



MT News International

Newsletter of the International Association for Machine Translation

In this issue . . .

Spotlight on the News	1
Products	3
Conference Reports	5
Conference Announcements	9
Resources	12
Calendar	15

Of special note . . .

- **Mendez sale finalized**
(page 1)
- **Reports on ACL/EACL
main conference
(page 5) and the
Human Language
Technology and
Knowledge
Management workshop**
(pages 6)
- **Interview with LISA
Managing Director,
Mike Anobile** (page 7)

Spotlight on the News

Bowne edges out Lionbridge in final auction of Mendez

[Adapted from Press Releases]

Bowne & Co., Inc. (NYSE: BNE), the world's largest financial printer and a leader in globalization services through its Bowne Global Solutions division, announced its agreement with Belgium-based Lernout & Hauspie Speech Products N.V. to acquire Mendez S.A. and related assets, for the purchase price of \$44.5 million on August 2nd.

Bowne cleared the final hurdle to complete the acquisition when the Commercial Court in Ieper, Belgium approved the acquisition on August 16th. This followed the U.S. Bankruptcy Court in Delaware's approval of the acquisition on August 7th, after a court-mandated auction held on August 2, 2001 in which Bowne was the successful bidder. *(While MT community members may not be very familiar with Bowne & Co., it has been actively developing a globalization & localization business. Bowne Global Solutions, since*

1997, primarily through the acquisition and consolidation of a select group of international localization and translation agencies. Standard and Poors estimates that globalization represents 6% of Bowne & Co.'s US\$1 billion a year revenues. -Ed.)

"With the approval of the Belgian court, we can now look forward to officially welcoming the Mendez employees into the Bowne and especially the Bowne Global Solutions family," said Robert Johnson, chairman and chief executive officer of Bowne & Co., Inc. "We are very pleased to have this important step behind us so we can now focus on finalizing the transaction and moving forward with our new team. We are excited by the prospect of combining these two very complementary businesses -- together they will be an even stronger company with a more complete offering of globalization and localization services."

Continued on page 14 ►



MT News International

Issue No. 28 (vol. 10, no. 1)

Editorial Board

EDITOR-IN-CHIEF:

Laurie Gerber

Fax: +1 (619) 226-6472

E-mail: mtni@eamt.org

CONSULTING EDITOR:

John Hutchins

E-mail: info@eamt.org

CONTRIBUTING EDITOR:

Colin Brace

Fax: +31 (20) 685-4300

E-mail: webmaster@eamt.org

REGIONAL EDITOR, AAMT

Hitoshi Isahara

Fax: +81-774-95-2429

E-mail: isahara@crl.go.jp

REGIONAL EDITOR, AMTA

David Clements

Fax: +1 (619) 490-3684

E-mail: David.Clements@lhsl.com

REGIONAL EDITOR, EAMT

Jörg Schütz

Fax: +49 (681) 389-5140

E-mail: joerg@iai.uni-sb.de

Autodesk Implements Localized Tech Sup- port with SYSTRAN

Time magazine reported in July that Autodesk, a U.S. CAD software developer (www.autodesk.com), is now offering customer support documents translated into French and Spanish by SYSTRAN.

The pilot project is expected to cut customer support costs in half while improving support to European customers. The implementation is ideal for MT because Autodesk has roughly 5,000, regularly updated technical support documents, made available to customers who visit the company's self-serve technical support website. Historically, this feature was only available to English-speaking customers, requiring.

For more information, or to view the Time article, see www.systransoft.com



LogoMedia partners with Lingvistica

[Adapted from press releases]

In June of this year, LogoMedia Corporation introduced its first language translation products for the language pairs: English to and from French, German, Italian, Spanish, Portuguese, Russian, Chinese, Korean, and Japanese. The company also disclosed plans to roll out additional translation services over the coming weeks, including several online versions of their translation technology. These will include a low-cost individual service with links to human translation options, and a real-time, two-way translation service for web-based applications, such as CRM, online shopping, etc.

In July, LogoMedia Corporation an-

nounced the formation of a strategic partnership with Lingvistica (Canada & Holland - www.ling98.com) for the sale and development of Slavic language machine translation systems. Already available in LogoMedia's TRANSLATE™ package and online translation service (www.logomedia.net) is English to/from Russian from Lingvistica (commercially known as PARS). Currently under joint development are English to/from Ukrainian and English to/from Polish. The two companies plan to develop machine translation systems for additional Slavic languages in 2002. The TRANSLATE package brings together a selection of MT engines that will provide the user with coverage of the major language pairs. Some of the language pairs are based on LEC technology, others on Barcelona technology, in addition to the newly-added Lingvistica technology.

About LogoMedia Corporation

Established in 1999, privately held LogoMedia Corporation is headquartered in Belmont, Massachusetts with affiliates in Melbourne, Kiev, Beijing, Seoul, Taiwan, Buenos Aires, Tokyo and Amsterdam. LogoMedia, Corporation provides quality and affordable software and subscription ASP services that generate automatic language translations of web pages, files, e-mails and electronic documents via the Internet and desktop.

The company is headed by Dr. Glenn Akers, President, CEO and co-founder of LogoMedia. Dr. Akers provides veteran industry leadership for the new venture. He co-founded and served as CEO of Language Engineering Corporation, an R&D company that develops patented automated translation software for Japan and other global markets. He also pioneered language technology at Dialog Systems, Verbex, Bell-Northern Research, Kurzweil AI, LogoVista and Digital Equipment Corporation. Joining him on the management team is Richard Strauss, Chief Operating Officer. Strauss was co-founder and COO of TechOnLine, an Internet company serving the information needs of engineers.

For more information see: www.logomedia.net or contact Richard Strauss, COO, 617-489-4000 ext. 726, rstrauss@logomedia.net □

TSRALI—A Valuable Web Resource for Translators from the University of Montreal

[press release]

The RALI Laboratory at the University of Montreal has just announced the launch of TSrali.com at: www.TSrali.com, an enhanced version of its popular, Web-based, bilingual concordancer.

Users familiar with the original *Trans-Search* service will notice improvements in response time and in the amount of text available. The Hansard database of parliamentary proceedings has been updated to cover 1986 to the present, and a judicial database has been added containing selected decisions from the Supreme Court of Canada, the Federal Court of Canada and the Tax Court of Canada. Together, the two databases total more than 200 million words of English-French translation.

The system functions as an interactive, contextually aware, bilingual dictionary of actual translations. A user faced with a translation problem submits a query, which may be in the form of a word, phrase, or some other expression. The system searches its database for all matching occurrences of that query and displays them in their sentential context, together with their translations. Each of these represents a potential solution to the original problem, as previously formulated by another translator. Flexible search options allow matching of word-variants (eats/eating/ate etc.) and patterns ("take" followed by "for a ride" in the same sentence).

TSrali.com is a subscription service, affordable by individuals and offering substantial group discounts. A free 5-day trial of the system is also available.

For further information consult www.TSrali.com. For an overview of other natural language technology available from RALI, see: www-rali.iro.umontreal.ca. □

Carnegie Mellon "Pocket" Translator

On June 14, the New York Times reported on the fielding of an experimental portable speech-to-speech translator developed at Carnegie Mellon University. The system, which runs on a small laptop, supports simple question and answer dialog between Croatian and English with a vocabulary of 5,000 to 10,000 words for each language. The laptop is small enough to fit in the cargo pocket (8 by 10 inches) of an Army uniform. The project was funded by the U.S. Army. The language software was developed at Carnegie Mellon's Language Technology Institute by a team led by Prof. Robert Frederking. Lockheed Martin provided the system integration and interface. The system was particularly developed for Army Chaplains who are frequently in the field with a shortage of human translators.

