



Published on *Multilingual Online Translation* (<http://www.molto-project.eu>)

[Home](#) > [Printer-friendly](#) > Book > Export > Html > 1674 > D1.6 Periodic Management Report T30

D1.6 Periodic Management Report T30

| | |
|---------------------------------------|---|
| Contract No.: | FP7-ICT-247914 |
| Project full title: | MOLTO - Multilingual Online Translation |
| Deliverable: | D1.6 Periodic Management Report T30 |
| Security (distribution level): | Confidential |
| Contractual date of delivery: | M30 |
| Actual date of delivery: | 7 Nov. 2012 |
| Type: | Report |
| Status & version: | Final |
| Author(s): | O. Caprotti et al. |
| Task responsible: | UGOT |
| Other contributors: | All |

Abstract

Progress report for the fifth semester of the MOLTO project lifetime, 1 Mar 2012 - 31 Aug 2012.

1. Publishable summary

The project MOLTO - Multilingual Online Translation, started on March 1, 2010 and will run until 31 May 2013 with the task to develop tools for translating texts between multiple languages in real time with high quality. MOLTO grounding technology is multilingual grammars based on semantic interlinguas and statistical machine translation to simplify production of multilingual documents without sacrificing the quality. The specific interlinguas are based on domain semantics and are equipped with reversible generation functions: namely translation is obtained as a composition of parsing the source language and generating the target language. An implementation of this technology is provided by GF, Grammatical Framework, which in MOLTO is furthermore complemented by the use of ontologies, as in the semantic web, and by methods of statistical machine translation (SMT) for improving robustness and extracting grammars from data. GF has been applied in several small-to-medium size domains, typically targeting up to ten languages but MOLTO will scale this up in terms of productivity and applicability.

A part of the scale-up is to increase the size of domains and the number of languages. A more substantial part is to make the technology accessible to domain experts without GF expertise and minimize the effort needed for building a translator. Ideally, the MOLTO tools will reduce the overall task to just extending a lexicon and writing a set of example sentences.

MOLTO is committed to dealing with 15 languages, which includes 12 official languages of the European Union - Bulgarian, Danish, Dutch, English, Finnish, French, German, Italian, Polish, Romanian, Spanish, and Swedish - and 3 other languages - Catalan, Norwegian, and Russian. In addition, there is on-going work on at least Arabic, Farsi, Hebrew, Hindi/Urdu, Icelandic, Japanese, Latvian, Maltese, Portuguese, Swahili, Tswana, and Turkish.

While tools like Systran (Babelfish) and Google Translate are designed for consumers of information, MOLTO will mainly target the producers of information. Hence, the quality of the MOLTO translations must be good enough for, say, an e-commerce site to use in translating their web pages automatically without the fear that the message will change. Third-party translation tools, possibly integrated in the browsers, let potential customers discover, in their preferred language, whether, for instance, an e-commerce page written in French offers something of interest. Customers understand that these translations are approximate and will filter out imprecision. If, for instance, the system has translated a price of 100 Euros to 100 Swedish Crowns (which equals 10 Euros), they will not insist to buy the product for that price. But if a company had placed such a translation on its website, then it might be committed to it. There is a well-known trade-off in machine translation: one cannot at the same time reach full coverage and full precision. In this trade-off, Systran and Google have opted for coverage whereas MOLTO opts for precision in domains with a well-understood language.

MOLTO technology will be released as open-source libraries, accompanied by cloud services, to be used for developing plug and play components to translation platforms and web pages and thereby designed to fit into third-party workflows. The project will showcase its

results in web-based flagship demos applied in three case studies: mathematical exercises in 15 languages, patent data in at least 3 languages, and museum object descriptions in 15 languages. The MOLTO Enlarged EU scenarios will apply MOLTO tools to a collaborative semantic wiki and to an interactive knowledge-based system used in a business enterprise environment.

2. Core of the report

This section describes the progress of each workpackage and discusses changes to the workplan, if necessary.

