Terminology Tools –
A closer look at Terminology Management Systems

Terminology Summer School - Vienna
09 - 13 July 2012

Klaus-Dirk Schmitz
Institute for Information Management
Faculty 03
University of Applied Sciences Cologne
klaus.schmitz@fh-koeln.de
Overview

- Tools for managing terminology
  - Terminology extraction tools
  - Terminology management systems
  - Terminology control tools
In many application scenarios of terminology work, the extraction of terminology from (existing) textual material is recommended.

We can differentiate between the following extraction methods:

- **Monolingual** term extraction *(text in electronic form)*
- **Bilingual** term extraction *(parallel aligned texts, i.e. TMs)*
- **Manual (human)** term extraction
- **Computer-assisted** term extraction *(tools propose term candidates)*
  - With **statistical** methods *(for “all” languages, cannot use knowledge about syntax)*
  - With **linguistic** methods *(better results, but only for “important” languages)*
  - With **hybrid** methods *(combining statistical and linguistic methods)*
Features of (monolingual) term extraction tools:

- Common functionalities from concordance programs (e.g. WordSmith): identify words, word statistics, KWIC index, alphabetic/frequency order
- Reducing inflected word forms to the basic canonical form: needed for real statistics, needs morphological knowledge
- Filtering and ignoring function words (articles, conjunctions etc.) and general language words (but what is general language?)
- Filtering and ignoring terms that are already included in a term base
- Identifying multi-word terms, noun phases and verbal phases
- Identifying discontinuous elements and elliptical constructions
Terminology extraction tools

Improving and enriching term candidates with SDL MultiTerm Extract
Settings to improve the results of term extraction with SDL MultiTerm Extract
Terminology extraction tools

Benefits and problems of term extraction tools:

- Term extraction tools are helpful in preparing terminology for large translation projects (with several translators) and for an initial feeding of a term base (with company or subject specific terminology).

- Result of a term extraction is a list of term candidates; the list must be checked; but what about the texts (with possible not extracted terms)?

- Results are only terms (and context examples), but no other terminological information; it is a kind of a to-do list for the terminologist.

- The more linguistics the better the results; but what about “less common” and minority languages?
Why terminology management systems?

- Besides the necessity to have access to existing electronic dictionaries, terminology data collections and term banks,
- and the availability of internet access for all kinds of terminology research and improvement,
- we need a mechanism and a tool to manage our "own" terminology
Questions before

- How many users will enter data (add + edit) and how many will have access (consult) to the terminology management solution?
- Different users with different access rights?
- Are the users all in one office, one building, one city, one country?
- Interchange of data necessary?
- Usage of data in other applications?
Which TMS solution?

- with word processors, spreadsheets
  - available and usage well known
  - searching and sorting possible
  - “flexible” structure
  - create a table with several columns for data categories
  - slow if many entries
  - inadequate for many data categories
  - poor retrieval facilities
  - no systematic terminology work possible
  - no concept orientation and term autonomy
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFB Vorder Theologische Fakultät</td>
<td>m</td>
<td>AFB connector</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFB Vorder Text</td>
<td>m</td>
<td>AFB connector text</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktiv</td>
<td>m</td>
<td>short</td>
<td>v</td>
<td>intermittenly closed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Anlagen Ausführung</td>
<td>m</td>
<td>short for current execution</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Abwasserbearbeitung</td>
<td>m</td>
<td>short for current treatment</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Abwasserentfernung</td>
<td>f</td>
<td>short for current removal</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Abwasserentfernung Distanz möglich</td>
<td>f</td>
<td>short for current distance possible</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Abwasserentfernung</td>
<td>m</td>
<td>short for current distance</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aktuelle Abwasserentfernung</td>
<td>m</td>
<td>short for current distance</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abwicklung durch Glaschichten</td>
<td>f</td>
<td>glass layer packing compound</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ablauf</td>
<td>m</td>
<td>drainage</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ablauf</td>
<td>m</td>
<td>drop</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Which TMS solution?

- with word processors, spreadsheets (MS Word, MS Excel)
Which TMS solution?

- with word processors, spreadsheets (MS Word, MS Excel)
Which TMS solution?

