Case Studies in practical terminology management

Gerhard Budin
TermTrain
2006-03-28
Overview

• Case Studies in practical terminology activities
  – eGovernment – Administrative Nomenclatures for improving government services
  – Environment, Risk Management - Federated terminologies, dictionaries, thesauri
CEN/ISSS WS-ADNOM

Towards a European Network for Administrative Nomenclature

CEN Workshop Agreement (CWA)

Gerhard Budin
University of Vienna
Chair, CEN/ISSS WS-ADNOM
Goal: Implementing a framework for **Semantic Interoperability** across domains and languages in Europe

- Without high quality and standards-based **terminologies** it is impossible to reach **precision, efficiency, and transparency** within and across eBusiness, eGovernment, eHealth, eLearning, eCulture, eScience, etc. processes and systems

- **Problem situation**: lack of accessibility to high quality resources in many domains and languages, diversity of coding schemes and data organization -> **lack of interoperability** (syntactic, semantic, pragmatic) across existing data bases

- **Cultural differences** across language communities and domain cultures are sometimes so strong that citizens as well as experts are lost in dynamic communication situations -> administrative and legal language is the best example
Goal: Implementing a framework for **Semantic Interoperability** across domains and languages in Europe

- Such terminologies are organized and used for multiple purposes in the form of **dictionaries, data bases, thesauri, classification systems, nomenclatures, taxonomies, ontologies, indexes, term lists**, etc.

- ADNOM provides a coherent methodology for modeling, mapping, presenting, and accessing such resources in the context of the emerging **European Interoperability Framework** and in line with – and contributing to new – European and international **standards** (ISO/TC 37, JTC 1/SC 32, etc.).

- Approach: **federated registries** based on **ebXML** and **Topic Maps** implementations, **meta-data standards**, **terminology mark-up framework** -> **Semantic richness/complexity** is managed and **visualized** and not eliminated.
Achievements – Results - Deliverables

- **Survey** on existing administrative nomenclatures and similar terminological resources in Europe as well as on the organizations managing these resources.

- Procedural **methodology** described in the CWA (with principles and recommendations).

- **Demonstrator** showing the implementation of the ADNOM approach (ebXML + Topic Maps (XTM) + ISO 16642 + ISO 11179 + other standards) (will be operational online in April), showing the “ADNOM Seamless Knowledge Core model” with federated registries, navigation services, etc.

- Emerging **organizational network** of relevant institutions (stakeholders), integrating existing networks; dissemination efforts (conference on 1st of December 2005 in Brussels on the Communicative Government organized by NL-Term together with CEN/ISSS WS-ADNOM).
A concrete example of mapping multilingual administrative nomenclatures

- Simplified example on names of government ministries and agencies in Austria, France, and Germany linked to the different scopes and responsibilities of these administrative units as far as pension schemes are concerned

- The asymmetries are presented by a visualization of the conceptual map as implemented in a Topic Map linking the data described above

- It includes the meta-data level using the COFOG classification (incl. the terms in 16 languages for “economic affairs”)

- In the framework of the “ADNOM Seamless Knowledge Core model” for organizing conceptual hierarchies
Concepts of Risk – A Cross-disciplinary and Cross-lingual Analysis of Risk Concepts and their Terminological Documentation

Gerhard Budin
Toward understanding multiple dimensions of “risk“

• “risk“ as a “phenomenon”
• people increasingly conceptualizing it from different perspectives, with different interests, for different purposes, in different ways, in different socio-cultural situations
• -> we are perceiving different “phenomena” that we happen to call “risk”
• -> the social foundation interacting with the cognitive dimension of the language of risk
• Scientific research is investigating such phenomena from their disciplinary perspectives -> developing theories of risk, methods of assessing risk, building databases, writing books, discussing political measures to reduce risk in daily life, etc.
Analyzing the diversity of risk terminologies

- Adopting a comparative approach toward different conceptualizations of “risk” in different disciplines and different languages / cultures
- Conceptual analysis, explicitating the semantic dimension, i.e. the meaning of terms
  - documenting terminological usage in domain discourse (within and across disciplines, within and across languages)
  - fulfilling the function of terminology work: disambiguate polysemous terms, clarifying the meanings of terms by making them explicit
- Descriptive terminography as a pre-requisite for prescriptive terminology standardization and harmonization
- Corpus analysis as a basis for term identification and extraction
Deliverables of WIN WP 2200 “Human Language Interoperability”

