PROTEUS
Scalable online machine learning for predictive analytics and real-time interactive visualization
687691

D6.7 Report on project communication and engagement activities - v1

Lead Author: TREE
With contributions from: DFKI, AMIII, BU, LMDP, TRI
Reviewer: TRI

<table>
<thead>
<tr>
<th>Deliverable nature:</th>
<th>Report (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissemination level: (Confidentiality)</td>
<td>Public (PU)</td>
</tr>
<tr>
<td>Contractual delivery date:</td>
<td>May 2017</td>
</tr>
<tr>
<td>Actual delivery date:</td>
<td>May 2017</td>
</tr>
<tr>
<td>Version:</td>
<td>1.0</td>
</tr>
<tr>
<td>Total number of pages:</td>
<td>28</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Communication, clustering, engagement</td>
</tr>
</tbody>
</table>
Abstract

This document provides details about the communication and engagement activities and materials produced in the project during the first half of PROTEUS (month 1 to month 18), as part of the communication activities and engagement with relevant related initiatives.

It is important to highlight that the this deliverable is focusing on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3).

This document provides, initially, information about the general dissemination of the project, including the corporate image design, communication materials, and established communication channels and tools in use. And then a summary on clustering and engagement activities carried out so far in the course of the project.
Executive summary

This document provides details about the communication and engagement activities and materials produced in the project during the first half of PROTEUS (month 1 to month 18), as part of the communication activities and engagement with relevant related initiatives.

It is important to highlight that this deliverable is focusing on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3).

This document provides, initially, information about the general dissemination of the project, including the corporate image design, communication materials, and established communication channels and tools in use.

In terms of the corporate image, a logo was designed already at the proposal stage, and this has been kept during the project, guiding the overall corporate image for PROTEUS. On this basis, a number of templates were also designed to be used in communication activities; in particular these were implemented for both presentations and documents, assuring that specific rules related to communication (e.g., standard formats, stable brand identity, etc.) have been fulfilled. Following from this, a template for acknowledgements and disclaimers was also produced in order to assure that the H2020 dissemination guidelines in this regard have been followed.

About the communication materials, PROTEUS is currently running three different channels targeting different purposes. Firstly, for the general public and for general communication purposes, the project’s public website is available online (www.proteus-bigdata.com). Secondly, for a more interactive communication and presence on the social media, a Twitter account is being used (https://twitter.com/proteus_bigdata). And finally, since the PROTEUS results are being released as Open Source, a GitHub organisation is hosting the software results of the project in several repositories, constantly updated with the evolution of the project (PROTEUS GitHub at: https://github.com/PROTEUS-H2020).

<table>
<thead>
<tr>
<th>Channel</th>
<th>Main purpose</th>
<th>Target audience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Website</td>
<td>Share general information about the project</td>
<td>Members of the public, academics, industry, other projects</td>
</tr>
<tr>
<td>Social media</td>
<td>Two-way communication with interested stakeholders</td>
<td>Members of the public, academics, students</td>
</tr>
<tr>
<td>GitHub</td>
<td>Sharing of software results</td>
<td>Programmers and big data experts from industry and academia, students</td>
</tr>
</tbody>
</table>

In terms of clustering and engagement activities, during the first period of the project a number of actions have been taken. Several workshops and conference have been attended with the aim of engaging with related initiatives and projects exploring relevant areas. In particular, FiCloud, Big Data Spain, the Big Data Value Association or the EuroPro workshops were venues where PROTEUS was presented, interacted with other initiatives, and engaged with the relevant community and stakeholders.

In addition, other activities included public talks and presentations given by the project’s partners during the project meetings or in conjunction with close organisations (e.g. other academic bodies). The interaction with related initiatives has also been fostered during this period, especially with sister projects (those funded under the same call and topic; ICT-16-2015), but also other relevant initiatives and projects in the field of Big Data management. This document provides a summary for those activities. And finally, PROTEUS has also played in role in the co-organisation of meetups, in particular the Flink meetup series in Madrid (Spain) where initiated by Treelogic at the beginning of PROTEUS, in cooperation with other organisations; the series are currently ongoing and there are plans for its continuation during the second half.
Document Information

