H2020-ICT-2017 call

Big Data PPP topics

(ICT14, ICT15, ICT16, ICT17b, Big Data prize)

Kimmo Rossi
DG CONNECT G3 – Data Value Chain
NCP training 13/5/2016 Brussels

H2020-ICT-2016-1 (Big Data)

- ICT 14 Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation (IA) - Budget 27 M€
- ICT 15 Big Data PPP: large scale pilot actions in sectors best benefitting from data-driven innovation (IA) - Budget 25 M€
- ICT 16 Big Data PPP: research addressing main technology challenges of the data economy Budget 31 M€
- ICT 17 b) Big Data PPP: benchmarking and evaluation (1 RIA) - Budget 2 M€
- Inducement Prize: Ground-breaking Horizon Prize on Big Data technologies - Budget 2 M€

Legal base, background material

Work programme, main part:

http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016_2017/main/h2020-wp1617-leit-ict_en.pdf

Technical Background Notes and Info Day presentations (not legally binding):

https://ec.europa.eu/digital-single-market/en/news/big-data-information-and-networking-days-horizon-2020-topics

Big Data PPP: The overall challenge

- The main objective is to roll out an industrial strategy to develop Europe's data driven economy as outlined in the EC Communication 'Towards a thriving data-driven economy' COM(2014)442
- The WP2016-17 implements the Big Data PPP's Strategic Research and Innovation Agenda: http://www.bdva.eu/sites/default/files/EuropeanBigDataValuePartnership_SRIA__v2.pdf

ICT 14 cross-sectorial and cross-lingual data integration and experimentation – "Innovation Spaces"

Problem statement:

- no systematic transfer of knowledge and technology across different sectors
- lack of data sharing and linking culture – data remains within sectorial "silos
- lack of agreed standards and formats; low rates of publishing data assets in machine discoverable formats
- textual data appears in many languages creates an additional challenge for sharing and linking such data.
- lack of secure environments where researchers and SMEs can test innovative services and product ideas based on open data and business data.

ICT 14

Overall Objective

To foster the exchange, linking and reuse of data assets. To integrate data assets from multiple sectors across languages and formats in a safe environment for experimentations of innovative services and product ideas.

ICT 14 a) Data Integration activities

- Innovation Actions addressing **cross domain/cross-lingual data integration** challenges of EU industries arranged along **data value chains**.
- Wide range of technical issues to be tackled (i.e. data models, entity identifiers, standards, multilingual support, brokerage schemes, data quality, privacy, etc…)
- Indicative project size: 1-3 MEUR

ICT 14 a) - Impact

Expected Impact: a. Data integration activities

- Data integration activities will simplify data analytics carried out over datasets independently produced by different companies and shorten time to market for new products and services;
- Substantial increase in the number and size of data sets processed and integrated by the data integration activities;
- Substantial increase in the number of competitive services provided for integrating data across sectors;
- Increase in revenue by 20% (by 2020) generated by European data companies through selling integrated data and data integration services offered.

ICT 14 b) Data experimentation incubators

Incubator(s) addressing big data industrial challenges in a cross-sectorial, cross-lingual and/or cross-border set-up. Experimenters: SMEs and start-ups. At least 50% of experiments to be defined by data providers.

The incubator will offer access to cross-sectorial, cross language data pools, computing infrastructure and open software tools in addition to an organizational, legal, IPR support environment.

Cascading Grants scheme – indicative size: 7 MEUR

ICT 14 b) - Impact

- At least 100 SMEs and web entrepreneurs, including start-ups, participate in data experimentation incubators;
- 30% annual increase in the number of Big Data Value use cases supported by the data experimentation incubators;
- Substantial increase in the total amount of data made available in the data experimentation incubators including closed data;
- Emergence of innovative incubator concepts and business models that allow the incubator to continue operations past the end of the funded duration.