- Multilingual corpus collection (glossaries, texts) and analysis, terminology identification and extraction
- Conceptual and combinatory architecture: definitional harmonization in cooperation with terminologists and specialists of different domains
- Hypertext design for web-based teaching, training and professional use of specified terms
- Positioning the terminology in text-production and text-translation: to aid comprehension and efficiency
- Verifying the adequacy of terms in different discourse situations, in cooperation with the users
- Database for didactical applications
Functions and Benefits of Terminology Management for Risk Communication

- Increasing the transparency of terms
- Help negotiate a common understanding of terms in transdisciplinary and transcultural discourse
- Help increase the consistency of risk discourse (written and spoken) and increase understanding in target audiences
- Reduce unnecessary synonyms, disambiguate polysems, help separate homonyms
- Help create risk terminologies in many languages
- Support knowledge sharing and knowledge transfer in cooperative work environments
- Support cross-cultural discourse (e.g. translation and parallel texts)
Terminology Standardization in ISO for Risk Management

- Draft ISO Guide 73 – Risk Management – Vocabulary (both in English and French)
- Generic, definitions of risk, risk analysis, etc. differ: Guide 73: risk = combination of the probability of an event and its consequences; Guide 51: combination of the probability of occurrence of harm and the severity of that harm + safety = freedom from unacceptable risk + harm = physical injury or damage to the health of people, or damage to property or the environment
Comparative Analysis of risk terminologies

+ hazard = potential source of *harm*

ISO Guide 73: knowledge organization:
Risk management

- r. assessment
- r. treatment
- r. acceptance
- r. communication

- r. analysis
- r. evaluation

- r. avoidance
- r. optimization

- r. transfer
- r. retention

Source identification
r. estimation
The WHO initiative

- Consensus building project across organizations, domains and approaches
- 50 key terms – survey on conceptual characteristics and definition components
- Documenting expert opinion diversity
- Basis for harmonization effort
- Establishing best practice in quality terminology work
Types of concepts/terms

• Data-oriented terms
  – Hazard vs. risk
  – Dose vs. concentration
  – Effect vs. response
  – Safety and uncertainty

• Action-oriented terms
  – Risk assessment, hazard assessment
  – Risk management, risk analysis
Methods of terminological analysis

- Focusing on collocations (collocates representing conceptual relations such as *analyze, assess, characterize, communicate, estimate, evaluate, identify, manage, monitor*, etc.)
- Focusing on conceptual and term relations
- Organized in dynamic conceptual fields
- Extended comments on each term, with indications on the acceptability of definitions among the expert group involved in the survey
- Applying semantic componential analysis
Examples

- Hazard = inherent property of an agent or situation capable of having adverse effects on something. Hence the substance, agent, source of energy or situation having that property.
- Risk = the probability of adverse effects caused under specified circumstances by an agent in an organism, a population or an ecological system.
- Both very productive nodes in the terminology.
- Result of componential analysis for ‘hazard’ in the acceptability survey: \{inherent property\} of \{entity to be specified\} with \{potential\} of \{adverse effects\}.
Examples

• Risk assessment = a process intended to calculate or estimate the risk for a given target system following exposure to a particular substance, taking into account the inherent characteristics of a substance of concern as well as the characteristics of the specific target system. The process includes four steps: hazard identification, dose-response assessment, exposure assessment, risk characterization. It is also the first step in risk analysis
Corpus identification and analysis

- The following slides show examples from various initiatives:
  - Screenshot from an entry in a termbase, built in the Multiterm software by Trados, screenshot from a risk dictionary
  - Extract from a glossary on management of natural hazards
  - Extracts from a trilingual information sheet on the problems encountered with the fundamental concepts and terms of risk management as defined in EN 292 and ISO Guide 51 on a multilingual level
  - Extracts from a glossary by the German Bundesinstitut für gesundheitlichen Verbraucherschutz und Veterinärmedizin, 2001, an example of a terminological analysis based on different source documents, comparing the different definitions of the same terms in different contexts
  - Extract from a publication by Pierre Lewalle, 1999, on the WHO initiative on risk terminology harmonization
risk

Definition probability and extent of damage due to a particular hazard (Loat/Meier 2003)

Phrases to improve risk management applications

CrossRef see also risk assessment, risk analysis Grammar noun

Source ISDR 2004

Definition the estimated probability that damage will occur to life, property, or the environment if a specified dangerous event occurs (TESEC-EUR-OPA 2001)

Note the two definitions from ISDR 2004 and TESEC-EUR-OPA 2001 differ to a limited extent

