

## Making terminology work

Kara Warburton  
kara@termologic.com

Feb 18, 2014

## What is “terminology management”?

- *Terminology management* includes collecting terms, defining their meaning, indicating their correct usage, translating terms, and storing this work so that it can be useful to others
- In commercial environments the notion of “term” is purely pragmatic
- A “term” is ANY piece of text that might help a potential consumer of the termbase

We are an LSP, a translation company,  
not a terminology service.

We are doing OK with TM.

## According to ISO , an LSP is...

A person or organisation that, beside translation services, also offers value added services such as:

- desktop publishing
- **setup and management of terminology databases**
- content management
- consultancy
- localization services, etc.

87% of translators say their productivity would improve  
with access to terminology (Hurst)

Terminology resources raise productivity 5 to 9% in translation

A language unit can reduce costs by 10%

Terminology research is required for 4% to 6% of all words in a text

(Champagne, 2004)

The scope of potential needs...

*Terminology work is necessary*

*for between 4% and 6%*

*of all words in a text*

This means... every year... work is being done to figure out or fix...

- 21 million words for IBM
- 2.2 million words for the WTO
- 1.2 million words for the World Bank
- 400,000 words for the average mid-sized global company

## TM

versus

## Terminology

- Matches at the segment level
  - Is bilingual
  - Requires a CAT tool
  - Low levels of metadata
  - Typically organized in separate client- or project-based files
- Matches at the sub-segment level
  - Is multilingual
  - Can be accessed outside of a CAT tool
  - High levels of metadata
  - One consolidated system

## Pull vs Push

### Pull

- Translators look up a term ONLY when they feel the need to
- When they look up a term, they don't know if they will find it
- Result: only a subset of the available terms are actually used, and the translator wastes time looking up terms that aren't there
- Example: Web site, spreadsheets, glossaries

### Push

- Whenever a term is in the termbase AND in the text to be translated, it is automatically shown to the translator
- Result: all available terms are used with minimal effort of the translator
- Example: auto-lookup (automatic term recognition) in CAT tools using CAT termbase



## What an LSP needs in a termbase

ANY piece of text that might help translators. This can include more than just “terms” in the conventional sense.

Basically, ANY piece of text that meets any of the following criteria:

- occurs frequently
- can embed in longer TM segments
- should NOT be translated
- has more than one possible translation
- has more than one possible meaning
- has a company-specific meaning, usage, importance, or desired translation
- has a risk-associated significance (marketing, legal, safety)

## Most terms comprise more than one word

- noun
  - smartphone
  - spoiler
  - tablet
- noun + noun + ...
  - cloud computing
  - surround sound system
- adjective + noun
  - fraudulent message
  - smart phone
  - forward-thinking enterprise
  - illegal combattant
- adjective + noun + noun
  - hybrid cloud platform
- adjective + adjective + noun
  - exotic financial instrument
- past participle + noun
  - limited warranty
  - federated database
- present participle + noun
  - operating system
  - phishing attack
  - developing country
- adjective + present participle + noun
  - global positioning system
- verbs
  - build, deploy, leverage, tweet, text, browse, pool, benchmark, virtualize,

# Ter·mo·lo·gic

---

The World Bank offers support to developing countries through policy advice, research and analysis, and technical assistance. Our analytical work often underpins World Bank financing and helps inform developing countries' own investments. In addition, we support capacity development in the countries we serve. We also sponsor, host, or participate in many conferences and forums on issues of development, often in collaboration with partners.

# Ter·mo·lo·gic

---

Our industry-leading technology searches the lowest fares from hundreds of airlines, and gives you access to Web Fares. Choose your plane seat with Seat Maps, and know when fares to your favorite destinations drop with Fare Watcher.


Print out a custom mini-guide before you go using Frommer's Guides. Browse our deals pages for popular destinations, including insider tips, great deals, and travel basics.

## Terminology database (termbase)

- A database containing “terms” and information about them
- Recommended if you expect to collect more than a few hundred “terms”
- You can create one with an off-the-shelf generic database product such as MS Access or a specially-designed Terminology Management System (TMS)
- Some popular TMS:
  - SDL: MultiTerm
  - Interverbum: TermWeb
  - Across: crossTerm
  - Kilgray: qTerm
  - MultiCorpora: termBase
  - Star: WebTerm

# Ter·mo·lo·gic

Concept ID 1517

 Add Comment



**Dictionary** World Bank Terminology

**Section** MC TermBase

**Domain** Environment and Natural Resources > Pollution Management

**Source of entry** WDR10-TermBase export

Kyoto Glossary

WDR 2010 Glossary

Term ID 1517-1

## English

### greenhouse gas

**Part of speech** Noun

**Term type** Full form

**Definition** Any of the atmospheric gases that cause climate change by trapping heat from the sun in Earth's atmosphere—producing the greenhouse effect.

