WP10: Dissemination and Exploitation

Olga Caprotti
University of Gothenburg

Jeroen van Grondelle
Be Informed

20 March 2012, UPC Barcelona
MOLTO Second Year Review
The 3rd MOLTO project meeting was held in Helsinki on 31 August - 2 September 2011, organized by UHEL.
# MOLTO Open Day

2 September 2012

## Program

<table>
<thead>
<tr>
<th>timeslot</th>
<th>title</th>
<th>speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 - 9:30</td>
<td><a href="/node/1333">MOLTO Project: Overview and Half-Way Results</a></td>
<td>Aarne</td>
</tr>
<tr>
<td>09:30 - 09:45</td>
<td>The Mathematical Grammar Library</td>
<td>Jordi</td>
</tr>
<tr>
<td>09:45 - 10:00</td>
<td>GF web IDE with focus on Finnish (demo)</td>
<td>Thomas/Aarne</td>
</tr>
<tr>
<td>10:00 - 10:30</td>
<td>Invited talk: Helsinki Finite-State Technology</td>
<td>Krister Lindén</td>
</tr>
<tr>
<td>10:30 - 11:00</td>
<td>Coffee break</td>
<td></td>
</tr>
<tr>
<td>11:00 - 11:15</td>
<td>Term extraction from ontology with pre-constructed queries</td>
<td>Inari</td>
</tr>
<tr>
<td>11:15 - 12:00</td>
<td>TermFactory (demo)</td>
<td>Lauri C.</td>
</tr>
<tr>
<td>12:00 - 12:30</td>
<td>Machine translation evaluation</td>
<td>Maarit</td>
</tr>
<tr>
<td>12:30 - 13:00</td>
<td>Approximating term similarities with syntactic evidence</td>
<td>Seppo</td>
</tr>
<tr>
<td>13:00 - 14:30</td>
<td>Lunch at Café Arppeanum</td>
<td></td>
</tr>
<tr>
<td>14:30 - 15:00</td>
<td>Invited talk: HFST as an open source environment for commercial and academic language tools</td>
<td>Kimmo Koskenniemi</td>
</tr>
<tr>
<td>15:00 - 15:30</td>
<td>A Reason-able view of Linked Data for Cultural Heritage</td>
<td>Milen Chechev (Ontotext)</td>
</tr>
<tr>
<td>15:30 - 16:00</td>
<td>Native PGF</td>
<td>Lauri A.</td>
</tr>
<tr>
<td>16:00</td>
<td>Closing</td>
<td></td>
</tr>
</tbody>
</table>
MOLTO-EEU

Kick off meeting

Between January 12-13, 2012 at Chalmers, organized by UGOT
Project Meetings

4th Project Meeting

Took place in Zürich, organized by UZH, on March 7-9, 2012
**Book**

Google Scholar BibTex

**Conference Paper**

Google Scholar BibTex  
Download: PBML-2010-Gimenez.pdf (170.97 KB)

Controlled Language for Everyday Use: the MOLTO Phrasebook, Ranta, Aarne, Enache Ramona, and Détrez Grégoire , Controlled Natural Languages Workshop (CNL 2010), Marettimo, Italy, (2011)  
Google Scholar BibTex  
Download: everyday.pdf (647.5 KB)

Google Scholar BibTex

Google Scholar BibTex  
Download: wmt10-cgmac.pdf (67.6 KB)

Google Scholar BibTex  
Download: ranlpWS-final.pdf (421.76 KB)

Google Scholar BibTex


Multilingual Packages of Controlled Languages: An Introduction to GF, Ranta, Araine, CNL-2010 (2nd Workshop on Controlled Natural Language), 14 September, Maretto, Sicily, (2010) Google Scholar BibTex Download: cnl-2010.pdf (1.11 MB)


Smart Paradigms and the Predictability and Complexity of Inflectional Morphology, Détrez, Grégoire, and Ranta Araine, EACL (European Association for Computational Linguistics), 04/2012, Avignon, (2012) Google Scholar BibTex

Typeful Ontologies with Direct Multilingual Verbalization, Angelov, Krasimir A., and Enache Ramona, Controlled Natural Languages Workshop (CNL 2010), Marettimo, Italy, (2011)  Google Scholar BibTex Download: FinalSUMOCNL.pdf (252.38 KB)

Frontiers of Multilingual Technologies

August 15-26, 2011

Barcelona, Spain.
Forthcoming
FreeRBMT and CNL

Free/Open-source Rule-based Machine Translation
Gothenburg, Sweden, 13-15 June 2012

This workshop aims to bring together the experience of researchers and developers in the field of rule-based machine translation who have decided to get on board the free/open-source train and are effectively contributing to creating a commons of explicit knowledge: machine translation rules and dictionaries, and machine translation systems whose behaviour is transparent and clearly traceable through their explicit logic. Also hybrid systems combining rule-based and statistical methods are in scope.

Programme Committee chair: Aarne Ranta
Organizing Committee Chair: Olga Caprotti

Proceedings to appear in
CSLI Studies in Computational Linguistics ONLINE

CNL
Third Workshop on Controlled Natural Language
(CNL 2012)
29–31 August 2012
Zurich, Switzerland

A controlled natural language (CNL) is based on natural language but comes with restrictions on vocabulary, grammar, and/or semantics. The general goal is to reduce or eliminate ambiguity and complexity.

Some of these languages are designed to improve communication among humans, especially for non-native speakers of the respective natural language. In other cases, the restrictions on the language are supposed to make it easier for computers to analyze such texts in order to improve computer-aided, semi-automatic, or automatic translations into other languages. A third group of CNL has the goal to enable reliable automated reasoning on seemingly natural texts. Such languages have a direct mapping to some sort of formal logic and should improve the accessibility of formal knowledge representations or specifications for people unfamiliar with formal notations.

All these types of CNL are covered by this workshop.
Exploitation by Be Informed

WP12 is about adoption

Proof of the Exploitation Pudding
If successful, Be Informed’s verbalization component can be permanently moved to GF

Successful =
Quality texts in multiple languages
Modular integration with (multilingual) ontologies
Analysts are able to develop grammars
Future of GF (community) is secured

Adoption:
Be Informed will base its default verbalization on GF
Customers/Partners will develop custom grammars
Pre-requisites for industrial exploitation

Modular integration
   GF can be adopted as an add-on, no low level conformance or changes required to adopt GF

Analysts are able to develop grammars
   Grammar development discipline should align with one or more industry roles, preferably not linguistically trained

Future of GF (community) is secured
   In many adoption models, GF is treated as a toolset, changing/fixing it is beyond typical industry expertise

   Therefore, GF’s future development and maintenance must we well grounded in a community to be widely adopted
Adoption of the Software Components

But also

Verbalization service

Offering Newsfeeds from structured data

Objects, specifically (Social)Events

Translation services

Like phrasebook app, in other domains

Lexical services

Leverage all the knowledge in resource grammars

Developing/selling grammars

Potentially based on standard vocabularies etc.