JIAMCATT 2012 Agenda

JIAMCATT 2012

25-27 April 2012

Luxembourg

PROVISIONAL AGENDA

Wednesday, 25 April 2012

8:30 - 9:00

Registration of participants

9:00 - 10:30

Opening of the session by Mr. Johannes van Hoof, Director for Support and Technologies of the European Parliament, chair of JIAMCATT 2012

- **Opening addresses**, Ms Janet Pitt, Director-General of Translation, European Parliament (15 min)

Item 1: Adoption of the agenda (5 min)

Item 2: Administrative issues (1 hour)

- **Report of the JIAMCATT Liaison Committee**, Ms Ulla Falk-Petersen, European Court of Auditors, Liaison Committee Chair

- **Report of the Secretary of JIAMCATT**, Ms Marie-Josée de Saint Robert, United Nations Office at Geneva

- **The JIAMCATT Wiki**, Mr Olaf-Michael Stefanov and Mr Olivier Pasteur

- **Questions on item 2**

10:30 – 11:00

Health break

11:00 - 13:00

Item 3: Working Group on standardization and interoperability of systems, including terminology systems

- **Overview of new developments**, Mr Alan K. Melby, ISO TC37, Moderator of the Working Group (15 min)

  o Recent developments in TBX: TBX (ISO 30042) is a standard for representing terminological data in an intermediate format. Five use cases for TBX will be presented and three new TBX-related developments will be described.

  o The MultilingualWeb - LT: This is a new project announced by W3C, the World Wide Web Consortium, in March, 2012. It will standardize how to add metadata to content that is to be translated. Initially, it will focus on HTML5, but the metadata system applies to other types of content markup as well.

  o ISO/TS 11669, a forthcoming document which provides guidance for translation projects. ISO/TS 11669 is in final editing and should be published by June, 2012. It describes best practices for conducting translation projects and includes a framework for developing structured project specifications.

- **The Linport project**, Mr Alan K. Melby (15 min)

  o In Alan Melby's JIAMCATT 2011 presentation there was a brief mention of the Container Project, whose goal was to define a standard way to put everything needed for a translation project into one package. At the end of the presentation, Mr. Josep Bonet of the EC-DGT made a comment about a similar project within the European Commission called the MED (Multilingual Electronic Dossier) project. He put me into contact with Mr. Tomas Carrasco-Benitez, leader of the MED project. Discussions during May, June, and July led to a merger of the Container Project and the MED project into one project called Linport (Language Interoperability Portfolio). In September 2011, the EC-DGT hosted the first international Linport symposium in Luxembourg. This presentation will describe the Linport project, in particular recent efforts to incorporate the only other project in this same space, the TIPP project of the *Interoperability Now!* initiative. The objective is to avoid multiple competing formats for translation project packages that are similar but not identical. The Linport project has already incorporated the framework of structured translation specifications in ISO/TS 11669.

- **Interinstitutional Information Sharing and Communication in the Field of Terminology**, Mr Rodolfos Maslias, Head of the Terminology Coordination Unit, DG-TRAD, European Parliament (30 min)

  o The presentation will focus on the terminology wiki, a platform of exchange between terminology coordination units of all EU institutions

participating to IATE): exchange of consolidation projects (for IATE); posting of information about conferences, trainings, meetings, etc.; supplementary material, e.g. language-specific guidelines for IATE; exchange of best practices; seminars and workshops; cooperation in trainings for terminologists.

- **The FAO Term Portal** Ms Ingrid Alldritt and Ms Maria Elisa Rech (30 min)

  o FAO will give an overview of their practical experience and lessons learned in the development of the FAO Term Portal, to create, store, manage, disseminate and maintain their terminology records. The system aims to standardize terminology in an Organization dealing with hundreds of subject areas, working in 6 official languages and more. The developments follow best terminological practices, and include a fully collaborative workflow system, user feedback functions, specialized thematic collections and interoperability with computer-assisted translation systems and other FAO systems. Many terminology management functions have been incorporated to allow for maximum flexibility in filtering, importing, exporting and thus providing services to FAO's meetings management, translation, interpretation, document production workflow and technical areas of specialization. The system also feeds large data sets into some innovation projects such as statistical machine translation (SMT), Open Archives, and provides web services for administrative and statistical systems. The Term Portal concept is especially open to partnerships with technical divisions and research organizations with the aim of acquiring a tripartite validation process, i.e. terminologist, linguist, technical expert.

