IN THIS ISSUE:

...about CELCT
Cross Language Evaluation Forum
PatExpert
Multimodal Communication
Visitors to CELCT
ontoPrivacy
Italian Benchmark
EUROMATRIX
Document Understanding Conferences
Spoken Language Translation
Recognizing Textual Entailment
Jointly Organised Events
Upcoming Events and Initiatives...
CELC, the Center for the Evaluation of Language and Communication Technologies, is a joint enterprise established by ITC-Irst (Istituto Trentino di Cultura- Centro per la Ricerca Scientifica e Tecnologica) and DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz), with funding by the Autonomous Province of Trento.

The Center for Scientific and Technological Research (ITC-Irst) was founded in 1976. The Center conducts research in the areas of Information Technologies, Microsystems, and Physical Chemistry of Surfaces and Interfaces.

The German Research Center for Artificial Intelligence (DFKI), with sites in Kaiserslautern and Saarbrücken is the leading German research institute in the fields of Pattern Recognition, Knowledge Management, Intelligent User Interfaces, Deduction and Multiagent Systems, Language Technology, Intelligent Visualization, Robotics, Safe and Secure Cognitive Systems.

Currently, CELCT has a full-time staff of seven people, comprising researchers and administrators. Amedeo Cappelli is the Scientific Director. The Board of Directors of CELCT includes: Fabio Pianesi (President), Gianni Lazzari from ITC-Irst, and Professor Wolfgang Wahlster, from DFKI.

Periodically some researchers from ITC-Irst and DFKI cooperate with CELCT, moreover students can choose to join CELCT for funded internships.

CELC was founded with the motivation that evaluation plays an essential role in advancing and assessing technology by improving software quality. Systems that have participated in some evaluation campaigns, such as, for instance, the Text Retrieval Conferences (TREC), organized by the National Institute for Standards and Technology (NIST) since 1992, have doubled their performance. As a matter of fact, most commercial Web search engines nowadays include technologies first developed for the TREC annual evaluation conferences.

The activity of evaluation must have a strong relationship with the research community and technology developers, but, at the same time, it must be independent from them, in order to guarantee the objectivity of the results. This was another strong motivation which convinced ITC-Irst and DFKI to create a new independent center by allocating to it resources and expertise accumulated during many years of activity in this field.

CELC was also intended to operate at the European level, by offering a global service for the evaluation of Human Language and Multimodal Communication Technologies both for the research community and software producers.

The mission of CELCT is to set up infrastructure and develop skills in order to operate successfully in the field of the evaluation. In particular, its goal is to study and devise evaluation procedures and metrics for evaluating human language and multimodal communication technologies such as, for instance, question answering, speech-to-speech translation, spoken language technologies, cross-language information retrieval, word sense disambiguation, web page accessibility and usability, etc.

Other aims of CELCT are the promotion and the organization of evaluation campaigns, the release of guidelines and the creation of benchmark resources.

In pursuing its goals, CELCT collaborates with many national and international research organizations and networks of excellence.

CELC offers also services for private clients, both academic and industrial, for the evaluation of their products.

Current and recent activities of the Center are described briefly in this newsletter, the purpose of which is to provide an overview and introduction of CELCT.

CONTACT: Veronica Giordani - Secretary
vegiordani@celct.it
+39 0461 405300
The Cross Language Evaluation Forum (CLEF), www.clef-campaign.org, supports global digital library applications by developing an infrastructure for the testing, tuning and evaluation of information retrieval systems operating on European languages in both monolingual and cross-language contexts, and creating test-suites of reusable data which can be employed by system developers for benchmarking purposes. CLEF aims also at establishing strong links, exchanging ideas and sharing results, with similar cross-language evaluation initiatives in the US and Asia, working on other sets of languages. The final goal is to assist and stimulate the development of European cross-language retrieval systems in order to guarantee their competitiveness on the global marketplace.

In particular, the following activities have been performed:

- the creation of guidelines for the generation and evaluation of the test sets;
- the coordination of the exercise for the competition;
- the generation of the Italian test sub-sets;
- the evaluation of the results of the participating systems;
- the maintenance of the evaluation campaign website;
- the authorship of the final summary, including the editing of the workshop proceedings.
- the organisation of the QA sessions at the workshop.