2.1 Project objectives for the period

The main objective of this 5th semester has been to consolidate the project tools and technologies towards the production of the final deliverables. In order to focus the developments to clear goals, the Consortium has agreed to identify 9 "MOLTO flagships" that highlight the achievements of the project and combine what has been produced across work packages:

- an IDE for GF including the Cloud-IDE and the Eclipse-IDE
- a professional translator tool with integrated terminology extension, predictive parsing, syntax editing, robust parsing
- an ontology query system, based on a generic reusable grammar (carefully built manually) that can be instantiated with automatically derived domain-specific extensions; demoed with mathematics and museum artefacts
- a hybrid patent translation system, scoring better results than its competitors, if possible
- the Mathematical Grammar Library, MGL, supporting both scenarios of math documentation (Wiki) and interactive tutorial (Sage, word problems)
- a museum visitors' system generating descriptions and providing a query system, in 15 languages
- the multilingual semantic wiki integrating almost all aspects (even SMT? an open question)
- a commercial product from Be Informed that is useful for their customers (to be specified)
- a trasversal methodology of evaluations, for each of the above, showing that high quality can be reached with reasonable effort using MOLTO tools

2.2 Work progress and achievements during M25-M30

This section provides a concise overview of the progress of the work in line with the structure of Annex I to the Grant Agreement.

For each work-package, except project management, which is reported in Section 2.3, the work-package leader provides the following information:

- A summary of progress towards objectives and details for each task;
- Highlights of clearly significant results;
- A statement on the use of resources for Period 2, in particular highlighting and explaining deviations between actual and planned person-months per work package and per beneficiary in Annex I (Description of Work).

Moreover, if applicable:

- The reasons for deviations from Annex I and their impact on other tasks as well as on available resources and planning;
- The reasons for failing to achieve critical objectives and/or not being on schedule with remarks on the impact on other tasks as well as on available resources and planning;
- Corrective actions.

WP2 Grammar developer's tools - M30

Summary of progress

GF Eclipse plugin <http://www.grammaticalframework.org/eclipse/index.html> has grown to Version 1.5.1 by June. It has been adopted by Be Informed and Ontotext. Camilleri and Ranta gave a GF crash course at Be Informed using Eclipse. There are moreover two publications: one in EAMT (a poster), one at FreeRBMT (full paper).

The Resource Grammar Library has been enhanced by two languages as external contributions: Japanese and Latvian. The MOLTO Phrasebook has been extended to Latvian. Work on Chinese and Maltese is going on.

With the release of D2.3, Grammar Tools and Best Practices, the work in this WP finished on 30 June. But there is further work planned, as dissemination and exploitation.

A preview version of [libpgf](#), a C-based reimplement of the GF runtime, is available since July. When finished, it should make GF technology accessible to applications that cannot make use of the current Haskell- and Java-based runtimes either due to resource constraints or interoperability concerns. In particular, libpgf should be easier to access from non-JVM-based programming languages. Bindings for Python are available since September.

Highlights

GF Eclipse plugin <http://www.grammaticalframework.org/eclipse/index.html>

D2.3 <http://www.molto-project.eu/biblio/deliverable/grammar-tools-and-best-pr...>

Deviations from Annex I

The delivery date of D2.3 was postponed from M24 to M27 to be able to profit from the initial experiences with the new partners' scenarios.

WP3 Translator's tools - M30

Summary of progress

After drawing the specifications in [D 3.1](#), we published the prototype as [Deliverable 3.2](#).

A first prototype of the translation editor, one built using the Google Web Toolkit, was tested for integration into the GlobalSight translation manager. This first prototype turned out to be too fragile and has now been replaced by simpler version, the Simple Translation Editor, found in <http://cloud.grammaticalframework.org/translator>.

Work has progressed integrating the translation tools prototype with the TermFactory web based term ontology editor (TF) that is to be used by ontologists and terminologists to create and maintain MOLTO domain dependent vocabulary.

The milestone MS8 (Translation tool complete) is achieved in parallel with [WP5](#) (Statistical and robust t.) in the sense that the MOLTO GlobalSight translation project management system is set up and available for utilisation in other work packages, including the use cases and their evaluation in [WP9](#). Also, MS8 can be considered achieved in the sense that TermFactory is integrated with OWLIM ontology repository, completing the ontology backend of the tool.

Highlights

During the reporting period, a login and access control component (GateService) has been added to TF. This component is integrated with and maintained through the GlobalSight user manager.

The OWLIM ontology repository has been integrated with TF, using a common interface based on Jena assembler library. Besides OWLIM, Jena triple databases (TDB, SDB) and WebDAV ontology documents can be edited directly from TermFactory.

Deviations from Annex I

UHEL has unspent money, and has recruited an evaluation project manager to work with [WP3](#) and [WP9](#) until M36+3, beginning in November 2012 (start of M32).

