welcome!

Jörg Schütz

bioloom

Sven C. Andrä

andrä ag

group
Employing Machine Translation in Glocalization Tasks
Part 1: changing world
Part 2: how MT fits in
Part 3: holistic solution
Part 1: Enterprise 2.0

knowledge worker

- creativity
- emotions
- knowledge
- experience

Web 2.0

Social Software

Globalization
Post-Taylorism
Internationalization
(Machine) Translation
New Media
Localization
Transaction Costs
Collective Intelligence
Sharable Resources
Crowdsourcing

Real Time
specific views
- integration of social software
- bridge: self and alien organization
- shaping parameters
- analyses and evaluations
- motivation
- communities of practice
- person / community extension
- hedonism and incentives

computer & information sciences

psychology & sociology
most visible observations

more content

pro-active users
general observations

shrinking
accelerating
expanding

{ world
“open source software” development

Specific observations

- micro enterprises
- small and medium enterprises
- large enterprises

Added value through optimizing transaction costs

Added value = appreciation / esteem

Minimizing transaction costs
challenges

- platforms
- media
- data / information / knowledge
- languages and cultures
- copyright and ownership
- safety and security
- ...

platforms
media
data / information / knowledge
languages and cultures
copyright and ownership
safety and security
...
necessary changes (transformations)

multiple levels
- data / information / knowledge
- process / workflow
- enterprise / market
- society / politics
- from insufficient automation to **interoperability**
- from broken process and data streams to **continuity**
- from manual, disruptive tasks to **coherent flows**
- from nontransparent scoring and evaluation to **valuable assets**
- from markup to **self-describing data**
- from ontology to **pattern in data**
- from time-delayed to **real time**
- from objects to **information shadows**

**semantic web & collective intelligence**
Part 2: MT as a »gadget«

how MT fits in: the MT biased use case
challenges

- Ubiquity of I&C media across domains and cultures
- Anytime available information in multiple languages
- New forms of data quality and data integrity
- Data stream translation
- Data and service interoperability
- Information and knowledge sharing
necessary changes (transformations)
services

- cloud-based
- accessible
- multiple formats
- standards
feedback cycles

- open
- responsive
- adaptable
- interoperable
self-learning

- stimuli sensitive
- context aware
- emergent
- stigmergent
- self-sustaining
• from enterprises to *transcultural communities* with micromanagement, agility and responsibility
• from numbers to *values* with streams and continuity
• from platforms to *social networks*
• from links to *relations*
general impact: real time "glocalization"

"good or best practices" vs. "next practices"
Part 3: towards a holistic MT solution

- inspired by biological immune system
- highly distributed
- feature extracting
- adapting
- self-organizing
- learning
- memorizing
- web / net services
- terminology biased
- pattern recognition
- context accumulation
- learning capability
- evaluate function
terminological data

- hidden in enterprise documents, databases and communication streams
- searchable but often not findable
- linked but not semantically related
Aufgrund seiner Schwerhörigkeit war es nicht möglich, sich mit dem Patienten zu verständigen.

Für Patienten, die neben ihrem Husten über ein kratziges Gefühl im Hals klagen, sind Hustensaft und Lutschpastillen geeignet.

Ambroxol, Lidocain oder Benzocain wirken lokalanästhetisch, Cetylpyridiniumchlorid und Dequaliniumchlorid sowie Chlorhexidin und Hexetidin desinfizierend.
### numbers

<table>
<thead>
<tr>
<th>Power</th>
<th>Term</th>
<th>German Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^6$</td>
<td>one million</td>
<td>eine Million</td>
</tr>
<tr>
<td>$10^9$</td>
<td>one billion</td>
<td>eine Milliarde</td>
</tr>
<tr>
<td>$10^{12}$</td>
<td>one trillion</td>
<td>eine Billion</td>
</tr>
<tr>
<td>$10^{15}$</td>
<td>one quadrillion</td>
<td>eine Billiarde</td>
</tr>
<tr>
<td>$10^{18}$</td>
<td>one quintillion</td>
<td>eine Trillion</td>
</tr>
<tr>
<td>$10^{21}$</td>
<td>one sextillion</td>
<td>eine Trilliarde</td>
</tr>
<tr>
<td>$10^{24}$</td>
<td>one septillion</td>
<td>eine Quadrillion</td>
</tr>
</tbody>
</table>

### term selection – terminology ranking
pattern recognition

learning phases:
- grow
- adapt
- shrink

pattern translation ranking

a picture of Monnet
a picture from Monnet
a picture by Monnet
a picture owned by Monnet

ein Bild von Monnet
We don't like involved discussions with the people involved.

learning phases:
- grow
- adapt
- shrink

- differing contexts
- divergent translations
- cultural (mis-)matches

translation context ranking
translation process evaluation with key performance indicators

- BLEU
- variants
- ... standards

- term filtering, ranking and scoring
- pattern and context ranking
- discrimination between streams and documents
- multiple MT engines
- artificial immune system
- different scoring and ranking algorithms
- learning and memorizing
- feedback interfaces and hooks to permit learning and adapting to changing contexts within specific environments
conclusio
MT visions

“… the spontaneous perception of connections and meaningfulness in unrelated things.”

—William Gibson
thank you!