

Anita Nuopponen, University of Vaasa, Finland

TERMINOLOGICAL INFORMATION AND ACTIVITIES IN WORLD WIDE WEB

1. Introduction

In the sixties Eugen Wüster (e.g. 1969: 6) saw computer as a tool for the terminography work when compiling dictionaries, as a translation tool giving translation equivalents, and further as a storage and retrieval tool for the purposes of coordination of terminology work. Today different kinds of software are used for vocabulary compilation and vocabularies or term banks can be sold on a diskette or a CD-ROM. Terminological databases can be integrated to translation software. Programmes for extracting terminological information from whole texts are being developed. Furthermore, computer networks have appeared and data collections can be reached from all over the world, as Wüster in the sixties - long before the Internet - was looking forward.

In this paper I shall concentrate on the terminological information and activities found in the Internet - more specifically in the World Wide Web (WWW, Web) - and the possibilities yet to be utilized.

The paper is based on the results and observations from two year's regular browsing of the Web. I have been searching for pages that are interesting for terminological research and terminology work¹ and following the evolving of new ways to utilize the new communication channel. Actually, the whole Web can be used as an encyclopaedia when looking for information about special field knowledge and terminologies in many different languages. Here I shall concentrate in comparing the Web with the traditional means of terminological information and documentation (I&D) and other terminological activities.

In order to structure the use of the Web for terminological purposes I distinguish between Information and Documentation (I&D) and Activities. Sometimes this is difficult because the Web is mixing up these two concepts. With I&D I refer to collecting and disseminating documents and e.g. bibliographical information. As activities I consider different kinds of attempts to utilize the Web more widely as an interactive information and communication tool as well as a co-operative environment.

2. Information and Documentation

Documentation, collecting and disseminating information, has been one of the central activities in the field of terminology. Felber lists the following tasks for terminological documentation:

- "1) collecting, recording, analyzing, ordering and storing of
 - terminological documents
 - terminological data (on individual concepts)
 - factographic data (terminological organisations, commissions, experts, data banks etc.)
- 2) providing pertinent documents and data, world wide,
- 3) disseminating terminological information via information systems or in form of secondary literature (bibliographies, directories etc.) and tertiary literature (bibliographies of bibliographies)." (Felber 1984: 334)

Traditionally this kind of information has been disseminated through dictionaries, bibliographies, newsletters, periodicals and other printed publications. The problem often is that the information given in a book or even in a newsletter is often out of date already when it reaches those who need it. Since sixties the information technology has been utilized for storing and retrieval of terminological data (i.e. translation equivalents, definitions etc.) in local term banks with an eventual modem access for personnel or paying members or clients. Other kind of information about terminological activities has mostly been distributed as printed publications. Some local data bases exist, but they have been only for local use (e.g. the data collections of Infoterm/TermNet).

The introduction of the Gopher and the Web made it possible to make the information freely available online. The Web is said to be already the world's largest collection of knowledge. With a little practice one can find very fast documents e.g. about one's own research field and thus get an overall picture what is happening on that specific field. Different kinds of searching tools, directories and indexes make the searching relatively easy.

I divide the terminological documentation found in the Web in primary, secondary and tertiary sources even though the borders are not always easy to draw (see fig. 1). (See Arntz & Picht 1989: 276; Felber 1984: 341.)