We are asking to move Deliverable D3.3 (Translation tools / workflow manual) to M33. The D3.3 manual is expected to include documentation of the workflow with help from other work packages and input from the new recruit, that have all been subject to a 3 months shift.

WP4 Knowledge engineering - M30

Summary of progress

During this reporting period, a new version of D4.3 Grammar Ontology Interoperability was submitted elaborating on the automatization of the grammar creation for controlled language (CL) to RDF interoperability. A new way of building the transformations between CL and SPARQL, treating SARQL as another concrete grammar along with the language specific ones was suggested.

We have designed and implemented the Query Grammar Helper Builder tool to connect a SPARQL end point and support an inexperienced user to create GF grammars - Abstract, English and SPARQL. As a part of the work in [WP2](#) we have developed an Eclipse plugin wrapper of this tool.

We have also started designing the needed improvements to the interoperability approach in order to productize it as an outcome of the project and in alignment with the exploitation objectives.

We have also introduced several new members to the Ontotext team - Laura Tolosi, Maria Mateva, Ilia Trendafilov and Georgi Georgiev.

Highlights

- improvements on Deliverable 4.3 and opening potential for better CL to RDF interoperability
- improvement of the autocomplete suggestion in the KRI-prototype
- design, implementation, testing, bug fixing of the Query Grammar Builder Tool
- MOLTO-KRI prototype(<http://molto.ontotext.com>) – improvements on the site

Deviations from Annex I

None

WP5 Statistical and robust translation - M30

Summary of progress

The milestone [MS8](#) has been achieved in M30 (Translation tool complete), which for [WP5](#) meant to have a complete system integrating the grammar and STM. Although the system is already available on Gothenburg's server we are still working on improvements.

The work done during the fifth period has been focused on 4 of the 6 tasks of the workpackage:

5.3 Robust Parsing First efforts to include the robust parsing work done in the previous semesters into the hybrid systems are being done. The work is in progress and the final idea is to be able to use GF's robust parsing to deal with the chunks instead of relying on Genia.

5.4 Baseline systems Refinements on the French GF grammar have been done in order to improve the performance. The German grammar has been done from scratch and it is now comparable to the French one.

5.5 Hybrid Models The new grammars have been integrated in the final hybrid system. Different versions of the previous hybrids are now available. In particular, a new system considers different probabilities for the GF translations according to the confidence in obtaining them. This information can be also used in the development step of the statistical system. A one-click system has been developed with the most promising hybrid system. This system will be updated with new hybrids whenever we obtain a better translation performance.

5.6 Systems evaluation A wider evaluation of the baseline systems has been done by including syntactic and semantic metrics into the evaluation. Also, the comparison with external translation systems such as Google and Bing has been redone in order to reflect the improvements of these systems during the last year. A comparison with Pluto is also done. However, we realized that since we share some data there is the possibility that our test sets are in their training data. We plan on using confidence estimation measures in order to be able to test on different patents for which none of us have translations such as American patents.

Highlights

- A GF grammar for patents has been developed for German and improved for French
- A hybrid system with the new grammars have been evaluated and a new one which takes into account probabilities for the GF translations has been built
- The work on robust parsing has resulted in two submissions to the Coling 2012 conference
- A one-click system for the hybrid translator has been build and is now available as a shell command on the server in Gothenburg. Partners wishing to test the system should contact UGOT to obtain access to the server.

Deviations from Annex I

There are no deviations from Annex I and at M30 the workpackage has produced a hybrid system for patent translation for English-to-French and English-to-German. For the opposite directions we use SMT as fallback.

However, we plan to continue the work on hybrid systems by improving the current German translator and the integration with robust parsing. With this, D5.3 will be postponed till January, when new hybrid systems will be also finished and prepared to be evaluated within [WP9](#).

WP6 Case study: mathematics - M30

Summary of progress

- Created a prolog-based reasoner to deal with elementary problems in arithmetic, implementing partition by subclasses and decomposition.
- Created GF grammars to express some word problems in English and Prolog

- Started integration with UZH AceWiki/gfsservice.

Highlights

- An AceWiki/gfsservice fork that allows entering the problem sentence by sentence in natural language and then automatically computes the associated mathematical model.