For further information, contact Robert Frederking Robert_Frederking@lti.cs.cmu.edu □

Products

Breakthrough in Interspecies Communication!

News scientist.com reports that Takara Co., a Japanese Toy manufacturer, will release a Dog-to-Japanese translator next year. The product, consists of a microphone-equipped dog collar/transmitter and a receiver/interpreter which translates the dogs utterances into some 200 words and sentences. The "Bow-lingual" translator will be released in Japan in February 2002. Future enhancements will include the ability to transmit translations directly to mobile phones. □

ALIS INTEGRATES THE IBM WEBSHERE TRANSLATION

[Based on a press release]

Alis Technologies Inc. announced on August 13 that it will incorporate the IBM WebSphere Translation Server into Gist-In-Time's server farm delivering on-line language comprehension solutions to B2B exchanges.

The Gist-In-Time server farm is currently used by clients like Netscape (Netscape 6.1 browser), Sympatico-Lycos (the largest French consumer portal in North America), Copernic.com (the search tools suite Copernic 2001 with over 15 million users worldwide) and BellZinc (Canada's largest B2B destination for small and medium size business), and gives visitors and users "gists" of Web pages. A gist allows users to comprehend electronic documents such as Web pages in their preferred language in just seconds.

The IBM WebSphere Translation Server will provide Alis' Gist-In-Time Internet solution with the following language pairs: English to/from Italian, English to German, English to Japanese, English to Simplified Chinese, English to Traditional Chinese and English to Korean.

The agreement is an extension of a relationship announced in March when Alis announced that it was offering the IBM WebSphere Translation Server to large corporations (Global 2000) wishing to store the solutions on their own servers. The IBM WebSphere Translation Server offers real-time automated translation for 12 language pairs and a module that lets organizations build specialized dictionaries.

For further information see: www.alis.com or www.ibm.com/software/voice, or contact Thierry Gauthier, Director, Marketing and Communications, Alis Inc. E-mail: tgauthier@alis.com, Tel: 514-747-2547 x.227; or Geraldine Kan, Public Relations, IBM Voice Systems, E-mail: gkan@us.ibm.com, Tel: 914-766-1715. □

SDL Localization Suite

[Based on a press release]

SDL International(R), a leading provider of globalization products and services, announced SDL Localization Suite on August 29. SDL Localization Suite offers localization agencies significant cost and time savings, giving them a seamless set of products to assist in the translation and QA process for software, document translation and editing through to test validation.

SDL Localization Suite is targeted at sophisticated localization agencies and corporate localization departments that want to make the best use of their resources as well as make the localization process more efficient. This unique localization solution provides products for key members of the localization group, including project managers, software engineers, localizers, translators, authors and QA engineers enabling them to improve productivity. The SDL Localization Suite includes Translation Memory product SDLX(R); software editor SDLinsight; machine translation product Transcend(R); and testing tools HelpQA, HtmlQA and ToolProof, the user interface bug detector.

Tools & Features

SDLX - Translation Memory tool that extracts previously translated terms and phrases from a database for re-use. SDLX can reduce translation time by up to 40 percent.

Transcend - machine translation engine that provides an instant first-pass translation of text. The engine can be refined with use to improve accuracy over time.

SDL TermBase - a terminology database of concepts that are represented in one or more languages.

HelpQA - an RTF tool that provides testing validity of help files.

HtmlQA - an HTML tool for testing the validity of HTML-based online help files.

SDLinsight - an advanced localization environment for Windows(R) resource

scripts and 32-bit Windows executables.

ToolProof - an automatic user interface bug detector.

With the purchase of the SDL Localization Suite, customers are entitled to a one-year membership in the SDLX Localization Premier Club, which provides discount benefits.

Language Support

The SDL Localization Suite Translation Memory and QA tools support any language including double-byte and bi-directional languages. Transcend machine translation supports the following language directions:

English to French English to German French to English English to Spanish English to Italian Spanish to English English to Portuguese Portuguese to English German to English

The suite is priced at US\$2,495.00 and will be available in September 2001. The SDL Localization Suite runs on Windows 95, 98, NT and 2000.

For more information see: www.sdlintl.com, or contact Suzanne A. Bellemore, SDL International, Tel: 603-262-6336, E-mail: sbellemore@sdlintl.com □

IBM Websphere Transcoding Publisher

[Based on a press release]

IBM WebSphere Transcoding Publisher is a server-based, software that dynamically translates Web content and applications into multiple markup languages and optimizes it for delivery to mobile devices, such as mobile phones, and handheld computers. This award winning solution helps to bridge data across multiple formats, markup languages and devices. Transcoding Publisher adapts, reformats, and filters content to make it suited for pervasive computing, giving companies get better access to customers, business partners and mobile employees on a variety of devices. IBM WebSphere Transcoding Publisher allows organizations to communicate with their mobile customers and employees using information that already exists in their information sys-

tems. There is no need to create and maintain multiple sets of Web pages and since transcoding is done in the real time changes to the original content are automatically reflected in all available formats. Version 4.0 was announced by IBM on August 31.

What is new in Version 4.0

Additional transformation functions allowing to further extend Web content include: HTML to VoiceXML, which makes Web pages accessible from a voice-only phone; HTML to PalmOS HTML for Palm.Net devices; Machine Translation transcoder, which transforms content into ten languages when working with Machine Translation Server.

Additional/enhanced content customization capabilities include: XSL Stylesheet Editor with robust search and sort management capabilities; External Annotation Editor; Capability to import user preferences from an external repository; Improved administration and monitoring tools; Centralized administration through external data sources, like LDAP; Request viewer for remote monitoring of WTP server deployed in a proxy or reverse proxy configuration;

Other enhancements include: Integration with WebSphere Edge Server Caching Proxy, which provides native cache support, SSL support, and HTTP 1.1 support; Automated migration from previous version; Significant performance improvements; WebSphere Application Server support & WebSphere Everyplace Suite integration.

For more information, visit www.ibm.com. □

SYSTRAN Personal for Linux

[Based on a press release]

Editor's note: The press release that follows appeared only in French. The version here was translated using SYSTRAN F->E on Babelfish.

Continued on page 14 ►

Conference Reports

ACL-EACL 2001

Pierre Isabelle

The 39th meeting of the Association for Computational Linguistics (the ACL) was held jointly with the 10th Conference of the European Chapter of the ACL in Toulouse (France) between 6-11 July 2001, with Bonnie Lynn Webber (Univ. of Edinburgh) serving as General Conference Chair and Patrick St-Dizier (Université Paul Sabatier, Toulouse) serving Local Arrangements Chair.

The overall schedule was a bit unusual: workshops first, then tutorials and finally main conference. Those who wanted to attend a workshop and the main conference had to choose between attending some tutorials (there was a nice choice) or to spend a day idle in Toulouse (which wasn't a bad alternative, give that the place is really very pleasant)?

Workshops

The first two days of the event were devoted to some 11 different pre-conference workshops, with Rebecca Bruce (University of North Carolina) serving as Workshop Chair:

- SENSEVAL (July 6-7)
- Computational Natural Language Learning -- CoNLL-2001 (July 6-7)
- 8th European Workshop on Natural Language Generation (July 6-7)
- Evaluation Methodologies for Language and Dialogue Systems (July 6-7)
- Human Language Technology and Knowledge Management (July 6-7)
- ARABIC Language Processing: Status and Prospects (July 6)
- Open-Domain Question Answering (July 6)
- Temporal and Spatial Information Processing (July 7)

- Data-driven MT (July 7)
- Sharing Tools and Resources (July 7)
- COLLOCATION: Computational Extraction, Analysis and Exploitation (July 7)

While MT practitioners could no doubt find some relevant and useful substance in any one of these workshops, the *Data-driven MT* was clearly right on target for them. I will only report specifically on that one.

