



<i>Project Acronym</i>	<i>SoBigData</i>
<i>Project Title</i>	<i>SoBigData Research Infrastructure Social Mining & Big Data Ecosystem</i>
<i>Project Number</i>	<i>654024</i>
<i>Deliverable Title</i>	<i>SoBigData e-infrastructure and VRE release</i>
<i>Deliverable No.</i>	<i>D10.7</i>
<i>Delivery Date</i>	<i>January 2019</i>
<i>Authors</i>	<i>Massimiliano Assante, Leonardo Candela, Roberto Cirillo, Luca Frosini, Lucio Lelii, Francesco Mangiacrapa, Pasquale Pagano</i>



DOCUMENT INFORMATION

PROJECT	
Project Acronym	SoBigData
Project Title	SoBigData Research Infrastructure Social Mining & Big Data Ecosystem
Project Start	1st September 2015
Project Duration	48 months
Funding	H2020-INFRAIA-2014-2015
Grant Agreement No.	654024
DOCUMENT	
Deliverable No.	D10.7
Deliverable Title	SoBigData e-infrastructure and VRE release
Contractual Delivery Date	30 November 2018
Actual Delivery Date	23 January 2019
Author(s)	Massimiliano Assante (CNR), Leonardo Candela (CNR), Roberto Cirillo (CNR), Luca Frosini (CNR), Lucio Lelii (CNR), Francesco Mangiacrapa (CNR), Pasquale Pagano (CNR)
Editor(s)	Massimiliano Assante (CNR), Beatrice Rapisarda (CNR)
Reviewer(s)	Roberto Trasarti (CNR), Beatrice Rapisarda (CNR)
Contributor(s)	N/A
Work Package No.	WP10
Work Package Title	JRA3_SoBigData e-Infrastructure
Work Package Leader	CNR
Work Package Participants	USFD, UNIPI, Fraunhofer, UT, IMT LUCCA, LUH, KCL, SNS, AALTO UNIVERSITY, ETHZ Zurich
Dissemination	Public
Nature	Other
Version / Revision	1.0
Draft / Final	Final
Total No. Pages (including cover)	28
Keywords	Virtual Research Environment, Resource Catalogue,

DISCLAIMER

SoBigData (654024) is a Research and Innovation Action (RIA) funded by the European Commission under the Horizon 2020 research and innovation programme.

SoBigData proposes to create the Social Mining & Big Data Ecosystem: a research infrastructure (RI) providing an integrated ecosystem for ethic-sensitive scientific discoveries and advanced applications of social data mining on the various dimensions of social life, as recorded by “big data”. Building on several established national infrastructures, SoBigData will open up new research avenues in multiple research fields, including mathematics, ICT, and human, social and economic sciences, by enabling easy comparison, re-use and integration of state-of-the-art big social data, methods, and services, into new research.

This document contains information on SoBigData core activities, findings and outcomes and it may also contain contributions from distinguished experts who contribute as SoBigData Board members. Any reference to content in this document should clearly indicate the authors, source, organisation and publication date.

The document has been produced with the funding of the European Commission. The content of this publication is the sole responsibility of the SoBigData Consortium and its experts, and it cannot be considered to reflect the views of the European Commission. The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

The European Union (EU) was established in accordance with the Treaty on the European Union (Maastricht). There are currently 27 member states of the European Union. It is based on the European Communities and the member states’ cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice, and the Court of Auditors (<http://europa.eu.int/>).

Copyright © The SoBigData Consortium 2015. See <http://project.sobigdata.eu/> for details on the copyright holders.

For more information on the project, its partners and contributors please see <http://project.sobigdata.eu/>. You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed. You are permitted to copy this document in whole or in part into other documents if you attach the following reference to the copied elements: “Copyright © The SoBigData Consortium 2015.”

The information contained in this document represents the views of the SoBigData Consortium as of the date they are published. The SoBigData Consortium does not guarantee that any information contained herein is error-free, or up to date. THE SoBigData CONSORTIUM MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, BY PUBLISHING THIS DOCUMENT.