<table>
<thead>
<tr>
<th>IST Project Number</th>
<th>687691</th>
<th>Acronym</th>
<th>PROTEUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Title</td>
<td>Scalable online machine learning for predictive analytics and real-time interactive visualization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project URL</td>
<td><a href="http://www.proteus-bigdata.com">www.proteus-bigdata.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU Project Officer</td>
<td>Martina EYDNER</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>D6.7</th>
<th>Title</th>
<th>Report on project communication and engagement activities - v1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work Package Number</td>
<td>WP6</td>
<td>Title</td>
<td>Communication, exploitation and dissemination</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date of Delivery</th>
<th>Contractual M18</th>
<th>Actual M18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>version 1.0</td>
<td>final ■</td>
</tr>
<tr>
<td>Nature</td>
<td>report ■</td>
<td>demonstrator □ other □</td>
</tr>
<tr>
<td>Dissemination level</td>
<td>public ■</td>
<td>restricted □</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Authors (Partner)</th>
<th>TREELOGIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsible Author</td>
<td>Name: Marcos Sacristán</td>
</tr>
<tr>
<td></td>
<td>Partner: TREELOGIC</td>
</tr>
</tbody>
</table>

Abstract (for dissemination)
This document provides details about the communication and engagement activities and materials produced in the project during the first half of PROTEUS (month 1 to month 18), as part of the communication activities and engagement with relevant related initiatives.

It is important to highlight that the this deliverable is focusing on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3).

This document provides, initially, information about the general dissemination of the project, including the corporate image design, communication materials, and established communication channels and tools in use. And then a summary on clustering and engagement activities carried out so far in the course of the project.

Keywords
Communication, clustering, engagement

Version Log

<table>
<thead>
<tr>
<th>Issue Date</th>
<th>Rev. No.</th>
<th>Author</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-March-2017</td>
<td>0.1</td>
<td>Marcos Sacristán (TREE)</td>
<td>Initial ToC</td>
</tr>
<tr>
<td>28-March-2017</td>
<td>0.2</td>
<td>Marcos Sacristán (TREE)</td>
<td>First draft contents</td>
</tr>
<tr>
<td>07-April-2017</td>
<td>0.3</td>
<td>Marcos Sacristán (TREE)</td>
<td>Adding abstract, summary.</td>
</tr>
<tr>
<td>03-May-2017</td>
<td>0.4</td>
<td>Marcos Sacristán (TREE)</td>
<td>Version ready for review</td>
</tr>
<tr>
<td>09-May-2017</td>
<td>0.5</td>
<td>Rachel Finn (TRI), Marcos Sacristán (TREE)</td>
<td>Refinement and improvement after review</td>
</tr>
<tr>
<td>15-May-2017</td>
<td>1.0</td>
<td>Marcos Sacristán (TREE)</td>
<td>Final release</td>
</tr>
</tbody>
</table>
# Table of Contents

Executive summary .................................................................................................................. 3  
Document Information ........................................................................................................... 4  
Table of Contents .................................................................................................................... 5  
List of figures and list of tables ............................................................................................... 6  
Abbreviations ......................................................................................................................... 7  
1 Introduction ......................................................................................................................... 8  
2 Communication activities, channels and tools .................................................................... 9  
  2.1 Corporate image ............................................................................................................. 9  
      2.1.1 PROTEUS logo ...................................................................................................... 9  
      2.1.2 Corporate templates ............................................................................................... 10  
      2.1.3 Acknowledges & disclaimer .................................................................................... 12  
  2.2 Communication materials ............................................................................................... 12  
      2.2.1 Promotional video................................................................................................... 12  
      2.2.2 Corporate presentation ......................................................................................... 13  
      2.2.3 Promotional leaflet ............................................................................................... 14  
      2.2.4 Corporate poster .................................................................................................... 15  
  2.3 Communication channels and tools .............................................................................. 17  
      2.3.1 Public website ....................................................................................................... 17  
      2.3.2 Twitter account ..................................................................................................... 20  
      2.3.3 GitHub organisation .............................................................................................. 20  
3 Clustering and engagement activities .................................................................................. 23  
  3.1 Joint workshops and conferences .................................................................................... 23  
      3.1.1 Big R&I workshop at FiCloud 2016 ...................................................................... 23  
      3.1.2 Big Data Spain 2016 .............................................................................................. 23  
      3.1.3 BDVA Summit 2016 .............................................................................................. 24  
      3.1.4 EuroPro Workshop at the EDBT/ICDT 2017 ......................................................... 24  
  3.2 Other activities ................................................................................................................ 24  
      3.2.1 Public talks and presentations .............................................................................. 25  
      3.2.2 Interaction with sister projects and related initiatives ............................................. 25  
      3.2.3 Co-organisation of meetups ................................................................................... 26  
4 Conclusions ........................................................................................................................ 27  
References .............................................................................................................................. 28
List of figures and list of tables

