

3 years already and still a long way to go...

Roberto Cencioni

INFSO - Information Society & Media
Digital Content & Cognitive Systems

Language Technologies & Machine Translation

infso-e1@ec.europa.eu



Agenda

- o So far
- o From now on

A long-term commitment

- EC has supported Language Technologies (LT) since 1980s
- a **fresh start** since mid-2008
 - renewed political commitment
 - explosion of online content, esp. the social web
 - promising technical advances
- **~120 M for Research & Innovation** in 2009-2011
 - 90 M from the ICT **research** programme
 - 30 M from the ICT **innovation** programme

Early years... until today

- first, **put LT back** in the EC loop
- to do so, provide a **broad framework**
- to boost **cross-border partnerships**
- and assess the **response** from the field
 - from 27 proposals in 2009 ...
 - ... to 200 proposals in 2011
- outcome: 50 projects by Q2 2012
- let's call it a large-scale "**community building**" exercise!

Early years...

- broader participation, old hands w/newcomers
 - ICT-7: **90** submissions,
400+ unique participants requesting 260 M
 - ICT-SME: **109** “language” proposals,
400+ unique participants requesting 170 M
 - 40+ countries, year after year
- balanced academic/industrial partnership
 - ICT-7: enterprises account for **39%** of the submitters
 - ICT-SME: **57%** of the proposers are SMEs

To do what? ICT-7 call, 50 M

“... effective solutions that support business and inter-personal communication and enable people to make sense of online content and services in Europe's many languages.”

- global content processing
eg **automated translation**
- mining of unstructured information
eg **text/audio analytics**
- natural interaction
eg **speech understanding & dialogue**

To do what? ICT-7 call, 50 M

- 16 projects, 100+ partners from 24 countries
- who is involved?
 - enterprises 39%
 - universities 38%
 - public/private laboratories, others 23%
- popular themes?
 - various forms of computer-aided **translation**
 - **text mining**: information extraction, summarisation, aggregation... in multiple languages
 - **speech recognition**, typically coupled with transcription & translation

With whom?

- **researchers** (linguistics, statistics, semantics...)
- **developers** of LT products (mostly SMEs)
- **vendors** of LT-enabled products
 - text and content analytics
 - enterprise search & content/media management
 - edu-entertainment & mobile services...
- providers of language **services**
- **early adopters**, private and public, small and big

Cross-cutting issues: language resources

- **1/3rd of our projects by Q2 2012...**
- FLARENET – charting the landscape
- ICT projects – “tools”, automating acquisition & cleanup
- CIP projects – “content”, making existing resources accessible & fit for reuse
- META-SHARE – “service” framework for sharing resources
- ongoing ICT-SME call – more (quality, clean) data for better technology:
 - high-tech, high-growth SMEs
 - “adding value to data”: pooling, trading, reusing

Groundwork for the coming years

- unifying **research roadmap**
 - META-NET
- **pooling & trading** R&D assets
 - META-SHARE + ICT-SME projects (selection in Nov)
- **LT within global Web**
 - MULTILINGUAL-WEB
- **technology transfer & uptake**
 - session w/business stakeholders, t/ow @ 10:00
- interplay between **national & EU actions**
 - exploratory work after the summer break

- o The immediate future (2012-13)
- o The near future (2014+)

and it's **not** going to be business as usual!

Bringing FP7 to its end

- FP7-ICT is not over yet, another 1.5 B to go...
- one or two more calls in 2012-2013
- work programme to be finalized mid-2012
- so we need early next year
 - a credible **strategy** & plan, and
 - the “**troops**” behind it
- consultations from Oct until Jan (online + face-to-face + inputs from ongoing actions)
- strategy depends heavily on & must prepare later operations under the new programme, 2014+

EC LT Days event scheduled for mid-March, 2012

Towards a new LT agenda

- **Questions will include:**
 - should LT continue as a **standalone** activity?
 - any research themes to be **discontinued**?
 - right balance between **research & innovation**?
 - how to best support **new commercial players**?
 - if we concentrated our efforts on a few **tightly integrated clusters** of projects
 - what would be the objectives & features of such clusters?
 - how do we set them in motion before the start of the new programme?
 - right balance between the above actions & open, **bottom-up, frontier-type research**?
 - what sort of broad-based **infrastructures**?

Common Strategic Framework (2014-2020)

- integrated approach to research & innovation
- one big package encompassing many programmes & structures
- three building blocks:
 - excellence in the research base (incl. ERC)
 - societal challenges (eg health, energy)
 - leadership in enabling technologies (**most of ICT**)
- coherent set of rules/instruments/fundings:
 - roadmap-based actions: strategic, high impact
 - open scheme: flexible, reactive
 - infrastructures: broad, shared

Timeplan

MFF	Financial Framework: resources for EU policies, including R&I
CSF	R&I Framework Programme + its constituent programmes
June 29	MFF proposal
Nov 30	CSF proposal

by mid-2012 strategic orientations are set,
specific budgets & work programmes will
follow ...

Conclusion

- **race is on** for future support in EU programmes
- time matters – we must have a clear strategy & supporting coalition **within 1 year**
- key conditions for success:
 - from **sponsorship** to **partnership**:
EC & MS + Industry & Academia
 - from one-off actions to a lasting scheme
 - delivering credible, useable results
 - combining science & market

Today's quotes

Nelson Mandela:

- If you talk to a man in a language he understands, that goes to his head. If you talk to him in his own language, that goes to his heart.

Former IBM CEO:

- If you aren't **fast** you're dead. If you aren't **right**, you're dead as well.

Thank you!

infso-e1@ec.europa.eu

