Appendix LDV

Stefanie Geldbach, Uta Seewald-Heeg

Appendix

The efforts toward standardization have sparked off numerous research projects and initiatives which have proposed and developed various standardized formats or lexicon models. While some of these standards (e.g. TMX) already have found wide acceptance and are supported by a growing number of commercial applications the fate of others is still undecided. Only future developments will show whether a given standard will actually be adopted by the language industry or rather be replaced by formats yet to come. At any rate, standardization of terminological and lexical resources is an important research field which will doubtlessly receive considerable attention also in the years to come. In order to facilitate the orientation within this dynamic field we decided to compile a short reference section which contains a – surely not exhaustive – list of past and ongoing standardization projects, file formats and organizations actively engaged in standardization issues. This list contains short definitions of important terms and projects discussed in this issue and provides links to websites for further reference.

1   Projects

ISLE (International Standards for Language Engineering). International project which aimed at improving the accessibility and availability of language resources. See *www.ilc.cnr.it/EAGLES96/isle/ISLE_Home_Page.htm*.

LIRICS (Linguistic Infrastructure for Interoperable Resources and Systems). This ongoing project aims at providing ISO ratified standards for language technology to enable the exchange and reuse of multilingual language resources. See *http://lirics.loria.fr/*.

MILE (Multilingual ISLE Lexical Entry). Model for the lexical representation which is based on an open core set of data categories. See for ex. *http://www.w3.org/2001/sw/Best-Practices/WNET/ISLE_D2.2-D3.2.pdf*.

SALT (Standards-based Access to Lexicographical and Terminological Multilingual Resources). Open-source project for creating open standards in order to facilitate the integration of terminological and lexicographical resources. See *http://www.loria.fr/projets/SALT/*.

2   Formats

MARTIF (Machine Readable Terminology Interchange Format). SGML-based format for exchanging terminological data between and among different terminology database systems published by ISO TC37/SC3 in 1999. See *http:// www.iso.org/*.

TBX (Term Base eXchange). Open XML-based standard developed by the Localization Industry Standards Association (LISA) to support the exchange of terminological data, see *www.lisa.org/standards/tbx*.

TMX (Translation Memory eXchange). Open XML-based standard exchange format for translation memories developed by LISA to allow easier exchange of translation memory data between tools and/or translation vendors, see *www.lisa.org/standards/tbx*.

OLIF2 (Open Lexicon Interchange Format). Open XML-based format developed by the OLIF Consortium to support the exchange of terminology and MT lexicons, see *www.olif.net*.

XLIFF (XML Localization Interchange File Format). Specifically designed to support the

localization of data. XLIFF has features for updating strings, revision control, marking different phases of the localization process, word count calculations, the provision of alternative or suggested language translations. XLIFF is an open standard. See *http://www.oasis-open.org/committees/xliff*.

3 Organizations

LISA (The Localization Industry Standards Association). Since 1990, the LISA Forums and Global Strategies Summits have been dedicated to delivering best practices and standards for facilitating international business. See *http://www.lisa.org/*.

ELRA (European Language Resources Association). Established as a non-profit organisation in 1995, ELRA's goal is to make available the language resources for language engineering and to evaluate language engineering technologies. See *http://www.elra.info/*.

ISO (International Organization for Standardization). Founded in 1947 as the successor of the National Standardizing Associations (ISA) established in 1926, ISO is a network of the national standards institutes of 156 countries. The organization develops International Standards required by business, government and society. See *http://www.iso.org/*.

OASIS (Organization for the Advancement of Structured Information Standards). Founded in 1993, OASIS is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. The OASIS Localization Technical Committees (TC) include the Translation Web Services TC and the XLIFF TC. See *http://www.oasis-open.org/*.

OSCAR (Open Standards for Container/Content Allowing Re-use). OSCAR is LISA's working group for the development and maintenance of open standards for the language industry. Founded in 1997, OSCAR is dedicated to the development of open standards. Standards worked out by OSCAR include: Translation Memory eXchange (TMX) – the certifiable standard for the exchange of translation memory data; – the standard for exchange of structured terminological data; and Segmentation Rules eXchange (SRX) – the standard for exchange of information about how translation tools segment text. See *http://www.lisa.org/sigs/oscar/*.