

Data Pitch

H2020-ICT-2016-1

Project number: 732506

D3.4

First Data Pitch Consultations

Coordinator: Johanna Walker, University of Southampton

With contributions from: Gareth Beeston, University of Southampton, Orsola DeMarco, Open Data Institute, Ed Parkes, Deborah Okenia

Quality reviewer: Elena Simperl

Deliverable nature:	Report (R)
Dissemination level (Confidentiality)	Public (PU)
Nature	Other
Document URL	
Work package	WP3 - Data and Data Providers' Liaison
Contractual delivery date:	31st March 2017
Actual delivery date:	5th May 2017
Version:	1
Keywords:	Consultation, Challenges, Data

Contents

Abstract

0. Executive Summary

1. Introduction

- 1.1 Aims of this Deliverable
- 1.2 The Data Pitch Concept
- 1.3 The Challenge Concept
- 1.4 Tracks
- 1.5 The Intersection of Tracks and Challenges
- 1.6 Developing Challenges
- 1.7 The Consultation

2. Background

- 2.1 Open Innovation
- 2.2 Open Innovation in Data Pitch
- 2.3 Data Incubators
- 2.4 Datathons

3. Consultation

- 3.1 High Impact Challenge Value Consultation
- 3.2 Consultation Strategy
- 3.3 Compliance
- 3.4 Consultation
- 3.5 Analysis
- 3.6 Workshop Process
- 3.7 Supporting Activities
- 3.8 Output

Abstract

This document outlines the work completed so far and planned to deliver the consultation on challenges for Data Pitch, including the public consultation, expert discussions and consultation with data providers.

0. Executive summary

Data Pitch is an innovation programme that aims to bring large organisations with data together with startups and small to medium enterprises with data skills to solve some of the problems that face society today.

To do this it must identify the highest impact economic and social problems that can be solved with data (“high impact challenges”), in three sectors (“tracks”), namely Smart Cities, Health and Wellbeing and Agriculture and Nutrition. These challenges are defined in part by companies willing to share their data in order to solve problems of concern to them (“data providers”) but also in part through a process of consultation with industry experts, startups and citizens.

A number of potential challenges which have been refined with expert input are presented in the consultation for views on their importance. The opportunity for experts, startups and citizens to provide their own ideas for challenges is offered. The consultation also seeks views on which data could be used to address these challenges.

The consultation will run for a month after which all input will be analysed at a workshop involving the Data Pitch team and invited experts. The final output of this consultation will be a number (yet to be determined) of clearly defined challenges. This will be included in the Call for Applications for Data Pitch Round 1. Startups will need to demonstrate that they have a clear plan to address these sustainably in order to join the Data Pitch Innovation Programme.

1. Introduction

A core part of the Data Pitch programme is to identify critical cross-sector and cross-border challenges from a wide array of Big Data Value stakeholders, including the data owners we partner with. These challenges will shape the innovation life cycle and determine the design of the innovation instruments used and the thematic scope of the incubator.

A challenge as defined by Data Pitch is a social or economic problem for which a solution (or solutions) can be developed with closed, shared or open data, or a combination of all three.

This document explains the challenges, describes the methods of collection of ideas for the challenges and presents some early ideas and plans for M1-M6.

1.1. Aims of this Deliverable

This deliverable describes both Data Pitch's process for, and initial results of, consultation for collection of input and ideas for our challenges.

This deliverable is related to Round 1 of the Data Pitch call. In the subsequent round we will publish a further deliverable building on the learning of Round 1 and gathering new input and ideas for Round 2.

This document is structured as follows: In section 2, we review the open innovation approach to solving high impact challenges adopted by Data Pitch. Section 3 outlines the specific steps of our consultative approach to identifying and defining the challenges and the output.

1.2. The Data Pitch Concept

Data Pitch brings together corporates and public sector organisations, that have data, with startups and SMEs which work with data. It is centred around an innovation competition with several tracks, which describe challenges set by data-owning organisations and through a method of crowdsourcing, and a 6-month incubation programme to help startups and SMEs develop solutions to meet these challenges.

The data sets can be open, shared or closed. The startups and SMEs will put forward proposals for creating high impact, innovative products and services in response to the challenges. The challenges will cover a number of topics including smart cities, food & agriculture, amongst others. Data Pitch manages the overall programme, organises the competition and selection process, funds the best ideas, and supports the best ideas through business incubation. The competition will be run twice, with multiple tracks and challenges in each call. Each call will draw upon challenges and data sets from several organisations.