The *Data-driven MT* workshop (*separate report to appear next issue—Ed*) was co-chaired by Jessie Pinkham (Microsoft Research), Kevin Knight (USC/ISI) and Franz Josef Och (RWTH Aachen). Fourteen refereed papers plus one invited paper were presented.

The invited paper was presented by Hermann Ney from the University of Aachen and was entitled *Stochastic Modelling: From Pattern Classification to Language Translation*. The author provided an excellent overview of the stochastic modeling approach to machine translation. He also pointed out that in the VERBMOBIL project evaluation his stochastic modeling approach resulted in lower error rates than any of the three competing approaches (semantic transfer, dialogue act based and example based).

The vast majority of the 14 refereed papers were concerned with knowledge acquisition from bilingual corpora, with special emphasis on learning transfer rules. Clearly this is a very hot issue!

Microsoft research was particularly visible at the workshop: in addition to having a seat on the program committee, they had no less than four refereed papers. Clearly, their involvement in MT is no joke!

Hermann Ney's group was also very visible with two refereed papers in addition to the invited one.

Tutorials

On Sunday July 8, five different half-day tutorials were offered:

- Statistical Analysis of Experimental Data for NLP (R. Harald Baayen, University of Nijmegen)

- Very Large Lexical Databases (James Pustejovsky, Brandeis University and Patrick Hanks, LingoMotors Inc)
- Recent Issues in Categorical Grammar (Christian Retoré, INRIA)
- Introduction to Bioinformatics (Rolf Backofen, University of Munich)
- Automatic Summarization (Inderjeet Mani and Mark Maybury, MITRE)

Main Program

The Program Co-Chairs were Norbert Reithinger (DFKI, Saarbrücken) and Giorgio Satta (University of Padua). Their program committee selected 68 refereed papers plus two invited talks.

The invited talks were as follows:

- *Interpreting the Human Genome Sequence Using Stochastic Grammars* by Richard Durbin (Sanger Centre, UK). This fascinating talk brought out the fact that grammatical description techniques based on familiar tools are proving useful too for languages that are very different from NL's.
- *Processing Broadcast Audio for Information Access* by Jean-Luc Gauvain (LIMSI, France). Excellent update on speech recognition technology. There has been very significant progress in recent years. So much so that large scale information retrieval in audio broadcasts is now becoming a reality.

Refereed papers were generally presented in three parallel sessions each featuring groups of 3 or 4 thematically related papers. Session themes included: Generation (2 sessions), Categorical Grammar, Parsing (2 sessions), Machine Translation, Grammars, Named Entity and Information Extraction, Statistical Parsing, Algorithms (2 sessions), Discourse, Lexical Semantics and Similarity, Segments and Segmentation, Statistical Machine Translation, Semantics, Question Answering, Evaluation, Language Models, Syntactic Analysis and Parsing, NLP for Information Processing, Morphology and Lexicon, Topic Spotting and Language Acquisition.

Continued on next page ►

ACL-EACL

...continued from previous page

Corpora, Disambiguation, Chunking and Tagging.

The Machine Translation session took place Monday the 9th in the afternoon. It featured the following papers:

➤ *A Machine Learning Approach to the Automatic Evaluation of Machine Translation* by S. Corston-Oliver, M. Gamon and C. Brockett from Microsoft. The authors show that it is possible to build a system that automatically classifies translations as being MT or human. This is not very surprising in itself; the interesting part comes when they show how this property can be used as a basis for improving MT systems.

➤ *Low-cost, High-performance Translation Retrieval: Dumber is Better* by Timothy Baldwin (Tokyo Institute of Technology). The author shows that for the purpose of retrieving translations in a Japanese-English translation memory system, indexing based on character bigrams proves more effective than linguistically more sophisticated methods.

➤ *Extracting Paraphrases from a Parallel Corpus* by Regina Barzilay and Kathleen R. McKeown (Columbia University). The authors consider cases where a single source language text possesses several different translations into the same language; they show that in such cases it is possible to extract word-level and phrase-level equivalences in the target language.

The Statistical Machine Translation Session took place Tuesday the 10th in the morning, and featured the following papers:

➤ *A Syntax-Based Statistical Translation Model* by Kenji Yamada and Kevin Knight (ISI). The authors show that a syntax-based model can be used to produce better alignments than simpler models. They have not applied their model to a translation task yet.

➤ *Fast Decoding and Optimal Decoding for MT* by U. Germann, M. Jahr,

K. Knight and D. Marcu. An award-winning paper (see below) that presents two new decoding algorithms and compares their performance against the traditional A* algorithm.

➤ *Towards a Unified Approach to Memory- and Statistical-Based Machine Translation* by D. Marcu. Interesting paper. Related to the previous one in that it introduces one of the two new decoding algorithms. Important because it shows a way to unify two things that are usually seen as unrelated.

➤ *Refined Lexicon Models for Statistical Machine Translation using a Maximum Entropy Approach* by Ismael García-Varea, Franz J. Och, Hermann Ney and Francisco Casacuberta. The authors show that using a maximum-entropy approach, they can develop richer lexicon models for their statistical MT system, with the result that corpus perplexity is reduced and translation quality is slightly better.

ISI's performance was truly remarkable, with three papers in the MT sessions, including an award-winning one, not to mention some other papers on different topics in the conference.

Best Paper Awards

For the first time in the ACL's history, the program committee granted a best paper award, under the sponsorship of the Elsevier Company. Two papers were selected and their authors received some prize money plus the opportunity to publish their article in the AI Journal. The two awards went to:

➤ Eugene Charniak for his paper entitled *Immediate Head Parsing for Language Models*. This paper reports a very important result: statistical parsers of the "immediate head" variety can be used as "language models" that can outperform traditional n-gram based language models. The paper is a model of clarity, precision and insight. A must read!

➤ Ulrich Germann, Michael Jahr, Kevin Knight and Daniel Marcu for their paper entitled *Fast Decoding and Optimal Decoding for MT*. This paper presents two new decoding algorithms for statistical MT and

compare their performance against the traditional A* algorithm. Important results that are presented in a clear and crisp fashion. It is good to see that some of the highest quality research in computational linguistics is focused on MT.

Conclusions

Overall, EACL/ACL 2001 proved to be a resounding success on all counts (with the possible exception of one of the caterers who failed to show up at the banquet!). The conference attracted 620 participants from all over the world, including 156 students. It featured some excellent workshops, including one on data-driven MT, and the main program included some great papers on MT, including an award-winning one. □

Workshop on Human Language Technology and Knowledge Management

Mark T. Maybury

Nearly fifty scientists and researchers gathered in Toulouse, France on July 6 and 7, 2001 for the workshop on Human Language Technology and Knowledge Management. Held in conjunction with the meeting of the Joint Association for Computational Linguistics and European Association for Computational Linguistics (ACL/EACL'01), this workshop focused on technologies that promise solutions to challenges in human computer interaction, information access, and knowledge management. The workshop was sponsored by ELSENET (Steven Krauwer, U. Utrecht, steven.krauwer@let.uu.nl) and The MITRE Corporation (Mark Maybury, maybury@mitre.org), the workshop brought together scientists from North America, Europe, Asia, Australia and South Africa. The focus of the group was on:

Expert Discovery: Modeling, cataloging and tracking of distributed

Continued on page 13 ►

People

A Chat with Mike Anobile, Managing Director of LISA

Interviewed by Laurie Gerber

LISA—the Localization Industry Standards Association—has been a defining force in the development of an international localization industry over the last 10 years. Practices pioneered in localization have in turn redefined the translation business, and have led the way in adoption of translation technology, particularly translation memory and terminology management tools. I spoke to Mike Anobile just after the recent LISA Forum which was held in Chicago August 29-31.

