

Technical Background Notes for Horizon 2020 Objective ICT-15-2016-2017 Big Data PPP: Large Scale Pilot actions in sectors best benefitting from data-driven innovation

DG CONNECT/G1

CNECT-G1@ec.europa.eu

<https://ec.europa.eu/digital-single-market/en/big-data>

<https://ec.europa.eu/digital-single-market/language-technologies-and-big-data>

This document is intended to provide background information and technical commentary on Topic ICT-15 2016 published as part of the 2016-17 Horizon 2020 work programme:

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

The official work programme text is the only legally binding source of information on the topic. Should any inconsistency between the present explanatory document and the official text be detected it is always to be resolved in favour of the work programme text.

Motivation of the Topic and Scope of this Document

The official text of the work programme specifies that proposals under this objective should be **Innovation Actions**, defined as:

An action primarily consisting of activities directly aiming at producing plans and arrangements or designs for new, altered or improved products, processes or services. For this purpose they may include prototyping, testing, demonstrating, piloting, large-scale product validation and market replication.¹

This means that research (and publication) activities should **not** be the primary focus of proposals under this objective.

The topic addresses the following challenges:

Specific Challenge: European research and development in data technologies produces promising results, but these are not yet deployed at large scale in a systematic manner. The challenge is 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. The Large Scale Pilot actions are meant to serve as best practice examples to be transferred to other sectors and also as sources of generic solutions to all data intensive sectors (page 43)

¹ http://ec.europa.eu/research/participants/portal/desktop/en/support/reference_terms.html

The motivation behind this topic is to demonstrate in concrete and verifiable terms how data assets and technologies can be used as the foundation of the (re-)engineering of existing or novel industrial activities. Examples of this are becoming very visible: in the summer of 2015 a European industry consortium concluded a multi-billion Euro deal to acquire a very large scale, very high quality data asset which all members of the consortium will jointly exploit to completely transform their business. This is one example of the type of transformative effect that large scale pilots funded under objective ICT-15-2016-17 are expected to have in any European economic sector that can be made more productive by the pervasive use of data technologies.

To quote again from the official text of the topic, if this is accomplished the expected impact is:

- *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). (page 44)*

Doing this will require:

1. Assembling a consortium where industrial partners are primarily represented by professionals who work in core business operations (as opposed to research laboratories)
2. Developing a plan of action that is consistent with the business strategy of the industrial partners concerned (e.g. avoids committing to technologies that the decision makers in the respective companies have no intention of deploying)
3. Developing a rigorous regimen of measurement, testing and reporting so as to establish proof that the innovations developed in the pilot really work in actual operating conditions and are consistent with important business parameters such as reliability, accuracy, cost structure.
4. Developing a plan of action that is consistent with the European, national and sectorial regulations the consortium members are subject to and whose main ideas and lessons learned may be replicated beyond the pilot itself.

The text of the work programme clarifies the scope of the objective:

Scope: Large Scale Pilot actions should address domains of strategic importance for EU industry and carry out large scale sectorial demonstrations which can be replicated and transferred across the EU and in other contexts. (page 43)

In this context, "*strategic importance*" means that activities undertaken in these domains can be rationally expected to increase the productivity and competitiveness of European enterprises **after one takes into full account the business plans and resources of non-European enterprises.**

Proposals that demonstrate poor awareness or unrealistic assessments of the global competitiveness landscape of the respective industries **will be penalised accordingly.**

Similarly, "*large scale sectorial demonstrations*" means demonstrations that will work at the scale of the relevant industrial processes and data volumes/intensity not as they are today, but as they are expected to be at the end of the pilot.. This requires for the industrial partners of the consortium to provide **credible quantitative projections** on the future scale of their data operations together with detailed and technologically credible plans for how work in the pilot will produce systems that meet those requirements.

Proposals that fail to provide such projections **will be penalised accordingly.**

The work programme provides some examples of possible industrial sectors that have been identified through sustained consultations and dialogue with industry²:

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.(page 43)

As stated in the official text of the work programme, these examples are indicative and not exclusive. This means, in particular, that a high quality proposal that comes from an industrial domain **not** included in the list above will be preferred to a mediocre proposal from a domain included in the list.

