

**TBX/OLIF Workshop
Heidelberg, Germany
November 4, 2002**

MINUTES

Discussions at the TBX/OLIF Workshop began at 9AM. Attendees included:

OLIF2 Consortium members:

Daniel Grasmick, SAP
Hubert Lehmann, Linguattec
Christian Lieske, SAP
Susan McCormick, SAP/OLIF Consultant; Co-Chair
Gregor Thurmair, Compendium
Michael Wetzel, Trados (also for TBX)

OSCAR/Terminology SIG for TBX:

Alan Melby, BYU; Co-Chair
Raphael Prono, Xerox
Paolo Vanni, J.D. Edwards
Ulrich Wachowius, SDL
Kara Warburton, IBM

- Participants introduced themselves and briefly described their interest in/use of TBX and OLIF:

- Gregor Thurmair – Compendium works on NLP applications, including MT ; OLIF useful for MT, as well as retrieval and ontology work.
- Alan Melby – OSCAR representative for TBX work; a descendent of work in the Text Encoding Initiative, TBX was developed under the auspices of SALT.
- Hubert Lehmann – Linguattec develops and markets language products, notably Personal Translator (PT); has reviewed OLIF and feels it's compatible with its PT lexicons.
- Kara Warburton – Head of terminology management at IBM; is interested in TBX and OLIF re work on integrating term management across systems in the workplace; IBM uses TransLexis for bilingual term management and English source term management in North America, and multilingual term management in Boeblingen, Germany; TransLexis export based on TIF, precursor to MARTIF; plan to support TBX

Kara is also currently head of the terminology SIG for LISA.

- Ulrich Wachowius – SDL uses variety of term management systems depending on customer needs; interested in term standards since terminology is highest priority for quality control; SAE provides measure of translation quality; Transcend is used by SDL for MT; pursuing work on controlled authoring.
- Christian Lieske – Lead for technical WG of the OLIF Consortium; works on NLP (e.g., term extraction) and XML (e.g., in the context of OASIS).
- Susan McCormick – Coordinator of OLIF Consortium; continuing work on OLIF implementation within SAP; provides support to potential OLIF users; has worked in cooperation with TBX development to provide mapping of OLIF data categories.

- Daniel Grasmick – SAPterm, the SAP proprietary termbase has 1.3 million main entries; SAP uses 4 MT systems and a controlled language tool; OLIF seen as exchange option for SAPterm and other language tools at SAP; SAP requires that all tool providers support OLIF exchange.
- Raphael Prono – Xerox Global Solutions, translation and terminology solution; exchange of terminology, distribution of terminology over the Web; customers include Ford and Skania.
- Paolo Vanni – Develops J.D. Edwards proprietary system for terminology; will implement a filter for TBX; may eventually have to support tools such as TermStar and SDLX.
- Michael Wetzel – MultiTerm iX is latest version of MultiTerm. Trados is ready to implement OLIF and TBX conversion (demo for OLIF is available).; considering using API; foreseeing plug-in architecture that allows hooking in converters. Questions remain on full implementation, since customers haven't expressed need for e.g., TBX.

- General points were made by Susan and Alan that OLIF and TBX are XML-based formats, with OLIF targeted to NLP applications and lexicons, and TBX targeted to terminology exchange.

- Kara reported on the Terminology SIG survey results (see www.lisa.org/2001/termsurveyresults.html) as they affect the TBX/OLIF exchange:

- Generally, there were few reported users of TBX or OLIF
- Most users still maintain terms in Excel files or simple text files; term exchange is often a matter of 'source term/target term/subject area.'
- Confusion among users re TBX and TMX

- Daniel briefly reported that the Global Content Creation Survey conducted by OSCAR yielded different results, specifically it showed that, of 150 respondents:

- 22 companies reported using OLIF
- 27 companies reported using TBX
- Again, the caveat that there appeared to be some confusion of TBX with TMX

Meeting participants were generally skeptical that the actual level of OLIF and TBX use is as high as this survey indicated.

- Question as to whether it's reasonable to discuss integrating TBX/OLIF when it's not clear that they're being used as separate formats now; points to need to more clearly identify the user need/demand regarding the two formats. Ideas for clarification of user needs:

- Need to develop use cases to justify further integration work
- Bowne is already working on TBX/OLIF mapping; could provide feedback
- Sun is using OLIF for glossaries; could provide feedback
- Alan suggested that Terminology SIG could help by providing sample terminology entries; Christian pointed out that sample OLIF entries are available from the OLIF Web site.
- Kara noted that the Terminology SIG has surveyed 8 companies for data categories used; could serve as input

- Susan pointed out that user feedback may support a minimal integration, where OLIF datcats are available to TBX users; Sue Ellen Wright has already included many OLIF datcats in the SALT DCR and Susan has reviewed the mappings.

Action Item: Establish a working group to conduct a language industry user survey designed to elicit information on terminological/lexical data exchange. Survey should aim to provide meaningful data on e.g., data categories used, explicitly what kind of data exchange is required (between which users, tools), what plans users have for term/lex exchange. Survey should be conducted with an eye to elucidating need for TBX and OLIF.

Working group will consist of following members:

OSCAR/Terminology SIG for TBX: Alan Melby, Kara Warburton, Paolo Vanni, Ulrich Wachowius, Raphael Prono

OLIF: Susan McCormick (Coordinator), Daniel Grasmick, Christian Lieske, Hubert Lehmann, Michael Wetzel, (perhaps) Gregor Thurmair

- Susan will be in touch with all prospective members in approximately 2 weeks to get planning started.
- Alan provided a demo of the TBX checker

The meeting adjourned at 12:20 PM.

Susan McCormick
November 17, 2002
Baltimore, MD USA