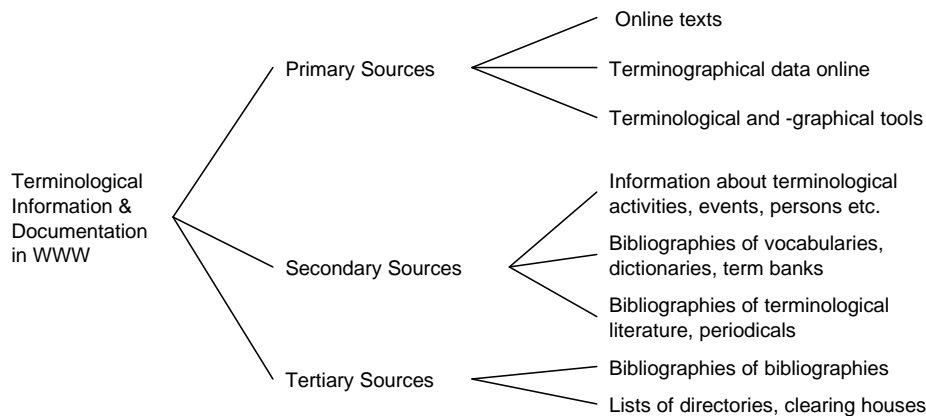


Fig. 1 Terminological Information & Documentation in the Web

2.1 Primary Sources

Publishing in the Web differs in many ways from the traditional publishing. It is faster and cheaper but also uncontrolled. The decision about the quality and origin of the discovered information is left to the information seeker. At the moment there are not yet many articles or longer texts about terminological research or terminology work available in the Internet², but I hope terminology people will soon "find" this media as a publication possibility too. Especially I would prefer to have the classical texts to be online, e.g. Wüster's dissertation. There is always the problem of copy-right: some of the articles must be withdrawn from the Web when they are published in printed form.

As soon as Gopher and the Web became popular, glossary activities started to flourish online. There are plenty of general language dictionaries³ but the amount of special language glossaries and vocabularies⁴ is growing even more rapidly. The first ones online were dictionaries like *Webster*⁵ and *Roget Thesaurus*⁶. Both private and cooperative efforts are made to publish online special language vocabularies. Here, again, the user has to evaluate the reliability of the source very carefully. Vocabularies are collected by one person, university institutions, research institutes or research groups, associations or companies. The Web seems to have made the so called hidden glossaries easily available, e.g. product and company specific glossaries. Companies are more and more starting to provide glossaries of the central terminology together with the product descriptions in the Web. This gives an added value also to the company's or association's home page.

The established "old" terminological data banks except *EURODICAUTOM*, the term bank of the European Union, are still missing in the Web. *EURODICAUTOM* could be reached freely over Internet very early. It has received an experimental WWW gateway last year as an unfunded project at the University of Frankfurt⁷. The data base is not

transferred to a WWW data base, but instead, the researchers are developing a link, an interface, between the original data base and the Web. Here are still many problems to be solved and many of the features of the database can't be offered through the Web interface. (ibid.)

The traditional terminological data banks contain an abundance of information about the terms and concepts as well about the concept and term relations. They use a data base and an information retrieval system. Most of the new terminological data collections in the Web are not terminological data bases or term banks in this sense because they contain very little data about the term and the conceptual relationships. Mostly they give just an entry term and a definition - both in one language. They are very often term-oriented and could be compared with printed dictionaries and even might have been that earlier. Actually many of them are simple html files with entries and definitions listed in an alphabetical order, e.g. *Medical Dictionary by MedicineNet*⁸. An alphabetical index, or sometimes also a search programme can be used for finding the definition. The more complex glossaries use *CGI (Common Gateway Interface)*⁹ program to interlink the Web and a separate data base. Sometimes they include cross references as links to the related terms. As an example of a more advanced online dictionary of this type is BioTech's free online biotechnology dictionary¹⁰. Hypertext possibilities are not yet used very much.

In addition to the world wide availability, terminologists have been dreaming for a long time of the possibility to add picture, video, animation and sound in a terminological data bank¹¹. In printed technical or other specialized vocabularies drawings, pictures, figures etc. have always been an essential component, because they can replace long definitions. In the Web it is possible, but at the moment there are not yet many multimedial data bases. BioTech's dictionary¹² and *Illustrated Glossary of Geologic Terms*¹³ try to utilize also the possibility of linking pictures to the entries. Multimedia is specially valuable for sign language dictionaries because it enables to show the necessary facial expressions (e.g. movements of eyebrows) and body language. An interesting project is *German Sign Language Dictionary on Psychology* compiled by Sign Language Research Center at Hamburg University¹⁴.