- with data base management systems
  - available and usage (well) known
  - powerful data modeling and retrieval
  - create and link relations (tables)
  - not ideal for linguistic data, but ok.
  - not appropriate in handling, but user interface programmable
  - no data consistency check (values allowed, mandatory values), but programmable
  - concept orientation and term autonomy possible, but expert data modeling needed
<table>
<thead>
<tr>
<th>Nr:</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>English:</td>
<td><strong>adaptor</strong></td>
</tr>
<tr>
<td>Synonyms:</td>
<td>adapter</td>
</tr>
<tr>
<td>German:</td>
<td><strong>Paßstück</strong></td>
</tr>
<tr>
<td>Synonyms:</td>
<td>Zwischenstück</td>
</tr>
<tr>
<td></td>
<td>Verbindungsstück (-teil)</td>
</tr>
<tr>
<td></td>
<td>Übergangsstück (-teil)</td>
</tr>
<tr>
<td>Note:</td>
<td></td>
</tr>
<tr>
<td>English Definition:</td>
<td>&gt;device for connecting sth. or making sth. &gt;fit</td>
</tr>
<tr>
<td>English Context:</td>
<td>Join these tubes of different diameters using adaptors.</td>
</tr>
<tr>
<td>German Definition:</td>
<td></td>
</tr>
<tr>
<td>German Context:</td>
<td></td>
</tr>
</tbody>
</table>
Which TMS solution?

- with terminology management programs
  - 😊 exactly adjusted to terminology work
  - 😊 powerful data modeling and retrieval
  - 😊 concept orientation and term autonomy provided or definable
  - 😊 elaborated user management, consistency procedures, interfaces to other applications and interchange options
  - 😞 not well known and not cheap
  - 😞 not all products fulfil all terminological needs
Terminology management systems

Terminology management systems are software applications that are designed to manage terminological data.

They support tasks related to terminology work and store the results: Terminological data can be entered, edited, deleted, retrieved and filtered.

Most of the systems available on the market are based on (relational) data base systems (MS-Access, SQL, Oracle).

Can be seen as a kind of CAT-Tools (CAT=computer assisted translation).

Tables in word processing or spreadsheet programs are not adequate for terminology management!
Terminology management systems

Classification of terminology management systems:

- **Complexity (languages):** monolingual / bilingual / multilingual
- **Entry structure:** predefined / free-definable / hybrid
- **Autonomy:** autonomous / CAT tool component / hybrid
- **Software technology:** stand-alone / client-server / browser-based
- **Business aspects:** proprietary / commercial / open source

E.g. SDL MultiTerm 2009
Designing a terminology management solution

Before designing a terminology management solution and choosing, adapting or programming a terminology management application:

- Analyze the needs and objectives
- Specify the user groups, tasks and workflow
- Define the terminological data categories needed
- Take into account the basic modelling principles
- Model the terminological entry
- Select, adapt, develop the software

Meta data

K.-D. Schmitz, IIM, FH Köln
<table>
<thead>
<tr>
<th>Entries</th>
<th>Hits Status</th>
<th>Ready</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE Zündkerze</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source PS-WbAuto, 423</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EN spark plug, sparking plug</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source DicAuto, 179; Gm23, 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Context:**

Kontext: The sparking plug ignites the compressed fuel-and-air mixture in the cylinder of a petrol engine by means of a spark which leaps from the central electrode.

Quelle: Gm23, 24
TWIN V.1.1 Usr - [German Key Search] Term Base: C:\PROGRA~1\TERMIN~1\TWIN\TERMS.MDB

Keyword: Schaltfläche PT

German Term: Schaltfläche
English Term: push button
French Term:
Spanish Term:
Russian Term:
Italian Term:
Portuguese Term:
Dutch Term:
Catalan Term:

Syst./Product: 
Subject Label: GUI
Subject Code: E4003
UDF 1:
UDF 2:
Pool: DTEN
Quality: M

Author: HERB Site: HERB
LUN_ID: 200478 Date: 04.07.92 01:0

Term Details: <PgDn>

German
Source: MSSYST, 4/92

English
Source: MSSYST, 4/92

Show All
Show Duplicates

Record: 1 of 14
TMS with defined (“fixed”) entry structure

• complex multilingual TMS

TWIN (Termbase for Windows, Blaha)
TMS Types: Samples

- TMS with defined (“fixed”) entry structure
  - complex multilingual TMS
    • with term autonomy

CrossTerm (across)
TMS with 2 versions and hybrid approach

MemoQ and qTerm (Kilgray)

TMS Types: Samples

1. IN WHICH WE ARE INTRODUCED TO WINNIE-THE-POOH AND SOME BEES, AND THE STORIES BEGIN
   
   QUO IN CAPITE NOBIS OSENTANTUR WINNIE ILLE PU ATQUE APESE NONNULLAE ET INCIPIUNT FABULAE

   ECCE EDUARDO USUS SCALIS NUNC TUMP-TUMP-TUMP OCCIPITE GRADUS PULSANTE POST CHRISTOPHORUM ROBINUM DESCENDENS.