Deviations from Annex I

We are running out of time to fully develop the prototype. The promised system would work in two modes: *Author mode* for entering a problem and *Student mode* for attempting to solve it. The first mode is more or less working but the last will take a lot of time. The scheduled Deliverable D6.3, **Assistant for solving word problems**, due by 1 September 2012 has been postponed to 1 December 2012. Some resources of UPC originally planned in [WP8](#) have been moved to this workpackage in order to accomplish the promised tasks.

WP7 Case study: patents - M30

Summary of progress

Due the incremental development of the prototype, most of the tasks have span till M30, when the final prototype is being completed. The next lines describe the progress of the following tasks:

- 7.2 Patent Corpora
- 7.3 Grammars for the patent domain
- 7.4 Ontologies and Document Indexation
- 7.5 Patents Retrieval System
- 7.6 Machine Translation Systems
- 7.7 User Interface

In relation to Task 7.2, the patents downloaded from the EPO website have been automatically translated and semantically annotated. The complete collection of files is available in the MOLTO repository, and it consists of 1) the original patent documents, 2) the English version of the patent documents having the semantic annotations, and 3) the automatic translations of claims, abstracts and descriptions. These documents constitute the main content of the retrieval databases.

As for Task 7.4 and Task 7.5, the ontologies, indexes and databases have been updated with the new dataset of documents.

Regarding Task 7.6, we designed process for patents translation that allows for building a translated document having the same XML structure as the original patent. As a result, the interface of the prototype can show the translated patents using the same user-friendly view as for the original ones. The translation of the documents consists of a pipeline involving the following 5 steps: First, the patent files are preprocessed in order to extract the text contained into the sections in a structured manner (step 1). Then, the formatting marks inline with the text are replaced by placeholders (step 2). And then, the resulting text is segmented and tokenized as required by the translation system (step 3). Soon after, the raw text is translated using the SMT system (step 4). The translated text is post-processed in order to recover the original structure of the document (step 5), including original formatting, claims enumeration and images.

Regarding Task 7.3, the query grammars have been refactored using the set of primitives defined in the Query Library work conducted in [WP4](#). In consequence, the English and French version of the patents query grammar were adapted to the new structure, and the German version has been developed from scratch. The new grammar is equivalent to the old one. The difference is the fact that it relies on the primitive query building functions defined in the Query Library. Developing a grammar using the Query Library requires less linguistic knowledge, but just selecting the right set of primitives that would be right for the task. In comparison to the previous patent query grammar, now it has fewer constructions, because of the fact that it is developed on top of the Query Library. As a consequence, the constructions are also more natural and the number of malformed constructions have decreased considerably. The current grammar consists of 31 patterns and it is able to parse/generate 359 query constructions in English, 111 in French and 147 in German.

Finally, regarding Task 7.7, the interface has been updated with the German version of the query grammars. Also, some basic tests have been carried out at two levels in order to assess the prototype functionalities. First, some deficiencies have been corrected regarding the usability of the interface, i.e., examples of the main page, the language selection and the visualization of the results in French and German. In addition, we studied the inherent logic of the queries and the expected results, so that the system returns results that can now be considered more appropriate or accurate.

The Deliverable 7.2 gives a detailed description of the modules and their functionalities.

Highlights

- A fully functional version of the final prototype is available by M30.
- The demo allows for querying the system in English, French and German.
- The patents in the database has original text in English, French and German and the translated documents.
- A fully completed pipeline for patent document translation.
- The new query library and its application to the patents use case has been presented at the Third Workshop on Controlled Natural

Language (CNL 2012), being held in Zurich at the end of August. <http://attempto.ifi.uzh.ch/site/cnl2012/>

Deviations from Annex I

In general lines, we are achieving the objectives related to [WP7](#). However, the Deliverable 7.2, planned for M27, has been delayed to M30 due several issues related to the gathering of the corpora, the pro/post process of the documents and the integration of the new query library. Also, we carried several basic tests in order to assess the behavior of the prototype in terms of query results and user interaction, which reported several deficiencies that have been corrected. Since D72 has been postponed, D73 is delayed accordingly from M33 to M36.

WP8 Case study: cultural heritage - M30

Summary of progress

The data collection (D8.1) and first prototype of grammars (D8.2) were delivered on time. The grammar prototype has six languages, but is being extended to 15. It implements the generation of describing texts from facts in the database. The final system (delivered as a part of D8.3) will also allow natural language queries about museum objects, applying the technology developed in [WP4](#).