The first Data Pitch call opens at 12:00 noon Central European Summer Time (CEST) 1st July 2017 and will remain open until 12:00 noon Central European Summer Time (CEST) 30th September 2017. Selected ideas will receive up to €100,000 equity-free investment and will be entered into the 6-month incubation program.

1.3. The Challenge Concept

Through open innovation, Data Pitch supports the development of credible, sustainable data products or services. However, in order that these new products or services are directed at solving the most important and valuable problems, socially and economically, it is necessary to identify and formulate the most high-impact challenges that can be addressed with data. Challenges can be formulated in a number of ways, but there are key best practices that must be adhered to:

Challenges should:¹

Be Inclusive and Engaged

Not only should a competition be open to as many applicants as is manageable but participation should be actively encouraged – as should be collaboration. Offering a forum (such as an online platform) around a challenge that fosters discussion brings together potential solution creators and encourages collaboration around the challenge that may not take place if the challenge creator does not create the space for conversation.

Target Market Failures

To really create scaled impact, XPrize focuses heavily on designing prize challenges that aim to create markets where there previously were none. An example of this is oil slick clean-up. After the notorious Deepwater Horizon oil spill in 2010, XPrize issued an Oil Cleanup challenge based on research that showed that surface oil cleanup technology had not advanced in terms of efficiency for some 30 years, resulting in a new skimmer technology four times as efficient as old technologies, unsurprisingly eliciting great interest from oil companies the world over.

Have Clear (Quantitative) Requirements

To best demonstrate effectiveness of designed solutions, it is necessary to ensure that results are measurable. This measurability of results should be embedded in the design of a challenge, and also make the job of fund managers easier when it comes to assessing competing solutions. “Develop improved weather forecasting model based on open data sets,” for example could read “Develop a weather forecasting model that predicts daily rainfall rates to within an accuracy of 95 percent in a 5-kilometer radius based on open data sources.”

Be Solution-Agnostic

By clearly defining a challenge and some key parameters of a solution without being overly prescriptive, providing pages and pages of technical requirements, funders of challenges allow for the organic growth and realization of ideas owned by challenge participants—those that will ultimately take their ideas forward into the marketplace.

Data Pitch consortium partner, The Open Data Institute has in-depth experience of developing and running challenges. Data Pitch will be led by their experience. Their key best practices, as defined in the publication ‘What Works in Open Data Challenges’ include:

- Setting clear objectives that reflect the primary interests of all core stakeholders. Our consultation described by this deliverable is a key part of this.

¹ As recommended by DAI Global, a company with 50 years experience in tackling social and economic development issues.

- Designing a bespoke challenge structure that reflects these primary objectives. Our twin tracks - exploratory and provider-driven - enable this.
- Committing to open design principles and being prepared to iterate or adjust plans. To achieve this we are in close collaboration with our data providers and encourage them, where appropriate, to test their data and challenge ideas in datathon settings, which are described below. Further, our reaching out to a wide variety of citizens and communities is very much in line with the principles of open.
- Optimising return on investment by running multiple challenges over an extended period. Our community of innovators, policy experts and entrepreneurs will leverage across challenges.

1.4. Tracks

In Call 1 we are focusing on challenges in three tracks: smart cities, including transport, manufacturing, retail, energy, environment and tourism, food and agriculture and health and wellbeing. This is for the following reasons:

- Impact: These are extremely high impact areas
- Accessibility and speed: We believe multiple datasets that are currently closed can be accessed in these tracks, which are not as subject to tight regulatory processes as other potential tracks
- Opportunity to leverage network effects: By focusing on tracks that are high priority for governments at local, regional, national and international levels, and that affect the continued functioning of society for all people, Data Pitch's work can engage with other efforts in these arenas to increase impact

In cities, there is intense correlation between various activities. Traffic is a key part of transport planning; it also plays an important part in environmental concerns, due to noise and air pollution, and also in Tourism. The category of Smart Cities, therefore, offers great opportunities for input from multiple directions into exploratory challenges.

More efficient land use is extremely important to ensure not only secure food provision but the livelihoods of farmers. There is also great interest around the ability to use satellite images to better understand flood risk and land use. Water management, as in cities, remains vital for the security of the future.

As population profiles change, managing and promoting health and well-being become increasingly important, especially across intersections with other facets of life.

1.5. The Intersection of Tracks and Challenges

Each Track will contain a number of Challenges pertaining to that particular sector. SMEs can apply through a particular track and address a specific challenge within it. The challenges will either be 'Exploratory', a more open challenge, or 'Provider-Driven', related to a specific challenge defined by our data providers (see section 3.7.2 below). The development of these challenges

forms the substance of this document.