2. HERE is Edward Bear, coming downstairs now, bump, bump, bump, on the back of his head, behind Christopher Robin.
   
   IT is, as far as he knows, the (1) only way of coming downstairs, but sometimes he feels that there really is another way, if only he could stop.
<table>
<thead>
<tr>
<th>M ID</th>
<th>English</th>
<th>Latin</th>
<th>Modified by</th>
<th>Modified at</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Winnie-the-Pooh</td>
<td>Winnie ille Pu</td>
<td>KB</td>
<td>2009.05.18 11:24</td>
</tr>
<tr>
<td>1</td>
<td>Christopher Robin</td>
<td>Christophorus Robinus</td>
<td>KB</td>
<td>2010.02.06 10:22</td>
</tr>
<tr>
<td>2</td>
<td>honey</td>
<td>mel</td>
<td>KB</td>
<td>2010.02.06 10:25</td>
</tr>
<tr>
<td>3</td>
<td>balloon</td>
<td>folliculum</td>
<td>KB</td>
<td>2010.02.06 10:27</td>
</tr>
<tr>
<td>4</td>
<td>cloud</td>
<td>nubecula</td>
<td>KB</td>
<td>2010.02.06 10:28</td>
</tr>
</tbody>
</table>
TMS Types: Samples

- TMS with 2 versions and hybrid approach
- MemoQ and qTerm (Kilgray)
K. D. Schmitz, IIM, FH Köln

- TMS with 2 versions and hybrid approach
- MemoQ and qTerm (Kilgray)

TMS Types: Samples
TMS Types: Samples

<table>
<thead>
<tr>
<th>Field name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creator</td>
<td>Administrator</td>
</tr>
<tr>
<td>CreateTime</td>
<td>12/26/2010</td>
</tr>
<tr>
<td>Note</td>
<td>Whorey</td>
</tr>
<tr>
<td>Language</td>
<td>Czech</td>
</tr>
<tr>
<td>Term</td>
<td>Účetní komise při Správní komisi sociální zabezpečení migrujících pracovníků</td>
</tr>
<tr>
<td>Case sensitive</td>
<td>Permissive</td>
</tr>
<tr>
<td>Forbidden</td>
<td>False</td>
</tr>
<tr>
<td>Prefix matching</td>
<td>50% minimum</td>
</tr>
<tr>
<td>Language</td>
<td>English</td>
</tr>
<tr>
<td>Term</td>
<td>Audit Board attached to the Administrative Commission on security for migrant workers</td>
</tr>
</tbody>
</table>
wastewater treatment plant

wastewater treatment works

HAS SUBTYPE: physical treatment plant
HAS SUBTYPE: biological treatment plant
HAS SUBTYPE: chemical treatment plant
HAS SUBTYPE: industrial wastewater treatment plant
HAS SUBTYPE: sewage treatment plant

Created: 2008-06-18 21:18:52; Created by: Bodil Nistrup Madsen (bnm); Modified: 2009-10-13 15:20:09; Modified by: Anna Odgaard (aeo);
Diagram: Wastewater treatment

Wastewater treatment plant

1. PURPOSE:
wastewater treatment

METHOD OF TREATMENT

1.1 physical treatment plant
METHOD: physical

1.2 biological treatment plant
METHOD: biological

1.3 chemical treatment plant
METHOD: chemical

PROCESS

1.2.1 activated sludge plant
METHOD: biological
PROCESS: activated sludge

1.2.2 trickling filter
METHOD: biological
PROCESS: activation of microorganisms

1.2.3 stabilization pond
METHOD: biological
PROCESS: oxygen dissolution

TYPE OF WASTEWATER

1.4 sewage treatment plant
TYPE OF WASTEWATER: produced by humans

1.5 industrial wastewater treatment plant
TYPE OF WASTEWATER: produced by industrial or commercial activities

K.-D. Schmitz, IIM, FH Köln
burden of proof  

**Subst. 1-1: d-e lat.**

**Definition** [...] is employed to signify the duty of proving the facts in dispute on an issue raised between the parties in a cause. [...] The burden of proof always lies on the party who takes the affirmative in pleading. [...] In criminal cases, as every man is presumed to be innocent until the contrary is proved, the burden of proof rests on the prosecutor, unless a different provision is expressly made by statute.

