



**Roberto Cencioni**

**European Commission**

**DG Information Society & Media**

*Digital Content & Cognitive Systems*

*Language Technologies*

*Brussels, 19-21 June 2012*



- the **European Union** brings together **500 million people** speaking **many languages...**
  - half of the EU adult population don't understand English
  - no such thing as a digital lingua franca since the advent of the social Web

### **EU policy framework: "*Digital Agenda for Europe*"**

- ✓ build a ***digital single market*** where content & services can flow freely
  - ✓ support cross-border exchanges between public online services
    - ✓ enable internationalisation of SMEs...

*but what if you don't understand what's being displayed?*



## **EC has supported Language Technologies since 1980s**

- a fresh start since 2008
  - renewed political commitment
  - explosion of online content, esp. social networks
  - sizeable industry of language services
  - promising technical advances
- **approx 150 M funding in the last 5 years**
- **nearly 60 projects underway**
- next *call for project proposals* in July

## main features of our projects

- multi-party, multi-nation partnerships
  - member states, associated countries, 3<sup>rd</sup> countries
- academia + industry + "launching" users
- focus on enabling technology
- multiple languages
- impact
  - target groups, application scenarios, demonstrators
- **significant involvement of commercial players**



## main themes of our projects

*"... 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."*

- content processing, eg automated translation
- mining of unstructured information, eg text analytics
- natural interaction, eg speech understanding & dialogue

in and between multiple languages



we struggle since 2009 to

- boost cross-national & -disciplinary **partnerships**
  - from 27 project proposals in 2009 ...
  - ... to 200 proposals in 2011 ...
- address **gaps & weaknesses**
  - in particular duplication & fragmentation
    - common **research vision**: needs, application scenarios, core technologies
    - **reuse** of basic R&D assets (data & tools)
    - technology **transfer & deployment**



- **research:** towards a federating vision & roadmap
  - META-NET (2010-2012)
- **business:** ease market uptake & deployment
  - LT-INNOVATE (2012-2014)
  - forum for business stakeholders esp. vendors
  - also, market data, value chains, non-technical barriers, demand & supply
- **public sector:** promote the use of LT by our own institution
  - inter-service group with major EC departments

## genesis & rationale of LT-INNOVATE

- what's the purpose of pouring money into research if results don't reach the market?
  - useful **research** i.e. research for a known purpose
  - yielding useable although possibly imperfect **technologies**
  - which are **exploited**, including for commercial purposes
- so, what are the roadblocks, the technical & non-technical barriers; how to connect the dots?
- how can we best integrate research & innovation within our programmes?



## from LT research to global markets?

- plenty of teaching & learning; a lot of local, narrow, publication-driven research; small teams, limited ambition
- national public funding either inexistent or aimed at ICT at large, culture, audiovisual, ...
- national funding favours research over engineering & experimentation
- no public procurement
- ~500 SMEs (est.) acting as technology/solution vendors, sometimes with insufficient access to market or financial muscle
- few pure LT applications, mostly embedded technologies, hence little visibility vis-à-vis a majority of buyers, analysts, sponsors/investors ...



## **completing the current programme (2012-13)**

- last series of calls in July, 4 topics:
  - 1) (cross-media) content analytics
  - 2) (high-quality) automated translation
  - 3) (speech-enabled) multimodal interaction
  - 4) industrialisation of promising technologies

## **launching the new programmes (2014-2020)**

- a. Horizon 2020 (H2020), research & innovation*
- b. Connecting Europe Facility (CEF), pan-European service platforms*
  - ➔ concrete orientations by Q1 2013

## last series of calls under the current programme:

- Challenge 4 – *Digital Content & Languages* of the ICT work programme 2012-2013, released in July
  - Call-10: open Jul 12; close Jan 13; budget: 27 M
  - SME call: same schedule; budget: 15 M (data + language)
- **Call-10**, R&D, objective 4.1:
  - a) cross-media **content analytics**
  - b) high-quality **machine translation**
  - c) speech-enabled **multimodal interaction**
- **SME-Call**, innovation, objective 4.3:
  - b) **industrialisation** of promising ("intelligence") technologies, field validation, business models & exploitation paths

## SME call:

### 4.3.b Easing transfer and take-up of LT

... **bringing** [selected] **LT closer to commercial maturity through an "industrialisation" process** including but not limited to:

- (i) *engineering* of promising but commercially untested technologies
- (ii) *integration* within existing or upcoming products and services;
- (iii) *first-use experimentation* and validation;
- (iv) *in-depth assessment* along technical, user related and economic dimensions;
- (v) *identification* of possible exploitation paths and viable business models, and of suitable sources of funding.

## **new programmes, 2014+**

- research (H2020)
  - innovation (H2020)
  - infrastructure (CEF)

## ***H2020 2014-2020:***

- **research & innovation**; budget (EC prop): 80 B,  
~20% for ICT
  - science (including FET)
  - *technologies* (most of the mainstream ICT themes)
  - societal challenges (e.g. health, ageing, education...)
- **on-going discussions with Council & EP**
  - outcome depends on concomitant discussion of *MFF*  
(= money for all EU policies)
- **meanwhile, early priorities & operations by Q1 2013**
  - first series of calls towards Q4 2013

## ***CEF 2014-2020:***

- **TEN model:** pan-European, public-interest infrastructures
- 3 pillars: Transport, Energy & ICT (9 B proposed)
- for ICT: (i) Broadband & (ii) ***Digital Service Platforms***
  - eGovernment services, eHealth, eCulture...
  - **Multilingual Access to online Services** (*new*)
- aim is to "enable service providers to offer their content & services in as many languages as possible..."
  - core platform (hub) + general-purpose tools & data sets + services co-developed with industry & government
- still a long way to go vis-à-vis Member States, update in the autumn

## Conclusions:

- heavy competition with other policy areas
- emphasis on competitiveness, growth & jobs
- so to be successful:
  - clear & credible **Res & Inn agenda**
  - well understood **impact** on economy & society
  - well connected & mobilized **community**
  - broad (incl national) recognition & **support**