Challenges can, but do not have to be linked to an organisation or set of organisations and/or to one or more data sets. Challenges can be cross-sector or sector-specific. The SMEs will be tasked with answering a particular challenge within their proposal and selecting it from the list within the application form.

Challenges will be divided into two groups: *exploratory* and *provider-driven*.

Exploratory challenges revolve around a general open problem; applicants will be asked to propose a solution to the problem using any combination of data available in the experimentation platform, as well as other datasets, open or closed, they might have access to. The main goal is to foster innovation and creative thinking applied to a less well-known problem, for example by identifying the relevant sub-problems and tackling them individually. Exploratory challenges will be associated by directional indicators of expected impact in order to help contestants position their solution against the expectations in the competition.

Provider-driven challenges are well-defined, focused problems specified by data providers. Both types of challenges will have with clear benchmarks and key performance indicators.

In some cases, we can also envision starting with an exploratory challenge in the first round of the call, to landscape the solution space, followed by more specific challenges with clear KPIs in the second round.

After the first round of the call, we will revisit the challenges to assess that they are still relevant, i.e., not displaced by other challenges or already covered in present or near-future plans and recent technology developments or changes in policies and regulations.

1.6. Developing Challenges

Traditional open data challenges often help demonstrate the power of data that has already been released. However, the way they are structured is generally unsuitable for effecting the release of new data. In Data Pitch, we work to define the challenges alongside the obtaining of the data. This means we may;

- Define a provider-driven question with a Data Provider, who then shares the data to enable a solution;
- Identify an important exploratory question from experts or the wider community and then source data to enable the solution;
- Receive data, and then consult with stakeholders to identify appropriate concepts for the exploration of value;
- Receive data from Data Providers, with whom we then work to define a provider-driven question based on their insights and our knowledge of high impact areas.

This in turn means that in developing the challenges there is a requirement to move from the unknown to the concrete. Not all challenges are in the same stage of development at the same time. Our participation in third party datathons (see section 2.4) will offer an opportunity for the testing and refinement of challenges in practice in order to ensure they offer the clearest route to high impact results.

1.7. The Consultation

The challenges will be developed during an extensive process of consultation across sectors and stakeholders. In this way we will identify and develop high impact challenges that meet economic and social needs. This process has already begun and will continue until the second Call for Applicants is launched.

Our consultation stakeholders fall approximately into three categories, for whom we have developed approaches as outlined in section 3.

1.8.1. Data Providers

During discussions for the provision of data sets to the Data Pitch programme, we will assist data providers in developing and defining key challenges within the appropriate tracks.

1.8.2. Industry Experts

We are seeking ideas from experts via our consultation, attending events and our Advisory Board.

1.8.3. Citizens and Communities

An important stakeholder in the creation of questions for challenges is the public, whether individuals, in scientific communities or innovation networks. This enables us to make sure we prioritise problems that matter to a variety of communities and demographics.

2. Background

In this section we introduce some background concepts, including open innovation, incubators and datathons.

2.1. Open Innovation

Open innovation, whereby organisations incorporate external ideas into their own products and services, is particularly valuable to large organisations for whom it can be difficult to remain agile and responsive to technological change. Furthermore, open innovation provides an inclusive framework through which organisations can operate in unison to tackle pressing issues in society and particular challenges faced in industry. Originally described by Chesbrough (2003), open innovation describes the process through which firms capitalise on external ideas, and create additional value for certain markets. Open innovation as a process can be performed in one of three approaches: outside-in, inside-out and coupled (Gassman & Enkel, 2004). The outside-in approach is defined as the method through which organisations leverage external innovations to increase their internal Research & Development (R&D) efforts and increase the value of their offering. Alternatively, the inside-out approach relates to the sharing of assets, selling Intellectual Property (IP), or sharing technology to facilitate the innovation process. The coupled approach combines the two former approaches into a single approach that utilises the methods of each for a more robust model of open innovation.

2.2 Open Innovation in Data Pitch

Data Pitch focuses on utilising open innovation for data value chains. In Data Pitch, the coupled approach to open innovation is used, as the project leverages the sharing of data assets by the data providers and provides an opportunity for data providers to share the usage of their hosting services for SMEs to gain access to the data through Application Programming Interfaces (APIs). Additionally, the coupled approach used by Data Pitch facilitates the submission of external innovation ideas, aiming to address the challenges faced by other organisations. Moreover, a crowdsourcing model is leveraged to source some of the challenges that the SMEs will need to address. Data Pitch not only enables the safe and effective sharing of data but also ideas, knowledge, and technologies between data owners from public and private sectors, open innovation/business incubation/acceleration experts, startups, and SMEs. This exchange gives rise to true innovation, and through the provision of a wide range of technical and non-technical services alongside commercially valuable data, early-stage entrepreneurs will be given the opportunity to focus on what matters most - nurturing their idea, testing it through experiments and working together with the data owners to build a sustainable relationship - and after six-months of incubation, will stand stronger in a global market.