MTNI: Historically, machine translation technology seems to have found only limited application within the localization community, specifically with a few multinational manufacturers who also invested heavily in controlled language as well as customized MT systems. Beyond this, localization consultants, and recently IDC analyst Steve McClure, have declared that MT is not of sufficient quality to be used in localization processes. However, this year we have seen two major localization houses acquire two of the major professional-level MT systems. What's going on? Is localization ready for MT? What has changed? How do you think that these localization providers are planning to make use of MT?

MA: Today's drive towards "making MT work" has more to do with Multilingual Information Management than with software localization. As more and more companies are confronted with the challenge of managing businesses across multicultural and multiple language environments, there is increased recognition in MT's value proposition

as well as greater incentives to make it work.

In well-defined application areas MT systems have been efficiently deployed for years. It has always been a function of understanding one's requirements, setting the right expectations, and committing the resources to properly implement an MT solution. Presently, the overwhelming need to cost-effectively manage multilingual data (i. e., in vertical sectors like Finance; Manufacturing; Retailing; IT; etc., and for horizontal applications like enterprise communications, e-commerce and customer/technical support, among others) positions MT as a viable investment. Today's demand by far exceeds the capacity of traditional translation service providers. The adoption/and investment in MT systems and related technologies by the leading localization companies underscores this fact.

MTNI: LISA just held one of its annual USA Forum conferences in Chicago. (*Note: LISA holds forums four times a year once in Asia, Europe, USA, and a developing market.*) What are the current trends and mood in localization?

MA: Economic considerations are making localization clients cautious. At the same time, consultancy is rising. Companies are taking a much more serious look at the investment in the localization of products, services, and back office components of localization. Clients are coming to understand the enterprise-wide impact of localization. PriceWaterhouseCoopers, Anderson Consulting, IBM Global Services, and Hewlett Packard, among others, are examples of emerging localization consulting services.

The series on IAMT and its regional divisions will resume with EAMT in the next issue. An interview with John Hutchins, who is both the current EAMT president, and outgoing IAMT president appeared in MTNI #24. Back issues of MTNI are available from your regional association headquarters.

Consolidation within the industry will continue, and not just among the large players. We will be seeing a lot of strategic alliances among smaller players to allow them to expand their offerings.

MTNI: Does the rise in consultancy reflect some other fact – for example has the wave of risk-tolerant early adopters passed, so that companies looking at localization now are just more conservative?

MA: Not at all. Many of the technologies being considered today are still in the early adoption stages: For example multilingual content management, and the combination of quality assurance tools with automated authoring tools.

Although the U.S. is still the largest market for localization services, U.S. companies tend to be somewhat fearful of language related issues. In addition, companies are not as prepared as you might expect to adopt new technology. Being close to the technology we are in a privileged position where we're very aware of the technology itself and the issues surrounding its use. I remember attending my first ASLIB conference in the 1980s marketing ALPS translation technology. I expected the market to be absolutely ripe for the technology we had to offer, but they were still struggling with word processing issues. I soon learned that companies only buy what they truly understand. It is an education issue. This means they must first understand their own requirements – how to apply the technology to solve specific problems, and what these solutions will bring to their businesses.

MTNI: What does localization need from language technology tools in general? If the localization world could have any technology/tools support it wanted, (ignoring the limitation of what is known to be available or feasible today), what would it want? Better versions of the same kinds of tools that are available now? Better integration? New functionality? Or something completely different?

Continued on next page ►

Mike Anobile

...continued from previous page

MA: Integration of tools that assist the entire workflow process and that are production and quality oriented. This includes terminology management tools, machine translation, translation memory, controlled language tools and others.

The key for successful adoption of these technologies though, is selecting the right solution for the domain a company deals with and having reasonable expectations. By expectations, I mean they need to understand the level of commitment necessary to use the tools effectively and allocate the resources to follow through. Problems in implementation include training, terminology development, and language competency. Companies preparing to localize have to have the people and money and project management in place to succeed. A good way to get a clear picture up front is to have an "information audit" by an outside consultant capable of assessing a company's internal and external multiple language communications requirements. (*Editor's note: An information audit is a process which provides a 'snapshot' of an organization's use of information.*) This is important to the localization process because it establishes a framework within which organizations can evaluate costs versus short, medium and long-term gains.

MTNI: There are three companies in particular (GlobalSight, Idiom, and Uniscape) who have developed very elaborate multilingual content management systems that include workflow and job management and quality control components to support localization projects (particularly website localization). But these systems are more elaborate and expensive than most companies need. What about companies that need simpler solutions?

MA: There are many providers of smaller solutions. The summary of the LISA Forum USA in Chicago includes

the presentation of a case study by Gayle Wooster of ESCA Alstom who evaluated the available options and chose a system called "Glides" for multilingual content management.

Consultancy for this kind of project is very important and very much in demand. People are always contacting LISA and saying that they are tired of hearing just from solutions providers and service vendors who want to promote their own products. Companies need help sorting through the options from impartial experts with no axe to grind. In fact, I recommend that translation technology providers should have a consulting arm that is prepared to recommend whatever solutions fit best, not just their own products.

MTNI: In addition to providing an additional income stream, it would give them access to much more information about what potential clients really want and need – information they couldn't necessarily get in a sales context.

When I attended the LISA forum in Washington D.C. in 1997, I was surprised to find that localization really meant adaptation of software to regional markets. But it seems that now localization has expanded its focus to all kinds of products.

MA: It's not just products, its every aspect of business that is affected by the decision to do business internationally, including back office functions.

MTNI: LISA has positioned itself to meet the needs of the industry leaders who occupy the top tier market segment internationally - companies that can afford to invest heavily in corporate image across many languages as they develop their commercial presence. But there are a huge number of companies for whom the idealized vision of localization defined by LISA members is either unaffordable, or more involved than they feel they need. Yet they want to go global. Does LISA have a role to play with this larger (in terms of number of companies), yet less ambitious globalization movement?

MA: Particularly since 1998 we have made deliberate moves away from the image that LISA is just for very large companies and organizations. In 2000 we restructured our Executive Committee to address the needs of all vertical industries (i.e., not just software) and added a new membership category – the Localization Industry Professional – aimed at individuals who contribute to the GIL process and want to take part in the benefits that being part of our associate provides, at a fraction of the corporate membership fee. We know that the strength of the organization is built on the contributions of a wide variety of member perspectives, and we are attracting a broader cross-section within the localization community.

LISA dues are reasonable in light of today's economic climate and the value we provide to all levels of the industry. The tiered fee structure allows us to offer significant services to all members, both large and small, as well as individuals skilled in the practices of localization workflow, technologies, and business processes. We have been a non-profit organization for 10 years, but we have always operated the association as a business. This means we can have a full-time staff, though we outsource many of the specialized functions such as the Newsletter and Website. Our membership (versus public) pricing for conferences, exhibitions and services are heavily discounted to the companies and individuals who support the association. It's a viable business model.

MTNI: Reports about localization companies/services in the press talk about "notoriously thin margins" and limited profitability. What keeps margins small in localization?

MA: The better-run companies are profitable. Our research data indicates that companies in the US\$750K – US\$ 2.5 million revenue range are often looking at 12-15% profit or greater. Localization companies larger than that have probably been involved in the consolidation frenzy and have likely acquired non-value added-resources that will take some time to sort out.

