

2010

TermWiki: A new collaborative terminology management solution

Uwe Muegge, *Monterey Institute of International Studies*



SELECTEDWORKS™

Available at: http://works.bepress.com/uwe_muegge/11/

TermWiki: A new collaborative terminology management solution

The development of TermWiki provides organizations with an open-source, easy-to-use environment for managing terminology. Uwe Muegge explains the benefits of this system and how it works.

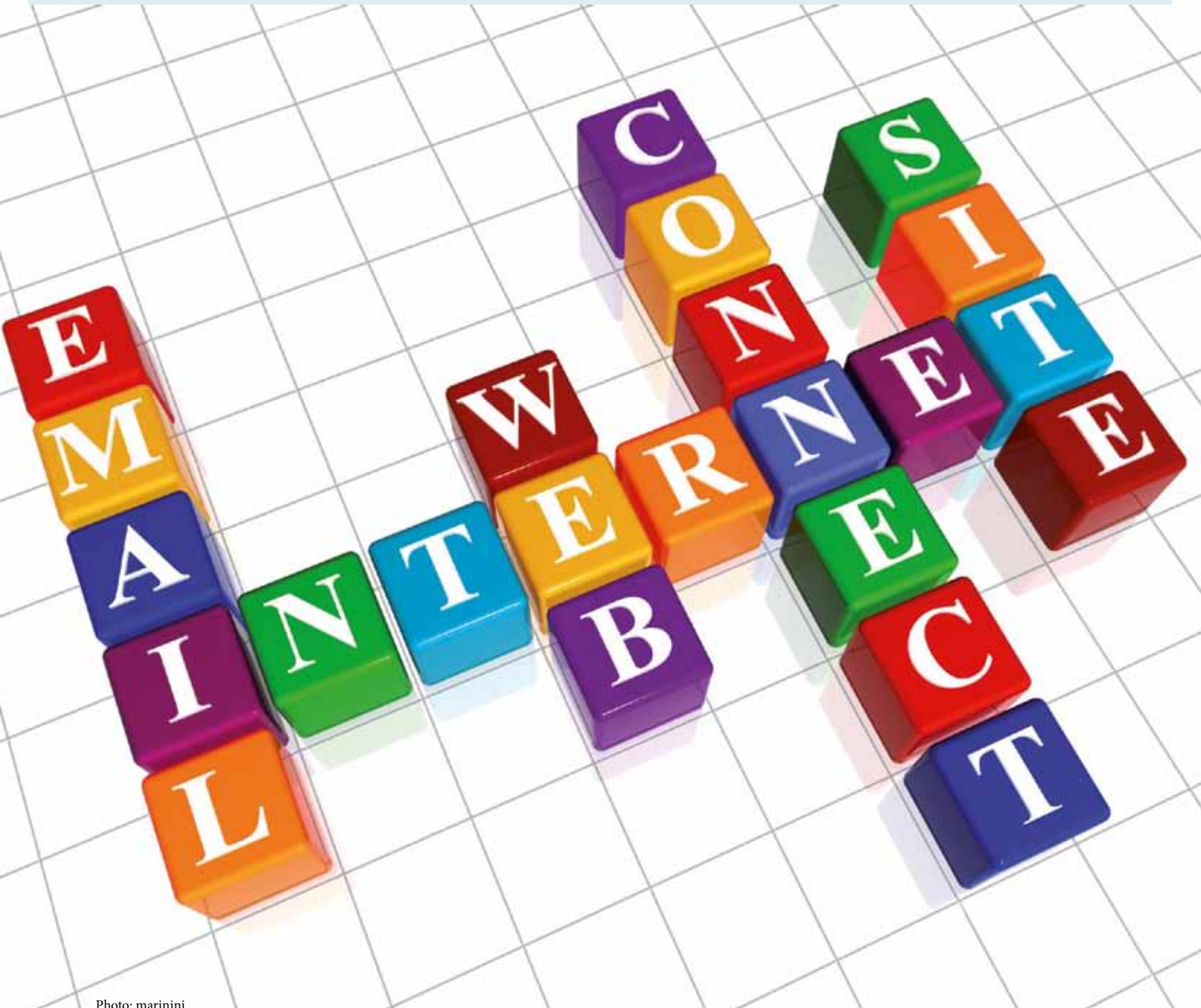


Photo: marinini

The top five issues with unstructured terminology management

Gaps in coverage/translation missing

While most organizations don't have a chartered, centralized terminology program in place, many groups within larger organizations have created word lists and glossaries to address the need for standardized terminology usage among