- **Terminology extraction tool for interpreters, terminologists and documentalists at the ITU**, M. Nelson Verástegui (ITU) (15 min)

  o Presentation of a terminology extraction tool developed at ITU to be used by terminologists, reference clerks, translators and interpreters in their daily work. The application is very dependent of the specific software and environment used at ITU, but it could give ideas to other organizations for developing similar tools in their context. Please notice that its goal is not to find new terms but terms already existing in the ITU terminology database.

  o **Presentation of the Survey on tools for interpreters**, by Mss Anne Aboh-Dauvergne, International Criminal Court, and María Angeles Cuallado, SCIC, European Commission (15 min) [to be confirmed]

13:00 - 14:45

Lunch

14:45 - 16:00

Item 3 (continued)

- **Legal terminology management**, Ms Caroline Reichling, Court of Justice of the European Communities (30 min)

    o [*to be translated in English*] Après un bref rappel du projet «Vocabulaire juridique multilingue comparé» , la présentation portera sur la description d'un projet terminologique pilote et sur la méthodologie suivie pour gérer certaines difficultés inhérentes à la terminologie juridique et à la spécificité des textes à traduire à la DGTrad de la Cour de justice.

Item 4: Working Group on machine translation

- **Overview of new developments**, Mr Nelson Verástegui, ITU, Moderator of the Working Group (15 min)

    o [*add summary of presentation here*]

- **Machine Translation at ECA**, by Mr José Ortiz Pintor, European Court of Auditors (30 min)

    o In January 2012 the Translation Directorate of the European Court of Auditors organised a conference on machine translation with the title of "Automated Solutions in Translation, Paving the Way", the purpose of which was to raise awareness on the potential strengths and limitations of machine translation through insights given by industry experts, examples of its use in the EU policy context and a presentation of current developments within the European Commission. The conference was considered a success and its conclusions attracted immediate interest especially from translation requesters at the Court, some of whom considered MT as a powerful tool capable of immediately supplying unlimited quantities of translation. This brought the challenge of how to meet this demand and how to integrate MT in the range of services our Translation Directorate offers. On 15 March 2012 a pilot stage with the European Commission's DGT Resources Directorate started with a view to clearly understand the needs of requesters and to developing best service practices. In this context, the aim of the pilot phase is to evaluate the quality – in relation to requesters' needs – of the translations provided and to fine-tune the solutions being implemented, as well as to create the necessary procedures and workflow patterns relative to MT as a service of our DTR. The presentation will explain the Automated Solutions conference and its main conclusions and also describe the very first impressions of the pilot with the Commission to implement MT at the Court of Auditors.

16:00 - 16.30

Health break

16:30 - 17:30

Item 4 (continued)

- **Machine Translation at UNHQ**, by Ms Cecilia Elizalde, Coordinator of the UN Technology Advisory Group, United Nations New York (30 min)

  o The Spanish Translation Service of the Documentation Division in New York and the Global Databases Service (GDS) of the World Intellectual Property Organization have collaborated in an experimental project on statistical machine translation for English-Spanish. The project consisted at feeding the system already developed by GDS for the automatic translation of patents with documents translated at the United Nations. Automatic and human evaluations indicate that the tool yields very good results in terms of adequacy and fluency. The system, which will be available for the duration of the project, is being used by translators and revisers with very positive feedback. These results have prompted translators and managers at the United Nations to include machine translation in the scope of gText, the current global project of the Department for General Assembly and Conference Services.

- **The MT@EC project - Overview of the situation**, Mr Josep Bonet, Head of the IT unit, and Mr Andreas Eisele, MT@EC Project Manager, Directorate-General for Translation, European Commission (30 min)

  o In the framework of the MT@EC project, aimed at providing an automated machine translation service to the European Commission and, most likely, the other European institutions, DGT has developed, since JIAMCATT 2011, translation engines for 52 language pairs, covering all language combinations with English as source or target language, as well as some other combinations. A maturity test showed in June 2011 that 10 combinations having English as source language were deemed satisfactory. For these combinations, all incoming requests are automatically preprocessed. For the remaining combinations, a translation can always be requested manually. New versions of the engines have been developed in 2012, while the system which will dispatch translation requests, unformat and reformat documents, tokenise and detokenise text, etc. will be in pilot phase in summer 2012. The complete system is planned to become operational in summer 2013.

