A Web-based Terminology Management System and the Translation Market

Balázs Kis
Kilgray Translation Technologies
Gyula, Hungary
`balazs.kis@kilgray.com`

Peter Reynolds
Kilgray Translation Technologies
Warsaw, Poland
`peter.reynolds@kilgray.com`

Abstract

This paper introduces a new multilingual terminology management system that suits translators and translation organizations, as well as larger enterprises that are currently the largest purchasers of commercial terminology management systems. The paper outlines the principles behind translation-related terminology management, and demonstrates how the new terminology management tool integrates with the translation environment and the translation workflow.

1   Introduction

There is a latent conflict between terminology management and translation. Terminology management is considered as the means of standardization in order to facilitate unambiguous communication in specific fields. Therefore, terminology management systems tend to focus on creating highly structured, very detailed and elaborate data structures so that all information related to concepts and terms can be entered, looked up, and shared.

As a result, terminology management systems integrate poorly with translation environments. This is not apparent from the technical implementation since commercial examples of integration do exist. The conflict is not in the technicalities but the use case: translators have a different focus that apparently falls outside the attention of creators of legacy terminology management tools. When we created our company, and started selling our translation environment, we were surprised how scarcely terminology tools are used among translators and (smaller) translation organizations (also in Fulford-Zafra 2005).

2   The Business Problem Statement

The main problem of translators is that the time constraints of a translation task do not allow for extensive terminology research. Of all the sophistication of terminology management, translators seldom need more than fundamental dictionary functionality: they immediately need the target-language equivalents while editing the translation, and they also need to be able to add new terms with the smallest possible effort, without leaving the editing environment. This observation led us to the principles upon which we have originally built the terminology component in our translation environment.

As our business was growing, so was the size of our customers. As opposed to small or medium translation organizations, larger enterprises need sophisticated terminology management – it is a core component of their research, development, and communication infrastructure. Likewise, larger translation organizations working for enterprises need the same level of terminology management. Our company adopted the organic-growth approach both in its business and its development strategy. As a result, we have recently encountered customers who needed more than the simplistic terminology management that our translation environment offered. This prompted us to create the terminology management system we are demonstrating in this workshop.

It is not necessarily the technical implementation or the fundamental concepts of the terminology management system that represent a novelty. It is rather the approach to terminology management and the resulting integration of systems that is new in the field, although the basic principles were outlined as early as in 1980 by Martin Kay (Kay 1980).

3 Simplistic Terminology Management in a Translation Tool

The translation environment we mentioned earlier is an integrated desktop application that offers all features in a single program. Features include terminology management in a simplistic way – let us summarize them in a few words:

- Term bases follow a proprietary database format (no relational database systems are used).

- The internal structure of entries complies with legacy terminology systems: we have implemented the three-level structure of (1) concepts (entries), (2) languages (indexes), and (3) terms, with meta-data attached to each level.

- We use a restricted set of meta-data, and the entry structure cannot be modified (although the internal database representation is flexible). If one uses the translation environment without the new terminology system, this limitation still applies.

- On the translation user interface, the structure of the term base entries is hidden: users primarily see source-language terms and their target-language equivalents, both in a match list and highlighted in the source-language text. There are simple commands (clicks and key shortcuts) to insert target-language equivalents in the translation. In addition, simple commands are available to add new terms.

The translation environment has had a server component from the start. Translation memories, term bases, documents (projects), and other resources can be published on a server, and they can be used the same way as in the desktop tool – full transparency is provided.

With server-based term bases, we had also introduced a simple collaborative workflow. A term base can be moderated, which means there are two user roles: translator and terminologist (or proof-reader). Translators can add new terms to the term base, but they are not visible to other translators until the terminologist approves them.

4 The New Development and its Integration

The new terminology management system aims at overcoming the limitations outlined in Section 3. We have planned the development in several stages, of which the first stage is completed as of now.

The system eliminates the restriction of meta-data at all levels of entries (concepts, indexes, and terms). It offers a web-based collaborative editing interface. Full support for importing and exporting TBX files is also included. The system facilitates read-only end-client access, and sophisticated access permissions in general.

The real challenge of this development lies in maintaining the compatibility with the existing translation environment. We have decided to integrate the terminology management with the translation resource server as closely as possible. All functionality related to term base management is still implemented in the translation resource server. The data model required no change because it was originally designed with flexibility in mind, and it was recently scaled up to ca. 2 million entries per term base.

We have also made effort to keep the system transparently interoperable with the translation environment itself. There are two types on the translation resource server: qTerm term bases and legacy term bases. From the translation environment, both types can be queried the same way. However, qTerm term bases can be edited from the new system only. Any legacy term base can be converted into a qTerm term base.

5 Conclusion: Further Development

As we have mentioned above, only the first stage of development is completed. The real novelties in the area of collaborative terminology management and feedback/approval workflow are yet to come in the next 12 months. We plan to conduct extensive beta tests, and one of our reasons to present the tool in this workshop is to invite professionals to give us feedback as the development progresses. Our aim is to create a tool that meets the requirements of the strictest terminology management scheme, yet it retains its simplicity and value for translators and translation organizations.

References

Heather Fulford, Joaquín Granell-Zafra. 2005. Translation and Technology: a Study of UK Freelance Translators. *The Journal of Specialised Translation*, Issue 4, JoSTrans, http://www.jostrans.org/issue04/art_fulford_zafra.pdf

Kay, Martin. 1980. The Proper Place of Men and Machines in Language Translation. *Xerox report CSL-80-11*, Xerox Palo Alto Research Center.