GLOSSARY

ABBREVIATION	DEFINITION
Transnational Access	<p>Transnational Access ensures free of charge access to the best European research infrastructures. Apart from the cost of research itself, travel, accommodation and subsistence costs maybe reimbursed by the projects budget.</p> <p>This opportunity is open to all European researchers and to some extent to researchers from non-EU countries.</p>
TNA	Transnational Access
Research Infrastructure	<p>Facilities, resources and services that are used by a research community to conduct research and foster innovation in their fields. Include: major scientific equipment (or sets of instruments), knowledge-based resources such as collections, archives and scientific data, e-infrastructures, such as data and computing systems and communication networks and any other tools that are essential to achieve excellence in research and innovation. They may be 'single-sited', 'virtual' and 'distributed'.</p>
RI	Research Infrastructure
VA	Virtual Access
Virtual Access	<p>Open and free access through communication networks to resources needed for research, without selecting the researchers to whom access is provided.</p>
Virtual Research Environment	<p>Innovative, web-based, community-oriented, comprehensive, flexible, and secure working environments conceived to serve the needs of modern science.</p>
VRE	Virtual Research Environment

TABLE OF CONTENT

DOCUMENT INFORMATION	2
DISCLAIMER	3
GLOSSARY	4
TABLE OF CONTENT	5
DELIVERABLE SUMMARY	6
EXECUTIVE SUMMARY	7
1 Introduction	8
2 SoBigData Portal software specification and release	10
2.1 SoBigData Portal Software specification	10
2.2 SoBigData Portal Software release	13
3 SoBigData Catalogue software specification and release	15
3.1 SoBigData Catalogue Software specification	15
3.2 SoBigData Catalogue Software release	17
4 SoBigData Virtual Research Environments Software Specification and Release	18
4.1 Exploratories VREs	18
4.2 Applications VREs	19
4.3 Project Internal VREs	21
4.4 SoBigData Lab VRE	22
4.5 E-Learning Area VRE	23
4.6 SoBigData Virtual Research Environments Software release	24
5 Conclusion	27
REFERENCES	28

DELIVERABLE SUMMARY

This deliverable describes the software that has been deployed to serve the needs of the SoBigData community, by delivering the platform and the VREs planned in “D10.4 SoBigData e-Infrastructure release plan 3” [7]. In particular, it reports on how such software has been exploited to make available the envisaged components, i.e. the SoBigData portal (and the underlying Virtual Organisation), the SoBigData Catalogue and the SoBigData Virtual Research Environments, together with the list and pointers to the software packages produced by the project and implementing such components, whose operation today constitutes the SoBigData e-infrastructure accessible from <http://sobigdata.d4science.org>.

EXECUTIVE SUMMARY

This deliverable describes the software deployed in the SoBigData e-Infrastructure to serve the needs of the SoBigData community, delivering the platform and the VREs planned in “D10.4 SoBigData e-Infrastructure release plan 3” [7]. It provides software specifications for the (i) SoBigData portal, available at <https://sobigdata.d4science.org/> which is representing the gateway for accessing and exploiting the SoBigData e-infrastructure and its services, (ii) the SoBigData Catalogue, available at <https://sobigdata.d4science.org/catalogue> which offers capabilities of searching and browsing the SoBigData e-Infrastructure published products, and (iii) the thirteen (13) Virtual Research Environments that have been deployed by the project at the time of writing, which have been classified following their offering type. Specifically, five (5) Exploratory VREs, three (3) Application VREs, two (2) Project Internal VREs, one Lab VRE and one Training VRE.

For each of the above-mentioned software specifications the deliverable also provides the reader with a comprehensive list and pointers to the software packages produced by the project and published on the GitHub platform.

1 INTRODUCTION

This document describes and delivers the platform and the VREs planned in D10.4 SoBigData e-Infrastructure release plan 3. The deliverable documents the software released as due by MS2, 3, 4 and 6 and shows how this software has been exploited to provide the SoBigData community with a set of facilities and Virtual Research Environments.

The enabling software initially described in D10.5 [1] has been improved during the period and it is now part of the gCube 4.13 release¹. It includes the following major subsystems:

- *Resource Catalogue*: it is a set of portlets and software libraries to discover, index and search all the datasets accessible through the D4Science infrastructure. It is based on CKAN technology (<http://ckan.org>);
- *Portal (or Gateway)*: it is a set of software libraries to realise a Liferay based web-portal customized to interface with the D4Science infrastructure and equipped with gCube portlets. Once installed and configured it will act as the “one stop shop” for the entire SoBigData e-Infrastructure. Through it users will have access to the resources and Virtual Research Environments created to serve the needs of the SoBigData community and scenarios;
- *Accounting Framework*: a set of services and portlets supporting the collection and consumption of resources usage metrics, i.e. it is a distributed system comprising an array of services automatically collecting per-resource usage metrics, integrated into the gCube SmartGears container that enables automatic accounting of user calls, plus a user interface for their visualisation and analysis;
- *VO Management Framework*: it is a set of basic services and portlets created and operated in the context of D4Science to serve the needs of SoBigData. This Virtual Organisation (VO) realizes the actual operational context for implementing and operating the SoBigData e-Infrastructure and its resources in autonomy with respect to the other communities and initiatives supported by D4Science;
- *VRE Management Framework*: it is a set of services and portlets supporting the creation and operation of VREs.