Figure 1: PROTEUS logo ...................................................................................................................... 10
Figure 2: PROTEUS corporate document template (cover) ................................................................. 11
Figure 3: PROTEUS corporate presentation template (cover) ............................................................. 12
Figure 4: PROTEUS acknowledges and disclaimer ............................................................................ 12
Figure 5: PROTEUS video, snapshots .................................................................................................. 13
Figure 6: PROTEUS presentation, cover and index snapshots ............................................................... 14
Figure 7: PROTEUS leaflet, main page ................................................................................................. 15
Figure 8: PROTEUS poster .................................................................................................................. 16
Figure 9: PROTEUS website, overview ............................................................................................... 18
Figure 10: PROTEUS website, analytics (overview) .......................................................................... 19
Figure 11: PROTEUS website, analytics (most visited pages) .............................................................. 19
Figure 12: PROTEUS Twitter account, snapshot ............................................................................... 20
Figure 13: PROTEUS GitHub organisation, home page (screenshot) ...................................................... 21
Figure 14: PROTEUS GitHub organisation, analytics ......................................................................... 22

Table 1: Summary of interaction activities with PROTEUS’ sister projects (ICT-16-2015) ............... 25
Table 2: Summary of interaction activities with other projects or initiatives .................................. 26
Abbreviations

**BDVA**: Big Data Value Association

**PEICH**: the PROTEUS Elastic Cache

**SOLMA**: the PROTEUS library for Scalable Online Machine Learning and Data Mining Algorithms
1 Introduction

This document provides details about the communication and engagement activities and materials produced in the project during the first half of PROTEUS. It is produced under the course of the task T6.5.-Communication activities and engagement with relevant related initiatives.

It is important to highlight the context of this deliverables D6.7 with regard to other deliverables in WP6. In particular, it is to be noted that this deliverable is focusing on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3).

This document is structured as follow:

- Section 2 provides information about the general dissemination of the project, including the corporate image design, communication materials, and established communication channels and tools in use.
- Section 3, focuses on the clustering and engagement activities carried out so far in the course of the project, providing a summary of the actions taken during the first half of the project.
- Finally, Section 4 will close the document with some conclusions.
2 Communication activities, channels and tools

In this first section the document focuses on some core communication elements for the project: firstly, the corporate image established for PROTEUS (including logo, branding details, corporate colours, etc.); secondly, several communication materials (assets) that have been produced as promotional resources; and finally, the main communication channels and tools that have been established and are being used for communication purposes of the project at the time of this report.

PROTEUS aims to reach a wide range of stakeholder categories, from the specific scientific fields tackled by the project, to the end user community, and finally to players who will ultimately contribute to delivering results to the market. The consortium will target the following stakeholder categories:

- **End users**, including industrial players, in different application domains.
- **Industry** and technical experts, including large companies, SMEs.
- **Academia**, including universities, research institutes and think tanks involved in the fields of big data processing, streaming analytics, machine learning, distributed systems, and visualization techniques.
- **Related research projects**, including consortia working on projects related to big data, especially projects funded under the same call topic (ICT-16-2015) and the Big Data Value Association.
- **The media**, including local, national and international TV, radio and printed media.
- **The public**.

PROTEUS has enabled different materials, channels and tools for different purposes, aiming to address these specific categories. These will be detailed in the coming parts of this section.

In terms of activities, however, not all of the categories of stakeholders have been addressed in this document, considering the context of D6.7 with regard to other deliverables in WP6: this deliverable is focusing on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3). For this reason, out of the categories previously referred, this document is focusing on engaging with the media, the general public, and other related research project, while other deliverables are addressing end users and industry (D6.3), as well as academia, universities, research communities (D6.5).

2.1 Corporate image

At the beginning of the project, a corporate image was designed to reflect the branding elements to be considered for any communication from the project. The following key elements were agreed in terms of corporate design and branding rules for dissemination.

2.1.1 PROTEUS logo

The PROTEUS logo had already been designed at the proposal stage, and this has been kept as the final logo for the project. It consists on a purple degraded-to-white capital “P”, with a purple triangle on the top left corner, and the project acronym in capital grey letters below.

The logo design can be better appreciated in the figure:
2.1.2 Corporate templates

Following the styles, colors and font defined by the project logo, a whole corporate image was created. In particular and for practical reasons, several templates were produced for its usage in the project communications, including:

- Document template, for all project documents (e.g. particularly used for deliverables).
- Presentation template, for all project presentations.