Comment the two definitions from ISDR 2004 and TESEC-EUR-OPA 2001 differ to a limited extent

conventionally risk is expressed by the notation
Risk = Hazards x Vulnerability

some disciplines also include the concept of exposure to refer particularly to the physical aspects of vulnerability

Beyond expressing a possibility of physical harm, it is crucial to recognize that risks are inherent or can be created or exist within social systems. It is important to consider the social contexts in which risks occur and that people therefore do not necessarily share the same perceptions or risk and their underlying causes

Definition the probability of harmful consequences, or expected losses (deaths, injuries, property, livelihoods, economic activity disrupted or environment damaged) resulting from interactions between natural or human-induced hazards and vulnerable conditions

risk

Phrases améliorer les applications de la gestion du risque Grammar noun

German Risiko
Phrases die Risikomanagementanwendungen verbessern

Italian rischio
From the risk termbase towards a risk dictionary

MULTH-WIN-UMB Risk Assessment and Technology Term Glossary conceptually ordered

1. PRE-EVENT
   A. RISK ASSESSMENT and TECHNOLOGY

   Entry Number [A1]
   Management Cycle A. risk assessment and technology, B. public awareness and planning, C. forecasting and warning, D. events and response, E. damage assessment, F. recovery

   English  risk
   Graminfo  <noun, sg, pl>
   French  risque
   Graminfo  <nom, m, sg, pl>
   German  Risiko
   Graminfo  <Nomen, N, Sg, Pl -en>
probability and extent of damage due to a particular cf. hazard (Loat/Meier 2003)
the estimated probability that damage will occur to life, property, or the environment if a specified dangerous event occurs (TESEC-EUR-OPA 2001)
the probability of harmful consequences, or expected losses (deaths, injuries, property, livelihoods, economic activity disrupted or environment damaged) resulting from interactions between natural or human-induced cf. hazards and vulnerable conditions (ISDR 2004)

Comment
1. the two definitions from ISDR 2004 and TESEC-EUR-OPA 2001 differ to a limited extent
2. conventionally risk is expressed by the notation Risk = Hazards x Vulnerability of people, property, and the surrounding area; some disciplines also include the concept of exposure to refer particularly to the physical aspects of cf. vulnerability

Beyond expressing a possibility of physical harm, it is crucial to recognize that risks are inherent or can be created or exist within social systems. It is important to consider the social contexts in which risks occur and that people therefore do not necessarily share the same perceptions of risk and their underlying causes

Definition Fr
EXP : exprimé comme le produit de l’aléa et de la vulnérabilité; prend parfois la signification de “danger” dans certain termes complexes ou dans le langage courant. (Loat/Meier 2003)
Risque = aléa x vulnérabilité des personnes, des biens et de l’environnement
un risque est un péril mesurable dans son occurrence, menaçant des personnes, des biens, des activités ou l’environnement. Il survient généralement de façon inopinée. (PF)
le risque est une mesure de l’occurrence d’un événement indésirable et/ou la mesure associée à ses effets et conséquences. (ASR)
Definition Ge  Risiko = Gefahr x Schadensanfälligkeit von Personen, Gütern und der Umgebung.
Im weiteren Sinn: Möglichkeit, dass aus einem Zustand, Umstand oder Vorgang ein Schaden entstehen kann. Im engeren Sinn: Grösse und Wahrscheinlichkeit eines möglichen Schadens. (Loat/Meyer2003)

Semininfo  risk (R)
BE [DANGER [DAMAGE [SOURCE=], DEGREE=], TARGET=], COSTS=] ]

Related terms and expressions En / Fr / Ge
<nouns, noun phrases>
• Disaster<sg,pl>  cf Hazard<sg, pl>/ cf catastrophe<f,sg,pl>/ cf Katastrophe<F, Sg, Pl-n>
• cf risk awareness<sg>/ cf perception<f, sg>, conscience<f, sg>du risque / cf Risikobewusstsein<, Sg>