www.legallawterms.com/Legal.asp-Definition-BURDEN%20OF%20PROOF: 11.04.05

onus of proof  

**Subst. 1-1: d-e**

**Kontext/Beispiel** In civil cases, the onus of proof lies with the claimant who must prove his case by balance of probabilities.

www.sixthform.info/lew/03_dictionary/dict_no.htm: 11.04.05

onus probandi  

**Subst. 1-1: d-e**

**Kontext/Beispiel** The party on whom the onus probandi lies is entitled to begin, notwithstanding the technical form of the proceedings.

www.legallawterms.com/Legal.asp-Definition-ONUS%20PROBANDI: 11.04.05

Beweislast  

**Subst., f. 1-1: e-d**

**Kontext/Beispiel** Die Beweislast lag bei der Beklagten, die eine eindrucksvolle Streitmach von Juristen, unterstützt von Historikern und Politologen, als Gutachter ins Gefecht schickte und obsiegte. www.goethe.de/ms/buk/archiv/material/benz.doc: 11.04.05

**Definition** [...] die einer Partei im Zivilprozess obliegende Last, die ihr günstig, aber bestrittenen Behauptungen zu beweisen. Kann das Gericht die Wahrheit einer Behauptung nicht feststellen, so ist gegen die Partei zu entscheiden, die ihrer Beweislast nicht nachgekommen ist. Besteht zugunsten der an sich beweisbelasteten Partei eine gesetzliche Vermutung, so muss der Prozessgegner den Gegenbeweis führen (§ 292 ZPO). www.wissen.de/xt/default.do?MENUname=Suche&SEARCHTYPE=topic&query=Beweislast: 11.04.05

**Definition** Die Beweislast ist ein Begriff aus dem Verfahrensrecht. Sie bestimmt, wer in einem rechtlichen Verfahren eine Tatsache beweisen muss (formelle Beweislast) und zu wasseren Lasten entschieden wird, wenn der Beweis nicht erbracht werden kann (materielle Beweislast). Relevant wird die Beweislast jedoch nur, wenn entscheidungserhebliche Tatsachen streitig bleiben. Steht für die Beteiligten eine Tatsache fest, bedarf es keines Beweises mehr. Ist eine Tatsache streitig, die für den Ausgang des Prozesses keine Rolle spielt, ist eine Beweisaufnahme ebenfalls überflüssig. begriffsportal.de/Beweislast: 11.04.05

Deutsch

onus probandi  

**Subst., n. 1-1: d-e lat.**

**Kontext/Beispiel** Wer trägt das onus probandi? In der Politik und in der politischen Philosophie stellt sich diese Frage so: „Auf wen soll der Gesetzgeber rational die Beweislast übertragen – auf den potentiellen Akteur oder auf den Widersprechenden?“

www.radnitzky.de/pub/2002f.pdf: 11.04.05
**Regenerative brake**

Definition: A braking system which enables an electric locomotive or train to reduce its energy consumption by feeding back the traction supply power generated by the motion of the train when it is descending a gradient.

Source: [http://www.mdma.org.uk](http://www.mdma.org.uk), 8.10.2006

**Context:** The current sent back into the overhead wire can be used by other trains going uphill. This ingenious system of braking not only renders absolutely safe the operation of the trains going down steep gradients, but also reduces the amount of current which must be obtained from the power station. To operate the regenerative brake the driver merely has to move a hand switch; after that everything is automatic.

Source: [http://mikes.railhistory.railfan.net](http://mikes.railhistory.railfan.net), 12.10.2006
TMS with definable entry structure

MultiTerm 2011 (Trados) – Web Interface

TMS Types: Samples
Definition: warm-blooded egg-laying vertebrates characterized by feathers and forelimbs modified as wings
Source: wordnetweb.princeton.edu/perl/webwn 2009-08-18

Image: English

- bird

Chinese (Han (Simplified variant)): 鸟
French: oiseau
German: Vogel
Japanese: 鳥 (T-U)
Swedish: fågel
Swedish: pippa
Cloud-based terminology management

- Since terminology work is “expensive”, why not involve the Crowd to create and validate terminology?
- You need a tool for managing terminology in the Cloud!