The figures shows a snapshot of the cover for the presentation and document templates.
Figure 2: PROTEUS corporate document template (cover)
2.1.3 Acknowledges & disclaimer

For the avoidance of doubts, in order to fulfil the communication requirements of the Grant Agreement, default acknowledges and disclaimer were also established for all project communications. The figure shows an example of the acknowledges and disclaimer text, including the EU flag:

“This project is funded by the European Union (Horizon 2020, Ref: 687691)”

Figure 4: PROTEUS acknowledges and disclaimer

2.2 Communication materials

This section provides information and description about the communication materials that the project has produced to date, following the corporate image and providing the essential information required for the different communication activities carried out by the project (both digital on-line, and physical in face-to-face events, conferences and meetings).

2.2.1 Promotional video

A promotional video has been produced in the early stages of the project. The video lasts for 2 minutes, and provides the key information as an abstract of PROTEUS on a multimedia format. It has been conceived as a way to disseminate the basis of the project targeting the general public, the media, and the industry (end users) on multimedia channels, social media and the website, but it can also be used as a resource for any other purpose.

Just as examples, the figure shows a couple of snapshots taken from the video:
For sharing and dissemination purposes, the video has been uploaded to Youtube, currently available at: https://youtu.be/53_h79cMv50, and hosted under the Treelogic’s Youtube channel [1]. The video currently has 276 views, demonstrating a reasonable impact already in the big data and manufacturing space. In addition, the video is also embedded in the home page of the project’s website (see section 2.3.1 for further information about the project’s public website).

2.2.2 Corporate presentation

A corporate presentation of the project was also produced in the early stages in order to provide the basic information in the form of presentation for any public event attended to promote the project and create a stable, recognisable brand identity. The template allows users to specify and personalise their presentation for specific targets, and is presented in a format familiar to stakeholders from academia, research and industry.

As an example, a couple of snapshot of the presentation’s cover and index can be found in the figure below:
2.2.3 Promotional leaflet

A leaflet for PROTEUS was produced at the beginning of the project as part of the project’s communication materials. The aim is for the partners and the project as a whole to count on a factsheet to be distributed in a printed format when attending events or conferences, targeting mostly industry and research communities.

The PROTEUS leaflet is presented in a diptych form, with a cover page, a back page and the main contents inside, including the following sections:

- Cover page: including the project’s logo, name and acronym.
- Main contents:
  - Abstract.
  - Duration.
  - Contributions & impact.
  - A short project description.
- Back page: including the project’s logo, a list of the partners, and acknowledgements.

More detailed information on the leaflet was released as deliverable D6.10 PROTEUS factsheet leaflet. The leaflet itself can be found as a public document available at the PROTEUS website at: https://www.proteus-bigdata.com/app/download/9107917670/20150111_Proteus_Brochure.pdf. As an example and reference, the central part of the leaflet providing the main contents is shown in the figure:
2.2.4 Corporate poster

A corporate poster was also produced in the early stages to provide the basic information in the form of a poster for any public event attended to promote the project. The poster’s design is shown in the figure below:
**Figure 8: PROTEUS poster**

For the high quality pdf file, this is publicly available at the project’s website at: [https://www.proteus-bigdata.com/app/download/9235596270/Proteus_poster.pdf](https://www.proteus-bigdata.com/app/download/9235596270/Proteus_poster.pdf)
2.3 Communication channels and tools

While the previous sections provided information on the project’s communication material, this section provides information about the communication channels established by the PROTEUS project. In particular PROTEUS is using three online communication channels and tools for different purposes:

- A public website, targeting the general public and anyone generally interested in the project.
- A Twitter account, as an interactive social media channel for the more dynamic and daily activity.
- A GitHub organisation (including different repositories), to engage with the developers community and provide the Open Source contributions of the project.

The following sections provide further information about the purpose and use of each of these tools for PROTEUS communications.

2.3.1 Public website

This section provides a summary of the design and implementation of the PROTEUS project website. The actual website is available at http://www.proteus-bigdata.com/. The website is hosted by the project coordinator Treelogic, and includes the following sections:

- Home
  - Project fiche
- Project description
  - Hybrid processing engine
  - Scalable online machine learning
  - Real-time interactive visual analytics
  - Industrial validation
- Partners
- Results
  - Software
  - Public deliverables
  - Publications
  - Dissemination
- Blog (news and events)
- Contact.

For further information, deliverable D6.9 “PROTEUS project website” previously released in month3 of the project provided details about the project’s website. The figure provides an overview:
PROTEUS mission is to investigate and develop ready-to-use scalable online machine learning algorithms and real-time interactive visual analytics to deal with extremely large data sets and data streams.