<verb phrases>
• to enhance risk awareness / accroître la conscience du risque / Risikobewusstsein schärften
• to monitor, announce and protect oneself from a risk / surveiller annoncer et se protèger d’un risque / ein Risiko überwachen, ein Risiko voraussagen, sich vor einem Risiko schützen
• to take a risk into consideration / prendre un risque en considération / dem Risiko Rechnung tragen
• to run a risk / encourir un risk / ein Risiko eingehen
<table>
<thead>
<tr>
<th>Deutsch / German / Allemand</th>
<th>Englisch / English / Anglais</th>
<th>Französisch / French / Français</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gefährdung</td>
<td>hazard</td>
<td>phénomène dangereux</td>
</tr>
<tr>
<td>Eine potenzielle Quelle</td>
<td>a potential source of harm</td>
<td>source potentielle de dommage</td>
</tr>
<tr>
<td>eines Schadens</td>
<td>of harm</td>
<td>de dommage</td>
</tr>
<tr>
<td>Risiko</td>
<td>risk</td>
<td>risque</td>
</tr>
<tr>
<td>Kombination der</td>
<td>combination of the</td>
<td>combinaison de la probabilité</td>
</tr>
<tr>
<td>Wahrscheinlichkeit eines</td>
<td>probability of occurrence</td>
<td>de survenne d'un dommage</td>
</tr>
<tr>
<td>Schadeneintritts und des</td>
<td>of harm and the severity</td>
<td>et de la gravité de ce dommage</td>
</tr>
<tr>
<td>Schadensausmaßes</td>
<td>of that harm</td>
<td></td>
</tr>
<tr>
<td>Gefährdungssituation</td>
<td>hazardous situation</td>
<td>situation dangereuse</td>
</tr>
<tr>
<td>Gefährdungssereignis</td>
<td>hazardous event</td>
<td>événement dangereux</td>
</tr>
<tr>
<td>Identifizierung der</td>
<td>hazard identification</td>
<td>identification des phénomènes</td>
</tr>
<tr>
<td>Gefährdungen</td>
<td></td>
<td>dangereux</td>
</tr>
<tr>
<td>Risikoeinschätzung</td>
<td>risk estimation</td>
<td>estimation du risque</td>
</tr>
<tr>
<td>Risikobewertung</td>
<td>risk evaluation</td>
<td>évaluation du risque</td>
</tr>
<tr>
<td>Auf einer Risikoanalyse</td>
<td>judgement, on the basis of</td>
<td>jugement porté, à partir de</td>
</tr>
<tr>
<td>beruhende Beurteilung,</td>
<td>risk analysis, of whether</td>
<td>l'analyse du risque sur le</td>
</tr>
<tr>
<td>ob ein vertretbares Risiko</td>
<td>tolerable risk has been</td>
<td>caractère tolérable ou non du</td>
</tr>
<tr>
<td>erreicht wurde</td>
<td>achieved</td>
<td>risque auquel on est parvenu</td>
</tr>
<tr>
<td>Hazard (Schädigungspotential) (gesundheitsgefährliche Stoff-, Erreger-eigenschaften)</td>
<td>The potential of a risk source to cause an adverse effect(s) / event(s). [Inherent property of an agent or situation capable of having adverse effects on something. Hence, the substance, agent, source of energy or situation having that property]</td>
<td>EU_risk</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Gefahr</td>
<td>A biological, chemical, or physical agent in, or condition of, food with the potential to cause an adverse health effect.</td>
<td>CAC ProcM CCFH 31 HACCP</td>
</tr>
<tr>
<td>Gefahr</td>
<td>Ein in einem Lebensmittel vorhandenes biologisches, chemisches oder physikalisches Agens oder ein Zustand dieses Lebensmittels, der sich schädlich auf die Gesundheit auswirken kann. (Anm. &quot;Zustand&quot; bedeutet Fehlen von essentiellen Komponenten, z.B. von Spurenelementen, im Lebensmittel.)</td>
<td>CAC VerfH HACCP_d</td>
</tr>
<tr>
<td>Gefahrenquelle</td>
<td>Ein Agens oder ein Faktor von biologischer, chemischer oder physikalischer Natur mit der Eigenschaft, eine Gesundheitsschädigung hervorrufen zu können.</td>
<td>BgVV</td>
</tr>
<tr>
<td>Gefahrenquelle</td>
<td>Ein biologisches, chemisches oder physikalisches Agens in einem Lebensmittel oder ein Zustand eines Lebensmittels mit einem Potential, gesundheitsschädlich zu wirken.</td>
<td>SCOOP</td>
</tr>
<tr>
<td>Gefahr</td>
<td>Set of inherent properties of a substance, mixture of substances or a process involving substances that, under production, usage or disposal conditions, make it capable of causing adverse effects to organisms or the environment, depending on the degree of exposure; in other words: it is a source of danger.</td>
<td>IUPAC</td>
</tr>
<tr>
<td>Gefahr</td>
<td>Hazard Indicates the potential (i.e., the possibility) of inducing adverse health effects under appropriate conditions. It is a qualitative expression. Appropriate conditions could occur during everyday life with one substance, or may be reached only at extreme exposure scenarios with another agent. Hazard gives information neither on the probability of the induction of an effect under defined conditions of exposure nor on the likelihood that it occurs at all in humans.</td>
<td>TOX</td>
</tr>
<tr>
<td>Gefahr</td>
<td>... analyzing the potential food hazards in a food business operation ...</td>
<td>93/43 EEC</td>
</tr>
<tr>
<td>Risiko</td>
<td>... Analyse der potentiellen Risiken für Lebensmittel in den Prozessen eines Lebensmittelunternehmens ... hier falsche Übersetzung lt. UNTERMANN</td>
<td>93/43 EWG</td>
</tr>
<tr>
<td>Gefahr</td>
<td>... Analyse dieser Gefahren in den Produktions- und Arbeitsabläufen beim Herstellen, Behandeln und Inverkehrbringen von Lebensmitteln ... hier korrekte Übersetzung lt. UNTERMANN</td>
<td>LMHVO</td>
</tr>
<tr>
<td>Gefährdung</td>
<td>... reviewing the analysis of food hazards, ...</td>
<td>93/43 EEC</td>
</tr>
<tr>
<td>Gefahr</td>
<td>... Überprüfung der Gefährdungsanalyse für Lebensmittel, ... hier falsche Übersetzung lt. UNTERMANN</td>
<td>93/43 EWG</td>
</tr>
<tr>
<td>Gefahr</td>
<td>... Überprüfung der Gefahrenanalyse ... hier korrekte Übersetzung lt. UNTERMANN</td>
<td>LMHVO</td>
</tr>
<tr>
<td><strong>Risk</strong></td>
<td><strong>The probability and severity of an adverse effect / event occurring to man or the environment following exposure, under defined conditions, to a risk source(s).</strong></td>
<td><strong>EU_risk</strong></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>[The probability of adverse effects caused under specified circumstances by an agent in an organism, a population or an ecological system]</td>
<td></td>
</tr>
<tr>
<td>Risiko</td>
<td>A function of the probability of an adverse health effect and the severity of that effect, consequential to a hazard(s) in food.</td>
<td>CAC ProcM CCFH 31</td>
</tr>
<tr>
<td>Risko</td>
<td>Wahrscheinlichkeit des Auftretens einer durch das Vorhandensein einer oder mehrerer Gefahren in einem Lebensmittel bedingten gesundheitsschädlichen Wirkung und ihr Ausmaß.</td>
<td>CAC VerfH</td>
</tr>
<tr>
<td>Risiko</td>
<td>Eine Funktion der Wahrscheinlichkeit einer gesundheitsschädlichen Wirkung sowie des Schweregrads dieser Wirkung als Folge einer oder mehrerer Gefahrenquellen in Lebensmitteln.</td>
<td>SCOOP</td>
</tr>
<tr>
<td></td>
<td>Risk means a function of the probability of an adverse health effect and the severity of that effect, consequential to a hazard.</td>
<td>EFA</td>
</tr>
<tr>
<td>Risiko</td>
<td>Risiko bedeutet eine Funktion der Wahrscheinlichkeit einer schädlichen Gesundheitswirkung und der Schwere dieser Wirkung als Folge eines Gefahrstoffs im Lebensmittel.</td>
<td>EFA_d</td>
</tr>
</tbody>
</table>
Figure 1 Quelle: ENVIR [after Lewalle, 1999]

