



European Programmes at the intersection of Data & Language

H2020 2014-2015
A broad overview

Roberto Cencioni

Torino, 21 Jan 2014



- New **Horizon 2020** framework programme
- Where **Data & Language** fit
- Language-related **research & innovation**



Horizon 2020

- The new EU Research and Innovation programme (70+ B)
- **ICT** a well-recognized pillar cutting across the programme
- 3 main components:
 - Excellent **Science** (incl FET eg cognitive systems)
 - **LEIT** (the subject of this presentation)
"Leadership in Enabling & **Industrial Technologies**"
 - **Societal** Challenges (eg Culture - digital heritage)
- LEIT work-programme accounts for c 1.8 B in 2014-15
 - ✓ Data & Language placed under "Content technologies and information management"



Key Features

- Simplification & harmonisation: instruments, procedures, operations
 - front end: single participant portal
 - back end: single EC dept. dealing with Legal, Audit, IT
 - electronic signatures, cost claims & reports
- Parts of the programme implemented through **agencies**
- **Submission**: no major differences wrt. FP7
- **Evaluation**: ~ same process, more emphasis on impact
- **Grant awarding**: 3 months from notification to signature
 - no "negotiation" as such
 - hence proposal must be ~ grant-ready



Key Features

- 3 main project (now "action") types:
 - **Research & Innovation**
 - same as "RTD projects" in previous FPs
 - **Innovation**
 - little or no research; integration, validation, evaluation, uptake...; from lab to market
 - result oriented, market driven, business agenda
 - **Coordination & Support**
 - as in the past: support the programme & its participants
- no procedural difference between small & big projects

Research and Innovation actions

- Actions primarily consisting of activities aiming to establish **new knowledge** and/or to explore the feasibility of a new or improved technology, product, process, service or solution.
 - They may include basic and applied **research**, **technology development** and integration, **testing** and validation on a small-scale **prototype** in a laboratory or simulated environment.

Innovation actions

- Actions primarily consisting of activities directly aiming at producing plans and arrangements or designs for new, altered or improved **products, processes or services**.
 - They may include prototyping, testing, **demonstrating**, piloting, **large-scale validation** and market replication.

Coordination and Support actions

- Actions consisting primarily of accompanying measures such as **standardisation, dissemination, awareness-raising and communication, networking, coordination or support services**, policy dialogues and mutual learning exercises and studies, including design studies for new infrastructure.
 - They may include complementary activities of networking and coordination between programmes in different countries.



Key Features

- Multi-party/nation **consortia**, no change here
 - international partners; low-income countries, others
- **Funding:**
 - Research & Innovation
 - **100%** direct costs + 25% overheads
 - Innovation
 - in general **70%** direct costs + 25% overheads
 - but **100%** for non-profit participants
 - Coordination & Support
 - **100%** direct costs + 25% overheads



- Data & Language managed by a **single unit** (G3) within the Media & Data directorate
 - Data (big, open, linked, public...) a powerful **carrier** for language research & innovation work
 - Data is *everywhere*, also within H2020...!
- **InfoDay** in Lux on Jan 15 (400+ participants)
<http://ec.europa.eu/digital-agenda/en/news/information-and-networking-days-h2020-work-programme-2014-2015-connecting-europe-facility>
- Plenty of **useful information on-line**:
 - work programme excerpts
 - slideshows, explanatory notes
 - participants, presentations at networking sessions



- **Two calls, many themes**
- **Very competitive**, due to i) high demand, ii) low budget & iii) new emphasis on impact/innovation:
 - **1st call open**, closing April 23, evaluation in June, notification in Sept, grants by year-end
 - #15 Big Data, innovation & take-up (**50 M**)
 - #17 Language, machine translation (**15 M**)
 - #22 Multimodal Interaction, 3 different flavours (**31 M**)
 - **2nd call due in Oct**, closing April 2015
 - #16 Big Data, research (**39 M**)



Language within Data

- **NLP**, information analysis, extraction & management
 - **no longer a standalone topic**
 - integral part of Big Data calls (#15, #16)
 - "cross-sector, cross-border, cross-lingual"
 - "structured & unstructured content"
- **MT**, multi-language technologies
 - **#17 Cracking the language barrier**
- **Speech**, interactive systems
 - **#22 Multimodal interaction**
 - conversational systems (Q&A, digital assistants, smartphones, human robotics...); games & interactive media; special needs & disabilities



#17 Cracking the language barrier

- **Open now, modest budget, 15 M in total**
 - a small number of research (1), innovation (3-4) & coordination (1) actions
 - overall, attempt to achieve significant improvements in *translation quality* in *fully automatic machine translation*
 - special focus on difficult & poorly served EU languages, esp. as target languages
 - reuse of and close collaboration with existing & emerging, European & national platforms



#17 Cracking the language barrier

- **MT research**
- 1 broad-based multi-disciplinary action (4 M funding)
- relatively **long-term view**, towards a paradigm shift?
- **address current gaps & weaknesses**, particularly
 - insufficient quality
 - difficult adaptation
 - lengthy improvement cycles
 - scarcity of sufficiently large corpora...
- explore **new avenues**, test against agreed **benchmarks**
- connect to and exploit existing/emerging platforms & partnerships



#17 Cracking the language barrier

- **MT innovation**
- 3-4 focused actions (10 M in total)
- significant improvements in real, common use situations
 - **quality, linguistic & domain coverage**
 - ease of configuration & customisation
- target domains are not prescribed but
 - compelling problem statement & business case
 - clear market needs & user groups
- objective driven, little or no research *per se*



#17 Cracking the language barrier

- **MT coordination**
- 1 action (1 M)
- one or more of the following dimensions:
 - coordinated infrastructure for MT **benchmarking**
 - optimum (re)use of **LR** from various sources
 - **collaboration/exchanges** in the MT & LR areas (national, H2020, CEF...)
 - systems & data **interoperability, metadata** harmonisation
 - observatory in the language-technology area...
- building on and extending existing forums, networks & associations



Thank you!

roberto.cencioni@ec.europa.eu

Unit G3 mailbox:
CNECT-G3@ec.europa.eu

[http://ec.europa.eu/research/participants/portal/
desktop/en/home.html](http://ec.europa.eu/research/participants/portal/desktop/en/home.html)



RESEARCH & INNOVATION

European Commission

Participant Portal

European Commission > Research & Innovation > Participant Portal > Home

HOME

FUNDING OPPORTUNITIES

HOW TO PARTICIPATE

EXPERTS

SUPPORT ▾

LOGIN

REGISTER



On this site you can find and secure **funding** for research & innovation projects under the following EU programmes:

- **2014-2020** Horizon 2020 - research and innovation framework programme
- **2007-2013** 7th research framework programme (FP7) and Competitiveness & Innovation Programme (CIP)

Non-registered users

- search for funding
- read the funding guide & download the legal documents
- check if an organisation is already registered
- contact our support services or check our FAQs

Registered users

- submit your proposal
- sign the grant
- manage your project throughout its lifecycle



WHAT'S NEW?



FUNDING OPPORTUNITIES



HOW TO PARTICIPATE



WORK AS AN EXPERT



MY PERSONAL AREA



INFORMATION AND SUPPORT