The foundation for the PROTEUS advances is the use of an optimized implementation of combined batch and streaming processing and building around this later scalable real-time processes. The developed algorithms and techniques will form a library to be integrated into an enhanced version of Apache Flink, the EU Big Data platform. PROTEUS will contribute to the Big Data area by addressing fundamental challenges related to the scalability and responsiveness of analytics capabilities. The requirements are defined by a steelmaking industrial use case. The techniques developed in PROTEUS are however, general, flexible and portable to all data stream-based domains.

CONTRIBUTIONS & IMPACT

In particular, the project will go beyond the current state-of-art technology by making the following specific original contributions:

- New strategies for real-time hybrid computation, batch data and data streams.
- Real-time scalable machine learning for massive, high-velocity and complex data streams analytics.
- Real-time interactive visual analytics for Big Data.
- Implementation the new advances on top of Apache Flink.
- Real-world industrial validation of the technology developed.

The PROTEUS impact is manifested:

- strategically, by reducing the gap and dependency of the US technology, empowering the EU Big Data industry through the enrichment of the EU platform Apache Flink;
- economically, by fostering the development of new skills and new job positions and opportunities towards economic growth;
- industrially, by considering real-world requirements from industry and by validating the outcome on an operational setting, and
- scientifically, by developing original hybrid and streaming analytic architectures that enable scalable online machine learning strategies and advanced interactive visualisation techniques that are applicable for general data streams in other domains.

Figure 9: PROTEUS website, overview
The website was launched in December 2015 (already at month 1 of the project) and has been active since then. More than 3,000 visitors have been visiting the website, with more than 8,000 page views. Some analytics (access: May 9th-2017) are shown in the figure below:

![Figure 10: PROTEUS website, analytics (overview)](image)

<table>
<thead>
<tr>
<th>Your most-visited pages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Page</strong></td>
</tr>
<tr>
<td>PROTEUS /</td>
</tr>
<tr>
<td>PARTNERS /partners/</td>
</tr>
<tr>
<td>PROJECT DESCRIPTION /project-description/</td>
</tr>
<tr>
<td>Scalable Online Machine Learning /project-description/scalable-online-machine-learning/</td>
</tr>
<tr>
<td>Demo paper accepted in SIGMOD 2016 /2016/03/10/demo-paper-accepted-in-sigmod-2016/</td>
</tr>
<tr>
<td>Project fiche /proteus/project-fiche/</td>
</tr>
<tr>
<td>Real-time Interactive Visual Analytics /project-description/real-time-interactive-visual-analytics/</td>
</tr>
</tbody>
</table>

![Figure 11: PROTEUS website, analytics (most visited pages)](image)
2.3.2 Twitter account

Since February 2017, PROTEUS is also managing a Twitter account under the username @proteus_bigdata (https://twitter.com/proteus_bigdata).

This account has been conceived for sharing more dynamic and quickly changing information, as well as providing an interactive communication channel for engaging with relevant stakeholders (industry, researchers, academia, other projects and initiatives, etc.). This communication channel uses hashtags to ensure that our messages are amplified beyond our followers and show up on newsfeeds created and followed by users (e.g., #bigdata, #manufacturing, #industry40). This allows us to gain new followers and augment our network of interested parties. The project has already gathered 72 followers in the last three months, and this is set to increase as the release of results from the project enable PROTEUS to share more meaningful information.

![PROTEUS Twitter account, snapshot](image)

2.3.3 GitHub organisation

PROTEUS is a project releasing most of the results as public, including both document reports and prototypes (software). For this reason, and in order to further communicate and engage with the developers’ community, the project is managing a GitHub organisation named PROTEUS-H2020 (https://github.com/PROTEUS-H2020).
Figure 13: PROTEUS GitHub organisation, home page (screenshot)
The organisation is composed by a set of repositories where the different components of PROTEUS are being developed, released, and shared with the community. The most relevant ones include:

- The PROTEUS Engine: an overhauled version of Apache Flink supporting hybrid computation on batch datasets and data streams.
- The PROTEUS Language: a declarative language library for Scalable Data Analysis on top of Apache Flink.
- PEICH: the PROTEUS Elastic Cache.
- The PROTEUS Incremental Analytics: a backend module that implements incremental version (~O(1) computational cost using approximations) of the most common analytics operations.
- SOLMA: the PROTEUS library for Scalable Online Machine Learning and Data Mining Algorithms adapted to the data analytics platform Apache Flink.
- PROTEIC.JS: An HTML5 and CSS3 charts library, tailored for batch and streaming data visualisations.

In addition to these, some other utilities are also being released in the GitHub as the project advances, including test examples, demo setup configurations on a containers-based architecture (i.e. Docker), streaming dataset simulators for experimentation and testing, among others.