Figure 1 describes how the above-mentioned subsystem are linked together. It must be noted that all the software released are open source, EUPL licensed and available and published on

- Subversion Repository publicly accessible at <http://svn.research-infrastructures.eu/public/d4science/gcube/trunk/>)
- GitHub platform accessible at <https://github.com/gcube-team/gcube-releases/tree/v4.13.0>

¹ <https://github.com/gcube-team/gcube-releases/tree/v4.13.0>

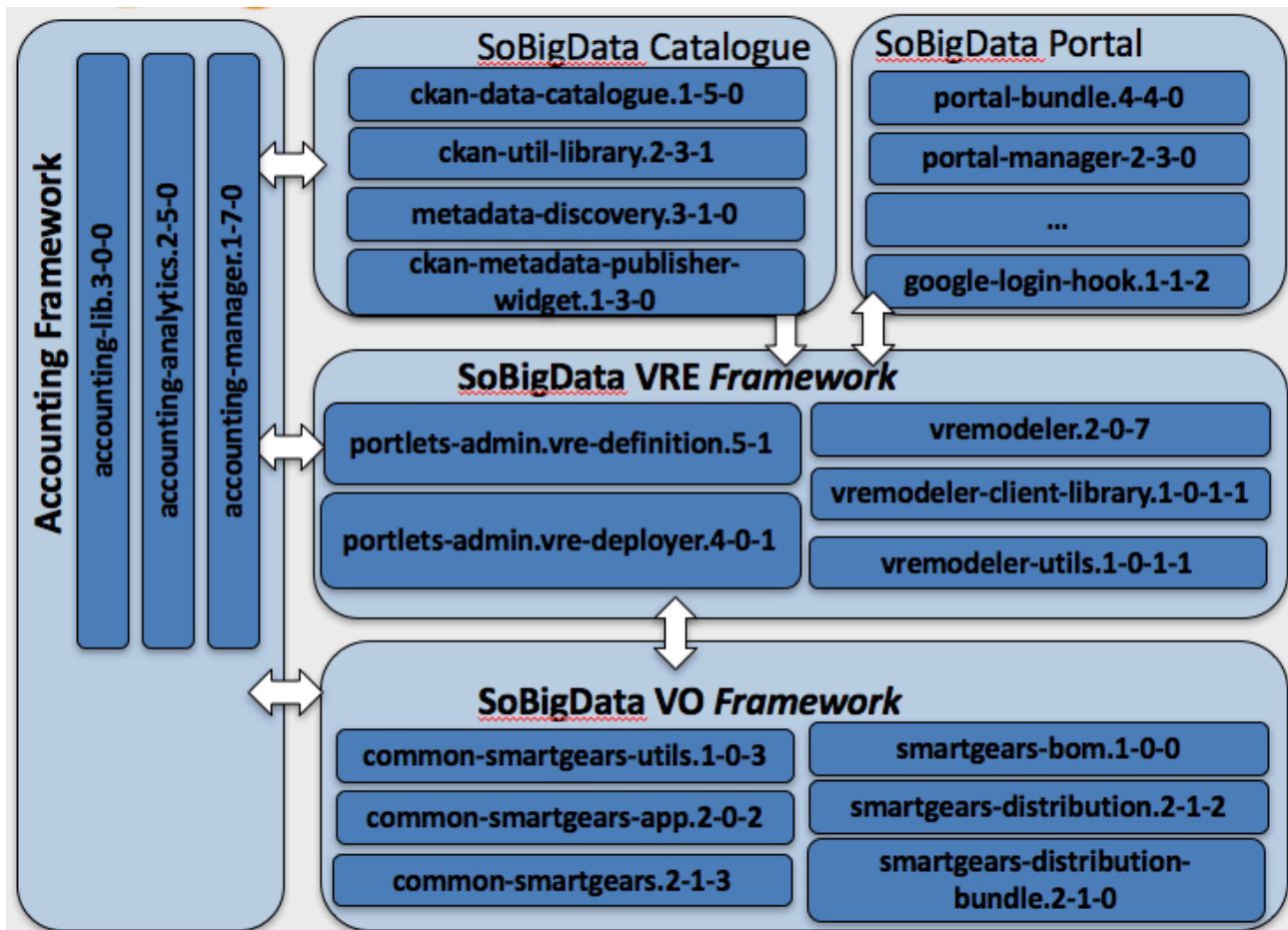


Figure 1. Software packages contributing by SoBigData and released in the gCube system

This software has been deployed and configured leading to three major services:

- the SoBigData Portal (cf. Sec. 2);
- the SoBigData Catalogue (cf. Sec. 2.2),
- the set of SoBigData Virtual Research Environments (cf. Sec. 3.2).