**Source of definition** Kyoto Glossary, WDR 2010 Glossary

**Context sentence** The most common greenhouse gases are carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), nitrous oxide (N<sub>2</sub>O), ozone (O<sub>3</sub>), and water vapor (H<sub>2</sub>O).

**Source of context** Kyoto Glossary, WDR 2010 Glossary

Term ID 1517-13

## English

### GHG

**Source of term** Kyoto Glossary, WDR 2010 Glossary, Acronyms in World Bank Documents

**Part of speech** Noun

Term ID 1517-11

## Spanish

### gas de efecto invernadero

**Source of term** Spanish-English Glossary

**Part of speech** Noun

**Term type** Full form

## Key principles to respect

- Term autonomy – each term is equally describable
- Concept orientation – an entry describes a concept and all its designations
- Data granularity – each field comprises only one type of information and only one instance of that information

A good starting point: TBX-Basic, available from [www.ttt.org](http://www.ttt.org)

## Concept orientation and term autonomy

English	French	Spanish
<b>1. Subject Field(s)</b> - Petroleum Deposits - Petroleum Asphalts - Crude Oil and Petroleum Products (General)	<b>Domaine(s)</b> - Gisements pétroliers - Bitumes - Pétroles bruts et dérivés (Généralités)	<b>Campo(s) temático(s)</b> - Yacimientos petrolíferos - Asfalto de petróleo - Petróleo bruto y derivados (Generalidades)
<b>bituminous sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR, STANDARDIZED	<b>sables bitumineux</b> <a href="#">Source</a> CORRECT, SEE OBS, MASC, PLUR, STANDARDIZED	<b>arenas bituminosas</b> <a href="#">Source</a> CORRECT, FEM, PLUR
<b>tar sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR, STANDARDIZED	<b>sables bitumeux</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR	
<b>oil sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR, CANADA, STANDARDIZED	<b>sables pétroliers</b> <a href="#">Source</a> CORRECT, SEE OBS, MASC, PLUR	
<b>asphaltic sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR, STANDARDIZED	<b>sables asphaltiques</b> <a href="#">Source</a> CORRECT, SEE OBS, MASC, PLUR	
<b>bituminous oil sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR		
<b>petroleum tar sands</b> <a href="#">Source</a> CORRECT, SEE OBS, PLUR		



## Concept oriented layout - IATE

en	
<b>Definition</b>	"tar sands" are a geological phenomenon found especially in Canada, in which bitumen-saturated sands are found near or at the surface of the ground; the hydrocarbon can be separated from the sand by mechanical or thermal processes. Heavy oil sands are similar to tar sands. Outside Canada the term "oil sands" generally refers to buried sandstone acting as an oil reservoir
<i>Definition Ref.</i>	WEC,Standard Terms of the Energy Economy
<b>Term</b>	<b>tar sands</b>
<i>Reliability</i>	3 (Reliable)
<i>Term Ref.</i>	WEC,Standard Terms of the Energy Economy
<i>Date</i>	24/09/2003
<b>Term</b>	<b>asphaltic sand</b>
<i>Reliability</i>	3 (Reliable)
<i>Term Ref.</i>	Customs Tariff Notex CCD 27.15
<i>Date</i>	24/09/2003
<b>Term</b>	<b>oil sand</b>
<i>Reliability</i>	3 (Reliable)
<i>Term Ref.</i>	WEC,Standard Terms of the Energy Economy
<i>Date</i>	24/09/2003

da	
<b>Term</b>	<b>asfaltsand</b>
<i>Reliability</i>	3 (Reliable)
<i>Term Ref.</i>	Toldtarif Notex CCD 27.15
<i>Date</i>	24/09/2003

## DO's

- Include:
  - part-of-speech
  - term type value
  - context sentence
- Put the term in “normal” form
  - singular nouns
  - infinitive verbs
  - normal capitalization

## Avoid gerunds as nouns

- ✗ blogging
  - noun
  - The act of creating entries in a weblog.
- ✓ blog
  - verb
  - To create entries in a weblog.