The figure below provides some analytics of the usage and activity within the PROTEUS GitHub organisation on the different repositories (access: May 9th, 2017):

![PROTEUS GitHub organisation, analytics](image)

Figure 14: PROTEUS GitHub organisation, analytics
3 Clustering and engagement activities

This section provides information on the activities carried out by PROTEUS aiming to link, interact, cluster and engage from a broad perspective, including other projects or activities, developers and technical communities, and the general public.

3.1 Joint workshops and conferences

In this section we present a list of joint workshops and clustering events attended or organised during the project, focusing on networking and engagement with other projects and initiatives (beyond the purely scientific dissemination and publications).

3.1.1 Big R&I workshop at FiCloud 2016

The 3rd International Symposium on Big Data Research and Innovation (BigR&I-2016) was organised as part of the FiCloud 2016 conference [2] (The IEEE 4th International Conference on Future Internet of Things and Cloud).

PROTEUS was participating in this workshop with a publication “An incremental approach for real-time Big Data visual analytics” [3], jointly co-authored by TREE and BU, Ignacio García, Rubén Casado (TREE) and (BU).

The workshop session enabled PROTEUS to interact with other existing initiatives and research in the field. The following list provides an overview of the workshop presentations and discussion:

- “Rule- and Cost-Based Optimization of OLAP Workloads on Distributed RDBMS with Column-Oriented Storage Function”. Takamitau Shioi and Kenji Hatano.
- “Infrastructure-Aware Functional Testing of MapReduce programs”. Jesús Morán, Bibiano Rivas, Claudio De La Riva, Javier Tuya, Ismael Caballero and Manuel Serrano.
- “An incremental approach for real-time Big Data visual analytics” Ignacio García Fernández, Rubén Casado and Hamid Bouchachia.

3.1.2 Big Data Spain 2016

The Big Data Spain [4] international conference is a business-oriented conference taking place yearly in Madrid around Big Data technology. 2016 edition brought together more than 1,000 attendees, and a successful program with more than 50 talks and workshops.

PROTEUS was presented in Big Data Spain 2016 twofold, presenting both a talk and a poster session for networking:

- TALK: “Advanced data science algorithms applied to scalable stream processing”, by David Piris and Ignacio García (TREE) [5].
- POSTER session: “A novel real-time processing and visualization big data solution focused on incremental processing”, by Ignacio García and Jorge Yagüe (TREE).
3.1.3 BDVA Summit 2016

The Big Data Value Association (BDVA) [6] is a fully self-financed non–for-profit organisation under Belgian law. The Big Data Value Association (BDVA) is the private counterpart to the EU Commission to implement the BDV PPP programme (Big Data Value PPP). BDVA has over 150 members all over Europe with a well-balanced composition of large and small and medium-sized industries as well as research organizations.

Even if PROTEUS was funded before the BDV PPP had been established and is not officially under the PPP umbrella, the links between PROTEUS and the BDVA are clear, and the association represents a natural target group for several purposes, including a research network, but also dissemination, communication and exploration activities.

PROTEUS was present in the BDVA Summit 2016 [7] (November 30th to December 2nd 2016, in Valencia - Spain) through project partners TRI and TREE who attended the event. In this occasion, the summit was used for networking and exploration of future cooperation and synergies, and even if PROTEUS was not presented, there are plans to exploit BDVA events in the future for a more active participation of PROTEUS.

3.1.4 EuroPro Workshop at the EDBT/ICDT 2017

The EuroPro workshop (1st International Workshop on Big Data Management in European Projects [8]) was organised as part of the EDBT/ICDT 2017 Joint Conference [9].

Together with other 7 EU-funded projects in the field of Big Data Management, PROTEUS participated to this clustering workshop with a presentation and further discussion searching for synergies, represented by Bonaventura del Monte (DFKI).

The list of participants and presentations follow:

- Christos Doulkeridis, Nikos Pelekis, Yannis Theodoridis, and George Vouros. “Big Data Management and Analytics for Mobility Forecasting in datAcron”.
- Philipp Grulich, Tilmann Rabl, Volker Markl, Csaba Sidló, and Andras Benczur. “STREAMLINE - Streamlined Analysis of Data at Rest and Data in Motion”.
- Bonaventura Del Monte, Jeyhun Karimov, Alireza Rezaei Mahdiraji, Tilmann Rabl, and Volker Markl. “PROTEUS: Scalable Online Machine Learning for Predictive Analytics and Real-Time Interactive Visualization”.
- Edwin Morley-Fletcher. “MHMD: My Health, My Data”.
- Claudio A. Ardagna, Paolo Ceravolo, Ernesto Damiani, and Marcello Leida. “Scouting Big Data Campaigns using TOREADOR Labs”.