In addition to Gothenburg City Museum, there has been interest to this WP from the Europeana project <http://www.europeana.eu/portal/>. A plan for later dissemination of the work includes the generalization of the results by making them available for Europeana. Also the Monnet project <http://www.monnet-project.eu/> has a common interest in this WP in the area of ontology localization and verbalization.

The work in this workpackage in particular addresses multilingual text planning, and was exploited in D8.2 and has resulted in two publications in 2012; see [WP10](#).

Highlights

- Delivery of D8.2, Multilingual Grammar for Museum Object Descriptions

Deviations from Annex I

The schedule for D8.3 is postponed from M30 to M36 due to a delay in the PhD defense of Dana Dannélls, one of the key persons of this WP; as a bonus, MOLTO can then fully profit from her thesis work funded from other sources.

WP9 User requirements and evaluation - M30

Summary of progress

To improve the communication between work packages, we have set up a bug tracking system in <http://tfs.cc/trac>, and assigned the flagship leaders to be in charge of their component. Everyone using the tools is encouraged to leave their comments and requests in trac.

Maarit Koponen's work in evaluating semantical aspects of machine translation quality is progressing. We have started recruiting people for translation quality evaluation, and got response from the University of Pisa, with Italian-English and Italian-German as possible language pairs.

Targeting towards D9.2, we have started gathering evaluations from the grammar writers. Individual use cases can be measured in terms of translation quality, but good grammar design principles will make any grammar easier to write and maintain. We evaluate the grammars in terms of D2.3, a best practices document.

Highlights

Maarit Koponen's article [Comparing human perceptions of post-editing effort with post-editing operations](#) was accepted in Seventh Workshop on Statistical Machine Translation (Montreal) and published in the proceedings.

Deviations from Annex I

No major deviations reported. Minor actions include the addition of grammar quality evaluation.

WP10 Dissemination and exploitation - M30

Summary of progress

New on the website, the publishing of news items from RSS feeds of MOLTO Consortium partners and from the GF source code repository, in the footer, and news items from MOLTO in the header, alongside the publications and the demos. A new collective demo

of the GF application grammars together with the novel GF cloud services is prominently featured on the website.

We are also collecting the Use of resources in a overall table (<http://www.molto-project.eu/workplan/resources>) that summarizes the data provided by the partners. Personal views (e.g. <http://www.molto-project.eu/workplan/resources/olga.caprotti>) and workpackage views will be available soon.

Two major events have been organized with the sponsorship of MOLTO: FreeRBMT 2012, and CNL 2012. Free Rule-Based Machine Translation, FreeRBMT 2012, took place in Gothenburg on 13-15 June, 2012 and was organized by UGOT (see <http://www.chalmers.se/hosted/freerbmt12-en>). A tutorial on the Apertium system followed as additional satellite event, it was attended by MOLTO partners from UPC and UGOT and resulted in the adoption of some of the Apertium lexicons in GF. Papers by J. Camilleri and by Cristina España-Bonet et al. presenting MOLTO results will appear in the online proceedings. Additionally the program included a series of presentations on MOLTO's current work in *GF resources and tools for machine translation* (see <http://www.molto-project.eu/freerbmt-program.html>). Many of the MOLTO talks have been streamed live from the moltoproject YouTube channel, <http://www.youtube.com/moltoproject>, where they can still be watched.

The Third Workshop on Controlled Natural Language, CNL 2012, took place on 29-31 August 2012 in Zurich, Switzerland. UZH has been organizing this meeting over the past years, and this time as a MOLTO activity. A few papers were presented by the MOLTO Consortium, listed below, but we also note a contribution by external researchers, Normunds Grūzītis, Pēteris Paikens and Guntis Bārzdīņš, *FrameNet Resource Grammar Library for GF*, using the MOLTO Phrasebook as case study in their work.

On 14 August, [Aarne Ranta](#) visited Lingsoft Inc in Helsinki. Lingsoft is "a full-service language management company", producing for instance the proofing tools for the Nordic languages and German in Microsoft Office products. Lingsoft is one of the most successful language technology companies, founded in 1986 and working with numerous partners and products. Recent products range from spell checking to language education tools, speech recognition, and translation. He was invited by the CEO Juhani Reiman and the Senior Advisor Simo Vihjanen to give a presentation of MOLTO's tools and discuss possible collaborations. MOLTO and Lingsoft share the belief in precise linguistic knowledge as a key to successful language processing. Lingsoft has now set up a group to explore the possibilities offered by MOLTO and GF. The focus is on machine-assisted translation for specific domains.