Examples:
- Wikipedia (www.wikipedia.org)
- TermWiki (www.termwiki.com)

Different approach to:
- web interfaces for TMS (e.g. MultiTerm-Web)
- web-based TMS (e.g. TermWeb)
diaphragm spring

Definition: conical disk spring with heat-treated fingers situated on the inside diameter used as the load bearing component of an automotive clutch assembly

Part of Speech: noun

Synonym(s): 

Sample Image: 

Pronunciation:

Industry/Domain: Automotive

Category: Automobile

Product:

Company:

Related News:

The Telecom Glossary
Decipher the language of the telecom world. Free PDF.
www.ciena.com/Acronyms_Guide
Cloud-/Crowd-based terminology validation
TMS functionality

- Definition of database (data categories, data model, languages, character sets etc.)
- User management (different users, access rights)
- Layout management (different layouts)
- Database management (several databases, local / LAN / Web, existing dictionaries)
- Retrieval functions (wildcards, fuzzy search, full text search, filters)
- Data entry functions (templates, pick lists, consistency control, user rights control)
- Data exchange functions (import, export, printing, TBX)
- Connectivity functions (TermExtract, WP, Translation Memory)
Terminology control

In many application scenarios of terminology work, the checking of correct and consistent use of terminology in documents (created by technical writers or translators) is recommended.

We can differentiate between the following control methods:

- **Monolingual** terminology control
- **Bilingual** terminology control (for translations)
- **Manual (human)** terminology control (part of proof reading & QA)
- **Computer-assisted** term control (tools analyze and check documents)
  - **Without linguistic** methods (for “all” languages, using the content of a term base)
  - **With linguistic** methods (better results, but only for “important” languages)
Features of terminology checking tools:

- Similar to spell checkers and auto correction
- Integrated into editors, authoring systems, CAT tools, but also as stand-alone programs
- Directly during the writing process of a document or translation, or as an autonomous process (when the document is finished)
- Connection to the term base entries (interactive or via export/import)
- Very often combined with grammar and style checking (controlled language)
- Using fuzzy search and/or linguistics (inflected terms in texts vs. canonical form of the terms in term bases)
- Deprecated terms must be maintained in the term base (no-terms)
Terminology control tools

Sample of a terminology check with acrolinx IQ Suite
Terminology control tools

- Dazu den 4-Wegehahn schließen, ...
- Dazu das Vierwegeventil schließen, ...
- Dazu den 4-Wege-Hahn schließen, ...
- Dazu das 4-Wege-Ventil schließen, ...
- Dazu das 4-Wegeventil schließen, ...

- Das Heizungselement muss vor dem Einbau gereinigt werden.
- Das Heizungselement muss vor dem Einbau gereinigt werden.
- Das Heizungs-Element muss vor dem Einbau gereinigt werden.
- Das Heizelement muss vor dem Einbau gereinigt werden.
- Das Heiz-Element muss vor dem Einbau gereinigt werden.

- Der Top-Hifi Verstärker ist ein Verstärker mit Frequenzweichen.
- Der Hifi-Verstärker ist ein Verstärker mit Frequenzweichen.
- Der Top-Hifi-Verstärker ist ein Verstärker mit Frequenzweichen.
- Der Hifi/Top-Hifi Verstärker ist ein Verstärker mit Frequenzweichen.
- Der Top-Hifiverstärker ist ein Verstärker mit Frequenzweichen.
Example

- USB stick **OK** label
- USB memory key **OK** docu
- USB flash drive **NO**
- USB memory stick **NO**
- memory stick **NO**
- pendrive **NO**
- thumbdrive **NO**
- key (OK)

After prescriptive terminology work:
Example: concept-oriented termbase

TermBase

TermEntry 123

EN Definition: xxx xxx xxx
• USB stick OK label
• USB memory key OK docu
• USB flash drive NO
• USB memory stick NO
• memory stick NO
• pendrive NO
• thumbdrive NO
• key (OK)

FR Definition: yyy yyy yyy yyy
• ...
Example: concept-oriented termbase

TermExtract

About data storage
- keys
- USB stick
- USB sticks
- USB memory key
- USB memory stick
- memory stick
- pendrive
- thumbdrive
- key

Example: concept-oriented termbase

TermBase

TermEntry 123
EN Definition: xxx xxx xxx
- USB stick OK label
- USB memory key OK docu
- USB flash drive NO
- USB memory stick NO
- memory stick NO
- pendrive NO
- thumbdrive NO
- key (OK)
FR Definition: yyy yyy yyy yyy
- ...