In the Web there are a several descriptions of the projects developing computer tools for terminology work. Some projects have worked on their own guidelines that they have published in the Web together with other information.¹⁵ This kind of information could be considered as a primary or secondary source depending on the contents. In software design it is usual to let the users to download and test the software freely in its earliest development phases. The same could be applied to the terminological and terminographical tools.

2.2. Secondary and tertiary sources

Secondary and tertiary sources collect information on the primary sources, e.g. terminological literature and vocabularies. *Infoterm*, *TermNet* and other different kinds of national, regional or international organisations that among other things carry out with terminological documentation. This is mostly done by publishing printed bibliographies and directories, for instance, *Infoterm* and *TermNet* publish a series of international bibliographies and directories including following titles:

International bibliographies of

- terminological papers and books (BT 1)
- periodicals pertaining to terminology (BT 7)
- terminological theses and dissertations (BT 12)
- reports and proceedings of terminological meetings (BT 11)
- standards and non-standardized guidelines for - terminology (BT 6)
- standardized vocabularies (BT 2)
- multilingual specialized vocabularies (BT 3)
- monolingual specialized vocabularies (BT 4)
- dictionaries of acronyms, initialisms and abbreviations (BT 13)
- collections of neologisms (BT 14)
- computer-assisted terminology (BT 9)
- bibliographies and catalogues containing terminological literature (BT 8)

World guide to terminological activities (BT 5)

Directory of terminological meetings (BT 10)

Who's who in terminology (BT 14)¹⁶

The bibliographies have been published either as whole books or as articles (see e.g. Manu 1980; Felber/Krommer-Benz/Manu 1979; Felber/Nedobity 1984; Felber 1979). Most of these bibliographies are now also available on a CD-ROM called *Sophia - European Databases in the Humanities* produced by the Library of the University of Vaasa¹⁷. The selected terminological bibliographies contain ca. 19 000 references with keywords on the CD-ROM. This kind of information must be an hopefully will be updated quite often.

Bibliographical information can also be found in the Web. I have been using the list of bibliographies above as a basis for my classification of Web sources (see fig. 1), because it covers the field of terminology quite well. In the Web there exist lists of printed vocabularies from different fields, e.g. Dykes Library at the University of Kansas has a list of medical dictionaries in its *Online Medical Reference System*¹⁸. The *International Organization for Standardization* (ISO) has lists of its publications in the Web including both vocabularies and guidelines for creating, compiling and coordinating terminologies¹⁹.

In addition to the bibliographies of printed vocabularies, there are also plenty of lists²⁰ of vocabularies, dictionaries and data bases that are freely or with restrictions available in the Web. During a couple of years' time I have been collecting links to both special language and general language vocabularies²¹. Both collections are growing gradually, but specially the number of vocabularies from different fields is increasing very. Directories and link collections like this may remind conventional bibliographies but the entries function not only as bibliographical facts but lead directly to the data base concerned. The Web gives an excellent possibility for frequent updates and immediate publishing of the updated pages. One problem of this kind of "bibliographic" work is that the links must be checked quite often because the vocabulary compilers may change the addresses for some reason or totally withdraw their database from the Web.

Directories like *Who's who in terminology*, *Directory of terminological meetings* and *World guide to terminological activities* represent information sources that have their natural place in the Web. Information in them is even more rapidly changing than in the bibliographies. In many fields there are such services. Information about scholars working with terminology, lexicology and translation in the French speaking world is found in *Marie Recherche* server maintained by Mme Danielle Collignon in Montréal²².

Terminological meetings are not yet covered, but I am adding information on terminological conferences in the *Terminology Forum* page in the case I happen to receive a call for papers. It is also interesting to add not only links to coming events but also to the past events. One channel for terminological conference information is a mailing-list called *Term-List* (term-list@uwasa.fi) that reaches quite a lot of people interested in practical and theoretical terminological activities. The *Terminology Forum* represents an equivalent of the *World guide to terminological activities*.