For each of the three major services above the document reports their software specification, including the software functionalities and a comprehensive list and pointers to the related software packages produced by the project and published on the GitHub platform.

2 SOBIGDATA PORTAL SOFTWARE SPECIFICATION AND RELEASE

This section delivers the SoBigData Portal software specification (cf. sec. 2.1) by describing how such software has been customised to meet the SoBigData community requirements and scenarios, and the SoBigData Portal software release (cf. sec. 2.2) providing the list and pointers to the software packages produced by the project and implementing the components described in the specification, whose operation today constitutes the SoBigData Portal of the SoBigData e-infrastructure, accessible from <http://sobigdata.d4science.org>.

2.1 SOBIGDATA PORTAL SOFTWARE SPECIFICATION

The SoBigData Portal is expected to be the end-user access point to the SoBigData services and Virtual Research Environments. A screenshot of the home page is in Figure 2.

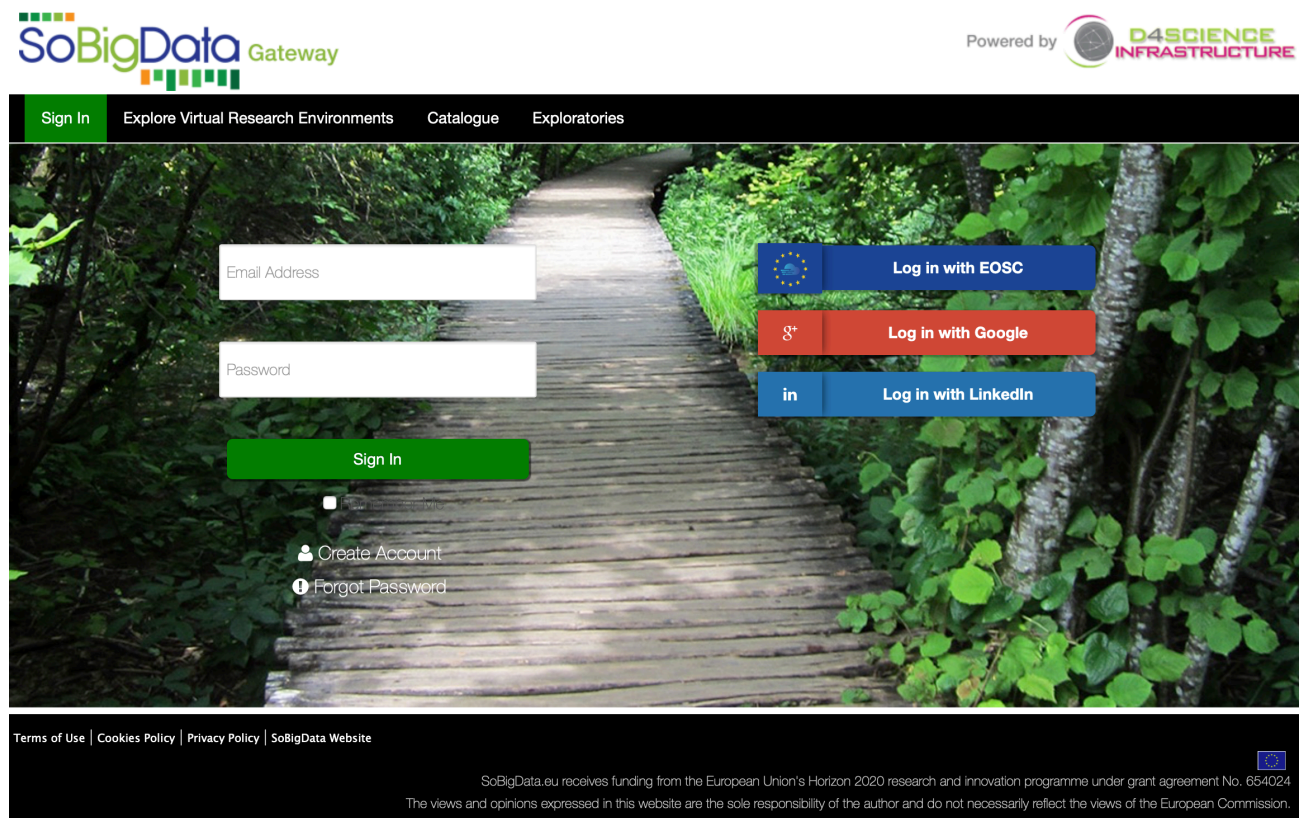


Figure 2. SoBigData Portal Home Page

Users are provided with login facilities and once logged in they are provided with a dashboard (Figure 3) realising a user-friendly working environment. The previous dashboard version has been heavily revised and redesigned following the SoBigData community feedback aiming at improving its usability. As shown in Figure 2, the dashboard provides users with an at-a-glance view of the whole SoBigData offering in one single place. The set of SoBigData Virtual Research Environments have been classified following the offering type, namely Exploratories, Applications and Training. A detailed explanation of these classes is given in Sec. 4 (SoBigData Virtual Research Environments) of this document.

Besides describing and providing access to all the VREs (Exploratories, Applications and Training classes) offering in the e-infrastructure, the dashboard provides users with:

- a) The SoBigData Resource Catalogue Widget (located in the top-centre of Figure 2.) providing possibility to search for catalogue content and statistics on the number of items, organisation, groups and types.
- b) A direct link to the SoBigData Lab VRE (beaker icon), allowing users to execute Methods and import their own in the e-Infrastructure by means of the Method Engine and the Method Importer software.
- c) A shared workspace (folder icon) for managing the objects shared in the VREs a user is member of as well as for managing his/her own objects organising them in folders;
- d) A direct link to the SoBigData Resource Catalogue (book icon);
- e) A communication area (Mail icon) for exchanging private messages with other users and co-workers.
- f) A Notification Widget (no. 20 next to the user's name) reporting the number of unread notifications and showing the most recent 3 when clicked as shown in Figure 4;
- g) A "Submit your Idea" widget, allowing users to suggest a description about the dataset/method they would like to have in the catalogue, that expands as an input form when clicked as shown in Figure 5;
- h) A "Upload your Method/Dataset" widget, allowing users to suggest a description about the dataset/method they would like to import in the catalogue, that expands as an input form when clicked as shown in Figure 5;
- i) A shared workspace Widget providing read only access to the objects shared in the VREs a user is member of as well as for managing his/her own objects organised in folders.

