H2020-ICT-2015

ICT 16
Big Data - Research

Francesco Barbato, Project Officer
European Commission – CNECT G3

ICT-16 Info Day,
Brussels 16/01/2015

The new Framework Programme

- H2020 is the financial instrument implementing the *Innovation Union*, a *Europe 2020* flagship initiative aimed at securing Europe's global competitiveness
- H2020 is NOT business as usual (i.e. a mere continuation of FP7)
- H2020 transitioned from R&D to Research and **Innovation** (R&I)
- H2020 is aimed at bringing great ideas from the lab to the **market**

Research and Innovation action (as described in the H2020 WP)

- Action 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. For this purpose 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.
- Projects may contain closely connected but limited demonstration or pilot activities aiming to show technical feasibility in a near to operational environment.

H2020-ICT-2015

ICT 16 Big Data Research

a) Research and Innovation Actions (Collaborative Projects) – a1) and a2)

b) Coordination and Support Actions

Budget

a) 36 M€

b) 1 M€

The call is open!

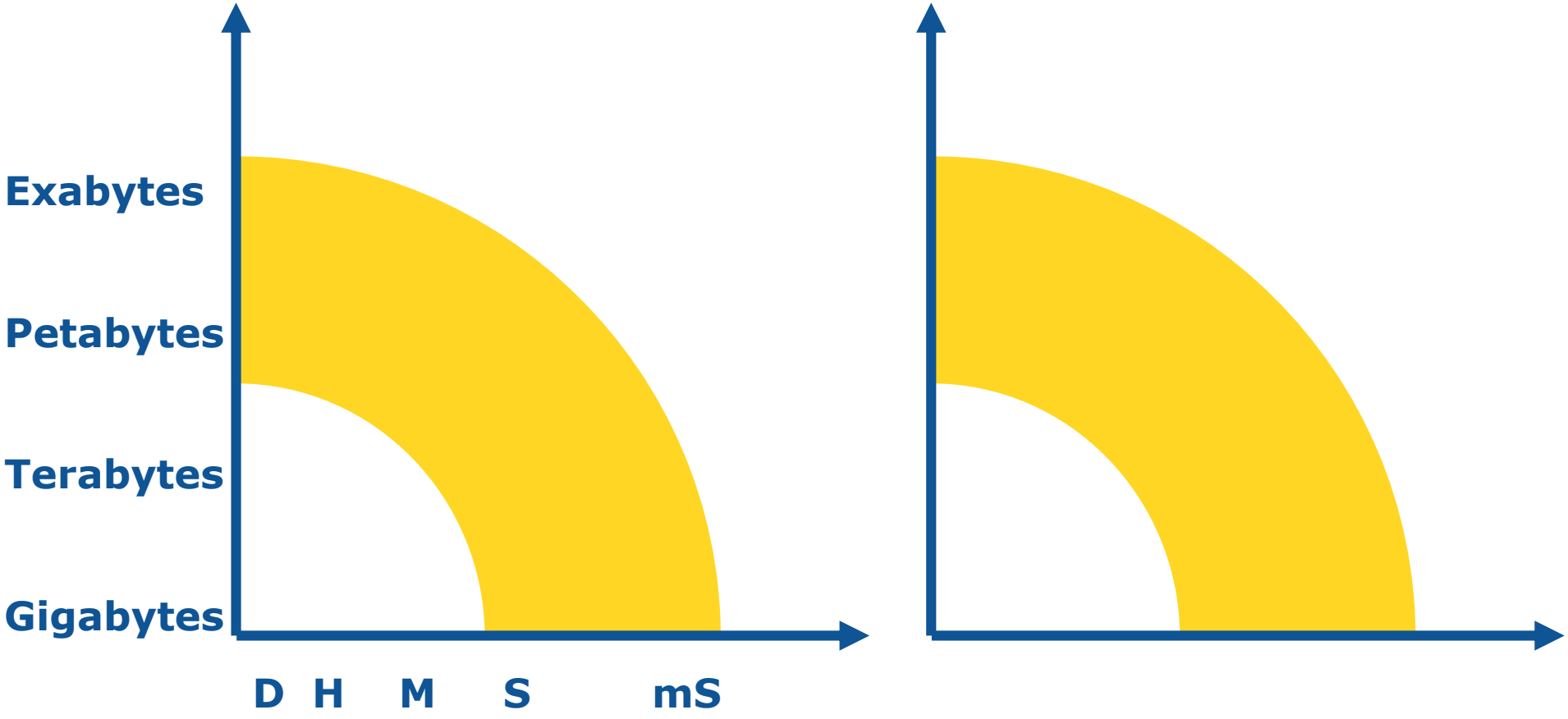
Deadline: 14 April 2015, 17:00 Brussels time