Moreover, CELCT has edited the CLEF proceedings published by Springer in LNCS (Lecture Notes in Computer Science) series since 2004.

The Center is also a sponsor of the annual conference together with the Network of Excellence for Digital Libraries (DELOS), Microsoft, the European Language Resources Association (ELRA) and Evaluation & Language resources Distribution Agency (ELDA).

For more information, see: clef-qa.itc.it/

CONTACT: Pamela Forner
forner@celct.it
+39 0461 405355

PatExpert

In late 2006, CELCT began collaboration with ITC-irst as a sub-contractor for the EU STREP project “Advanced Patent Document Processing Techniques” (PatExpert), whose partners are: Fundació Barcelona Media, Universitat Pompeu Fabra (Spain); Universität Stuttgart (Germany); Aristotle University of Thessaloniki (Greece); ITC.irst Center for the Scientific and Technological Research (Italy); Industrie Software Jochen Brügmann (Germany), Fraunhofer Gesellschaft (Germany), IALE Tecnología (Spain), and the European Patent Office (Germany).

This project overall focus is to change the approach to patent processing from textual to semantic. As described in the literature, “The project will develop a multimedia content representation formalism based on Semantic Web technologies for two technology areas and investigate the retrieval, classification, multilingual generation of concise patent information, assessment and visualization of patent material encoded in this formalism, taking the information needs of all user types as defined in a user typology into account.”.

CELCT has been participating in this project and its role is to carry out linguistic annotation of patent texts for both part-of-speech and phrase-level chunking.

For more information, see: www.patexpert.org

CONTACT: Cameron Shaw Fordyce
fordyce@celct.it
+39 0461 405332
Multimodal communication

CELC propose evaluation services and consulting for the design of web interfaces. The main aim of this activity is to propose an evaluation system which guarantees products that can be accessible and used by everyone.

For the planning stage, we use an approach focused on the user (User Centred Design – UCD) in which the user plays an important role not only in the formation of system requirements, but in all subsequent steps of development.

CELC’s evaluation activity is developed in five steps in which the end user is always first and foremost:

- requirements analysis: informational meetings with the purchaser and the system’s users;
- task analysis: the user actions are divided into hierarchically ordered tasks;
- preliminary planning: mock-up development, where the results of the task analysis are translated into a mock-up, i.e. a sample of the system before the real planning;
- formative evaluation: this step involves identification of the aspects of the site which can be improved and guidelines are supplied which explain the proposed improvements;
- summative evaluation: an evaluation of the resulting system’s quality, with the intention of understanding if the design has reached the desired goals.

During 2006 CELC completed studies of two web interfaces: the portal for the “Cassa del Trentino”, a new public financial institution, which will manage all the on-line activities; and the portal for Ontotext, a tool for knowledge-based automated word research, including the identification and function allocation, which has been developed by ITC-irst.

Figure 1. Example of Hierarchical Task Analysis (HTA).

CONTACT: Valentina Bruseghini
bruseghini@celct.it
+39 0461 405333

Visitors to CELCT

Figure 1. From the left: Carol Peters (CLEF Chair Person), Donna Harman (NIST-National Institute of Standards and Technology) at CELC’s inaugural ceremony.

Figure 2. Kerstin Klöckner (DFKI) during the Eyetracking Tutorial.

Figure 3. John S. Garofolo (NIST-National Institute of Standards and Technology) presenting at AISV 2006.
ontoPrivacy is a specialized ontology developed by CELCT in collaboration with the Institute of Legal Information Theory and Techniques of CNR-National Research Council (Project: Data Mining, Ontologies and Semantic Web), for the modeling of a glossary of keywords extracted from the Italian Personal Data Protection Code (D.lgs. 196/2003).

This glossary has been manually created by an expert in the legal field and it is made up of both specific terms of the Public Administration domain (e.g. atto amministrativo/administrative act) and generic words (e.g. razza/race). Synonymy, hyponymy and hypernymy relations are identified among such terms (e.g. administrative act and judicial act are kinds of act).