Highlights

YouTube videos of MOLTO related talks, <http://www.youtube.com/moltoproject>

Toward multilingual mechanized mathematics assistants, Saludes, Jordi, and Xambó Sebastian, EACA 2012 (Proceedings), 06/2012, p.163-166, (2012)

The Patents Retrieval Prototype in the MOLTO project, Chechev, Milen, González Meritxell, Màrquez Lluís, and España-Bonet Cristina, WWW2012 Conference, Lyon, France, (2012)

Multilingual Online Generation from Semantic Web Ontologies, Dannélls, Dana, Enache Ramona, Damova Mariana, and Chechev Milen, WWW2012, 04/2012, Lyon, France, (2012)

MOLTO Enlarged EU - Multilingual Online Translation, Caprotti, Olga, and Ranta Aarne, 16th Annual Conference of the European Association for Machine Translation, 05/2012, Trento, Italy, (2012)

The GF Mathematical Grammar Library, Caprotti, Olga, and Saludes Jordi, Conference on Intelligent Computer Mathematics /OpenMath Workshop, 07/2012, (2012)

Multilingual Verbalisation of Modular Ontologies using GF and lemon, Davis, Brian, Enache Ramona, van Grondelle Jeroen, and Pretorius Laurette, Third Workshop on Controlled Natural Language (CNL 2012), Volume 7427 LNCS, (2012)

General Architecture of a Controlled Natural Language Based Multilingual Semantic Wiki, Kaljurand, Kaarel, Third Workshop on Controlled Natural Language (CNL 2012), 09/2012, Volume 7427 LNCS, p.110-120, (2012)

Probabilistic Robust Parsing with Parallel Multiple Context-Free Grammars, Angelov, Krasimir A., COLING 2012, (Submitted)

How Much do Grammars Leak?, Angelov, Krasimir A., COLING 2012, (Submitted)

The GF Eclipse Plugin: An IDE for grammar development in GF, Camilleri, John, and Angelov Krasimir, 16th Annual Conference of the European Association for Machine Translation, 05/2012, Trento, Italy, (2012)

An IDE for the Grammatical Framework, Camilleri, John, Third International Workshop on Free/Open-Source Rule-Based Machine Translation (FreeRBMT 2012), 06/2012, (2012)

Deep evaluation of hybrid architectures: Use of different metrics in MERT weight optimization, Cristina España-Bonet, Gorka Labaka, Arantza Diaz De Ilarraza, Lluís Marquez and Kepa Sarasola, Third International Workshop on Free/Open-Source Rule-Based Machine Translation (FreeRBMT 2012), 06/2012, (2012)

Comparing human perceptions of post-editing effort with post-editing operations, Koponen, Maarit, Proceedings of the Seventh Workshop on Statistical Machine Translation, June, Montréal, Canada, p.181-190, (2012)

Future activities:

- forthcoming GF tutorials: at ICFP conference in Copenhagen <http://icfpconference.org/icfp2012/> (The 17th ACM SIGPLAN International Conference on Functional Programming), at Jiao Tong University of Shanghai <http://en.sjtu.edu.cn/> in October 2012.
- Seppo Nyrkkö will give a presentation about MOLTO in [Kites](#), a Finnish national association for multilingual communication and content management.
- planned a Chinese translation of the GF book <http://www.grammaticalframework.org/gf-book/>
- agreed on the location and local organization for the third GF Summer School to take place at Fraueninsel <http://www.frauenwoerth.de/>, Bavaria, Germany, on 18-30 August 2012

Deviations from Annex I

WP11 Multilingual semantic wiki - M9

Summary of progress

Three lines of work were followed: developing a multilingual ACE grammar (ACE-in-GF), extending the AceWiki system based on the GF technology (currently referred to as AceWiki-GF) and extending the Attempto reasoner RACE.

In collaboration with UGOT (John J. Camilleri) a GF-based multilingual grammar for ACE was developed. This grammar has the following properties:

- it covers the AceWiki subset of ACE, but is easily extendable towards full ACE
- it is available in 10 languages (in addition to ACE), but is easily extendable to other languages given that they are available in the GF resource grammar library
- its performance is similar to existing ACE parsers
- it is directly usable in AceWiki-GF

This resource is fully presented in Deliverable D11.1.