3. Terminological Activities

Instead of only distributing information about terminological activities the Web can be used to coordinate or conduct the activity itself, e.g. a project, a course, a conference, information service, etc. There are not yet, however many of terminology centres online. Also both Infoterm and TermNet are still missing totally from the Web. Fig. 2 suggests some possible developments for the terminological activities in the Web.

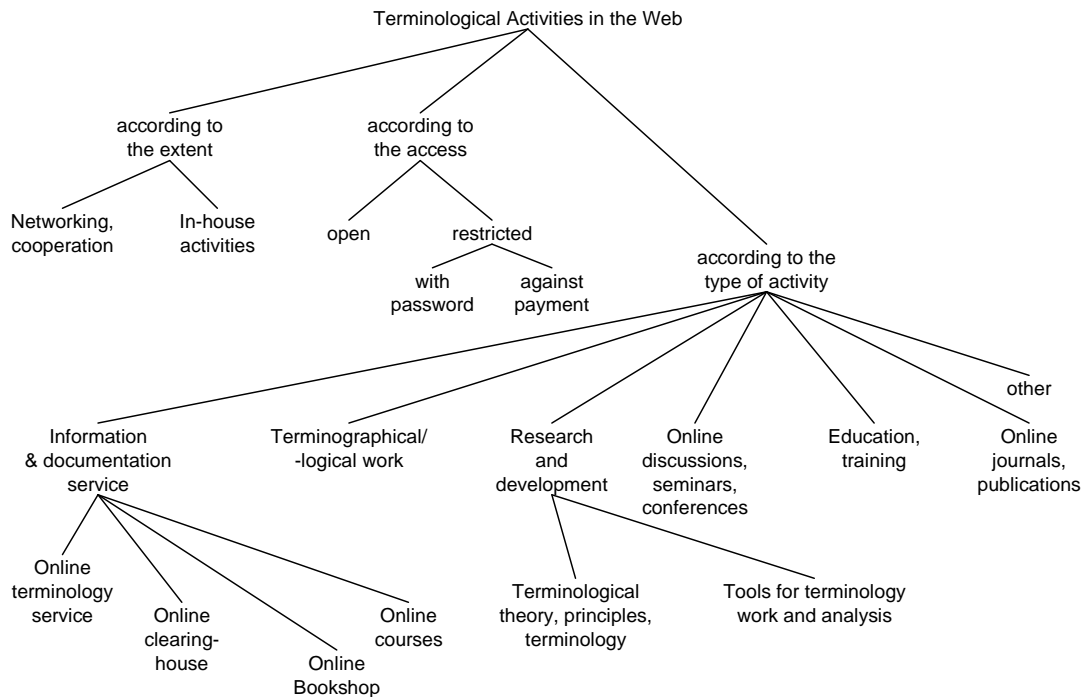


Figure 2. Some possible developments of terminological activities in the Web

3.1 Coordination and Information Service

The Web offers an ideal infrastructure specially for different kinds of cooperation and networking. It can be used to create new cooperation and intensify the old connections. The need for international cooperation networks have been recognized early on in the field of terminology. Already in the turn of the century some fields, e.g. electrotechnics, started to develop their terminology in a world-wide cooperation. Infoterm and TermNet are examples of the international networking activities in this field, both established in the seventies. It will be interesting to get these and other terminological organizations in the Web. At the moment coordination organisations like *Realiter* (*Red Panlatina de Terminología*) and *RITerm* (*Red Iberoamericana de Terminología*) have their home pages²³, but not yet any other activities in the Web.

Internet does not exclude anymore commerce and this opens many new possibilities for terminological service providers. For terminology centres, the Web can bring many new possibilities to give services restricted only to their members and other interest groups if they so wish. Journals and newsletters can be published electronically with a free or a password access. Publications and different kind of services (e.g. term banks) can be advertised and sold online. A good example is the newly opened page of *The Swedish Centre for Technical Terminology* (TNC) that has not only information about the TNC dictionaries and the TERMDOK CD-ROM database but also forms to order these by e-

mail²⁴. Recruiting new members can also be done by telling about the membership benefits.

