Terminology and LSP in a language infrastructure

Solstrand, 16 December 2008

Magnar Brekke & Gisle Andersen





LSP

- Language for Special Purposes
- Fachsprache
- Professional Language
- Language in the Workplace
- Symmetrical/asymmetrical communication

LSP Constrained

- Semantically
- Syntactically
- Morphologically
- Lexically
 - "Technical terminology"

Universe of Knowledge

- Mirrored by a universe of discourse
- The Concept: gateway
 - language independent
- Conceptual structure ("ontology")
 - hierachic
 - associative/satellite
- The Term: key
 - language specific <---->shared globally

Specialist corpora and terminology

Specialist Corpora

- Magnitude
- Breadth of coverage/Domains
- Availability
- Copyright issues

Specialist Terminologies

- Core/Domain-focal
- Concept oriented (i.e. → multilingual)
- Defined/delimited
- Availability/accessibility
- Need for metadata

Harmonisation of Terminological Resources

- Metadata for LRs require a common terminology
- Comprehensive, consistent, based on (some degree of) expert consensus
- We propose a joint effort to maximise the value of existing relevant terminology resources in the CLARIN member countries
- Initiative needed for LR purposes, but also other professional and technical domains
- Enables cross-linguistic communication, translation, MT/
 TM, knowledge management, ...

Proposal

- Develop language-independent methods and tools for constructing and consolidating multilingual corpora, termbases and ontologies
- Aim: integrating various terminology resources internationally
- CLARIN is the right international framework for this coordinated effort.
- The Norwegian language bank may be the right national framework for such a coordinated effort



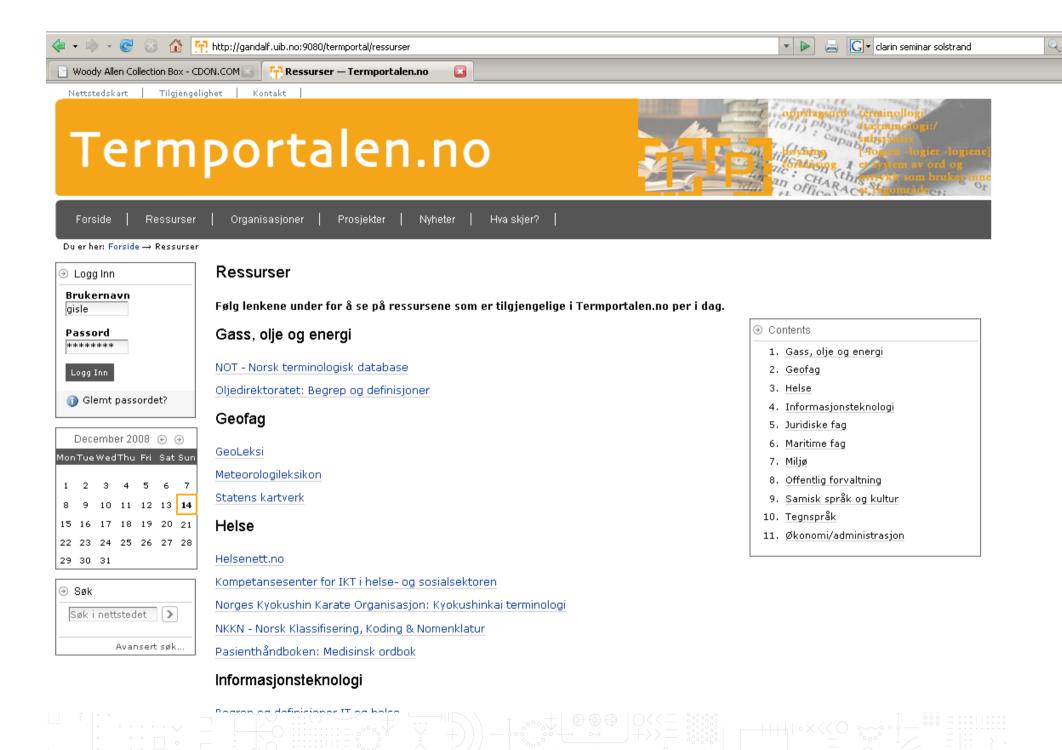
The challenge

- Technical and administrative challenges
- European (and Norwegian) terminology resources are
 - scattered
 - often not updated
 - in many different formats
- Consensus between experts needed
- Coordination between terminologists and domain experts needed
- A national responsibility
- But need for international coordination (multilingual LRs)

What needs to be done (nationally)

- 1. Survey of existing relevant resources (in progress)
- 2. Access to these resources
- 3. Updating the resources
- 4. National coordination with field experts, concept consensus
- 5. International coordination (CLARIN, ISO)
- 6. Interoperability, search across databases

The resources should remain distributed.



Terminology and language infrastructures

All scientific/professional fields

CLARIN: Humanities and Social Sciences

Language resources, language technology, linguistics

TERMINOLOGY