



# H2020-ICT-2015

## ICT 16 Big Data - Research

Francesco Barbato & Stefano Bertolo  
Project Officer

European Commission – CNECT G3

ICT16 InfoDay, Brussels  
16/1/2015

## a2) Innovation Actions: Benchmarks

The goal is to create big data analysis and prediction benchmarking environments of **industrial relevance**

- **Industrial relevance**

- Not (only) about 'cool' technology, but about problems (European) industry has
- Examples (to be learned from, not to be duplicated):
  - <http://www.tpc.org>
  - <http://prof.ict.ac.cn/BigDataBench/>
  - <https://amplab.cs.berkeley.edu/benchmark/#>
  - <http://wifo5-03.informatik.uni-mannheim.de/bizer/berlinsparqlbenchmark/spec/index.html>
  - <http://ldbc.eu>

## a2) Innovation Actions: Benchmarks

- **Sustainability strategy**
  - Benchmarks help industry make investment decisions
  - Benchmark organizations support decision process in long term
  - Don't duplicate/overlap existing benchmarking activities (but study their current activities to learn best practices)

## **a2) Innovation Actions: Benchmarks**

- **ICT-16.a2 is mutually exclusive of ICT-16.a1 (i.e. no research, no innovation allowed)**
- **Expected size of consortium: as compact as possible to get the job done. BUT**
- **It must contain obvious EU industrial champions in the domain of interest**
- **It must contain a 'balanced' list of industrial participants (to ensure that benchmark will be fair and not skewed in favour of anyone in particular)**

## **b) CSA – Challenges and Prize schemes (1M€)**

Support actions to define challenges and prize schemes for **verifiable performance** in tasks requiring extremely large scale prediction and deep analysis. **Compact consortia** are required to organise and run well-publicised fast turn-around prediction competitions based on European datasets of a significant size. **Proposals in this category are expected to be short in duration** and are not required to provide sustainability strategies past the end of the project.

## **b) CSA – Challenges and Prize schemes (1M€)**

### **Extremely large scale prediction/analysis**

- Identify prediction tasks that are relevant to those who benefit from the predictions (not only those who write the prediction tools)
- Examples on <http://www.kaggle.com>
- Competitions must be based on large scale datasets of European Interest

## **b) CSA – Challenges and Prize schemes (1M€)**

### **Motivate the 'what'**

- The challenge defines 'what' problem must be solved
- Contestants propose the 'how'
- Make it very clear who does (or should) care about the 'what' and for what reasons
- Talk to industry (decision makers and product lines, not only researchers)
- Talk to industry early
- Talk to industry often to refine scope of challenge

## **b) CSA - Challenges and Prize schemes (1M€)**

Verifiable performance

- Example <http://sortbenchmark.org>
- Must agree in advance on details of challenge/testing environment, data, scoring rules
- Must agree to the publishing of results



## **b) CSA - Challenges and Prize schemes (1M€)**

Housekeeping: implement and test (cloud) software infrastructure to

- Make datasets, computing resources available to contestants
- Receive submissions
- Score results
- Anonymously report results
- 'Turnkey' solution

A good precedent: <http://www.ebi.ac.uk/Rebholz-srv/CALBC/>



## **b) CSA - Challenges and Prize schemes (1M€)**

The prize money will be made available and the prize award conditions stated in the next H2020 Work Programme (2016-2017).

## b) CSA - Challenges and Prize schemes (1M€)

Examples in other domains: <http://ec.europa.eu/digital-agenda/en/inducement-prizes>

- **Affordable Food Scanner** (take sample of food, return nutritional properties, allergens, ...) 1M€
- **Collaborative Spectrum Sharing** (improvement in terms of capacity per single fibre, spectrum range and/or spectral efficiency and reach) 0.5M€
- **Breaking Optical Transmission Barrier** (maximise fibre transmission capacity per channel , spectrum range and/or spectral efficiency and reach) 0.5M€



# Further Information - Technical Background note

The ICT-16's Technical Background Note is available on the event website the Digital Agenda Website:

<https://ec.europa.eu/digital-agenda/events/cf/ictpd14/item-display.cfm?id=12590>



**Thank you!**

**Questions?**

**Francesco Barbato & Stefano Bertolo  
Project Officers  
European Commission – CNECT G3**