□

Conferences and Events

Postediting Workshop at ATA 2001

Los Angeles, California
October 31, 2001

How Can MT Possibly Help a Professional Translator? The half-day seminar with this question for a title, aims to provide an answer.

With the maturing of machine translation technology it has become a possible production aid for professional translators. Clearly it is not suitable for all translation tasks, but it is useful for some. A reality of using MT in a translation job is that the output must be edited to make it of publishable quality. This process, generally called post-editing, is related to revision as it has been practiced in professional translation for years, but it is not the same. The purpose of this workshop is to explore the relevant issues for post-editing of machine translation output in a professional setting. 5 presenters will bring a variety of experiences with post-editing to the workshop. Issues addressed will be: 1) When is MT a possible aid to a professional translator? 2) What skills are needed for post-editing? 3) Judging MT output. 4) Practical experiences in post-editing.

The seminar will be led by Winfield Scott Bennett; and Alan K. Melby, ATA director and professor, Brigham Young University, Provo, Utah; They will be joined by: Walter Hartmann, MT Consulting; Rick Woyde, Detroit Translation Bureau; Paul Wu, TheOne Technology Group.

Seminar hours: 9:00am-12:00noon, See the ATA conference website for more information: www.atanet.org/conf2001/

Note that the main ATA conference (November 1—3) includes a number of sessions related to machine translation and the use of CAT tools. See the sessions listed at: www.atanet.org/conf2001/hac.htm □

First Workshop on NLP and XML

Tokyo, Japan, November 30,
2001

Call for participation

XML, the universal structured data representation meta-language, has become the standard framework for publishing on the net, as well as the standard e-commerce language to build B2B and B2C Web services. A major concern for this scenario is the "point of creation" bottleneck, at which creating useful, well-structured XML data can consume an undue amount of time and effort. Hopefully, NLP can resolve this bottleneck by automating the conversion from unstructured or semi-structured text data into XML documents by extracting the richer structure hidden in the original NL descriptions. This is "NLP for XML" that can give some intelligence, or disambiguation capabilities to XML generating engines.

Conversely, XML can help NLP researchers, especially those with annotated corpus-based approaches, by providing them with knowledge representation frameworks for the morphological, syntactic, semantic and/or pragmatic information structure of NL resources. GDA, which stands for Global Data Annotation, is an XML based NL corpus representation language originated by Dr. Koichi Hashida, an invited speaker on "NLP and XML."

The workshop will provide a good opportunity to exchange ideas on a broad range of NLP and structured IT related topics including but not limited to:

NLP for XML

- Ontology extraction into XML-based structured languages using XML Schema

- Message Translation for multilingual B2B, B2C e-commerce applications
- Automatic XML to XML schema mapping by XML vocabulary translators with NL morphological analyzer
- Web (XHTML) resource discovery and indexing
- Automatic hyperlink (XLink) generation
- Multimodal techniques to take advantage of XML compound documents (e.g. search the key string in XHTML, MathML, SVG and SMIL components at the same time)

XML for NLP

- NL Corpora representation languages and the conversions among them, from and to RDB, and from raw text
- XML based Machine Translation / Interlingua
- XML based multilingual Web contents management system
- Tree transducers implemented by XSLT
- IR powered by both NLP and XML
- Task-oriented Summarization using XML Schemas
- Voice XML applications and the dialogue scenario generation
- Foreign language e-Education (CALL) material (texts, drills, grading systems etc.) generation by XML

See: hal2001.itakura.toyo.ac.jp/~chiekon/nlpxml □

For dates, locations, and contact points of other upcoming conferences related to MT and NLP, see the calendar on page 15.

LISA Forum Europe

Vienna, Austria
November 7-9, 2001

Call for Participation

Global Design For Mobile Markets: The Localization of Wireless Products and Services. Major Topics: Global Usability; Design and Markup; Multilingual Content Management; M-Computing in Localization (Impact and Localization Workflow); Localization of M-Commerce (Products, Services, Tools); New Business Models and emerging standards.

The growth of mobile products and services, M-Commerce, highlights global usability within localization practices and strategies. A paradigm has emerged within the market where user-centered design and cross-cultural usability testing have become an integral part of the localization workflow. Global usability engineering not only complements this process but also makes it more efficient and creates new potential for business in such things as side products and spin-offs, while also benefiting localized content management and outside support – and mobility is playing a greater role. Methodology standardization is rapidly catching up with the challenges of these developments and so it is important that M-Commerce professionals and managers also keep up, in order to cope with the intricacies of multilingual markets.

The LISA Forum Europe in Vienna will define the developments in localizing the mobile market and the benefits of global usability. It will provide an opportunity for all involved in the mobile market to review emerging best practices, new business models, and even new paradigms for the interfaces between globalization, localization, M-Commerce, and usability engineering in the context of developing new products and services in a globally wireless world.

Website: www.lisa.org/events/2001vienna/ □

PACLIC 16

Jeju Island, Korea
January 31- February 2, 2002

Call for Papers

The Korean Society of Language and Information is pleased to announce the 16th Pacific Asia Conference on Language, Information and Computation (PACLIC 16). The conference is an annual meeting of scholars in theoretical and computational linguistics from the Pacific Asia region. The aim of the conference is to provide a forum for recent work in various areas of linguistics and bring together researchers from various areas of theoretical and computational linguistics. (The conference site Jeju, is an island located south of the Korean peninsula and enjoys a semi-tropical climate and distinctive culture. It is known as the Isle of Gods. Following the conference, the island will also host World Cup 2002.)

Extended abstracts are invited on unpublished research on all aspects of both theoretical and computational linguistics, including, but not limited to: morphology, phonology, syntax, semantics, pragmatics, discourse analysis, typology, corpus linguistics, formal grammar theory, natural language processing, natural language systems and related computer applications.

Program chair, Prof. Ik-Hwan Lee, Yonsei University, Seoul, Korea. For further information contact pacllic16@pacllic.org. See: www.pacllic.org □

▲ PACLIC 16 Important Dates	
Submission deadline	September 30, 2001
Notification to authors	October 31, 2001
Final papers	November 30, 2001

TMI 2002 Important Dates →	
Submission deadline	October 15, 2001
Notification to authors	December 10, 2001
Final papers due	January 25, 2002

TMI 2002

Keihanna, Japan
March 13 - 17, 2002

Call for Papers

The ninth meeting of TMI (Conference on Theoretical and Methodological Issues in Machine Translation) will be held near the historic cities of Nara and Kyoto in Japan. The workshops and tutorials will be held jointly with the Natural Language Processing Society, Japan. TMI 2002 is supported by the NTT Communication Science Laboratories as part of their 10th anniversary celebrations, as well as AAMT, ANLP and IEICE.

Authors are invited to submit substantial, original, and unpublished research on any issues relevant to machine translation. Papers should be in English, not longer than 10 pages (around 5,000 words), including references. Topics of interest include, but are not limited to:

- MT for the Web
- Practical MT (multilingual eCommerce, localization, etc.)
- Methodologies for MT (statistical, example-based, KBMT, ...)
- Speech and dialogue translation
- NLP techniques for MT
- Knowledge acquisition for MT systems
- MT evaluation techniques and evaluation results
- MT for cross-lingual retrieval and question answering

General Chair: Sergei Nirenburg, Computing Research Lab, NMSU, USA; Program Committee Chairs: Teruko Mitamura and Eric Nyberg, Carnegie Mellon University, USA; Publicity and Local Arrangements: Francis Bond and Hiromi Nakaiwa, NTT Communication Science Laboratories, Kyoto, Japan

Website: www.kecl.ntt.co.jp/events/tmi/ □

COLING-2002

Taipei, Taiwan

August 24—September 1, 2002

Call for Papers and Workshop Proposals

In its 40 plus years of existence, the biennial COLING conference has been a productive forum for scholars all over the world to exchange original research papers on a broad range of topics in computational linguistics. COLING is an international forum for discussion and presentation representing the current state of the art and determining standards of computational linguistics research.