## DON'T

Create two separate entries (records) for the same concept such as:

- two entries containing the same term (same meaning too)
- an entry for a term, another entry for the same term in plural
- an entry for a term, another entry for a shorter form of the same term
- an entry for a term, another entry for some other variation of the same term
- an entry for a term, another entry for a synonym

# Ter·mo·lo·gic

## Exact same term

Term ID 7973-1  
**English**  
**benefit stream**  
Part of speech Noun  
**Spanish** flujos de beneficios  
**French** flux d'avantages  
**Russian** поток выгод  
**Chinese** 收益流  
**Portuguese** fluxo de beneficios

Term ID 14370-1  
**English**  
**benefit stream**  
Part of speech Noun  
**Spanish** serie de beneficios  
**French** flux des avantages

## Singular and plural

Term ID 5290-1  
**English**  
**operating loss**  
Part of speech Noun  
**French** perte d'exploitation  
**Russian** операционный убыток

Term ID 13682-1  
**English**  
**operating losses**  
Part of speech Noun  
**Spanish** pérdidas de explotación

## Different hyphenation

Term ID 19138-1  
**English**  
**off-balance sheet item**  
Part of speech Noun  
**Russian** забалансовая статья

Term ID 15664-1  
**English**  
**off-balance-sheet item**  
Part of speech Noun  
**French** poste hors bilan

**DON'T separate terms with the same meaning**

## Shorter and longer forms

**Definition** An econometric estimation technique involving minimizing the sum of the squares of the vertical distance between the observed data points and the regression line.

Term ID 13387-1

**English**

**least squares method**

**Part of speech** Noun

**Spanish** método de mínimos cuadrados

**French** méthode des moindres carrés

Term ID 18702-1

**English**

**least-squares regression method**

**Part of speech** Noun

**Russian** регрессивный анализ методом наименьших квадратов

**Russian** определение регрессивных зависимостей с помощью метода наименьших квадратов

## Synonyms

Term ID 16732-1

**English**

**avian flu**

**Part of speech** Noun

**Spanish** gripe aviar

**French** grippe aviaire

**Russian** птичий грипп

Term ID 20795-1

**English**

**bird flu**

**Part of speech** Noun

**French** grippe aviaire

**DON'T** separate terms with the same meaning

## Violation of concept orientation

**Concept Notes** See liquid organic cleaner

Created by Administrator (admin) Creation date 2012-10-06 06:14:14

### English

#### cleaner, liquid organic

Program/Project Indigo

Part of speech Unspecified

Form Full form

**French** nettoyant organique

**German** Reinigungsmittel, flüssig, organisch

**Spanish** limpiador, liquido orgánico

**Italian** smacchiatore per liquidi organici

**Japanese** 有機洗淨液

**Chinese** 有机液体清洁剂

**Definition** A cleaner made of organic materials that removes ink stains from cleaning cloths.

**Concept Notes** Maintenance

Created by Administrator (admin) Creation date 2012-10-06 06:14:14 Changed

### English

#### liquid organic cleaner

Program/Project Indigo

Part of speech Unspecified

Form Full form

**Contextual example** Remove ink stains with liquid organic cleaner.

**French** produit organique nettoyant

**German** flüssiges organisches Reinigungsmittel

**Spanish** líquido organico de limpieza

**Italian** smacchiatore per liquidi organici

**Japanese** 有機洗淨液

**Chinese** 有机液体清洁剂

## Data elementarity – common mistakes

- Two terms in one term field
- Other “stuff” in term field
- Content and source of content in one field
- Comments appended to other text



## Don't put two terms in one field

Created by Sysadmin (sysadmin)    Creation date 2008-11-18 15:30:49  
Change date 2008-11-18 15:30:54    Term ID 77-7

### English

**intermediate bulk container (IBC)**

Abbreviation false

**Swedish** IBC-behållare

**French** conteneur en vrac intermédiaire

Dictionary Multi-dictionary Demo

Section Juridikordbok - Swedish-English legal dictionary, demo

### English

**depreciation (decrease in value) of inventories**

**Swedish** Inkurans

Concept ID 1322

Dictionary Multi-dictionary Demo

Section Juridikordbok - Swedish-English legal dictionary, demo

Domain BUSINESS & ADMINIST > law

### English

**certify (attest, testify)**

**Swedish** Intyga

**Definition** Press job "loaded" into press memory

Created by Administrator (admin) Creation date 201

## English

**loading, loaded**

**Program/Project** Indigo

**Part of speech** Unspecified

**Form** Full form

**French** chargement, chargé

**German** Laden, geladen

**Spanish** carga, cargado

**Italian** Caricamento, caricato

**Japanese** ロード(済み)

**Chinese** 正在加载, 已加载

Don't put two terms in one field

## English

**automatic document feeder (ADF)**

**Program/Project** Chad 10

**Part of speech** Noun

**Form** Full form

**Contextual example** Use the following steps to scan additional items f

Automatic Document Feeder (ADF) or Automatic Photo Feeder (APF) when the on.