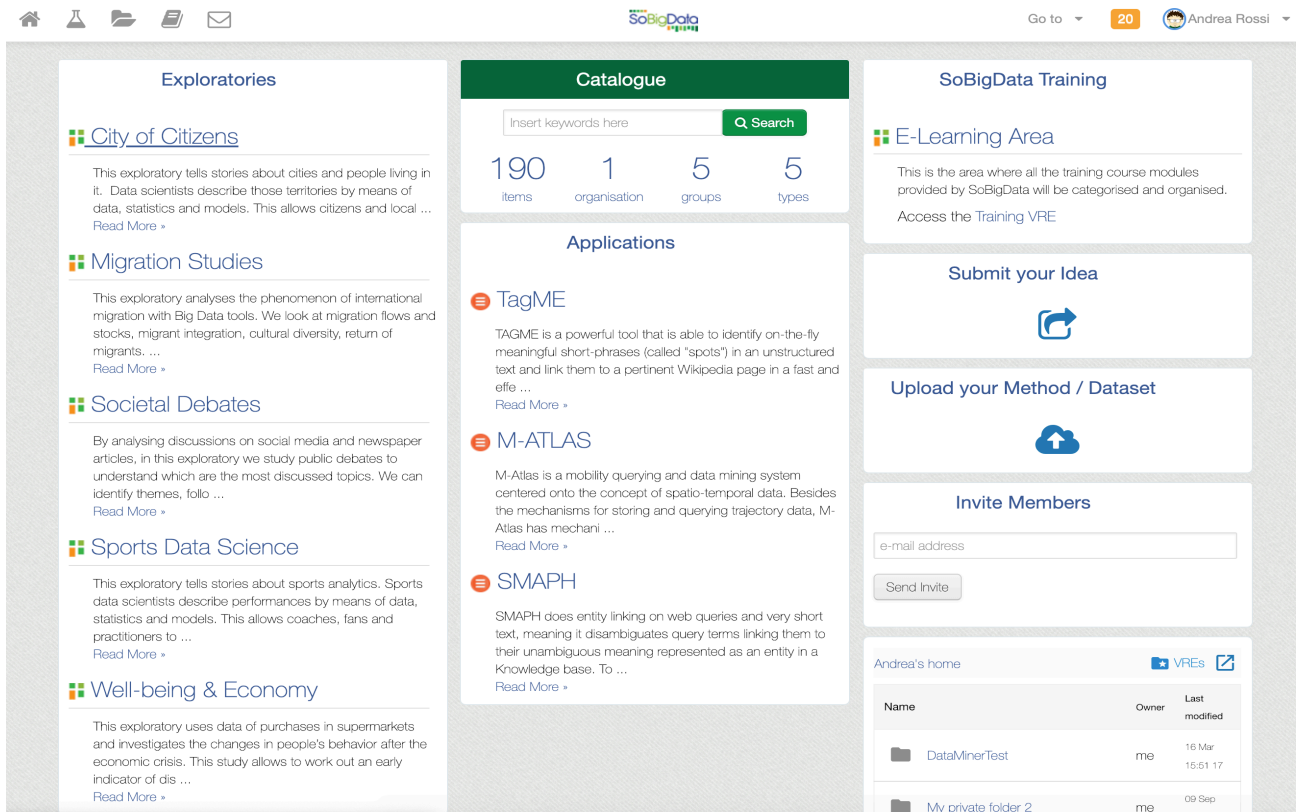


Figure 3. SoBigData Portal user dashboard

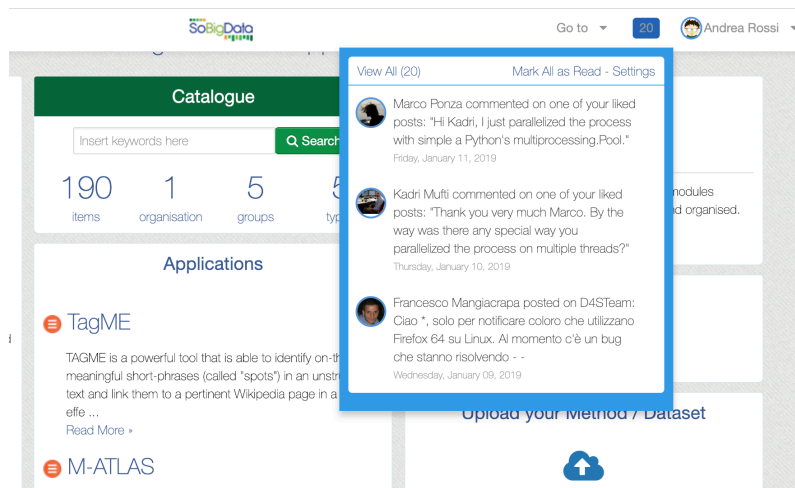


Figure 4. SoBigData Portal user dashboard notification widget

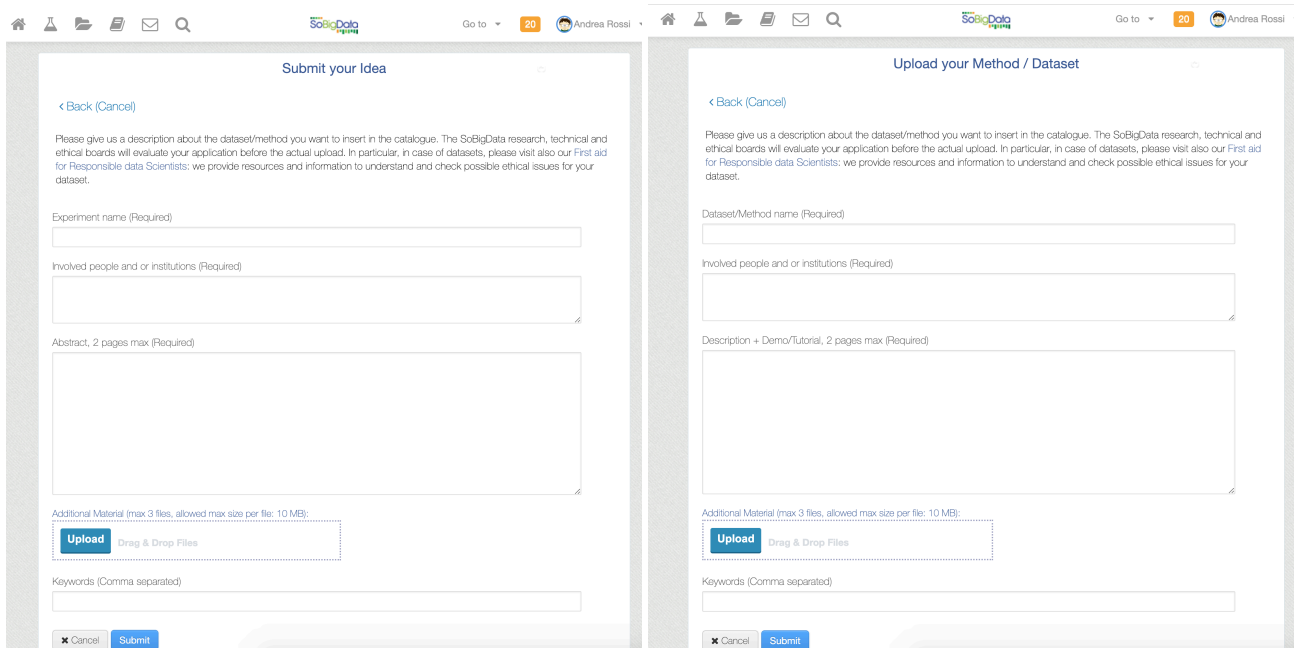


Figure 5. SoBigData Portal user dashboard Submit Idea and Upload Method Forms expanded

2.2 SOBIGDATA PORTAL SOFTWARE RELEASE

This section contains the list and pointers to the software packages produced by the project and implementing the components described in the specification above, whose operation today constitutes the SoBigData Portal of the SoBigData e-infrastructure, accessible from <http://sobigdata.d4science.org>.

All the software released are open source, EUPL licensed and available and published on

- Subversion Repository publicly accessible at <http://svn.research-infrastructures.eu/public/d4science/gcube/trunk/>
- GitHub platform accessible at <https://github.com/gcube-team/gcube-releases/tree/v4.13.0>

Subsystem	Description	Packages
SoBigData Portal	A Liferay based web-portal customized to interface with the D4Science infrastructure and equipped with gCube portlets. This portal acts as the “one stop shop” for the entire SoBigData e-Infrastructure. Through it, users have