Taking the relevant terms contained in the glossary as starting points, we have followed a bottom-up approach to create the ontology. The classes in ontoPrivacy have been defined by analysing the lexical entries of our glossary and the relations among them. Five main classes have been identified: (i) Events: something that happens at a given place and time; (ii) Scenes: the place and time where an event occurs; (iii) Non-Physical Objects: mental objects, like conceptualizations; (iv) Physical Objects: things that are observable and can be experienced through sensory input; (v) Roles: functions played by individuals and groups.

The idea for the future is to evaluate ontoPrivacy trying to apply the same conceptual organization to new regulated areas.

CONTACT: Rachele Sprugnoli
sprugnoli@celct.it
+39 0461 405374

I-CAB (Italian Content Annotation Bank), created thanks to the collaboration between CELCT and ITC-irst, is a corpus of Italian news annotated with semantic information at three different levels: temporal expressions, entities (i.e. person, organizations, locations and geo-political entities), and relations between entities (e.g. the affiliation relation connecting a person to an organization). So far I-CAB has been manually annotated with the first two levels, while the third one is a work in progress.

As we intend I-CAB to become a benchmark for various automatic Information Extraction tasks, we followed a policy of reusing already available mark-up languages. In particular, we adopted the annotation schemes developed for the ACE (Automatic Content Extraction) Entity Detection and Recognition and Time Expressions Recognition and Normalization tasks.

For our purposes, these standards turned out to be adequate, as they allow for a semantically rich and normalized annotation of: (i) different types of entities (i.e. objects or set of objects in the world), (ii) different types of entity mentions (i.e. any textual reference to an entity), (iii) different types of temporal expressions (e.g. absolute expressions, such as “March 13 2005”, and implicit expressions, such as “three days later”). We also follow the guidelines provided by the Linguistic Data Consortium (LDC) that develops linguistic resources to support the ACE program.

As the ACE-LDC guidelines have originally been developed for English, part of the effort consisted in adapting them to the specific morpho-syntactic features of Italian. We have also extended them to include a wider range of entities.

I-CAB consists of 525 news documents (around 182,500 words) taken from the local newspaper “L’Adige”. For the creation of I-CAB we have chosen the freely distributed annotation tool Callisto, developed at the MITRE Corporation.

In I-CAB, the manual annotation of the corpus is merged with automatic annotation of lower linguistic levels (tokenization, lemma, PoS, multi-words). All the different levels of annotations are delivered in the Meaning Annotation Format (MAF), an XCES and TEI conformant scheme which was developed within the EU-funded MEANING project and has now been extended to represent temporal expressions and entities.

CONTACT: Valentina Bartalesi Lenzi
bartalesi@celct.it
+39 0461 405352
This year CELCT joined three European universities, Edinburgh University (Scotland), Charles University (Czech Republic) and Saarland University (Germany), and two private companies MorphoLogic (Hungary) and GROUP Technologies AG (Germany) in beginning the EUROMATRIX project, contract number FP6-IST-5-034291-STP. This European Union FP6 research project will address European economic and social needs by aiming at the automatic translation of technical, social, legal and political documents across all pairs of EU languages, including the languages of new member states.

EUROMATRIX partners will research and apply the most advanced machine translation (MT) technologies systematically to all European language pairs to achieve this goal. For most translation directions, that involve languages of the new and near-term prospective member states, the project will provide machine translation systems for the very first time. For language pairs where MT systems already exist, EUROMATRIX will design and investigate novel combinations of statistical techniques and linguistic knowledge sources as well as hybrid MT architectures in order to combine the accuracy of rule-based techniques with the adaptivity of data-driven approaches.

CELC'T's role in this project is an important one in that the Center will provide all the evaluation services to the project partners and other research groups both industrial and academic and conduct analysis and interpretation of the results of the two scheduled competitive annual international MT evaluation campaigns. CELCT hopes to advance the state of the art in both machine translation and its evaluation by conducting additional research in evaluation procedures and automatic metrics and will also provide an up-to-date survey of machine translation evaluation metrics and methods in order to support the international research community.