3.2 Other activities

This sections summarises a list of other engagement, general communication, and clustering activities carried out by PROTEUS during the period. It includes public talks and presentations, interaction with related projects and initiatives, or meetups, among other activities.
3.2.1 Public talks and presentations

In addition to the workshops and conferences attended as listed in the previous section, PROTEUS has also organised smaller scale presentations and public talks linked to the project’s events (e.g. project’s meetings). The following lists some talks organised by PROTEUS:

- Public talk at the University of Oviedo during the kickoff meeting (January 22nd, 2016). “Introduction to Apache Flink”, by Asterios Katsifodimos (DFKI) and Rubén Casado (TREE).
- Public talk at the University of Oviedo by Ignacio García and Jorge Yagüe (TREE). “ProteicJS - innovando en la visualización de datos a través del software libre” [10] (in English: ProteicJS – innovating in data visualisation with Open Source software). October 14th, 2016.
- Public talk at the University of Bournemouth during a consortium meeting (October 20th, 2016). A guest speaker was invited together with a presentation of the visualisation library developed in PROTEUS, named PROTEIC.JS:
  - “The Cloudera Big Data ecosystem”, by Justin Hancock (CLOUDERA, guest speaker).
  - “Introducing Proteic.js, a novel visualisation library for Big Data”, by David Piris and Ignacio García (TREE).

3.2.2 Interaction with sister projects and related initiatives

PROTEUS has devoted special attention to interacting and clustering with the sister projects, those funded under the same call as PROTEUS (ICT-16-2015) researching on synergistic and complementary fields. The table presents a summary on the interaction and clustering activities performed with other projects and initiatives during the first half of the project:

<table>
<thead>
<tr>
<th>Project acronym</th>
<th>Project name</th>
<th>Website</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaVeL</td>
<td>Variety, Veracity, VaLue: Handling the Multiplicity of Urban Sensors Cloud-LSVA Cloud Large Scale Video Analysis</td>
<td><a href="http://www.vavel-project.eu/">http://www.vavel-project.eu/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td>SETA</td>
<td>An open, sustainable, ubiquitous data and service ecosystem for efficient, effective, safe, resilient mobility in metropolitan areas</td>
<td><a href="http://setamobility.eu/">http://setamobility.eu/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td>HOBBIT</td>
<td>Holistic Benchmarking of Big Linked Data</td>
<td><a href="https://project-hobbit.eu/">https://project-hobbit.eu/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td>STREAMLINE</td>
<td>Streamline</td>
<td><a href="https://streamline.sics.se/">https://streamline.sics.se/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Joint workshop (EuroPro, March 21st, 2017).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Bilateral f2f meeting (Berlin, 26th April 2017)</td>
</tr>
<tr>
<td>datACRON</td>
<td>Big Data Analytics for Time Critical Mobility Forecasting</td>
<td><a href="http://www.datacron-project.eu/">http://www.datacron-project.eu/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Joint workshop (EuroPro, March 21st, 2017).</td>
</tr>
<tr>
<td>SUMMA</td>
<td>Scalable Understanding of Multilingual Media</td>
<td><a href="http://summa-project.eu/">http://summa-project.eu/</a></td>
<td>• Social media (twitter) contact.</td>
</tr>
<tr>
<td>TOREADOR</td>
<td>TrustwOrthy model-awaRE Analytics Data platfORm</td>
<td><a href="http://www.toreador-project.eu/">http://www.toreador-project.eu/</a></td>
<td>• Email discussions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Joint workshop (EuroPro, March 21st, 2017)</td>
</tr>
</tbody>
</table>
PROTEUS

Deliverable D6.7

SEE.4C SpatiotEmporal ForEcasting: Coopetition to meet Current Cross-modal Challenges http://see4c.eu/ • Email discussions.

Cloud-LSVA Cloud Large Scale Video Analysis http://cloud-lsva.eu/ • Email discussions.