**French** bac d'alimentation automatique (ADF)

**Russian** устройство автоматической подачи документов (АПД)

**Danish** automatisk dokumentføder (ADF)

**German** automatischer Vorlageneinzug (ADF)

# Ter·mo·lo·gic

---

DON'T put anything with the term other than the term.

**English**

**Browse...**

**French** Parcourir...  
**Russian** Обзор...  
**Danish** Gennemse...  
**Czech** Procházet...  
**Catalan** Navega...

Term ID 5126-1

**English**

**conformed copy [of a contract]**

**Part of speech** Noun

**French** exemplaire certifié conforme

**English**

**(be) equilibrated**

**Status** Proposed

**Japanese** 平衡化する

**English**

**fitness (for a particular service)**

**Status** Proposed

**Japanese** (特定用途への) 適合性

DON'T put anything with the term other than the term.

Concept ID 1417

Dictionary Multi-dictionary Demo

Section Juridikordbok - Swedish-English legal dictionary, demo

Domain BUSINESS & ADMINIST > business adm

**English**

**blue collar (industrial, factory) worker**

**Swedish** Industriarbetare

Entry number: 4171

Entry class: Trilingual Glossary

Collection: GATT Glossary

**English**

**raise a plea (to)**

**Spanish**

**oponer una excepción**

**impugnar**

**French**

**soulever une exception**

**anglais**

Domaine(s) : - industrie du cuir

**anglais**

**rabbit (fur)**

Définition :

The dressed and dyed rabbit furskin.

The screenshot displays the SDL MultiTerm Server 2011 interface. On the left, a list of terms is shown, with "(continued)" highlighted. On the right, the details for entry number 11196 are displayed, showing translations for English and Japanese.

**SDL MultiTerm Server 2011**

English English

Entry number: 11196

**English**  
 Category: SW UI element  
 Domain: Software  
 Owner: 74 (Services)  
**(continued)**  
 Status: Approved

**Japanese**  
**(続き)**  
 Status: Approved  
 Product: ACE

% of Page  
 % of setpoint  
 % of setpoint (in K)  
 % of setpoint\*  
 % of Total Area  
 % syr vol  
 % syringe  
 %RD Actual  
 %RD Limit  
 %RSD  
 & Basic Safety & Op.  
 (add via 'UltraGrid Designer')  
 (be) equilibrated  
**(continued)**  
 (corrected for path length, attenuation)  
 (corrected to use same units as signal)  
 (Derivative)  
 (Description not specified.)  
 (Enter a value)  
 (Enter changed password)  
 (Enter final pressure for pos. 1)  
 (Enter non-compliant password)  
 (Enter the module name)  
 (Enter the value)  
 (Entered in Instrument Details section.)

## Mixture in term field

**Product group** Printers

**Created by** Administrator (admin) **Creation date** 2012-10-06 06:14:14

**English**

**cards (index, greeting)**

**Program/Project** APJ - CMS Glossary

**Part of speech** Unspecified

**Form** Full form

**Japanese** カード (インデックス、グリーティング)

**Chinese** 卡片 (索引卡、贺卡)

**Chinese** 卡片 (索引卡、贺卡)

Term ID 7973-1

**English**

**benefit stream**

Part of speech Noun

**Spanish** flujos de beneficios

**French** flux d'avantages

**Russian** поток выгод

**Chinese** 收益流

**Portuguese** fluxo de beneficios

Term ID 14370-1

**English**

**benefit stream**

Part of speech Noun

**Spanish** serie de beneficios

**French** flux des avantages


- accounting date
- accounting entry
- accounting judgments
- accounting ledger controls
- accounting period
- accounting period
- accounting policies and practices
- accounting policy
- accounting practice
- accounting prescriptions
- accounting price
- accounting principle
- Accounting Principles Board
- accounting procedure
- accounting rate of interest
- accounting rate of return
- accounting ratio
- accounting reconciliations
- accounting standard
- accounting standard
- accounting treatment
- accounting unit
- account receivables
- accounts payable
- accounts receivable
- accounts receivable
- account statement