In 2002, Taiwan will host the 19th COLING conference. This will be the first time that COLING is held outside Europe, North America, or Japan. It will be a chance for participants to experience the energy behind Taiwan's vibrant growth in knowledge technology, as well as the natural beauty of Formosa and its rich cultural heritage.

Submitted papers should describe original work, completed or in progress, rather than merely planned, and clearly indicate the current state of advancement of the work. No previously published papers should be submitted. Please see the conference website for submission details: www.coling2002.sinica.edu.tw.

The Program committee also welcomes submissions of proposals for workshops with focus on computational linguistics. For details on Workshop submission please contact Workshop Chair: Antonio Zampolli (Email: pisa@ilc.pi.cnr.it)

Conference Schedule:

•Tutorials: 24 - 25 August, 2002
(Academia Sinica)

•Conference: 26 August - 30 August,
2002 (Howard International House)

•Post-Conference Workshops: 31
August, 1 September, 2002
(Academia Sinica)

COLING-2002 is organized by Academia Sinica, ACLCLP, and Tsing Hua University under the Auspices of

the International Committee on Computational Linguistics.

General Chair: Chu-Ren Huang, Academia Sinica; Program Chair: Winfried Lenders, University of Bonn, Germany, Lenders@uni-bonn.de; Jen-Yi Lin, general contact for conference information, COLING02@sinica.edu.tw See www.coling2002.sinica.edu.tw/ □

COLING 2002 Important Dates

Workshop Proposals	January 15, 2002
Paper Submission	February 15, 2002
Notification to authors	May 1, 2002
Final papers	June 15, 2002

AMTA-2002

Tiburon, California
October 2002

First Announcement

AMTA is pleased to announce the site of its next biennial conference: AMTA-2002 will take place in the second week of October at the Tiburon Lodge, situated in the picturesque town of the same name, about 30 miles north (and just across the Bay) from San Francisco. For a glimpse of the hotel and the beautiful surroundings, you can check out the Lodge's Web site at tiburonlodge.citysearch.com/1.html; but don't pay attention to the room rates, since we will be offering participants a substantial discount.

Several key members of the Organizing Committee for AMTA-2002 have also been selected. Violetta Cavall-Sforza of San Francisco State University will be Local Arrangements Chair; Steve Richardson of Microsoft Research will be Program Chair; and Bob Frederick of Carnegie Mellon will be in charge of the workshops and tutorials.

More details on the conference will appear in upcoming issues of MTNI. Or contact AMTA President Elliott Macklovich, macklovi@iro.umontreal.ca □

In Future Issues...

Book Reviews:

•John Hutchins', "Early Years in Machine Translation: Memoirs and Biographies of Pioneers", Reviewed by Maghi King.

• Dan Melamed's, "Empirical Methods for Exploiting Parallel Texts", Reviewed by Graham Russel

Conference reports:

• ACL Data-Driven MT Workshop

• TALN Very Large Scale MT Workshop

• MT Summit (of course!)

Interviews:

• Bente Maegaard, Summit VIII Program Chair

• Junichi Tsujii, AAMT President

Errata

Elliott Macklovich's email address: The profile of Elliott Macklovich in MTNI #27 included an incorrect email address. The correct email address is: macklovi@iro.umontreal.ca

Last issue's feature on AMTA may have implied that MTNI has not previously included personal profiles or interviews. This was not intended and certainly is not true!

#24 Interview with John Hutchins, as incoming IAMT president.

#20 Profile of Hozumi Tanaka, incoming IAMT president

#17 Profile of Eduard Hovy, AMTA president.

Back issues of MTNI containing these features are available from your regional association headquarters. □

Censorship and Online Translation

Many schools and other organizations have resorted to filtering programs in order to keep Web surfers from retrieving pornography or other material they consider unsuitable. But at least one filtering program, sold as BESS. The Internet Retriever from N2H2 has the surprising property of systematically blocking translation sites.

The reason for this is described in detail at sethf.com/auticensorware/bess/loophole.php, the website of Seth Finkelstein, a computer scientist and political activist who was recently profiled in the New York Times following his receipt of the Electronic Frontier Foundation's Pioneer Award for his investigation of censorship online. archive.nytimes.com/2001/07/19/technology/circuits/19HACK.html

The following is excerpted from Finkelstein's description. "[N2H2]'s product (sometimes called BESS) is server-based. That is, the blacklist is kept on a remote machine, and all requests to read anything on the web have to first go through that machine (or similar) for approval. The server is a kind of choke-point, or censor..."

"The most interesting category of BESS turns out to be one which isn't mentioned in the documentation or PR. In particular, N2H2 has a category called LOOPHOLE which is, to put it politely, *undocumented*."

Sites that constitute loopholes for filtering software include those services which allow privacy or anonymity in viewing by encrypting content or transmissions, those that provide language translation, and even dictionary lookup. Interestingly, sites that allow checking of HTML for conformity to various encoding standards also are blocked. Possibly in response to Finkelstein's exposé, it appears that a work-around has been added to allow privacy administrators to designate site-specific unblocking, for example for amazon.com, whose contents are encrypted and would otherwise be blocked: www.n2h2.com/support/safeweb_instructions.php. □

Resources

Laurie's Links

Laurie Gerber

Möbius Links

Some months ago I was looking for health information online for a family member. I found a promising site that had "links to resources and information". Each link I followed from there, however had similar "links to sites with resources and information" but no real resources or information that I could find. Several more jumps took me back at the original site with no information to show for it. What was I researching? Alzheimer's disease.

Acronym Overload

In my recent transition from the purity of academic student life to the rough and ready wired world, I was baffled by the number of acronyms which everyone in the world must already know, since no one bothers to write them out or explain them. Here are just a few: EULA, ROI, TCO, SCM, CRM, DTD, SMS. (Test yourself! Then see the definitions below.) I found relief at Webopedia www.webopedia.com. In addition to a term lookup with clear and concise definitions, they offer a list of emerging new terms, so you can stay current and baffle your friends before they baffle you.

Guidelines for success with MT

MT system documentation usually gives some admonishments about writing short clear sentences for best results, but for users who want more help, I found two sources of very detailed and helpful instructions for optimizing MT performance for regular text, and for websites:

General writing guidelines for translation: www.languagepartners.com/reference-center/wri4tr.htm

While you're at the Language Partners website, take some time to look around. In addition to information on their regular business of software resale and translation/localization consulting, they have quite a number of useful white papers on

these and related topics.

Guidelines for building translatable websites: www-106.ibm.com/developerworks/usability/library/us-mt/ This offering comes to us from Teri O'Connell, whom many in the U.S. MT community will remember from the 1993-1994 MT evaluations done by PRC (together with John White.)