18:00 - 20:00

Cocktail at the Abbey

Thursday, 26 April 2012

8:45 - 9:00

Item 5: Working Group on translation support (tools and services)

- **Overview of new developments**, Mr Olivier Pasteur, WTO, Moderator of the Working Group (15 min)

    o [*add summary of presentation here*]

9:00 - 10:00

Address by Mr Miguel Angel Martínez Martínez, Vice-President, European Parliament, Focal point for multilingualism

---

10:00 - 10:30

Item 5 (continued)

- **The Need for an Open Source CAT Suite and the Olanto Project**, Mr Karim Benzineb, Olanto Foundation (30 min)

    o There are currently about sixty to seventy open source or free tools available on the Internet if we consider the combined fields of CAT tools, localization tools, dictionaries, ontologies, translation management tools and other language-related utilities. None of them offers a complete suite of translation assistance and management tools. They rather tend to solve a single problem, or to propose two or three integrated utilities. Additionally, most of them are not based on a native client-server architecture) and do not come with a professional installation, maintenance and support service. The goal of the Olanto project is to provide a complete suite of CAT tools already integrated into a natively web-based solution and to set up a network of certified partners which will be trained to provide professional installation, maintenance and support services. In summary, Olanto's mission is to streamline and professionalize the use of open source CAT tools in Translation Services.

10:30 – 11:00

Health break

11:00 - 13:00

Item 5 (continued)

- **New Toolbox at the Translation Bureau**, Mr André Guyon, Senior Language Technologies advisor, Translation Bureau of Canada (30 min)

    o In 2011, the Translation Bureau undertook consultations of its language specialists on an unprecedented scale to get a realistic global picture of their situation. The many findings included the need to

modernize the set of tools provided for translation. The 3 year project quickly developed considerable momentum, and has achieved cruising speed in under a year. Three packages have already been delivered. The components already delivered have provided improvements, and additional features are added with each deliverable. The approach adopted by the technical director of the project is to deliver quickly, deliver often and improve as you go. Help in revising, machine translation, speech recognition, dictation, ergonomics, architecture are all aspects implemented to really modernize, and even innovation. We would like to show you a little of what our people, coast to coast, can now access. We would certainly also be open to sharing similar efforts that take place in your environment.

- **The Mercury Project**, Ms Cecilia Elizalde (UN New York) and Mr Yaroslav Kurennoy (UNOV) (30 min)

  o Mercury is a web-based translation solution that the United Nations Office in Vienna is developing in the context of the gText Project of the United Nations Secretariat, a global project aimed to develop a central repository of United Nations documents and tools for language staff in all duty stations. Mercury provides translators a single entry-point to all the tools they need in their daily work, including monolingual and bitext references, terminology, computer-assisted translation and machine translation. The interface is user-friendly, which will minimize the time spent on IT training. Mercury is linked to the document management system and is accessible from any computer connected to the Internet, so it would facilitate flexible working arrangements, workload sharing among duty stations and contractual translation. It will also provide statistical information for planning purposes.

- **Modernization of the Languages Service of the United Nations Stabilization Mission in Democratic Republic of the Congo**, Ms Maria Rosaria Venosa (MONUSCO) (30min)

  o When I arrived in Kinshasa, my office had no dictionaries and an extremely poor internet connection that made it impossible to connect to the official online terminological resources. I noted that there was a significant lack of consistency within target texts: different circulars had the same acronym or expression translated in different ways. I also noticed that there was no network between the linguists of the mission and other UN entities. In order to resolve these issues, a linguistic unit from scratch had to be created. To accomplish this, the following actions were required: creation of an online translation request form to organize the workload; setting up professional rules; deontology: what a translator can/not do; establishing realistic deadlines; CAT tools market search to determine which programme would be most appropriate for the mission?; acquiring Text Bases from other UN organizations; training fellow translators at the use of CAT tools; managing and updating the tool on a daily basis. A number of problems incurred: confidentiality issues (i.e. SRSG's documents;

