

7.4 Challenge 4: Technologies for Digital Content and Languages

Digital content remains the material basis for a multilingual knowledge based society. However the explosive growth of digital content (both structured and unstructured) makes it important for European citizens and organisations to learn to manage it effectively and to extract from it maximum value in terms of private or public, personal or organisational decision making, planning and management.

The focus of Challenge 4 is on:

- exploring and testing new approaches, methods and techniques to extract, interpret and exploit information from unstructured multilingual and/or multimedia sources, yielding actionable knowledge;
- developing and testing in realistic operating conditions new algorithms and software frameworks to analyse, interact and visualize extremely large volumes of data in real time;
- supporting Small and Medium Enterprises (SMEs) developing innovative applications in structured and unstructured digital content management and, particularly, in the reuse of open data.

Support actions for road-mapping are also envisaged to prepare the constituencies for bringing together research and innovation aspects in Horizon 2020.

.....

Objective ICT-2013.4.3 SME initiative on analytics

Target Outcomes

Helping European Small and Medium Enterprises acquire the competences and resources they need to develop innovative content and data analytics services. Development of services based on the use of available data, particularly from public bodies, is specifically required for theme a) and encouraged for theme c).

a) Integrated Open Data Incubator

An Integrated Project to establish an environment and calling for efficient, small scale development of services of commercial interest based on the use of European open data by Small and Medium Enterprises (SMEs). The IP should:

- devote most of its resources to publish and manage regularly scheduled and well advertised calls for SMEs to submit mini-proposals to be funded for a period between six and twelve months.
- create a computing infrastructure where the winning mini-proposals will find accurate, up-to-date and (when useful and feasible) linked versions of the data they need for their services and, if they so wish, deploy the experimental version of their services.
- establish a mechanism for connecting open data demand and supply by systematically contacting European public bodies for their open data and assisting

them in the efficient and sustainable publication of such data, if needed with targeted engagements.

- solicit open data reuse ideas from the general public and conduct a European wide open data reuse information campaign.
- The IP will finally create a process to connect the most successful SMEs with sources of funding and business networks.

b) Easing transfer and take-up of language technologies

Language technologies are often deployed within products and services relating to web or enterprise intelligence, including text and audio mining, social media analytics and sentiment analysis, enterprise search and content management, online and cloud based translation, etc.

This action calls for focused user- and market-oriented projects in any of the above areas, with the overall goal of bringing language technologies closer to commercial maturity through an "industrialisation" process including but not limited to: (i) engineering of promising but commercially untested technologies, e.g. in terms of performance, robustness and coverage; (ii) integration within existing or upcoming products and services; (iii) first-use experimentation and validation in a clearly identified application domain; (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.

Consortia shall include players from the demand and supply sides, in particular SMEs, who have a clear interest in the exploitation of results.

c) Software components and intuitive end user applications based on reuse of open data

Development of software components supporting the whole life cycle of reuse of multilingual open data, particularly from public bodies. This includes:

- usable data publications methodologies and tools, adapted to the operating conditions of typical public bodies and rigorously tested for traceability, usability and sustainability in a public body environment;
- methods and tools for linking open data sets produced by public bodies;
- methods and tools for optimising open data applications based on public demand both in terms of content and in terms of functionalities/usability;
- cross platform development tools for delivering intuitive and responsive open data applications on multimodal devices and environments such as mobile, tablets as well as desktop.

Expected Impact

- A European open-source MT system becomes the most widely adopted worldwide; post-edited MT becomes the standard mode of translation within 5 years, increasing significantly (> 25%) the efficiency of human translation.
- Dozens of data application software components, used by hundreds of developers.
- Hundreds of applications, reusing billions of open data records, used by millions of end users around the EU

- Wider creation of valuable applications by integrating available public data with the users' own data, including contextual information available from mobile devices.

Funding Schemes

a): IP

b), c): STREP

Indicative budget distribution

- IP: EUR 5 million,

- STREP: EUR 15 million

Call:

FP7-ICT-2013-SME-DCA