If the URLs given in this column are sometimes too long and daunting to type out, the column is also posted online at: www.eamt.org/resources/index.html where you can use the hyperlinks instead

Publications from Michael Blekhan

Michael Blekhan, President of Lingvistica 98, Inc. will provide copies on request of:

- An informal comparative evaluation of three of the free online MT systems: LogoMedia, Reverso/PROMT, and SYSTRAN. (Request the "comparative analysis")

- A detailed history of the PARS machine translation system, interwoven with the history of machine translation itself. (Request "Machine Translation: Professional Experience")

Contact Michael Blekhan at: ling98@videotron.ca

News Alert Service

Northern Light www.northernlight.com has a free news alert service that I don't know how I ever lived without. Sign up on their website for the "free alert service", then set up as many keyword searches as you want. Any news items that match your boolean query will be emailed to you daily or weekly. One of the great features of the service is that you can test the results of your query on their content repository as you build it. I've tried other alert services—most of which I had to pay a lot for, and which didn't approach the relevance and timeliness I get from Northern Light. In addition to the alert service, you can also do various kinds of searches on their content repository anytime. Their indexed resources include lots of internet data, plus published items not usually available online. Some of the items are

free, others are available to buy for prices ranging from US\$1 for news articles to several thousand dollars for copies of strategic market research bulletins.

About those acronyms

EULA—End User Licensing Agreement (software); ROI—Return on Investment and TCO—Total Cost of Ownership (tools evaluation); SCM—Supply Chain Management and CRM—Customer Relationship Management (enterprise business software categories); DTD—Document Type Definition (XML); SMS—Short Message Service (mobile telephones) □

Workshop on Human Language Technology and Knowledge Management

...continued from page 6

organizations and communities of experts.

Knowledge Discovery: Identification and classification of knowledge from unstructured multimedia data.

Knowledge Sharing: Awareness of and access to enterprise expertise and know-how.

A range of human language technology areas were addressed including speech and language processing, translation, summarization, multimedia presentation, content extraction, dialog tracking. The workshop included reports on the application of human language technology to knowledge management and moderated group brainstorming sessions to create a road map for Human Language Technologies for the next decade. Some of the topics identified included the use of human language technology to enable:

Automated retrieval, extraction, and enrichment of information and knowledge from multimedia, multilingual, and multiparty information sources.

Translingual or crosslingual retrieval, presentation, and sharing of knowledge.

Automated detection and tracking of emerging topics from unstructured multimedia data (e.g., documents, web, video news broadcasts).

Use of knowledge sources to facilitate knowledge mapping and access (e.g., lexicosemantic such as WordNet, semantic such as geospatial Gazetteers, semistructured such as thesauri, encyclopedia, fact books)

Automated question-answering from heterogeneous sources

Intelligent tools that support the automated bibliometrics and document analysis/understanding in support of discovery of distributed experts and communities of expertise

Summarization and presentation generation of knowledge (e.g., knowledge maps, lessons learned).

Modeling of user knowledge, beliefs, plans, (dis)abilities and preferences from queries, created artifacts, and human computer interactions.

Two keynote address initiated each day of the workshop. Dr. Steffen Staab (ssi@aifb.uni-karlsruhe.de) Senior Researcher and Lecturer in the Knowledge Management Group at the Applied Computer Science Institute (AIFB) at University of Karlsruhe, Germany presented the opening keynote: "Knowledge Portals". Steffen described how the construction of intelligent access and the provisioning of information to knowledge portals remains a manually intensive task. He described the use of ontologies as a conceptual backbone for providing, accessing and structuring information in support of building and maintaining knowledge portals. Dr. John Domingue (j.b.domingue@open.ac.uk), Deputy Director of The Knowledge Media Institute at The Open University in England, opened the second day with his keynote "Supporting Organisational Learning through the Enrichment of Documents". John described knowledge sharing by enriching web documents with informal and formal representations, a process which captures the context in which a document is created and applied. John demonstrated how this enrichment facilitates retrieval and comprehension. Participants also enjoyed an invited talk from Hans Uszkoreit (DFKI

Saarbruecken, uszkoreit@dfki.de), Scientific Director at the German Research Center for Artificial Intelligence (DFKI), Head of DFKI Language Technology Lab, and Professor of Computational Linguistics at the Dept. of Computational Linguistics & Phonetics of Saarland University at Saarbrücken. Hans addressed "Crosslingual Language Technologies for Knowledge Creation and Knowledge Sharing", describing the role of information extraction, automatic hyperlinking and (human) inference in the process of decision support in applications such as call centers. Hans exemplified "automatic relational hyperlinking" using the example of the Hypercode system, developed for a large German bank to facilitate work with legacy code by densely interlinking source code and documentation.

A poster session and papers sessions were held reporting results in the areas of ontology construction, question answering, summarization, multilingual processing, multimedia processing and dialog. During brainstorming sessions, the group outlined key challenges and promising new approaches in these areas. For example, for ontologies the group emphasized the need for tools and tasks that were reusable across domains to create/populate ontologies, the need to integrate shallow and deep methods, and the need to collaborate with domain ontology creators. They outline challenges in ontology quality, ambiguity, and usability. The group highlighted the promise of the semantic web, the importance of information extraction "plugins", the possibility of organizing massive documents via domain specific ontologies, the opportunity to use a top or core ontology to bootstrap new domains, and the value of multidisciplinary (e.g., domain experts, linguists, knowledge engineers) collaborative teams.

With respect to summarization, the group outlined challenges as including the appropriate level/depth of analysis/representation (e.g., semantic relations, speech acts, rhetorical structure), summarization presentation/visualization, speech for presentation of short summaries, the appropriate use of indicative vs. informative summaries, and the need for

Continued on next page ►

Workshop on Human Language Technology and Knowledge Management

...continued from preceding page

action-oriented summaries (e.g., executive/management summaries). The group discussed a range of solutions encompassing the analysis of information, its transformation (including operations such as selection, aggregation, abstraction) and its presentation.

The group also identified a number of fundamental multilingual challenges including relations between cultures, languages, lexical resources, and ontologies, complexity, and the challenge of new application domains such as content-driven hypertextual authoring and cross-lingual news linking. The group identified resources (e.g., wordnet, euronet, application databases, text resources) as key to advancement, the interlingua approach as promising, and the importance of deeply annotated data combined with machine learning. Finally, the group turned its attention to multimedia challenges and opportunities. Challenges include the integration of multiple media, the nature of processing (is it centralized or mobile), the challenges of privacy, security, and scalability, the importance of both remembering and forgetting information, the need for multilingual and multisource information extraction, and the challenge of cross-document co-reference resolution. Location-based services were highlighted as a promising future area.

Two cross cutting enabling capabilities were identified for all of the addressed areas. First, the need for (intelligent) text annotation. Second, the need for large-scale annotated corpora to enable automated training and system evaluation.

ELSNET has captured the workshop input and will continue to revise a technology roadmap. A web site to share the materials and results of the workshop has been set up at www.elsnet.org/acl2001-hlt+km.html. □

SYSTRAN Personal for Linux

...continued from page 4

After having announced in April a new version of SYSTRANLinks, SYSTRAN launches in July SYSTRAN Personal for Linux, thus extending its range of consumer goods.

This software bound for the users of Linux makes it possible to translate any document quickly: e-mail, letters, files of work, articles of press, etc. It is the manner easiest to translate until 5KB text, and to simply use the translations in all the applications by simple copy-sticking. The access to the translation is done through the usual navigator of the user.

SYSTRAN Personal for Linux is available today in 3 European language pairs: French->English, French->German and French->Spanish. The range will be supplemented soon by the pairs French->Dutch, French->Italian, French->Greek, French->Portuguese.

This product is downloadable since the electronic store of www.SYSTRANSOFT.com at the cost of \$30 (download blind).