For more information, see: www.euromatrix.net/

CONTACT: Cameron Shaw Fordyce
fordyce@celct.it
+39 0461 405332

Document Understanding Conferences

CELC'T was involved in the 2005 edition of the Document Understanding Conferences (DUC) in the role of evaluating automatically created document summaries. In particular, CELCT participated in testing a new manual method for summarisation evaluation called "Pyramid Method", which was developed in an attempt to address a key problem in summarisation: namely, the fact that different humans choose different content when writing summaries.

The pyramid method addresses the problem by using multiple human summaries to create a gold-standard and by exploiting the frequency of information in the human summaries in order to assign importance to different semantic units (SCUs). The SCUs are then organized in a 'pyramid', which has at the top the most important information, i.e. that information which is present in all manual summaries.

The SCUs at the top are assigned the highest 'weight'. Through an annotation interfaces, the automatic summaries themselves are then divided into SCUs which are matched to the ones in the pyramid. In this way, they get scored, proportionally to the weight of the pyramid SCUs. CELCT will participate in the DUC2007 summarization campaign as a contributor to the creation of the pyramid and the annotation of the automatic summaries.

For more information, see: duc.nist.gov/
The past year saw the entry of CELCT into the organization of another major international campaign of evaluation in the field of machine translation.

Current members include ATR (Japan), CAS (China), CLIPS (France), CMU (USA), ETRI (Korea), ITC-irst (Italy), and UKA (Germany).

CELCT participated in the 2006 edition of IWSLT through the preparation of training and test text data for the Italian to English task for C-Star partner ITC-irst. For the 2007 edition, ITC-irst will be organizing the overall workshop and CELCT will be organizing the evaluation of the participant systems.

For more information, see: iwslt07.itc.it/

CONTACT: Cameron Shaw Fordyce
fordyce@celct.it
+39 0461 405332

Textual Entailment Recognition (RTE) is the task of deciding, given two texts, whether the meaning of one (H) can be plausibly inferred from the other (T).

Under the aegis of PASCAL, a Network of Excellence under the 6th EU Framework Program, www.pascal-network.org, RTE challenges have been carried out since 2004, gaining growing popularity among the NLP community.

Elements of this success are related to the fact that textual entailment may serve as a unifying generic framework for applied modeling of semantic inference, which captures generically a broad range of inferences that are relevant for different application, such as Question Answering (QA), Information Extraction (IE), Summarization, Machine Translation (MT), paraphrasing, and for certain types of queries in Information Retrieval (IR). More specifically, the RTE challenges aim to focus research and evaluation on this shared underlying semantic inference task and isolate it from other application specific problems.

In 2006, the Center began the organisation of the 2007 campaign, in collaboration with the previous campaigns’ partners. In particular, CELCT has been coordinating:

- the extraction of fragments of text (T) from the output of real NLP systems or from manually collected copyright-free documents (such as articles from Wikipedia and Wikinews);
- the creation of pairs where each (T) has been linked with a hypothesis (H), presenting some type of relations with it;
- the manual annotation of the pairs by at least three human assessors.

A development set was released in December, 2006. The actual exercise will take place in March 2007. The results will be automatically produced by comparing the systems output with the human assigned judgements. CELCT will also carry out the evaluation and the editing of the participants’ papers.

For more information, see: www.pascal-network.org/Challenges/RTE3

CONTACT: Danilo Giampiccolo
giampiccolo@celct.it
+39 0461 405354
**Jointly Organised Events**

**EYE TRACKING TUTORIAL, Trento, 8-12 November 2005:** CELCT has also promoted the development of local capabilities in the evaluation of interfaces utilizing Eye Tracking systems. In collaboration with DFKI, a week-long tutorial was held by Kerstin Klöckner of DFKI in the use of Gaze Tracker, a software package for the analysis of eye movements.