access to the resources and Virtual Research Environments created to serve the needs of the SoBigData community and	<p>Distribution</p> <ul style="list-style-type: none"> • distribution.portal--distribution.1-2-0 • maven-portal-bom.3-4-0 <p>Portal</p> <ul style="list-style-type: none"> • portal-bundle.4-6-0 • portal-manager-2-4-1 • portal.shibboleth-login-hook.1-1-0 • portal.linkedin-login-hook.1-3-0

	scenarios.	<ul style="list-style-type: none">• portal.google-login-hook.1-3-0• portal.login-hook.1-4-0• portal.notifications-common-library.1-4-0• portal.notifications-portlet.2-5-0• portal.social-mail-servlet.2-3-0• portal.social-networking-library.1-16-1• portal.social-networking-ws.2-1-1 <p>Portlets-user</p> <ul style="list-style-type: none">• portlets-user.catalogue-badge.1-0-0• portlets-user.workspace-widget-portlet.1-1-0• portlets-user.notifications.2-4-0• portlets-user.socialprofile.2-1-0• portlets-user.workspace-portlet.6-20-1• portlets-user.join-vre.3-6-1
--	------------	--

3 SOBIGDATA CATALOGUE SOFTWARE SPECIFICATION AND RELEASE

This section delivers the SoBigData Catalogue software specification (cf. sec. 3.1) by describing how such software has been customised to meet the SoBigData community requirements and scenarios, and the SoBigData Catalogue software release (cf. sec. 2.2) providing the list and pointers to the software packages produced by the project and implementing the components described in the specification, whose operation today constitutes the SoBigData Catalogue of the SoBigData e-infrastructure, accessible from <https://sobigdata.d4science.org/catalogue>.

3.1 SOBIGDATA CATALOGUE SOFTWARE SPECIFICATION

The SoBigData Catalogue's Home page Look & Feel is shown in the figure 6. The Groups and the (Resource) Types are easily accessible giving the end-users an overview of the catalogue content at a glance.

The screenshot shows the SoBigData Catalogue Home Page. At the top, there is a navigation bar with 'Sign In', 'Explore Virtual Research Environm...', 'Catalogue', and 'Exploratories'. Below this is a search bar and a statistics section showing 190 items, 1 organisation, 5 groups, and 5 types. The main content area is divided into 'Browse by Groups' and 'Browse by Types' sections, each with a grid of category cards. A 'More about Types' section provides detailed descriptions for Applications, Datasets, Deliverables, Experiments, and Methods. At the bottom, there are sections for 'Popular Formats' and 'Popular Tags'.