3.2 Projects and Terminology Work

In projects the Web can function as a collaborative forum giving up-to-date information about the participants and the progress of the process, to discuss, and to keep track of the different stages of the documents, to test and get and give feedback etc. Rules, principles, philosophy and theoretical background for the project can be explained in the Web. Proposals etc. can be submitted for discussion. Some of the activities and services in the Web may need to be restricted while other parts may be open for all.

Some of the activities can be realized totally in the Web. One such project is the *Virtual Hyperglossary* (VHG) that was started 1995 in the connection of the *Virtual course on Principles of Protein Structure* at Birkbeck College, Crystallography Department. It is a collaborative project with the ambition "to provide a distributed set of glossaries (terminological data bases) on the WWW".²⁵ According to the VHG philosophy "glossaries can be created by virtual teams"²⁶.

Terminology work is often centred around different kind of organisations and associations. The Internet and specially the Web gives an opportunity for the individual subject field specialists, researchers, translators, LSP teachers, technical communicators etc. to participate in a voluntary cooperation. Projects like *LOGOS*²⁷ and *Free On-Line Dictionary of Computing* (FOLDOC)²⁸. Both are freely accessible databases based on a voluntary compiling work. LOGOS is a multilingual dictionary updated and corrected by a network of professional translators meanwhile FOLDOC is a monolingual English definition data base compiled by 600 volunteers with a good knowledge of computing.

For very specialized subject fields producing an online glossary seems to be one way to connect the geographically widely scattered experts. Many glossaries are collected in connection with research projects or courses.

3.3 Conferences and Education

In many areas there are already courses and virtual colleges in the Web that could be used as models for terminological education and training, too. Several of the university institutions interested in terminology research and work have their home pages, e.g. Innsbruck, Surrey, Vaasa²⁹. I have been creating Web sites for my courses in order to distribute handouts, slides, all kind of information concerning the courses, defining the assignments and publishing the best ones. This is just a modest start but I am planning a Web based course for open and distance learning purposes. Universities and others

offering courses in terminology research and work could be cooperating by using the Web and other Internet tools. Some plans exist already.

The Web is nowadays often being used as an information tool throughout the whole process of organising a conference. The Web site goes through changes from the very first call for papers. The abstracts and the programme as well as other conference information can be conveniently distributed before the conference (e.g. this conference) and afterwards maybe the whole articles will be. As an example could be mentioned *The Fifth International World Wide Web Conference*³⁰. The Web site includes a link to the next conference page, paper proceedings, presentation slides, audio and video clips of speakers, overview of the event, programme, press releases, conference broadcast, information on the organizing committee, links to the sponsors, history of the conference series (links to the previous conferences' home pages), attendance information, registration, hotel and travel information etc. There could even be conferences that would be only organized in the Web.

4. Conclusion

The Web is an ideal platform for various kinds of terminological activities. It is not only an online archive or storage for electronic documents or a world-wide bibliography or encyclopaedia of terminological and terminographical information but it can play an important role as an interactive collaborative communication tool in terminological research, terminographical projects, as well as in education. Even though there is already plenty of terminological activities in the Web, the major part is still missing - hopefully not very long.

Notes

¹ See: Terminology Forum: <http://www.uwasa.fi/comm/termino/>

² See Terminology Forum: <http://www.uwasa.fi/comm/termino/>

³ See e.g.: <http://www.uwasa.fi/comm/termino/diction.html>.

⁴ See e.g.: <http://www.uwasa.fi/comm/termino/special.html>

⁵ See e.g.: <http://c.gp.cs.cmu.edu:5103/prog/webster>.

⁶ ARTFL Project's hypertext version: http://humanities.uchicago.edu/forms_unrest/ROGET.html

⁷ See; <http://www.uni-frankfurt.de/~felix/eurodicautom.html>. Note: this is under a private home page of Felix Gaetgens and is therefor subject to changes.