Although Large Scale Pilot actions are required to have a strong focus in a given industrial domain, they may involve cross-domain activities where these provide clear added value. (page 43)

The most visible European examples of the transformative effect of data technology on an industry are focused on clearly recognisable traditional industrial domains. Pilot proposals solicited under this objective, however, are free to bring together consortia where this effect is accomplished by horizontal or vertical integration.

Based on the trends clearly visible in global business, one can expect at least the two following patterns to generate convincing large scale pilot proposals fit for objective ICT-15-2016-17.

The first one is vertical integration of data assets along a supply chain.

² See the "European Big Data Value Strategic Research and Innovation Agenda" http://www.bdva.eu/sites/default/files/EuropeanBigDataValuePartnership_SRIA_v2.pdf produced by the Big Data Value Association, the European Commission's partner in the Big Data Value Public Private Partnership <http://ec.europa.eu/digital-agenda/en/big-data-value-public-private-partnership>

Imagine a company X that supplies an instrumented components C to a second company Y, which integrates it into its product P. Company X may be benefit from learning about the performance/behaviour of its component C, but it is company Y, and not company X, which has a relationship with the buyer of product P.

Another example might involve two distinct companies A and B that are both part of a distribution chain (e.g. a food chain) and decide to jointly track how items move down the distribution chain.

The second example is the reuse of data assets across a given industry sector. This could be accomplished by the creation of a safe environment within which members of a consortium share data assets to be used across business processes.

For example, a company that operates a fleet of vehicles delivering packages may decide to share data about such operations with a brick-and-mortar retailer that may draw business insights from those data (e.g. parts of town and times of the day with traffic patterns correlated to better in-shop sales).

In both cases, the implementation of data management tools based on existing data standards (and the development of new standards where needed) is likely to be considered an effective way of fostering industry-wide adoption of novel data processes and components.

These are just indicative and hypothetical examples: as explained, proposals will be judged by their prospects for making European companies more productive and in control of larger market shares through the use of data assets and technologies:

Large Scale Pilot actions will propose replicable solutions by using existing technologies or very near-to-market technologies that could be integrated in an innovative way and show evidence of data value (see the section "Expected Impact"). Their objective is to demonstrate how industrial sectors will be transformed by putting data harvesting and analytics at their core (page 43).

In order to do this, it is imperative that the proposals themselves be conceived and planned "at scale", both in terms of the infrastructures and data resources they make available:

Large Scale Pilot actions are expected to exhibit substantial visibility, mobilisation, and commercial and technological impact. Proposals should demonstrate that they have access to appropriately large, complex and realistic data sets. (page 44)

and in terms of the financial resources they mobilise:

The Commission considers that proposals requesting a contribution from the EU of between EUR 10 and 15 million would allow this specific challenge to be addressed appropriately. Nonetheless, this does not preclude submission and selection of proposals requesting other amounts (page 44)

In this context, it is important to observe that when the Big Data Value Public Private Partnership was established, the following commitments were made:

The European Commission and Europe's data industry have committed to invest €2.5 billion in a public-private partnership (PPP) that aims to strengthen the data sector and put Europe at the forefront of the global data race. [...]The EU has earmarked over €500 million of investment over 5 years (2016-2020) from Horizon 2020 which private partners are expected to match at least four times over (€2 billion).³

In order to stay on track towards this commitment, PPP large scale pilot proposals requesting 10-15 million Euros of Horizon 2020 funding are expected to demonstrate a **verifiable** commitment to additional investment of 50-75 million in data assets or technologies by the partners of the consortium.

Since the ability to leverage industry funding is highly indicative of the impact the industrial partners expect from the pilot, proposal evaluators will be instructed to pay particular attention to this feature of proposals.