hazard/risk assessment concept system

risk analysis

risk assessment

risk communication

risk management

risk management

risk evaluation

risk monitoring

emission and exposure control

hazard assessment

hazard identification (1)

risk characterization

dose-response assessment

exposure assessment

hazard identification

hazard evaluation

hazard characterization
Modeling risk assessment cycles
Floodsite: Language of risk – a conceptual risk model

Source
  e.g. rainfall, wind, waves

Pathway
  e.g. overtopping, overflow, flood plain inundation

Receptor
  e.g. property, people, environment

Harm
  e.g. loss of life, stress, material damage, environmental degradation
Environmental Thesauri

- Thesaurus in the sense of documentation thesaurus
- Purpose: providing an indexing and retrieval “language” -> for environmental information systems -> mono- and multilingual thesauri, federation of thesauri due to federation of the information systems they belong to
- International/inter-regional co-operation – GEMET General European Multilingual Environmental Thesaurus, EPA Thesaurus, etc.
- Terminology control -> normative approach -> input of standardized terminology
- Standard data model -> thesaurus structure
Brown trout
Salmo trutta
truite

UIN=6349

General European Multilingual Environmental Thesaurus (GEMET)

UN FAO Thesaurus and Ontology

Government State/Local Private Enterprise Academe

Terminology Sources