ICT 15: Large scale pilot actions in sectors best benefitting from data-driven innovation - "Lighthouse Projects"

Problem statement:

- European R&D results in data technologies are not yet deployed at large scale in a systematic manner.
- Need to stimulate effective piloting and targeted demonstrations in large-scale sectorial actions ("Large Scale Pilot actions"), in data-intensive sectors, involving key European industry actors.

ICT 15: objectives

- Large Scale Pilot Actions **in data intensive sectors** involving key European industrial actors.
- Their objective is to demonstrate how industrial sectors will be **transformed** by putting big data technologies at their core.
- The Large Scale Pilot actions are meant to serve as best practice examples to be **transferred** to other sectors.

ICT 15: characteristics & requirements

- Possible industrial sectors for Large Scale Pilot actions include (but are not limited to) health, energy, environment, earth observation, geospatial, transport, manufacturing, finance and media.
- Large Scale Pilot actions are expected to exhibit substantial **visibility, mobilisation, and commercial and technological impact**. Proposals must demonstrate that they have access to appropriately large, complex and realistic data sets.
- Indicative project size: 10-15 MEUR

ICT 15: characteristics & requirements

- A **Consortium** where **industrial partners** are represented by professionals who work in **core business operations** (as opposed to research laboratories)
- Develop a plan that is consistent with the business strategy of the industrial partners concerned (**e.g. avoiding committing to technologies that the decision makers in the respective companies have no intention of deploying**)

ICT 15 – Industrial Requirements

- Describe the **industrial strategy and development plans** of the commercial companies in your consortium.
- State **explicitly** and in quantitative, verifiable detail what amount of **own resources** the company intends **to invest to leverage the grant received**, if selected for funding.
- If a company has no concrete/verifiable plans to invest additional/own resources, note so explicitly.

ICT 15 - Impact

- Demonstrated increase of productivity in main target sector of the Large Scale Pilot Action by at least 20%;
- Increase of market share of Big Data technology providers of at least 25% if implemented commercially within the main target sector of the Large Scale Pilot Action;
- Doubling the use of Big Data technology in the main target sector of the Large Scale Pilot Action;
- Leveraging additional target sector investments, equal to at least the EC investment;
- At least 100 organizations participating actively in Big Data demonstrations (not necessarily as partners of the projects).

ICT 16: research addressing main technology challenges of the data economy

Problem statement:

Low value generation from (Big) Data assets because the **available software and IT architecture solutions** are not adapted to the processing, analysis and visualisation of data in a situation where the volume, velocity and variety of the data are increasing rapidly (=Big Data).

ICT 16: objective

To fundamentally improve the technology, methods, standards and processes, building on a solid scientific basis, and responding to real needs.

ICT 16: characteristics

Address cross-sector and cross-border problems or opportunities of clear industrial significance. These will include (but are not limited to):

- Software stacks designed to help programmers and big data practitioners take advantage of novel architectures in order to optimise Big Data processing tasks;
- Distributed data and process mining, predictive analytics and visualization at the service of industrial decision support processes;
- Real-time complex event processing over extremely large numbers of high volume streams of possibly noisy, possibly incomplete data.

Indicative project size: 2-5 MEUR

ICT 16: requirements

- All human factors claims (e.g. usability, maintainability) concerning software to be developed will need to be **rigorously tested by methodologically sound experiments** with clear plans to recruit **adequate numbers of appropriate experimental subjects**
- **Access to appropriately large, complex and realistic data sets.**
- Make best possible use of large volumes of diverse **corporate data** as well as, where appropriate, **open data** from the European Union Open Data portal and/or other European open data sources
- Make appropriate use of and/or contribute to data exchange and interoperability **standards**.

ICT 16: expected impact

- Powerful (Big) Data processing tools and methods that demonstrate their applicability in real-world settings, including the data experimentation/integration (ICT-14) and Large Scale Pilot (ICT-15) projects;
- Demonstrated, significant increase of speed of data throughput and access, , as measured against relevant, industry-validated benchmarks;
- Substantial increase in the definition and uptake of standards fostering data sharing, exchange and interoperability.

