The use of SMT in financial news sentiment analysis

Thomas Dohmen
SemLab

- SemLab founded in 2000
- Software applications for:
  - Finance
  - Healthcare
- ViewerPro: semantic analysis platform for financial news (DowJones, Reuters, etc.)
- NewsSentiment.eu for financial news sentiment
Research Projects

- SemLab participates in EU and nationally funded r&d projects
- Collaboration with leading research institutes
- Last decade: invested 5.8 million Euro in research
Research Topics

• Semantic analysis technology suitable for wide ranging r&d:
  • Ontology
  • Language
  • Finance Tools
  • Social Media
  • ...

[Logos and icons for TOWL, SOLLIM, Let's MT!, NORM, and EuroScope]
- Statistical Machine Translation (SMT)
- Focus: Small European languages
- Topic-specific translation systems
- E.g. Business and Finance language
- SemLab provides parallel corpora for system training
- Integration into news sentiment website
Electrocomponents plc, the operating company for Allied Electronics and RS Components, reports record sales and increased profit margins for the year.
Electrocomponents PLC - profit up

Electrocomponents plc, the operating company for Allied Electronics and RS Components, reports record sales and increased profit margins for the year.
• UsingLetsMT engine for translating output
  – Translate news events
• Using LetsMT engine for translating input
  – E.g. Dutch financial news translated into English for semantic analysis
  – Evaluate translation quality from machine readable perspective
MT Standards

• Standard API
  – Easy integration of MT system
  – Current Standards (e.g. Google API) too basic
  – Not rich enough for domain-specific MT

• Data Collecting standards
  – Webcrawling: web polluted with bad MT data
  – Need for markup to distinct MT from human translations

• General translation standard
  – Accessibility
  – Quality assurance
Questions?