Definitions

collaborative software

type of software that enables work on a common project by several concurrent users that may be geographically distributed

glossary

collection of words that have special meaning in a project

term

word that has a special meaning in a given subject field

termbase

database that contains a collection of words that have special meaning in a given subject field

terminology

collection of words that have special meaning in a given subject field

terminology management

effort to control the usage of words that have special meaning in a given subject field

terminology management system

type of translation software that enables users to efficiently collect, process, and present terminology

wiki

type of collaborative software program that typically allows web pages to be created and collaboratively edited

themselves. For instance, the hardware engineering department might have an Excel spreadsheet that lists the key terms the members of this team have agreed upon. Having such a glossary is a valuable asset, as this word list can serve as a guideline for each member of the group who might use his or her own preferred term in the absence of such a resource. Maintaining a glossary is typically not found in many job descriptions or business objectives, the glossaries that exist tend to be less complete than desirable. Busy knowledge workers will find it difficult to continuously update these resources, and therefore these glossaries are typically highly incomplete: the more recent a name of a feature or function, the more likely it is that the term is missing.

Content distributed/ No complete view of all available assets

In the absence of a central repository for terminology data, Excel spreadsheets residing on individual disk drives or buried in a network file system are the typical answers to the corporate terminology challenge. In this type of fragmented environment, where various groups within an organization maintain their own termbases for their own purposes, the vast majority of employees in the organization does not benefit from these glossaries at all since they have no knowledge that these assets even exist.

Inconsistencies across terminology collections

Let's assume that the communicators in the various units of an organization do the right thing and not only maintain glossaries but also do so in a team environment, i.e. each contributor is aware of the other's efforts. Now the challenge is to normalize these collections. If each owner of a glossary maintains a separate document – and it doesn't matter if it's a simple word list in a comma-separated format, an Excel spreadsheet, or a terminology database – there will always be inconsistencies between these collections. And these inconsistencies, if undetected and uncorrected, will result in terminological inconsistencies between marketing collateral and product labeling, user manual text and illustrations, or user interface items and online help. And that's a big problem.

Content outdated/not validated

Translating terminology typically adds a layer of complexity to a terminology management

project. Glossary translations are often provided by professional translators who may or may not be familiar with the specific terminology of a sub-domain, e.g. a translator who is a medical doctor may be assigned to work on a highly specialized glossary of radiology terms. In many cases, translators don't have access to the foreign language terms that have been used in the overseas subsidiary of the client sponsoring the terminology project, because those terms have not been captured in a glossary. So for multilingual glossaries that have been created using external resources, which is the typical case today, validation by in-country subject-matter experts is highly recommended. The problem with in-country review is that qualified bilingual resources are few and far between. This means that the logistics of getting glossary documents to the right person in a foreign country, having that person review the document and send it back can be a nightmare. Which is why in the document-centric environment, many glossaries have never been validated, which in turn means that the reliability of the translated terms is questionable at best.

Content not integrated in processes such as editing, review

Many, if not all, corporate communication processes benefit from the availability of a well-maintained terminology collection, e.g. writing, editing and reviewing of technical documentation; writing, editing and reviewing of marketing collateral; translating, editing and reviewing of technical documentation; etc. With glossary documents in various stages of approval floating around in the organization, it is difficult, to say the least, to design a workflow that provides approved terminology to the various types of communicators when they need it.

The benefits of a centralized, wiki-based terminology management solution

Terminology management in the cloud

The biggest difference between a traditional documents-based terminology management approach and TermWiki's centralized web-centric paradigm is the fact that with TermWiki, e-mailing Excel documents back and forth between terminology stakeholders is a thing of the past. Any

term that a user enters into the system is available to all other users in the entire organization. And the value of giving visibility to all terminology assets an organization owns by storing them all in one place cannot be overemphasized: Centralized management of terminology ensures that all internal users (e.g. employees) and external users (clients, partners, and service providers) of terminology have easy access not only to the complete dataset, but also to the latest version of it.