Notice

The information requested for the proper assessment of the work described in a proposal including

1. Data assets made available to the consortium
2. Industrial projections of increased productivity and market share
3. Industrial commitment to additional investments (as required by the terms of the Big Data Value Public Private Partnership agreement)

must be presented in the main body of the description of work (what is known as Part B Section 1) and not in the section (Part B Section 2) describing the members of the consortium.

³ http://europa.eu/rapid/press-release_IP-14-1129_en.htm

Appendix: a list of questions that proposals must answer in order to be in scope of objective ICT-15 2016-17 of Horizon 2020

This appendix contains a list of simple questions that a consortium should ask about the proposal to be submitted. If the proposal as submitted does not contain a clear answer to the majority of the relevant questions it places itself at a serious disadvantage in a very competitive selection process (because the evaluators will be specifically instructed to look for the answers to these and other questions)

1. Is the large scale pilot going to use data in order to **improve an existing industrial/commercial process** or to create the conditions for **novel industrial/commercial processes**?
2. If the large scale pilot is going to use data in order to **improve existing industrial/commercial processes**, please provide **explicit, quantitative** data (ideally displayed in a dedicated table, which the evaluators will be instructed to look for and whose absence to note) on:
 - a. their cost structure⁴
 - b. their current technological constraints/limits
3. If the large scale pilot is going to use data in order to **improve existing industrial/commercial processes** please provide **explicit** and **verifiable** details (ideally displayed in a dedicated table, which the evaluators will be instructed to look for and whose absence to note) on:
 - a. How you plan to measure and report changes in the cost structure of the processes at the end of the pilot (if funded)
 - b. How you plan to measure and report changes in the technological constraints by the end of the pilot (if funded)
4. If the large scale pilot is going to use data in order to create the conditions for **novel industrial/commercial processes**, please explain:
 - a. The exact characteristics of the novel processes (including their expected performance requirements and cost constraints)
 - b. How they are consistent with the known industrial strategies and existing or planned technological infrastructure of the partners that intend to deploy them in their business
5. If the large scale pilot is going to use data in order to create the conditions for **novel industrial/commercial processes**, please provide **explicit** and **verifiable** details (ideally displayed in a dedicated table, which the evaluators will be instructed to look for and whose absence to note) on:
 - a. How you plan to measure and report the cost structure of these processes at the end of the pilot (if funded)
 - b. How you plan to measure and report on the technological performance parameters by the end of the pilot (if funded)
6. Please provide a detailed account of the **ownership and user right structure of the data assets** to be used and/or produced during the pilot (if selected for funding).

⁴ https://en.wikipedia.org/wiki/Market_analysis#Industry_cost_structure

Provide a detailed account of who will own or have user rights to said data assets after the end of the pilot (if selected for funding).

7. Please describe the **industrial strategy and development plans** of the commercial companies in your consortium and for each of them (ideally in a dedicated table, which the evaluators will be instructed to look for and whose absence to note) state **explicitly** and in **quantitative, verifiable** detail what amount of **own** resources the company intends to invest to leverage the grant received for this proposal, if selected for funding. If a company has no concrete/verifiable plans to invest additional/own resources, note so explicitly.
8. Using a credible competitiveness analysis framework of your choice⁵, provide a credible analysis as to why the changes in industrial/commercial processes introduced by the result of your pilot will make the members of your consortium (and other European industries in the same sector) more competitive **after taking full account of global competition**. For each commercial partner of the consortium provide (in a dedicated table, which the evaluators will be instructed to look for and whose absence to note) a motivated **quantitative** estimate of the resulting increase in market share (together with an equally **quantitative** estimate of the total size of the market by the end of the pilot).
9. If the large scale pilot foresees the creation of data supply chains, please describe what **standards** will be used (or developed) and what **quality assurance methods** will be deployed to ensure that data assets of the appropriate quality to support the intended industrial/commercial process are exchanged with minimal friction across different companies.
10. Please describe in detail any **legal constraint** that the specific industry sector addressed by the large pilot imposes on the collection and management of the data assets to be used. Similarly for national legislation that affects consortium members.

⁵ E.g. https://en.wikipedia.org/wiki/Porter_five_forces_analysis