META strive:
Making a paradise garden
out of the jungles

Serge Gladkoff
GALA Standards Director
President, Logrus International
META Forum, Brussels, June 2012
GALA: The association for Europe

- 170 European companies
- 55% of GALA membership
- 32 European countries
- Most active companies
- LSPs, technology, clients

www.gala-global.org
Current [dizzzoo]array

- 100 types of DATA
- 1100 CMS systems
- 50 TMS systems
- 10 TM tools
- 10 major MTs and 2000 Moses systems 😊
- 50 LANGUAGES
  = 27,500,000,000 problems

What interoperability?
Current disconnect:

- Process friction is high and only seems to be getting worse;
- MT is not ready for deployment, requires technological level beyond LSP capabilities;
- Technologies are like zoos; standards are not implemented;
- Clients, however, demand price level as if HQMT is real;
- LSPs are forced to juggle with skills and tools that have no relation to translation, language and SME knowledge;
- No unified technological platform exists;
- Technical overhead is too high; skilled translators and PMs spend time on operations which are not paid for by the clients;
- Technologies in general are either too expensive or proprietary;
- Tools developers seem to be absorbed with their own vision;
- Data (content) creation process lacks understanding of localization needs;
- Lack of rich metadata to ease multilingual translation process.

www.gala-global.org
LSP are trying to survive in a stormy ocean of incompatible technologies and processes, trying to swim with whatever tools are within reach.

The most advanced develop their own technologies.

Let’s have a look at some shiny bits and pieces we have now:
“Thinking about what customer wants and exceeding expectations.”
La Vanguardia: Daily Quality Publication based on MT

“Rule-based linguistic system with a deep analysis and a transfer approach.”

Environment
- Load-balancing configuration
- High reliability
- High performance
- High availability

Hard Figures
Volume
- 70,000 words/day
Tight timeframe
- 5 hrs
Team
- 8 linguistic post-editors
- 20 journalists

SUCCESS: Increase + 4% of copies, + 7% of readers
All other newspapers in Spain: -2% to -16%

www.gala-global.org
**Drehfeld**

 examines, which is something that is rotating for these purposes. Drehfelden werden erzeugt, um die Wellen von Drehstrommotoren und selbstständig anlaufenden Wechselstrommotoren anzuheben. Das Drehfeld dient den Kanälen auf der Welle des Motors, bestehenden Motor magnetisch mit.

A rotating magnetic field is a magnetic field which changes direction at (ideally) a constant angular rate. This is a key principle in the operation of the alternating-current motor.
How SYSTRAN innovates in language technology?

- **Systran Hybrid MT engine**: SYSTRAN introduced the first product (SYSTRAN ENTERPRISE SERVER 7) powered by a hybrid machine translation (MT) engine. Why investing on Hybrid Technology? Combining Rules-based Linguistic processing with Corpus-based Statistical processing gives the ability to get top quality translations.

- **Systran Training server**: by collecting training data and running initial trainings, Systran helps enterprises to build their own customization resources (translation memories, dictionaries, glossaries). Train the software to any specific domain or business objective allows Systran customers to achieve cost-effective publishable quality translation.

Beyond MT: the multilingual tool box!

- Based on linguistic descriptions of 27 different source languages, Systran is now developing new derived applications:
  - Multilingual sentiment analysis
  - Multilingual name entity recognition
  - Multilingual word sense disambiguation
  - Multilingual speech translation
Logrus: TM+MT workflow center
The FUTURE?

How do we make sure that those (and other) shiny variety of nuts and bolts will not turn into rust, but will be used to build the machine of the future?
What next?

- We need removal of duplication of effort that will engage the most advanced ideas of today (cloud);
- With so many various technologies clearly there’s a big and concrete need for innovative language technologies;
- We need serious open investment of effort into the translation technology of the future and collaboration;
- Standardization and interoperability are key;
- Data is the key; we need to change the authoring process; add massive metadata; standardize; create common, non-proprietary technology platform for productive engagement.
META attempt

- META-NET is the attempt to bring together a fragmented field (including LT researchers, language communities, language professionals, LSPs, translators, funding agencies, politicians, etc.) under one umbrella organization: META (Multilingual Europe Technology Alliance)

1. GALA will bring to META knowledge of the industry – practical problems and needs – and will help to identify research themes for future research and development.
2. GALA will help to prepare the architecture for a future platform of seamless multilingual linguistic services for the Europe and beyond.
3. LSPs are the source of information about real production problems: what metadata is required to ease the process; what standards are working, call for common, non-proprietary technology platform for productive engagement.
4. GALA has the experience of collaboration to share.
Thank you!

www.gala-global.org
standards@gala-global.org
sgladkoff@logrus.net

About GALA
The Globalization and Localization Association is the largest global non-profit association within the language industry, providing resources, education, ideas and research for companies working with translation services, language technology and content localization. GALA member companies are vendors and buyers of language services and technologies. They deploy sophisticated multilingual strategies and proven tools to take content and products to markets around the world.