ICT 17 b) – benchmarking and evaluation – problem statement

Need for widely recognised benchmarks and performance evaluation schemes to **avoid fragmentation or overlaps**, and to allow **measuring progress** in (Big) Data challenges by solid methodology, especially in emerging areas where the significance of Big Data is rapidly increasing.

ICT 17 b) – benchmarking and evaluation

• The benchmarking action will identify data management and analytics technologies of European significance:

Define benchmarks and organise evaluations that allow following their certifiable **progress on performance parameters** (including energy efficiency) of industrial significance,

Liaise closely with data experimentation/ integration (ICT-14) and Large Scale Pilot (ICT-15) projects **to respond to key European industries real needs**, and to provide a basis for measuring success of the PPP.

ICT 17 b) – benchmarking and evaluation

Indicative project size: 2 MEUR

"European significance" of industry/technology sectors should be determined by objective criteria such as turnover, world-wide market share and growth rates of the European companies who provide or use such technologies.

Action shall address areas of activity **that do not yet have a benchmarking/evaluation scheme**. We already have:

- LDBC: benchmarking for Graph Databases
- HOBBIT: benchmarking for Linked Big Data

ICT 17 b) – Requirements

- Benchmarks proposed **must be of industrial relevance**, i.e. relevant for European developers and providers of data technologies,
  - European developers should continuously improve their performance (=competitiveness) as **measured against benchmarks**,
  - Consortia should identify **industrial actors that have expressed interest** in the technology for very specific business reasons, and involve them in the definition of benchmarks and performance goals.

ICT 17 b) – Requirements (II)

• Benchmarks are invited for **those technologies for which they do not exist** already.

• Technology benchmark initiatives must give reasonable **guarantees that they will continue to exist** throughout the entire life-cycle of the relevant technology,

- Consortia should **secure industrial involvement**,
- Benchmarks are expected to help European technology providers to become globally competitive, consortia should **involve potential clients** in the definition of the benchmarks.

ICT 17 b) – what is expected

(see background note)

- Provide a precise definition of the Big Data technology for which you intend to develop a benchmark,
- Develop benchmarks for specific technology which deserves it,
- Ensure the sustainability of the benchmarking activities past the end of Horizon 2020 funding.

ICT 17 b) – Expected impact

- Availability of solid, relevant, consistent and comparable metrics for measuring progress in Big Data processing and analytics performance;
- Availability of metrics for measuring the quality, diversity and value of data assets;
- Sustainable and globally supported and recognized Big Data benchmarks of industrial significance.

Inducement Prize: Ground-breaking Horizon Prize on Big Data technologies

- A "place holder" for this prize can be found on page 130 of the current WP2016-17
- A more detailed description will appear in the summer update of WP2016-17
- Disclaimer: the following two slides present provisional information, subject to the adoption of the summer update of WP2016-17. In case of any inconsistency, the most recent adopted work programme will prevail.

The problem: becoming very accurate and efficient in predicting the future based on past data

How the prize scheme works for contestants:

- Extremely large amounts of past data about EU weather, energy production/consumption will be made available to train your algorithm(s)
- You submit your fully implemented prediction algorithm to a platform prepared by CSA SEE.4C-688356 (as many submissions as you want)
- The platforms automatically scores the performance of your submission on unseen data based on a public and verifiable success metric
- Best score in category (accuracy, energy consumption, others to be announced…) wins the prize

Indicative Budget: €2M over several categories

Stages	Indicative date and time
Opening of the contest	Second quarter 2017
Deadline for submission of application	Fourth quarter 2017
Award of the Prize	Second quarter 2018

Eligibility criteria: The contest will be open to any legal entities (including single persons) or groups of legal entities.

Exclusion criteria: listed in Financial Regulations Art. 105, 106, 108
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:298:0001:0096:EN:PDF

Thank you for your attention!

kimmo.rossi@ec.europa.eu