Data Pitch will offer the infrastructure to stimulate open innovation in the data economy by offering a collaborative space for data owners, startups, SMEs and end-users to come together to work on solving meaningful data challenges that will be defined through discussions with the data providers and other stakeholders. Using our experience from ODINE, Data Pitch will provide world-leading support regarding financial, legal, marketing, data and computing resources. Ultimately, we aim to ensure that winning entries are able to go on and make their solutions commercially viable companies and create further value from Big Data. This is in line with recommendations of the 2015 Open Innovation 2.0 conference, which emphasised the need for extensive case studies and best practice sharing as the only way to move forward towards a European Innovation Ecosystem augmenting the Digital Single Market. As the global digital economy continues to expand, other regions/groups are seeking to develop similar Big Data capacity.

Open innovation literature recommends the use of SMART criteria (specific, measurable,

attainable, relevant, and time-bound) to make sure the problem scope and the range of solutions expected are clearly defined. While defining the challenges that will drive the Data Innovation Lab, we take into consideration the available information on infrastructure and industrial strategies of the data providers to guarantee that services/products issued from our incubation programme satisfy a need of the industry. Both the exploratory, but most likely the performance-driven challenges can be defined in the context of a particular enterprise IT system or Big Data software suite in order to make sure that the solutions will be easier to assess and deploy in industrial processes and to foster collaborations between data providers and startups and SMEs funded by Data Pitch.

2.3 Data incubators

Data Pitch offers through its incubator a wide array of support measures to startups and SMEs that presented the most promising and sustainable solutions to the challenges identified by the process defined in this document. The incubator will support the development of emerging data-driven products, services, and business ideas around the challenges into successful commercial or social enterprises.

The programme is set up to help micro-through-to-SME businesses working with data to meet the challenges. The most promising data innovators that graduate from the incubator will be further assisted to build a successful business that addresses high impact challenges by connecting them with potential partners, advisors, business angels, and investors. To achieve this objective, we are liaising with accelerator programs run by Beta-i (currently 5 accelerator programs) as well as with investors, regional and international startup communities, and business schools, including SeedCamp, TechCity and the Founders Institute.

2.4 Datathons

In addition to the defining of the tracks and challenges, the challenges specifically will be tested for their popularity and value during the course of a series of datathons. Data Pitch will leverage the reach and attendance of existing datathons and hackathons, such as Pixels.camp and Hack Zurich, across the EU, by partnering with them.

The datathons will:

- Enable Data Pitch to test the challenges in practice
- Gain an understanding of the most popular challenges and the reasons why
- Answer any queries relating to the formal call application process

3. The Consultation Approach

This section will outline our consultation on high impact challenges and, the process we will be following. The result will contribute to the Call as outlined in section 3.8.

3.1. Core Activity – High Impact Challenge Online Consultation

The aim of the Consultation is to engage stakeholders and seek their views on three key issues. This will enable both crowdsourcing of ideas from interested parties such as startups alongside input from industry and data experts. The key areas are:

1. What they think should be achieved by data sharing;
2. What they believe the high impact challenges are
3. How they would like to be involved in Data Pitch

3.2. Consultation Strategy

Our consultation strategy follows the process below:

1. Map Stakeholders - these have been described above. We have also sought to know more about these by including demographic questions in the consultation.
2. Determine Methods and Tools – Initial desk research was conducted to identify high impact challenges from other projects globally. This was then developed and extended using external expertise to create a full set of consultation questions, including demographic information.
3. Create Consultation Webpage - This is done in line with the strategy and practices described in *Data Pitch Deliverable 6.1, Online presence and Marketing Tools*
4. Announce and Communicate – The consultation is widely publicised using social media, traditional media and media partnerships.
5. Run Consultation – The consultation will be available from May 2 – June 4 2017.
6. Analyse Content – The content will be analysed during a workshop with invited experts from across Europe in June. The workshop process is described in greater detail below.
7. Provide Synopsis of Consultation Results – The results will inform the Call published for Round 1 of Data Pitch in July 2017.

3.3. Compliance

Our consultation is compliant with the European Commission Minimum Standards

Clarity: Clear content of the consultation process. We have followed the European Commission format, and included supporting documentation.