Instantly familiar user interface

As one of the goals of the TermWiki project was to make the user experience of this new terminology management system as intuitive as possible, building on the proven MediaWiki platform was an obvious choice. Users of TermWiki are immediately familiar with the wiki interface that features a task-oriented layout where all major items are in plain view and immediately accessible. In TermWiki, any user will be able to make a valid contribution within minutes of their first use.

Access terminology from any computer

With TermWiki, it doesn't matter if users have a PC, MAC, Linux or other operating system, and there is no need to install and maintain any special software. The only software required for searching, entering, editing, translating, approving, and downloading terminology is a web

browser. This makes TermWiki a truly universal terminology management tool. And not only are all terms visible to every employee in the organization (depending on their user privileges), but everyone works on the same data set and in real time. With TermWiki, organizations can use glossaries as the powerful, comprehensive knowledge bases that they are.

Predefined data categories support consistency

TermWiki has been designed by members of CSOFT's engineering team. CSOFT is the largest language service provider in Asia, and therefore, these developers are intimately familiar with the language needs of global organizations large and small. This is why TermWiki goes far beyond the simple word pair model used in many glossaries that often leaves open more questions than it answers. At the same time, the developers of TermWiki also wanted to create a system that doesn't require its users to have a Ph.D. in linguistics. This is why TermWiki offers a small number of predefined data categories that have been carefully selected for their relevance to global communication projects, e.g. Definition, Part of Speech, Usage Status, Company, Product, etc. Many of these data categories even have pick lists for users to choose a value from, e.g. Part of Speech has Noun, Verb, Adjective, Adverb, Proper

Noun, and Other to speed up data entry and to ensure consistency across entries.

TermWiki editing environment

Complete audit trail

While TermWiki is a tool that makes collaboration among large and geographically distributed teams as easy as possible, that doesn't mean there are no checks and balances in place. In fact, the opposite is true. With the history feature, the system creates a permanent record of when a term is added or changed and by whom. And since each transaction is logged, it is easy to undo edits and to revert to a previous version – regardless of how long ago an edit was made. Also, supervisors can track the quantity and quality of individual contributions. Likewise, the information available under the history tab can be used as an objective source for rewarding team members who have made outstanding contributions to their corporate termbase.

Simplified validation of terms

Only a validated term is a good term! As the review of termbase entries by a subject matter expert has traditionally been a soft spot in many corporate terminology management systems, offering better validation support was high on the

The screenshot shows the 'Edit Term: EN:hardware' page in the TermWiki system. The interface includes a top navigation bar with links for 'en', 'discussion', 'edit', 'edit source', 'history', 'move', and 'watch'. The main content area contains the following fields:

- Definition:** type of computer component that consists of physical technology
- Part of Speech:** noun
- Usage Status:** Approved
- Synonym:** computer hardware
- Sample Image:** Hardware.png (with an 'Upload file' button)
- Pronunciation:** (with an 'Upload file' button)
- Industry:** A dropdown menu with options: Business services, Candy and confectionary, Computer (selected), and Construction.
- Product Category:** A dropdown menu with options: Tablet PC, Desktop PC, Servers, and Workstations.
- Product:** iPad

On the left side, there is a sidebar with the TermWiki logo and navigation links: home (Main Page, Current events), support (TermWiki Help, FAQ, Support Desk), search (Go, Search), and termwiki (Dashboard, Browse Terms, Add a New Term, Language Home, Industry Home, Product Category).

priority list for TermWiki. Fortunately, it is very easy to implement a validation process in a web-based system. Since subject matter experts have easy access to terminology, just like any other user of TermWiki, the only special features needed for validators were a) a mechanism for rating entries, and b) a forum for providing feedback/soliciting additional information. The rating mechanism was implemented as a picklist for Usage Status, i.e. the validator simply selects the appropriate value from a picklist of statuses such as New, Preferred, Admitted, Obsolete, etc. If the validation process should require an exchange of information, each entry has a discussion page associated with it, which terminology stakeholders can use to inter-actively debate the validity of an entire entry or parts thereof, provide supporting documentation, and maintain a record of the process for future reference.