SYSTRAN and Linux

Linux has a strong potential of deployment of Internet infrastructures for large companies. SYSTRAN offered until now of the solutions servers. Google and Altavista in particular retained it for their platform of translation, which supports a few million pages seen per day! Linux represents currently a little less than 2% of the operating systems of the computers, but it is in full progression. Its share of market on the Internet servers knows a true explosion (27% in 2000 against 15.8% in 1998 according to IDC). Since many years SYSTRAN supports its technology on Unix systems. They make it possible to fulfill the requirements of the management of the enormous SYSTRAN knowledge bases, in particular in terms of performances, while keeping an accessibility for the linguist-computer

specialists who enrich the translation systems day after day.

For further information, including the full French press release, see: www.systransoft.com □

Bowne edges out Lionbridge

...continued from page 1

About Bowne & Co.

Bowne & Co., established in 1775, is the global market leader in the field of empowering information by combining superior customer service with appropriate new technologies to manage, repurpose and distribute a client's information to any audience, through any medium in any language, anywhere in the world. The world's largest financial printer, Bowne is also the leading provider of localization services to the software industry, and is among the leading providers of outsourcing services and digital print-on-demand solutions.

About Bowne Global Solutions

Bowne Global Solutions, www.bowneglobal.com, is a provider of comprehensive globalization solutions -- enabling its customers to reach their customers in any part of the world. Bowne's global team delivers strategy consulting, technical communication, product localization and multi-language web content creation and management services. With this suite of integrated globalization services, Bowne Global Solutions enables industry leaders in the technology product, corporate online learning and digital entertainment industries to compete and grow global market share through locally-relevant, culturally-connected products and services.

See the Bowne website for full-text press-releases: www.bowne.com. Contacts: William J. Coote Vice President and Treasurer Bowne & Co., Inc. +1 212-886-0614 bill.coote@bowne.com; Donald J. Plumley Chief Marketing Officer Bowne Global Solutions, Inc. +1 323-866-1104 donald.plumley@bowneglobal.com □

Calendar

2001

October 30-31: SIIA (Software and Information Industry Association) 2001 European Forum. October 30-31, 2001. Landmark Hotel, London. Conference theme: Software distribution models/ What works in Europe today? See: www.siiia.net/euroforum/overview.html.

October 31-November 3: 42nd Annual Conference of the American Translators Association, Regal Biltmore Hotel, Los Angeles. See: www.atanet.org.

November 6-10: IEEE International Conference on Software Maintenance 2001: Systems and Software Evolution in the Era of the Internet, Florence, Italy. See: www.dsi.unifi.it/icsm2001.

November 7-9: LISA Forum Europe, Vienna, Austria. Global Design For Mobile Markets: The Localization of Wireless Products and Services. See: www.lisa.org/events/2001vienna

November 15-16: Localisation 2001 - eLocalisation. The 6th Annual Localisation Conference and Industry Showcase organized by the LRC (Localization Research Center), University of Limerick, Ireland. Conference topics: The business of eLocalisation - the localisation portal; Workflow management and transaction automation; Integration and interoperability of technology; eContent - a new challenge? See: lrc.csis.ul.ie/LRC/Conf2001/index.htm

November 27-29: NLPRS 2001, 6th Natural Language Processing Pacific Rim Symposium. National Center of Science, Tokyo, Japan. See: www.r.dl.itc.u-tokyo.ac.jp/NLPRS2001.html

November 27-29: ICCS 2001, International Conference on Chinese Computing, Singapore. See: www.colips.org/conference/iccc2001/en

November 29-30: Translating & the Computer 23, London. Organized by the Association for Information Management (Aslib) and supported by the Institute of Translation and Interpreting, the British Computer Society, and the European Association for Machine Translation. Program Chairs: Daniel Grasmick, SAP; Ruslan Mitkov, Univ. Wolverhampton; Chris Pyne, Lionbridge Deutschland; and

Olaf-Michael Stefanov, United Nations. See: www.aslib.co.uk/conferences/index.html. Contact: Nicole Adamides, Aslib, Staple Hall, Stone House Court, London EC3A 7PB, tel: +44 (0) 20 7903 0000; fax: +44 (0) 20 7903 0011; e-mail: nicole.adamides@tinyworld.co.uk.

November 30: Twelfth CLIN Meeting (Computational Linguistics in the Netherlands) Department of Computer Science, University of Twente. Hosted by the Parlevink language engineering group at the University of Twente. The languages of the conference will be Dutch and English. See: parlevink.cs.utwente.nl/Conferences/clin2001.html

November 30: The First Workshop on NLP and XML, Tokyo, Japan. See: hal2001.itakura.toyo.ac.jp/~chiekon/nlxml

December 11: 5th Australasian Natural Language Processing Workshop, University of Adelaide, Adelaide, Australia. A one-day workshop on Natural Language Processing held in conjunction with the Australian AI conference (AI'01) See: www.comp.mq.edu.au/hlt/events/anlp5

2002

January 21-25: 1st International Wordnet Conference. Organized by Global Wordnet Association together with CIIL Mysore, India. The conference will be held at the Central Institute of Indian Languages, Mysore (Karnataka State) See: www.hum.uva.nl/~ewn/gwa.htm

January 28-February 1: 20th International Unicode Conference. Washington, DC, USA Organized by the Unicode Consortium. See: www.unicode.org/iuc/iuc20

January 31-February 2: PACLIC 16. The 16th Pacific Asia Conference on Language, Information and Computation. Jeju Island, Korea. Hosted by the Korean Society of Language and Information. See: www.paclic.org

February 17- 23: CICLing-2002. Mexico City, Mexico. Second International Conference on Intelligent Text Processing and Computational Linguistics. **Deadline for paper submissions October 10, 2001; for short papers, November 5, 2001.** See: www.cicling.org/2002.

February 28 - March 2: II International Conference on Specialized Translation, Barcelona, Spain. See: www.upf.es/dtf/activitats/2icst/2ctradsp.htm. Conference Secretariat: Dept. Traducció I Filologia, Universitat Pompeu Fabra, La Rambla 30-32, E-08002 Barcelona, tel: +34 (93) 542-24-09 / 542-22-75; fax: +34 (93) 542-16-17; e-mail: 2congres@grup.upf.es.

March 13-17: TMI 2002. The 9th Conference on Theoretical and Methodological Issues in Machine Translation. **Deadline for paper submissions October 15, 2001** Keihanna, Japan. See: www.kecl.ntt.co.jp/events/tmi/

May 9-11: The Fifth Symposium on Natural Language Processing 2002 + Oriental COCOSA Workshop 2002. Hua Hin, Prachuapkirikhan, Thailand. Organized by: Thammasat University, Patumthani, and National Electronics and Computer Technology Center (NECTEC), Thailand. **Submission Deadline: February 1, 2002.** See: kind.sit.tu.ac.th/snlp-o-cocosa2002

July 7-12: ACL-2002. University of Pennsylvania, Philadelphia, PA, USA. Organizers: Martha Palmer (Local Arrangements).

August 7-10: 16th World Congress of the International Federation of Translators, "New Ideas for a New Century" Submission deadline, September 1, 2001. See: www.stibc.org/fit2ki.htm.

August 24-September 1: COLING 2002. Academia Sinica, Taipei, Taiwan. Organizers: Chu-Ren Huang, Academia Sinica (Conference Chair); Winfried Lenders (Program Chair), University of Bonn, Germany Lenders@uni-bonn.de; Antonio Zanpolli (Workshop Chair) pisa@itc.pi.cnr.it. Deadline for Workshop Proposals: January 15, 2002 Deadline for paper submission: February 15, 2002. Conference schedule: Tutorials: August 24 - 25 (Academia Sinica); Conference: August 26 - 30 (Howard International House); Post-Conference Workshops: August 31 - September 1, (Academia Sinica). See: www.coling2002.sinica.edu.tw