Targeting: In order to ensure that all relevant parties have an opportunity to express their opinions, we are targeting the following groups in the following manner (non-exhaustive list)

Publication: The consultation will be located online at datapitch.eu, in line with single access point best practice.

Time Limits: The consultation will be open for the month of May. Given the highly targeted approach to consultation this will be sufficient time.

Feedback: The final output will be included in the Call for Applications published on July 1, 2017.

3.4. Consultation

The Consultation can be found at: <https://datapitch.eu/news/challenge-ideas-survey/>

3.5. Analysis

The results of the consultation will be continuously reviewed and any suggestions regarding the provision of data will be followed up appropriately as soon as they are received.

Once the consultation is completed, the input received requires thorough analysis, which will happen at a workshop.

The aim of the workshop will be to prioritize challenges. This will be based on the popularity of votes, on the viability of suggested challenges and the suitability of data, which will be determined by expert review. It is anticipated that some challenges may have to be held over until the second Call.

3.6. Workshop Process

The workshop will be attended by members of Data Pitch, invited experts and the facilitator who developed the consultation with Data Pitch. The workshop will follow 3 phases:

Phase 1:

Phase 1 will begin with a review of voting results and analysis by descriptive data such as stakeholder category (start up/SME, corporation, public authority, citizen) and distribution across Member States. It will also include a determination of most popular challenges by Track (Smart Cities, Nutrition and Agriculture. Health and Wellbeing).

This will be entered into a framework to show the high-impact challenges in each area and, if appropriate, to each stakeholder group.

Phase 2:

Phase 2 will involve the review of suggested challenges. These will be refined by uniqueness, completeness, level of expertise and practicability. They will also be compared to the provider-driven challenges to check for repetition between the two.

Those that are suitable and unique will then also be entered into the framework.

Phase 3:

The final phase will involve the cross-referencing of the most high-impact challenges with data availability to select final challenges. Data Pitch will not necessarily be able to address all these questions in the first round. In some cases, identifying the data provider or sharing the data may mean a longer process than we currently have. It is anticipated that, given the datasets, size of call and nature of the challenges, 10 secure, high impact challenges will be identified for the Call in July 2017.

3.7. Supporting Activities – Events and Data Provider Interviews

3.7.1. Expert Discussions

We are talking to a wide range of experts. Our experts are expert on the industry, which means they may work in the industry, consult to the industry, research the industry or regulate the industry. We are attending events that take the form of roundtables – specific topic discussions where all participants are given equal opportunity to participate – where attendees are people with a wide variety of technical, entrepreneurial, and data governance backgrounds, in order to capture further insights not included in the consultation.

We have piloted an informal roundtable in Wales, accessing exploratory challenges. A nominal group technique was used to ensure all attendees presented ideas. Twenty four participants attended from business, start ups, journalism, government and academia.

Event name	location	start date	end date	event URL
Future Cities	London	16/05/17		
London Tech Week	London	12/06/2017	16/06/2017	https://londontechweek.com/
EU startup conference	Berlin	13/04/2017	13/04/2017	http://www.eu-startups.com/eu-startups-conference-2017/
tictec	Florenz	25/04/2017	27/04/2014	https://www.mysociety.org/research/tictec-2017/
Datasummit	Berlin	28/04/2017	29/04/2017	
the next web	Amsterdam	18/05/2017	19/05/2017	http://thenextweb.com/conference/
pioneers	Vienna	01/06/2017	02/06/2017	https://pioneers.io/festival2017/

3.7.2. Interviews - Data Providers

As part of our negotiations with data providers, including organisations and EU projects, we have been elucidating key challenges that they wish to address. These have largely taken place during videoconferences. The full list will be available when the call is published. However, all exploratory challenges will be considered for inclusion, whether or not the organisation comes on board in the first round. During the negotiations that precede the incubation, we will define with data experimenters and providers the quantitative metrics that will determine the success of the experiment (e.g. 20% faster execution time; +5% in confidence of estimation compared to state of the 50 users for the new app; sales adding up to 50K€, etc).

3.8. Call Output

The proceeds of the Consultation will be a minimum of 10 clear challenges, related to datasets, that will be published in *Data Pitch Deliverable 4.1 Call Definition*.

References

- [1] Henry Chesborough (2003) *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, Cambridge, MA
- [2] Oliver Gassman & Ellen Enkel, (2004) *Towards a Theory of Open Innovatio: Three Core Process Archetypes*. Proceedings of R&D Management Conference, 06/06/04 Lisbon, Portugal