



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

[Home](#) > [Printer-friendly](#) > Book > Export > Html > 2049 > D1.7A Advisory report

---

## D1.7A Advisory report

<b>Contract No.:</b>	FP7-ICT-247914 and FP7-ICT-288317
<b>Project full title:</b>	MOLTO - EEU Multilingual Online Translation
<b>Deliverable:</b>	D1.7A Advisory report
<b>Security (distribution level):</b>	Consortium
<b>Contractual date of delivery:</b>	M39
<b>Actual date of delivery:</b>	27 May 2013
<b>Type:</b>	Report
<b>Status &amp; version:</b>	Final
<b>Author(s):</b>	Stephen Pulman (University of Oxford), Keith Hall (Google Research)
<b>Task responsible:</b>	UHEL
<b>Other contributors:</b>	

### Abstract

This annex to Deliverable D1.7 is the final report written by the MOLTO Advisory Panel after attending the last project meeting in Barcelona on 23 May 2013.

## Advisory Panel Report

The original aims of the MOLTO project were to use the GF approach to provide high quality translations of texts in limited domains in real time, enabling an author to simultaneously produce versions of a text in multiple languages. These aims also included expansion to cover new languages, to enhance lexical and other resources, and to develop frameworks and training to simplify and make more productive the tasks of grammar development. Other aims included an investigation of the role of controlled languages in interaction with ontologies and other types of reasoning and knowledge representation systems, and to explore hybrid approaches to machine translation - to combine the precision of GF based translation with the coverage and robustness of

statistical machine translation methods. The project also hoped via its industrial partners to demonstrate that GF applications could be of commercial value. Three case studies were envisaged: mathematical exercises (using the Sage framework) in 15 languages, patent data in at least 3 languages, and museum object descriptions in 15 languages.

Over the course of the MOLTO project much progress has been made in achieving the above goals. One of the challenges to building and maintaining hand-built grammar systems is managing the grammar development. The MOLTO project delivers a new set of development tools ranging from the cloud-based grammar editor to an integrated development environment plugin (Eclipse). The Grammatical Framework summer school continues to train new members of the community in developing the grammars using these tools. The development of a new translation system takes a matter of days; adding a new language to a system takes hours (once the resource grammar for the language exists). The most labor intensive part of the work is developing a resource grammar, which takes on the order of 3 to 9 months; once completed it is easily exploited through the MOLTO tools and existing systems based on MOLTO technology can utilize the new language (for translation, generation, or information access).

In order to ease the development of multilingual grammars (and domain specific resource grammars) MOLTO delivers tools and techniques to expand the lexicons. One approach is to utilize translations of wordnet which maintain links across the lexical entries. This allows the grammar developer to write sense-disambiguated translation lexicons.

In the final year of the project, a robust parser for the Grammatical Framework grammars was developed. This statistical parser bridges the gap between brittle hand-coded grammars and data-driven statistical parsers. The parser is capable of generating parse fragments even when a complete analysis is not available under the defined grammar. Performance is competitive with widely used systems like the Stanford parser.

MOLTO explored a variety of applications of the rich grammar formalism of GF along with the development tools. One focus was on the application to multilingual information access: the Ontotext ontologies, the Cultural Heritage retrieval and verbalization, and ACEWiki inference. While another focus was to expand the translation capabilities of GF by exploring the integration of GF and statistical machine translation techniques. Leveraging the strong syntactic typing from GF, the GF/SMT translation system was able to perform state-of-the-art translation for patents.

One of the industrial partners, beInformed, has successfully deployed systems based on MOLTO technologies to model their information and general customer-facing documents in multiple languages. The other industrial partner, OntoText, has detailed plans to include MOLTO technologies in its product line.

We were particularly impressed by the effort devoted to evaluation, particularly of translation, but also of other MOLTO applications such as business logic modelling. For translation, the original promise of the MOLTO project was to provide high quality precise translations, within limited domains. The various tests carried out with human judges largely seem to confirm that this goal has been achieved: whereas existing commercial systems provide wider coverage than the MOLTO tools, the quality of the results is not as high. Similarly, the comparison carried out by Be Informed of the MOLTO tools against their existing solution (Velocity) seems to confirm their superiority.

Overall we believe the project team are to be congratulated on what they have achieved over the

course of the project: we regard the project as having successfully accomplished all of the goals it originally set for itself.

---

**Source URL:** <http://www.molto-project.eu/wiki/living-deliverables/d17a-advisory-report>