Osswald
Catherine Pelachaud
Volha Petukhova
Laurent Prévot
Stephen Pulman
James Pustejovsky
Laurent Romary
Purificação Silvano
Manfred Stede
Thorsten Trippel
Carl Vogel
Menno van Zaanen
Annie Zaanen
Heike Zinsmeister

Table of Contents

The MEET Corpus: Collocated, Distant and Hybrid Three-party Meetings with a Ranking Task
Ghazaleh Esfandiari-Baiat and Jens Edlund 1

MSNER: A Multilingual Speech Dataset for Named Entity Recognition
Quentin Meeus, Marie-Francine Moens and Hugo Van hamme 8

Attitudes in Diplomatic Speeches: Introducing the CoDipA UNSC 1.0
Mariia Anisimova and Šárka Zikánová 17

Automatic Alignment of Discourse Relations of Different Discourse Annotation Frameworks
Yingxue Fu 27

A New Annotation Scheme for the Semantics of Taste
Teresa Paccosi and Sara Tonelli 39

What to Annotate: Retrieving Lexical Markers of Conspiracy Discourse from an Italian-English Corpus of Telegram Data
Costanza Marini and Elisabetta Jezek 47

Lightweight Connective Detection Using Gradient Boosting
Mustafa Erolcan Er, Murathan Kurfalı and Deniz Zeyrek 53

Shallow Discourse Parsing on Twitter Conversations
Berfin Aktas and Burak Özmen 60

Search tool for An Event-Type Ontology
Nataliia Petliak, Cristina Fernandéz Alcaina, Eva Fučíková, Jan Hajič and Zdeňka Urešová 66

Tiny But Mighty: A Crowdsourced Benchmark Dataset for Triple Extraction from Unstructured Text
Muhammad Salman, Armin Haller, Sergio J. Rodriguez Mendez and Usman Naseem . 71

Less is Enough: Less-Resourced Multilingual AMR Parsing
Bram Vanroy and Tim Van de Cruys 82

MoCCA: A Model of Comparative Concepts for Aligning Constructicons
Arthur Lorenzi, Peter Ljunglöf, Ben Lyngfelt, Tiago Timponi Torrent, William Croft, Alexander Ziem, Nina Böbel, Linnéa Bäckström, Peter Uhrig and Ely E. Matos 93

ISO 24617-8 Applied: Insights from Multilingual Discourse Relations Annotation in English, Polish, and Portuguese
Aleksandra Tomaszewska, Purificação Silvano, António Leal and Evelin Amorim 99

Combining semantic annotation schemes through interlinking
Harry Bunt 111

Fusing ISO 24617-2 Dialogue Acts and Application-Specific Semantic Content Annotations
Andrei Malchanau, Volha Petukhova and Harry Bunt 122

Annotation-Based Semantics for Dialogues in the Vox World
Kiyong Lee 133

Annotating Evaluative Language: Challenges and Solutions in Applying Appraisal Theory
Jiamei Zeng, Min Dong and Alex Chengyu Fang 144

Attractive Multimodal Instructions, Describing Easy and Engaging Recipe Blogs
Ielka van der Sluis and Jarred Kiewiet de Jonge 152

Workshop Program

09:00 **Opening; Session 1**

The MEET Corpus: Collocated, Distant and Hybrid Three-party Meetings with a Ranking Task
Ghazaleh Esfandiari-Baiat and Jens Edlund

MSNER: A Multilingual Speech Dataset for Named Entity Recognition
Quentin Meeus, Marie-Francine Moens and Hugo Van hamme

Attitudes in Diplomatic Speeches: Introducing the CoDipA UNSC 1.0
Mariia Anisimova and Šárka Zikánová

10:30 **Coffee break**

11:00 **Session 2**

Automatic Alignment of Discourse Relations of Different Discourse Annotation Frameworks
Yingxue Fu

11:30 **Session 3: Flash presentations**

A New Annotation Scheme for the Semantics of Taste
Teresa Paccosi and Sara Tonelli

What to Annotate: Retrieving Lexical Markers of Conspiracy Discourse from an Italian-English Corpus of Telegram Data
Costanza Marini and Elisabetta Jezek

Lightweight Connective Detection Using Gradient Boosting
Mustafa Erolcan Er, Murathan Kurfalı and Deniz Zeyrek

Shallow Discourse Parsing on Twitter Conversations
Berfin Aktas and Burak Özmen

No Day Set (continued)

Search tool for An Event-Type Ontology
Nataliia Petliak, Cristina Fernandéz Alcaina, Eva Fučíková, Jan Hajič and Zdeňka Urešová

Tiny But Mighty: A Crowdsourced Benchmark Dataset for Triple Extraction from Unstructured Text
Muhammad Salman, Armin Haller, Sergio J. Rodriguez Mendez and Usman Naseem

Less is Enough: Less-Resourced Multilingual AMR Parsing
Bram Vanroy and Tim Van de Cruys

MoCCA: A Model of Comparative Concepts for Aligning Constructicons
Arthur Lorenzi, Peter Ljunglöf, Ben Lyngfelt, Tiago Timponi Torrent, William Croft, Alexander Ziem, Nina Böbel, Linnéa Bäckström, Peter Uhrig and Ely E. Matos

12:15 **Poster visits**

13:00 **Lunch break**

14:00 **Session 4**

ISO 24617-8 Applied: Insights from Multilingual Discourse Relations Annotation in English, Polish, and Portuguese
Aleksandra Tomaszewska, Purificação Silvano, António Leal and Evelin Amorim

Combining semantic annotation schemes through interlinking
Harry Bunt

Fusing ISO 24617-2 Dialogue Acts and Application-Specific Semantic Content Annotations
Andrei Malchanau, Volha Petukhova and Harry Bunt

Annotation-Based Semantics for Dialogues in the Vox World
Kiyong Lee

No Day Set (continued)

16:00 **Tea break**

16:30 **Session 5**

Annotating Evaluative Language: Challenges and Solutions in Applying Appraisal Theory
Jiamei Zeng, Min Dong and Alex Chengyu Fang

Attractive Multimodal Instructions, Describing Easy and Engaging Recipe Blogs
Ielka van der Sluis and Jarred Kiewiet de Jonge

17:30 **Closing**