Rich media experience

The old adage that a picture is worth more than a thousand words certainly holds true in a terminology management environment. A photo of an item can easily be the most informative part of a termbase entry. With digital images being so easily available from so many sources, e.g. internal corporate media repositories, collections of royalty-free work such as Creative Commons, and commercial collections such as Corbis, Fotolia, and iStockphoto, adding an image to an entry is easier and cheaper than ever. In TermWiki, all it takes to upload an image is clicking a few buttons. And by the way: TermWiki also supports uploading audio files. This is a helpful feature if you wish to assist users with the pronunciation of a new or unusual English term or any foreign-language term.

Support for standards ensures easy data exchange

TermWiki supports key terminology standards such as the LISA termbase exchange standard TBX, CSV and MultiTerm XML, the XML localization interchange format XLIFF, and the ISO standard for terminology data categories ISO 12620. TermWiki's standards compliance enables terminology users to integrate this terminology management solution into their existing IT infrastructure: TermWiki plays nice with authoring tools, content management systems, and of course, translation memory and translation management systems. And for those organizations that have invested in an older proprietary terminology management system, TermWiki preserves the financial and intellectual

investment by offering a simple data migration path. Open Source, community, and hosted editions give organizations of any size the professional terminology management solution they need.

Free open source edition is fully customizable

TermWiki is designed to meet the terminology needs of large organizations, and the functionality of the standard edition of TermWiki is probably more than adequate for many if not most corporate users. However, if an organization should have special requirements, such as the need for different or additional data categories, the architecture of TermWiki is flexible enough to allow for easy modifications of the source code. It is also possible to develop additional features that are currently not available in the standard edition, such as automatic notification of subject matter experts/reviewers when new terms are added. And since TermWiki is an open source project, users of TermWiki can either have their own internal IT resources perform this type of development work, or outsource it to CSOFT, the developer of TermWiki. To access the open source download page for TermWiki, go to the following link: http://www.cssoftintl.com/solutions_termwiki_download.php

Community edition gives freelancers a powerful terminology management tool

The easiest way for individuals or groups of language professionals to manage and share terminology, is to use the community edition of TermWiki, which is available free of charge at www.termwiki.com. Using the community edition of TermWiki, anyone with an interest in terminology management has immediate access to a structured environment for storing, sharing, and finding terms, definitions, and related information. Best of all: The community edition of TermWiki is not only free like the open source edition, but there is not even any software to download and install since the community edition is a hosted, web-based service. The community edition of TermWiki is the right platform for people who want to launch an individual or group terminology project without having to commit any financial and/or IT resources to that project first.

Value-added services available

Organizations that are looking for a customized solution or additional functionality, complete privacy of their data, short deployment time, and ensured maintenance, have the option of

using CSOFT's development and hosting services. CSOFT is offering various packages that include software hosting, data migration, additional modules (e.g. notification module, data import/export modules, etc.), and consulting/training services.

The bottom line

TermWiki is an innovative answer to an old and persistent question: How to provide a structured environment for managing terminology in organizations that typically don't have a high level of linguistic expertise? TermWiki offers the powerful multilingual management and tracking features large organizations require and yet is so easy to use that anyone with the most rudimentary computer skills will be comfortable with this system within minutes. And now that there are free open source and instantly available community editions, TermWiki is probably not only the most flexible but also the most cost-effective web-based terminology management system available today.

References

- DEPALMA, D.A. TermWiki Lowers the Barrier to Terminology Management. Global Watchtower [online]. 2010, no. 14 January 2010. Available from World Wide Web:<<http://www.globalwatchtower.com/2010/01/14/cssoft-termwiki/>>.
- MUEGGE, U. Disciplining Words: What You Always Wanted to Know About Terminology Management. *tcworld*, 2007, vol. 2, no. 3, p. 17-19.

contact



Uwe Muegge is the Director of MedL10N, the life science division of CSOFT. He is currently a member of TC37 at the International Organization for Standardization (ISO) and teaches graduate courses in

Computer-Assisted Translation at the Monterey Institute of International Studies.

uwe.muegge@medl10n.com
www.medl10n.com
www.muegge.cc