H2020-ICT-2015

- “big data” is when the size of the data itself becomes part of the problem

- “big data” is data that becomes large enough that it cannot be processed using conventional methods

Volume, Velocity, Variety



a1) Research and Innovation – Collaborative Projects

- **Data structures**
  - Examples (not to be duplicated): HyperLogLog, CountMinSketch, DHT
  - Memory hierarchy conscious data structures
  - Energy conscious data structures
  - Streams
- **Algorithms**
  - Energy conscious algorithms (reversible computing?)
  - Memory hierarchy conscious data structures
  - Parallelism (exploitation of multi cores; functional programming)
  - Streaming analytics
  - Real time
  - Data curation
  - Prediction

a1) Research and Innovation – Collaborative Projects

- **Visualization**
  - Not just pretty pictures
  - Usability (on various platforms)
  - Measurable impact on business processes
  - Importance of experimental protocol, population of subject matter experts

What we will be looking for: user-defined challenges

- **FP7:**
  - first researchers identify topic interesting to them,
  - then they look for narrative/qualitative validating use cases
- **H2020:**
  - first industry identifies a problem they have and cannot solve with current methods,
  - then they recruit for scientists to find one or more solutions,
  - then they define expected impact as a measurable outcome
  - then they try the solutions (run rigorous experiments)
  - then they report the measured outcomes (good or bad)
  - Usefulness of system proposed should be an hypothesis to be experimentally validated

What we will be looking for: data management plan

- **How to manage proprietary data**
  - Make sure industrial partners are comfortable sharing
  - Ensure protection of personal data
- **How to manage/share non-proprietary data**
  - In H2020 non-proprietary data under Open Access regime
  - Choose formats wisely for maximal reuse
  - Establish credible post-project future for datasets

What we will be looking for: availability of data sets

Proposal must clearly state (ideally in a dedicated, easy to find, table):

- Which datasets: what type of data do they contain? how big are they now? How fast do they grow?
- When: must be available at the very beginning of the project. Additional data at M6, M12, M18…?
- From whom: which partner(s) of the consortium have which access rights? (relevant if customer data)

What we will be looking for: European data sets

- Can I use Twitter, Flickr, …? No. Not primarily
- Ecologically valid (i.e. operational) data from European companies
- Data sets from EU open data portals
- Data sets from other EU public sources, e.g. Copernicus http://www.copernicus.eu see http://europa.eu/rapid/press-release_IP-13-1067_en.htm

Looking for inspiration?

- The PPP's Strategic Research and Innovation Agenda is available online at

http://www.bigdatavalue.eu/

- Technical Priorities
- Industrial Sectors

Ongoing FP7 Big data Project

In particular Projects selected from

- ICT Call 8
- ICT Call 11

- ICT SME-DCL
- ICT SME-DCA

http://cordis.europa.eu/fp7/ict/content-knowledge/projects_en.html

Expected Project size

- Projects shall be fit for purpose!
- The resources and budget requested shall be appropriately justified in the proposal

Further Information - Technical Background note

The ICT-16's Technical Background Note is available on the Info Day website :

https://ec.europa.eu/digital-agenda/en/news/horizon-2020-ict-16-big-data-networking-day

Thank you!

Francesco Barbato, Project Officer
European Commission – CNECT G3