TermEntry 234
EN Definition: xxx xxx xxx
- keyboard key OK
- key (OK)
**Example: concept-oriented termbase**

**TermExtract**

**About data storage**

Xxxxx xx xxxxxxx keys xx
xx xxxxx xxxxx Xxxx xxxxx
xxxx xx xxxxxx xxxxx xxxxx
USB stick xxx
xx xxxxx xxxxx xxxxx xx
xxxx xxxxx xxxxx xxxxx
USB sticks xxx
xxxx xxxxxx xxxxx xxxxx
USB memory key xxx xxxxx xxxxx
xxxx xxxxxx xxxxx xxxxx
pendrive xxx x xxx xx
xxxxxx XXX press the key
xxxx xxx xxxxxxx xxx xx
xxxx xxxxxxxx xx

**TermBase**

**TermEntry 123**

**EN** Definition: xxx xxx xxx
- USB stick OK label
- USB memory key OK docu
- USB flash drive NO
- USB memory stick NO
- memory stick NO
- pendrive NO
- thumbdrive NO
- key (OK)

**TermEntry 234**

**EN** Definition: xxx xxx xxx
- keyboard key OK
- key (OK)

**LexPrint/MT**

... USB memory key IT, USB stick, key, Def: xxx xxx xxx
...
...
... USB stick IT, see USB memory stick
...
...
key IT, 1) data storage: see USB memory key, 2) input device: see keyboard key
Example: concept-oriented termbase

TermExtract

About data storage
Xxxxx xxx xxxxxxx keys xx
xxxxxx xxxxxx Xxx xxxxx
xxxxx USB stick xxx
xx xxxxxxx xxxxxx xx
xxxxx USB sticks xxx
xxxxx xxx xxxxxxx xxx
xxxxx USB memory key xxx xxxxxx xxx
xxxxx xxx xxx xxxxx press the key
xxxxxxx xxx xxxxxxx xxx
xxxxxxx xxx xxxxxxx

TermBase

TermEntry 123
EN Definition: xxx xxx xxx
• USB stick OK label
• USB memory key OK docu
• USB flash drive NO
• USB memory stick NO
• memory stick NO
• pendrive NO
• thumbdrive NO
• key (OK)
FR Definition: yyy yyy yyy yyy

TermEntry 234
EN Definition: key (OK)
• USB stick, key, Def: xxx xxx xxx
• USB stick IT, see USB memory stick
• key IT, 1) data storage: see USB memory key,
  2) input device: see keyboard key

LexPrint/MT

...
In real terminology practice (in industry):

- Term extraction is used only:
  - for an initial termbase feeding
  - for the preparation of huge translation projects

- Lexicographical print products are rarely needed:
  - sometimes for web glossaries
  - for rule-based MT dictionaries (statistical MT ???)
Example: concept-oriented termbase

TermEntry 123
EN Definition: xxx xxx xxx
- USB stick OK label
- USB memory key OK docu
- USB flash drive NO
- USB memory stick NO
- memory stick NO
- pendrive NO
- thumbdrive NO
- key (OK)
FR Definition: yyy yyy yyy yyy
- ...

TermEntry 234
EN Definition: xxx xxx xxx
- keyboard key OK
- key (OK)

K.-D. Schmitz, IIM, FH Köln
Concept-oriented terminology management

In real terminology practice (in industry):

- Concept/term creators (experts, techWriters etc.) need access to a concept-oriented termbase with term autonomy, when they create new concepts and new terms (ROI)

- Termbase users (techWriters, translators etc.) need access to a concept-oriented termbase with term autonomy when they use terminology (ROI)

- Other tools (CMS, KB, ERP, CAD, CAT etc.) need access to a concept-oriented termbase with term autonomy
Conclusion 1

- Modern terminology management systems allow for professional “real-time” terminology work with terminology producers and terminology users around the world.

- Client-server architecture is a precondition for this and has many advantages, but increases efforts and costs of installation, maintenance and support.

- Not all systems are optimized for specific application scenarios (e.g. small groups, teaching).
For (computerized) terminology management and termbase design:

- Consult literature and guidelines for terminology management (e.g. Wright/Budin: Handbook of Terminology Management Vol I / II and training course material)
- Follow (ISO) terminology standards (e.g. ISO 704, ISO 1087, ISO 12620, ISO 26162)
- Create guidelines and quality procedures for your own terminology work and your own terminology management solution!
Thank you for your attention

Prof. Dr. Klaus-Dirk Schmitz
Fachhochschule Köln
Fakultät 03 - ITMK/IIM
Mainzer Str. 5
50678 Köln
klaus.schmitz@fh-koeln.de