⁸ MedicineNet: <http://www.medicinenet.com/index.html>

⁹ See more: The Common Gateway Interface: <http://hoohoo.ncsa.uiuc.edu/cgi/>

¹⁰ BioTech: Biotechnology Dictionary: <http://biotech.chem.indiana.edu/pages/dictionary.html>.

¹¹ E.g. Norevik i: Nordterm 1983. Seminarium: Datorstött terminologiarbete 3-5.5.1993. Swedish Centre for Technical Terminology, Sweden.

¹² BioTech: Biotechnology Dictionary: <http://biotech.chem.indiana.edu/pages/dictionary.html>.

¹³ Illustrated Glossary of Geologic Terms: http://www.public.iastate.edu/~geology_100/glossary.html.

-
- ¹⁴ German Sign Language Dictionary on Psychology:
<http://www.sign-lang.uni-hamburg.de/english.html>
- ¹⁵ See: <http://uwasa.fi/comm/termino/research.html/>
- ¹⁶ Picht & Draskau 1985: 205f.; Infoterm 4-82: 3.
- ¹⁷ See: <http://www.uwasa.fi/kirjasto/sophia.html>.
- ¹⁸ Dykes Library Online Medical Reference System, University of Kansas:
<http://www.kumc.edu/service/dykes/refassist/refhome.html>
- ¹⁹ ISO vocabularies: <http://www.iso.ch/cate/01.html>
- ²⁰ See <http://www.uwasa.fi/comm/termino/others.html>.
- ²¹ See: <http://www.uwasa.fi/comm/termino/>
- ²² Marie recherche: <http://www.refer.qc.ca/donnees/M10/pcsearch.htm>
- ²³ Realiter: <http://www.iula.upf.es/cpt/cptit.html>; RITerm: <http://www.iula.upf.es/riterm/ritempo.html>
- ²⁴ TNC: <http://www.tnc.se/>
- ²⁵ The Virtual Hyperglossary: <http://www.dl.ac.uk/CBMT/glossary/intro.html>.
- ²⁶ Ibid: <http://www.dl.ac.uk/CBMT/glossary/philosophy.html>.
- ²⁷ LOGOS: <http://www.logos.it/query.html>
- ²⁸ FOLDOC: <http://wombat.doc.ic.ac.uk/foldoc/editing.html>.
- ²⁹ See more: <http://www.uwasa.fi/comm/termino/>
- ³⁰ The Fifth International WWW Conference: <http://www5conf.inria.fr/>

Bibliography

- FELBER, H. (1979). Draft - Project recording of bibliographic data of mono- and multilingual specialized vocabularies (BT 3/4). Wien: Infoterm, 3 p.
- FELBER, H./ KROMMER-BENZ, M./MANU, A. (1979). International Bibliography of standardized vocabularies. Infoterm Series 2. München; New York; London; Paris: K. G. Saur. 540 p.
- FELBER, H./ NEDOBITY, W. (1984). International Bibliography of standards and non-standardized guidelines for terminology (BT 6). Wien. Infoterm. 13 p.
- INFOTERM 4-82. Infoterm Annual Report 1981. 19 p.
- KROMMER-BENZ, M. (1984). International bibliography of computer-assisted terminology. Wien: Infoterm. 40 p.
- KROMMER-BENZ, M. (1985). World guide to terminological activities. 2nd completely revised and enlarged edition. Infoterm Series 4. K. G. Saur. 158 p.
- MANU, A. (1980). Internationale Bibliographie mehrsprachiger Wörterbücher auf dem Gebiet des Bauwesens. International Bibliography of multilingual specialized vocabularies in the field of construction. Wien Infoterm, 3 p.
- PICHT/DRASKAU (1985). Terminology: An Introduction. University of Surrey, England.
- WÜSTER, Eugen (1969). Die vier Dimensionen der Terminologiearbeit. Mitteilungsblatt für Dolmetscher und Übersetzer, Nr. 2/15 (März), 1-6.