Priority Theme 1: Translation in the Cloud

Andrejs Vasiljevs

Tilde, Latvia
andrejs@tilde.com

META-NET FORUM 2012
A Strategy for Multilingual Europe
Brussels, Belgium, June 20/21

Co-funded by the 7th Framework Programme and the ICT Policy Support Programme of the European Commission through the contracts T4ME, CESAR, METANET4U, META-NORD (grant agreements no. 249119, 271022, 270893, 270899).
Vision: Applications needed by EU citizens and businesses

- Europe-wide social and business networking in native language
- Mobile and internet services in native language for e-Commerce, education, travel, entertainment, etc.
- eGovernment reaching all linguistic groups and enabling political discussion across borders
- Unlimited TV/movies cross-language subtitling/interpretation
- Ever present Personal Interpreter
- Translingual Spaces: dedicated locations for ambient interpretation
Vision: Services for the EU Society and Citizens

- A ubiquitous online platform combining automatic translation, language checking, post-editing, as well as human creativity and quality assurance
  - for generic and special-purpose services
  - free for small volume use and for high-volume baseline quality
  - involve providers of computer supported HQ human translation
  - business opportunities for a wide range of service and technology providers
- Assured privacy, confidentiality and security provided by trusted service centers
- Quality upscale models: instant quality upgrades
- Domain and Task specialization
Vision: Services for the EU Society and Citizens

- Ubiquitous translation services for a full range of quality levels, fast, affordable
- Covers written and spoken language from formal language to chats and social networking
- Multi-media multi-language content delivery
- On mobiles, tablets, PCs, etc.
A Translation Cloud Scenario

Service Broker

Specialized MT/LT Web Services
- PR Brochures
- Informal Language
- Int. Company Names
- Automatic Summarization
- Patents
- Times and Places
- Human Post-Editing
- Annual Reports

http://www.meta-net.eu
Impact

- Breaking language barriers facilitating job opportunities, mobility, social cohesion and regional parity
- Business opportunities
  *Global market of language services 25 bln EUR, 2011*
- Regaining global technological leadership in translation technologies
- Enabling access to the rich multilingual European culture
- Improved LT for under-resourced European languages (cf. META-NET Language White Papers)
- Enabler of cross-lingual “knowledge management” and advanced technologies like cross-lingual content analytics, IE, summarization, sentiment analysis

http://www.meta-net.eu
Current State of the Art in Research

- Open-source MT Toolkits (e.g., Moses)
  - Maturing and scaling to reach the requirements of massive production use
  - No general interoperability with LT tools (e.g., parsers, NE recognisers, etc.)
  - Confidence estimation (let alone quality estimation) is an open issue

- Growing availability of multilingual data
  - Available for many languages, yet in a few domains only
  - META-SHARE contribution

- Experimental prototypes for AV media translation
  - Typically re-speaking scenarios (involving human translators)

- Support of online media very limited

- Promising research results from FP7 and ICT-PSP projects
  - Fragmentation, sustainability, scalability issues

http://www.meta-net.eu
Novel Research and Targeted Breakthroughs

- Overcoming obstacles for high quality
- Quality assessment models and new automatic metrics (task and application oriented)
- Computer-supported creative (human) translation/authoring
- Semantic translation paradigm by including linked data, ontologies, semantic and inference models into MT
- Inclusion of specialized analysis, generation and transfer models (register, style, situation)
- Improvement of speech recognition/analysis (specifically for MT and MT environments)
Organisation of Research and Innovation

- Large cooperating projects
  - Sharing infrastructures: resources, evaluation
- Smaller projects – providing building blocks
  - National languages (resources, technologies)
  - Component technologies
- Combined funding (EC, national, private)
- Inclusion of industry and translation professionals in the entire research and innovation process
- Solving legal hurdles on using data for research
- Connection to CEF
- Infrastructural support (selected areas)
  - Resources, evaluation suites, organisation of challenges
Thank you
andrejs@tilde.com

office@meta-net.eu

http://www.meta-net.eu
http://www.facebook.com/META.Alliance