President's letters); wrong perception of the tool by non linguists colleagues, i.e. this is a memory tool (MultiTrans) and not an online translation services (like Google translation or similar ones). Some solutions were found, i.e. not importing confidential documents in the CAT tool and training on the difference between MultiTrans and Google translation was undertaken. Other areas of concern were also dealt with: how the CAT tool was accepted in a non computational working environment; why were some translators not using the tool? How was the tool perceived in the Mission: enemy or friend? Benefits for the mission included: increased terminological consistency within target texts; easier access to UN officially recognized terms; creation of a shared database; decreased mistranslation.

- **Interinstitutional Language Technology Watch: language technology monitoring at the EU institutions**, M. Ian Morris, Council of the European Union (30 min)

  o The presentation will briefly explain how the II LTW group was established in 2009 and show where it stands within the framework of the EU institutions' Interinstitutional Committee for Translation and Interpretation, before describing the group's working methods, outlining the technology areas covered to date and suggesting some possible next steps.

13:00 - 14:30

Lunch

14:30 - 16:00

Item 6 : Workshops

- **Machine translation: brainstorming next steps after testers' results, including common evaluation criteria**, moderated by Nelson Verástegui, ITU (Room José Ensch – Plenary)

- **Pre-translation at the European Parliament**, Mss Laura von Weissenberg, Karolina Zielinska et Mimi Renneboog, Pre-Translation / Euramis Unit, DG TRAD (Room A11 – first floor)

  o The workshop will focus on the implementation of workflows based on the systematic re-use of previously translated texts and describe how DG TRAD has moved from manual and semi-manual procedures and only haphazard use of CAT tools in 2007 to the virtually complete automation of pre-processing and consistent use of CAT tools in 2012. Taking as a starting point an explanation of the Euramis workflows, which are based on the separation of the storage and the working memory, the tools developed for facilitating the manual pre-treatment of translation requests, the tools suite applied for automation purposes and the actual functioning of automation will be the object of a detailed

presentation. The 'political' aspects of the project will be also addressed: the need to overcome resistance to change by setting mandatory upload targets, investing in training and establishing and implementing safe working protocols for the use of CAT tools.

- **Tools for the Arabic language**, moderated by Ms Simar Proust, ILO (Room A13 – first floor)

16:00 - 16:30

Health Break

16:30 - 18:00

Item 6 (continued): Workshops

- **The UN global terminology project - outcome and challenges**, moderated by Ms Danielle Henripin and Mr. Michal Ziemski, United Nations, New York (Room José Ensch – Plenary)

- **The Desktops: workflow management systems (DGT)**, moderated by Mr Josep Bonet, DGT, European Commission (Room A11 – first floor)

    o As a follow-up to the JIAMCATT 2010 presentation on Mandesk, an information system for translation unit and translation demand managers, DGT will present the progress made so far. The system is scheduled to go live in the second half of 2013. An advanced prototype of the workflow system will be shown, including the Calendar function, already operational and allowing the manager to visually identify the workload and availability of his/her collaborators. Tradesk, the tool for translators which grants simple and seamless access to all documents related to translation dossiers, will also be shown. Tradesk has been operational since 2010.**Recap of global terminology project (history/structure, workshops and their outcome); presentation of phase I outcome: UNTERM global search portal (with a bit of technical data); summarize the development/testing process, with acknowledgment of support from JIAMCATT community; analyze the impact on local/global terminological practice; emphasize the importance of communications/collaboration/coordination (and how we ensure those aspects); present some of the lessons learned, as well as the road ahead and the challenges we anticipate in phase II.

Friday, 27 April 2012

8:30 - 10:30

Summary of workshops and discussion

10:30 - 11:00

Health Break

11:00 - 13:00

Closing of the Plenary

- **Feedback on the session**
- **Agenda for JIAMCATT 2013**
- **Closing address**

13:00 - 14:30

Lunch break

14:30 - 17:30

Brief presentations by exhibitors of their products and discussion with participants

- MultiCorpora
- SDL
- Terminotix
- Transperfect
- Wordbee