AceWiki-GF was further developed by adding preliminary support for multiple grammars, multiple articles, ambiguity management, and grammar editing. A large number of AceWiki-GF demo wikis have been made publicly readable/editable on the Attempto website. Most of these wikis are based on grammars developed in MOLTO and in previous GF-related projects. Some of the simpler grammars can also be edited. The current work on AceWiki-GF and its underlying ideas were published at CNL 2012 and presented as both a talk and a demo.

The Attempto reasoner RACE is currently extended to handle arithmetic, linear equations and text problems. This work – not being part of the actual MOLTO tasks – aims at providing AceWiki-GF with an alternative reasoning capability that covers the complete first-order subset of ACE. The current, still incomplete, version of RACE was demonstrated at CNL 2012.

Organization of meetings and conferences:

- 7-9 March the 4th MOLTO project meeting was organized by [Norbert E. Fuchs](#) at the University of Zurich
- 29-31 August the 3rd Workshop of Controlled Natural Language (CNL 2012) was organized by Tobias Kuhn and [Norbert E. Fuchs](#) at the University of Zurich

Highlights

- Multilingual GF-based ACE grammar (ACE-in-GF)
- A demonstration version of AceWiki-GF is publicly available
- RACE extensions
- CNL 2012

Links to online resources

- CNL 2012 workshop: <http://attempto.ifi.uzh.ch/site/cnl2012/>
- ACE-in-GF project website: <https://github.com/Attempto/ACE-in-GF>
- AceWiki-GF demo wikis: <http://attempto.ifi.uzh.ch/acewiki-gf/>
- AceWiki-GF source code: <https://github.com/AceWiki/AceWiki/tree/gfservice>
- RACE web-interface: <http://attempto.ifi.uzh.ch/site/tools/>

Deviations from Annex I

A small deviation is expected for the deliverables of this workpackage to allow time for Tobias Kuhn's contributions, who is a leading developer of AceWiki currently on a researcher's visit abroad. The revised schedule is as follows:

- D11.2 planned for M33 (end November 2012) is postponed to M34 (end December 2012)
- D11.3 planned for M37 (end March 2013) is postponed to M38 (end April 2013).

WP12 Interactive knowledge-based systems - M9

Summary of progress

During this period two major topics were addressed from the adoption phase as described in the DoW.

A GF bootcamp was held at Be Informed at June, 4-6 in cooperation with UGOT. During this bootcamp the Be Informed team first received an in-depth introduction to grammar building using the Grammatical Framework. Building on that knowledge, several workshop sessions were held to discuss the theory and practice of (semi-)automatically converting Be Informed business modeling "language" (Be Informed meta models) and "speech" (Be Informed models) to Grammatical Framework constructs. Furthermore, discussions on a number of technical issues concerning the integration of Grammatical Framework technology into the Be Informed Business Process Platform.

Also in this period, we tried to capture requirements from a large number of perspectives. Some requirements apply to the verbalization component to be developed in WP12, but many also apply to the functionality that can be based on this component. Requirements were derived from business usage scenario's:

- Review, Validation and Feedback of Models
- Text based Editing of Models
- Self documenting Models
- Textual User Interfaces for Model Driven Applications
- Communicating Model Based Decisions Also on the non-functional part of a commercial implementation of GF in the Be Informed Business Platform requirements like portability, modularity and graceful degradation performance models.

A full overview of these requirements are presented in Deliverable D12.1.

Deviations from Annex I

none

2.4. Deliverables and milestones tables

| Number | Title | Attached files |
|--------|---|---|
| D 1.5 | Periodic Management Report T24 | D1.5.pdf |
| D 2.3 | Grammar Tools and Best Practices | MOLTO_D2.3.pdf |
| D 3.2 | D3.2 The Translation Tools Prototype | D3.2.pdf |
| D 5.2 | Description and evaluation of the combination prototypes | D.5.2.pdf |
| D 6.2 | Prototype of commanding CAS | D6.2.pdf |
| D 7.2 | Patent MT and Retrieval Prototype | D72.pdf |
| D 8.2 | Multilingual grammar for museum object descriptions | d8.2-grammars.tar.gz , D8.2.pdf |
| D9.1 A | D9.1A Appendix to MOLTO test criteria, methods and schedule | D9.1A_2012-Apr-5.pdf |
| D 11.1 | ACE Grammar Library | d11_1.pdf |
| D 12.1 | Requirements for GF based Verbalization in Be Informed | D12.1.pdf |
| D X.2 | Annual Public Report M24 | DX.2.pdf |

The only milestone due in this period is that of **WP5** and **WP3**, **Translation tool complete**, which has been met by its due date, 1 September 2012. The next milestone **MS9**, **Case studies complete**, involves the work-packages on mathematics, patents retrieval, and cultural heritage. We are delaying the work on cultural heritage and therefore we will have to shift part of this milestone too.