Besides sister projects, PROTEUS have interacted with additional projects or initiatives in the field of Big Data research during the period. The table provides a summary of these activities:

Table 2: Summary of interaction activities with other projects or initiatives

<table>
<thead>
<tr>
<th>Project acronym</th>
<th>Project name</th>
<th>Website</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDVA</td>
<td>Big Data Value Association</td>
<td><a href="http://www.bdva.eu/">http://www.bdva.eu/</a></td>
<td>• TRI and TREE attended the BDVA Summit (November 30th 2016, Valencia)</td>
</tr>
<tr>
<td>My Health My Data</td>
<td>A new paradigm in healthcare data privacy and security</td>
<td><a href="http://www.myhealthmydata.eu/">http://www.myhealthmydata.eu/</a></td>
<td>• Joint workshop (EuroPro, March 21st 2017, Venice - Italy)</td>
</tr>
<tr>
<td>Big Data Europe</td>
<td>Empowering communities with data technologies</td>
<td><a href="https://www.big-data-europe.eu/">https://www.big-data-europe.eu/</a></td>
<td>• Joint workshop (EuroPro, March 21st 2017, Venice - Italy)</td>
</tr>
<tr>
<td>SUPERSEDE</td>
<td>SUpporting evolution and adaptation of PERsonalized Software by Exploiting contextual Data and End-user feedback</td>
<td><a href="https://www.supersede.eu/">https://www.supersede.eu/</a></td>
<td>• Joint workshop (EuroPro, March 21st 2017, Venice - Italy)</td>
</tr>
</tbody>
</table>

3.2.3 Co-organisation of meetups

PROTEUS has been involved as a co-organiser of the Apache Flink Meetup series in Madrid (Spain) [11] since the first edition. At the time of this report, three meetups have been organised in Madrid by Rubén Casado (Head of Big Data at TREE and technical manager of PROTEUS) and David Piris (also member of the PROTEUS team):

• 1st Flink Meetup in Madrid, 25th February 2016, Madrid (Spain). “Introduction to Apache Flink” [12]. Presentations by Rubén Casado (TREE) and Dr. Fabian Hueske, PCM Apache Flink and co-founder of DataArtisans.

• 2nd Flink Meetup in Madrid, 20th May 2016, Madrid (Spain). “Workshop Apache Flink” [13], a more practical meetup with presentations by Rubén Casado and David Piris (TREE).

• 3rd Flink Meetup Madrid, 15th December 2016, Madrid (Spain). “Apache Flink en acción” (Apache Flink in action) [14], with presentations by Rubén Casado (TREE) and Zhenhao Li (Accenture Netherlands).

There are plan to continue the meetups series in the future, and PROTEUS as a project will keep involved.
4 Conclusions

In this document, details have been provided about the communication and engagement activities and materials produced in the project during the first half of PROTEUS (month 1 to month 18), as part of the communication activities and engagement with relevant related initiatives. It is important to highlight that this deliverable has focused on the general communication activities (e.g. general public) as well as on engagement and clustering activities with other related projects and initiatives, while other deliverables are more focused on the scientific dissemination activities (i.e. see deliverable D6.5) and the technology transfer and community building activities (i.e. see deliverable D6.3).

PROTEUS has been active in communication activities, and is running at the time of this report three different channels targeting different purposes. Firstly, for the general public and for general communication purposes, a project’s public website is available online (www.proteus-bigdata.com), permanently being populated and updated with new information and contents (new, public results, events, etc.). Secondly, for a more interactive communication and presence on the social media, a Twitter account is being used (https://twitter.com/proteus_bigdata). And finally, since the PROTEUS results are being released as Open Source, a GitHub organisation is hosting the software results of the project in several repositories, constantly updated with the evolution of the project (PROTEUS GitHub at: https://github.com/PROTEUS-H2020).

In terms of clustering and engagement activities, during the first period of the project a number of actions have been taken. Several workshops and conference have been attended with the aim of engaging with related initiatives and projects exploring relevant areas. In particular, FiCloud, Big Data Spain, the Big Data Value Association or the EuroPro workshops were venues where PROTEUS was presented, interacted with other initiatives, and engaged with the relevant community and stakeholders.

Finally, other activities included public talks and presentations given by the project’s partners during the project meetings or in conjunction with close organisations (e.g. other academic bodies). The interaction with related initiatives has also been fostered during this period, especially with sister projects (those funded under the same call and topic, ICT-16-2015), but also other relevant initiatives and projects in the field of Big Data management. And in addition, PROTEUS has also played a role in the co-organisation of meetups, in particular the Flink meetup series in Madrid (Spain) initiated by Treelogic at the beginning of PROTEUS, in cooperation with other organisations; the series are currently ongoing and there are plans for its continuation during the second half.

As a final conclusion of this document, the consortium feels positive about the results and impacts achieved during the first period, in terms of general communication, clustering and engagement with other relevant initiatives. Several tools and channels have been established, and both the quantity and quality of actions undertaken are considered satisfactory. Having said that, these activities will be even strengthened in the second period, due to a higher level of maturity of the results enabling more intense and higher impact communications.
References

[1] Treelogic Youtube channel: https://www.youtube.com/channel/UC7YXVurU4sBwKaRTNhj124Ab