## Term autonomy violation

A term is documented in the entry of another term in the definition or any other text field



# Ter.mo.lo.gic Term autonomy violations

 <b>UNTERM</b>		
Language	Term	Acronym
English	sustainable development	
French	développement durable; développement viable	
Spanish	desarrollo sostenible	
Russian	устойчивое развитие	
Chinese	可持续发展	
Arabic	التنمية المستدامة	
Cross-Ref	<a href="#">environmentally sustainable development</a> ; <a href="#">unsustainable development</a> ; <a href="#">sustainable growth</a>	
Note Eng	<p>The management and conservation of the natural base, and the orientation of technological and institutional change, in such a man the needs of the present are met without compromising the ability to meet the needs of future generations. It conserves land, water, plant and animal genetic resources, is environmentally non-degrading, technically appropriate, economical socially acceptable.</p> <p>Relater term: sustainable development goals</p>	
Note Frn	<p>Développement qui répond aux besoins du présent sans compromettre la capacité des générations futures à répondre aux leurs. Variantes à éviter : (1) développement tenable; (2) développement soutenable</p>	

## Common field misuse

- Content in wrong field.
- Multiple instances of content in one field.
- Different things in a field.

## Content in wrong field

### English

**active inlet valve**

Origin 5988-4269

Comment Abbreviation: AIV

**French** vanne d'entré active  
**German** aktives Einlassventil  
**Italian** valvola di ingresso attiva  
**Spanish** válvula de entrada activa  
**Chinese** 主动阀  
**Japanese** アクティブインレットバルブ

### English

**ADME**

Origin 5909-4656

Comment Abbreviation for absorption, distribution, metabolism, and excretion

**French** ADME  
**German** ADME  
**Italian** ADME  
**Spanish** ADME  
**Chinese** ADME  
**Japanese** ADME

# Ter·mo·lo·gic Multiple instances in one field

Domaine(s) : - botanique

anglais

Rabbit brush

Définition :

1. Any of various plants of the genus *Chrysothamnus*, esp. : a common shrub (*c. graveolens*) of western North America that covers vast areas affording a retreat for jackrabbits. 2. Rabbitweed. ←

Sous-entrée(s) :

synonyme(s)

Rabbit bush

es

*Definition* →

1) El que "atiende a las necesidades del presente sin comprometer la capacidad de las generaciones futuras para atender a sus propias necesidades."  
2) "El que hace posible cumplir los objetivos de crecimiento económico, al mismo tiempo que garantiza la protección del medio."

*Definition Ref.*

1) Naciones Unidas, Asamblea General: Resolución A/RES/42/186, anexo, punto I.2  
<http://www.un.org/spanish/do...> ;  
2) Diccionario Clave, 1996 (voz "desarrollo") <http://clave.librosvivos.net...>

français

témoin de connexion

recommandé officiellement par la Commission générale de terminologie et de néologie (France)

Définition :

1. Appliquette envoyée par un serveur de la toile mondiale à un utilisateur, parfois à l'insu de celui-ci, au cours d'une connexion afin de caractériser cet utilisateur.  
2. Par extension, information que l'appliquette peut enregistrer sur le disque de l'utilisateur et à laquelle le serveur peut accéder ultérieurement. ←

## Different things in a field

Note Eng	Payment of <a href="#">interest</a> on a loan. Might be paid in currency, goods or services; often paid on a specified date. Related term: interest payment date
Note Frn	Montant des intérêts payés en devise étrangère ou sous forme de biens ou de services au cours de l'année considérée. Variante: versement d'intérêts

Note(s) :

Voir aussi : logiciel espion.

L'emprunt **cookie** est largement utilisé en français.

Article tiré du *Dictionnaire de la comptabilité et de la gestion financière*, version 1.2, reproduit sous licence.

## Key standards for termbases

- TBX-Basic – available from [terminorgs.net](http://terminorgs.net)
- ISO 12620 – Data categories (ISO TC37 DCR – [www.isocat.org](http://www.isocat.org))
- ISO 26162 – Design, implementation and use of terminology management systems
- ISO 16642 – Terminological Markup Framework
- ISO 30042 – TermBase eXchange (TBX)
- ISO 12616 – Translation-oriented terminography

## Conclusion

- Use standards when designing your termbase
- Keep the automatic term recognition in mind when entering terms
- Do not put a translation memory into your termbase