3. Project management during M25-M30

Project management during the period consisted mainly in maintaining the routine communication with the partners, by holding a monthly skype call, and in distributing the second installment of the funding.

UGOT received the 2nd interim payment from the EU and it distributed it to the partners on 15 August, 2012. Each partner received also the financial assessment from the EU and an overview of the payment that has been sent. The Consortium has now received 85% of the MOLTO total budget which is the maximum amount possible before the approval of the final reporting.

Followup actions after the annual review included discussions within the Consortium on how to organize a better showcase for the final results of the project and in addressing the reviewers' remarks and suggestions (see [Task 1.8](#)). Updated versions of some deliverables

were produced and made available on the website.

In terms of infrastructure, the svn repository is currently being used by a larger number of members of the Consortium and in addition there is a new bug-tracking system installed and running at UHEL.

4. Use of the resources

Tables on the usage of resources are not available for midterm reporting, however we have a rough initial estimate of persons' months by almost all nodes. Ontotext has not been able to provide the data.

Reporting period: 01-03-2012 - 31-08-2012

| Partner node | WP1 | WP2 | WP3 | WP4 | WP5 | WP6 | WP7 | WP8 | WP9 | WP10 | WP11 | WP12 | Employee |
|--------------|------|------|------|------|------|------|------|-----|------|------|------|------|----------------------|
| UPC | | | 0.25 | | 2.00 | | | | | | | | lluís.marquez |
| UPC | | | | | 4.00 | | | | 1.33 | | | | cristina.españa |
| UPC | | | | | 1.00 | | | | | | | | xavier.carreras |
| UPC | | | 0.75 | | 1.00 | | 3.75 | | | | | | meritxell.gonzalez |
| UPC | | | | | | | | | | 0.25 | | | Sebastian Xambo |
| UPC | | | | | | 3.00 | | | | | | | ares.ribo |
| UPC | 0.50 | | | | | 3.00 | | | | 0.25 | | | jordi.saludes |
| UGOT | 1.00 | | | | | 0.50 | | | | 1.50 | | | olga.caprotti |
| UGOT | | | 2.00 | | | | | | | 1.00 | | | Shafqat.Virk |
| UGOT | | 0.60 | 0.60 | | | | | | | | | | thomas.hallgren |
| UGOT | | 1.00 | | 0.50 | | | | | | | | | aarne.ranta |
| UGOT | | 0.80 | | | 2.00 | | 1.00 | | | | | 1.00 | ramona.enache |
| UGOT | | | 3.00 | | 3.00 | | | | | | | | krasimir.angelov |
| UGOT | | 3.00 | | | | | | | | 1.50 | 1.50 | | johnj.camilleri |
| UZH | | | | | | | | | | | 6.00 | | Kaarel.Kaljurand |
| UZH | | | | | | | | | | | 0.50 | | Norbert E. Fuchs |
| UHEL | | | 1.00 | | | | | | | | | | lauri.carlson |
| UHEL | | | 3.00 | | | | | | | | | | lauri.alanko |
| UHEL | | | 3.00 | | | | | | | | | | inari.listenmaa |
| UHEL | | | 3.00 | | | | | | | | | | chunxiang.li |
| UHEL | | | 3.00 | | | | | | | | | | seppo.nyrkko |
| BI | 0.10 | | | | | | | | | 0.20 | | 1.10 | Jeroen van Grondelle |
| BI | 0.15 | | | | | | | | | | | 0.75 | Joris van Aart |
| BI | | | | | | | | | | | | 1.10 | Herko ter Horst |
| BI | | | | | | | | | | | | 0.80 | Jeroen.Daenen |
| BI | | | | | | | | | | | | 0.40 | Henk.Kroeze |

Source URL: <http://www.molto-project.eu/wiki/living-deliverables/d16-periodic-management-report-t30>