For more information, see: [dfki.de/~kekl/](http://dfki.de/~kekl/)

**EACL 2006, Trento, 3-7 April 2006:** CELCT sponsored the Multi-lingual Question Answering (MLQA'06) workshop at EACL 2006, which addressed a number of challenging issues for research in multilingual QA, including searching in multilingual document collections, collecting and merging answers found in documents from different languages, using heterogeneous multilingual data collections (such as the Web, XML data or data in relational databases) for answer generation, and interpreting questions in different languages.

For more information, see: [mlqa06.program.html](http://nlp.uned.es/MLQA06/program.html)

**IWSLT 2006, Kyoto, 27-28 November 2006:** CELCT was happy to be among the supporting organizations of the 2006 edition of the International Workshop of Spoken Language Translation (IWSLT), an annual spoken language translation workshop organized by the C-Star Consortium.

For more information, see: [www.slc.atr.jp/IWSLT2006/](http://www.slc.atr.jp/IWSLT2006/)

**QA@CLEF 2006, Alicante, Spain, 20-22 September 2006:** CELCT, together with ITC-irst, organized and chaired the QA main and break-out sessions, and presented an overview of the QA track.

For more information, see: [clef-qa.itc.it/2006bis/CLEF-2006.html](http://clef-qa.itc.it/2006bis/CLEF-2006.html)

**AISV 2006, Trento, 29 November - 1 December 2006:** CELCT contributed to the organisation of AISV 2006, the annual conference of AISV (Associazione Italiana di Scienze della Voce), the Italian association for speech sciences. CELCT is actively involved in the administrative and secretariat management of AISV.

For more information, see: [aisv2006.itc.it](http://aisv2006.itc.it)

**Upcoming Events and Initiatives**

**Second Workshop on STATISTICAL MACHINE TRANSLATION, Prague (at ACL 2007), 23 June 2007:** This workshop focuses on statistical and hybrid methods for machine translation and includes a shared translation task. The evaluation of machine translation systems is a growing field and this workshop will also focus on determining the best methodology for evaluating translation quality with automatic metrics and subjective human evaluation.

For more information, see: [www.statmt.org/wmt07/](http://www.statmt.org/wmt07/)

**PASCAL Workshop on Textual Entailment and Paraphrasing, Prague (at ACL 2007), 28-29 June 2007:** For the first time one workshop will combine recognizing and generating textual entailment and Paraphrase, two technologies which address similar aspects of natural language semantics, in order to facilitate the convergence of efforts in such a promising field of research in Natural Language Processing and the exchange of ideas among a wider audience.

For more information, see: [nlp.cs.nyu.edu/WTEP](http://nlp.cs.nyu.edu/WTEP)

**EVALITA, Rome (at AI*IA 2007), 10 September 2007:** A new initiative, coordinated by CELCT and ITC-irst, devoted to the evaluation of Natural Language Processing tools for Italian. The general objective of EVALITA is to encourage the development of language technologies for the Italian language, providing a shared framework where different systems and approaches can be evaluated in a consistent manner.

For more information, see: [evalita.itc.it](http://evalita.itc.it)

**QA@CLEF 2007, Budapest (at ECDL 2007), 19-21 September 2007:** The workshop will propose a main task scenario, i.e event-targeted Question Answering on a heterogeneous document collection (news articles and Wikipedia), and many sub-tasks, both monolingual - where the questions and the target collections searched for answers are in the same language - and bilingual - where source and target languages are different. A number of pilot tasks are under consideration. The track is organized by several institutions (one for each source language) and jointly coordinated by CELCT and UNED, Madrid (ES).

For more information, see: [clef-qa.itc.it/](http://clef-qa.itc.it/)

**IWSLT 2007, Trento, 15-16 October 2007:** CELCT is organizing the 2007 International Workshop of Spoken Language Translation together with ITC-irst. The IWSLT workshop is an annual spoken language translation workshop and evaluation challenge organized within the framework of the C-Star Consortium. For more information, see: [www.slc.atr.jp/](http://www.slc.atr.jp/)

**CELCT’s CONTACT:**

Via Sommarive 18, 38050 Povo di Trento - Italy  
Tel: +39 0461 405300  
Fax: +39 0461 405372  
[www.celct.it](http://www.celct.it); [info@celct.it](mailto:info@celct.it)