Figure 6. SoBigData Catalogue Home Page

Specifically, the Groups have been mapped with the SoBigData Exploratories (cf. sec. 4) while for each different Resource Type (Dataset, Method, Training Material, Application, Experiment) a dedicated metadata schema was created in order to better capture the specificity of the different published resources. The catalogue indeed represents a core service where all the resources contributing to form the SoBigData e-Infrastructure are registered thus to make it possible for clients to discover them and be informed on their characteristics for, e.g. properly using them. This catalogue serves both (a) human users willing to know the offering of the e-Infrastructure in terms of datasets and services / methods and (b) other services willing to dynamically discover resources to consume / interact with to deliver their services.

The catalogue is conceived to be nicely integrated in the overall SoBigData Portal (cf. Sec. 2) and has been populated by relying on a software, namely Product Publishing Form (see Figure 7) specifically conceived to provide its users with facilities for creating, publishing and modifying product records.

Figure 7. Product Publishing Form

Users can publish and/or modify items by providing reach characterisations of them aiming at maximising their exploitation and reuse in accordance to Open Science practices. For each typology, a community specific set of attributes characterising them have been identified and documented by dedicated Wiki pages². Dataset attributes include: general description attributes, e.g. title and creator(s); accessibility properties, e.g. VA or TNA; technical properties, e.g. size and format; legal and ethical attributes, e.g. whether containing personal data; intellectual properties, e.g. licences. Method attributes include: general description attributes, e.g. title and creator(s); accessibility properties, e.g. VA or TNA; technical properties, e.g. programming language; intellectual properties, e.g. licences.

² <https://wiki.gcube-system.org/gcube/SoBigData.eu: Metadata Profile for gCube Data Catalogue>

3.2 SOBIGDATA CATALOGUE SOFTWARE RELEASE

This section contains the list and pointers to the software packages produced by the project and implementing the components described in the specification above, whose operation today constitutes the SoBigData Catalogue of the SoBigData e-infrastructure, accessible from <http://sobigdata.d4science.org/catalogue>.

All the software released are open source, EUPL licensed and available and published on

- Subversion Repository publicly accessible at <http://svn.research-infrastructures.eu/public/d4science/gcube/trunk/>
- GitHub platform accessible at <https://github.com/gcube-team/gcube-releases/tree/v4.13.0>

Subsystem	Description	Packages
SoBigData Catalogue	<p>A facility to discover, index and search all the datasets accessible through the D4Science infrastructure. It is based on CKAN technology (http://ckan.org/).</p> <p>This catalogue is expected to serve both (a) human users willing to know the offering of the e-Infrastructure in terms of datasets and services / methods and (b) other services willing to dynamically discover resources to consume / interact with to deliver their services;</p>	<p>Data-catalogue</p> <ul style="list-style-type: none"> • catalogue-ws.1-3-0 • ckan-util-library.2-8-0 • metadata-discovery.3-3-0 <p>Portlets-user</p> <ul style="list-style-type: none"> • gcube-ckan-data-catalogue.1-6-0 <p>Data-access</p> <ul style="list-style-type: none"> • ckan-connector.1-1-2 <p>Portlets-widgets</p> <ul style="list-style-type: none"> • ckan-metadata-publisher-widget.1-4-0 • catalogue-sharing-widget.1-0-1

4 SOBIGDATA VIRTUAL RESEARCH ENVIRONMENTS SOFTWARE SPECIFICATION AND RELEASE

At the time of writing this document thirteen (13) Virtual Research Environments (VREs) have been created and are operational. These VREs have been classified following the offering type, namely Exploratories, Applications, Lab, Training and Project Internal:

- 5 Exploratories VREs: the list of Exploratories is inherit by the WP11 definition and are: Societal Debates, Well-being & Economic Performance, City of Citizens, Migration Studies and Sport Data Science. Each of this corresponds to a VRE;
- 3 Applications VREs: the list of Applications which are present in the Catalogue have been organised in 3 main VREs according to the type of services provided and are: TagME, SMAPH, M-Atlas;
- 2 Project Internal VREs: SoBigData.eu VRE and SoBigData.it conceived to provide the SoBigData project members with a VRE-based working environment useful for the communications and collaboration among project and initiative members;
- 1 Lab VRE: the SoBigDataLab VRE linked on the top of the portal as well as in the catalogue if the method is integrated in the computational engine provided by SoBigData and described in the deliverable D10.9;
- 1 Training VRE: the e-Learning_Area VRE hosts training materials developed within the SoBigData project.

4.1 EXPLORATORIES VRES

The five Exploratory VREs structure are easy to explore and better linked to the catalogue, in fact as shown in Figure 8, the right side we can find a general description and the news feed as well as the links to the different stories, e.g. Polarised Political Debates and Monitor Topics across Time and space. On the left side of the VRE we find the direct link to the catalogue and the list of the Applications linked related to the exploratory, e.g. TagMe.

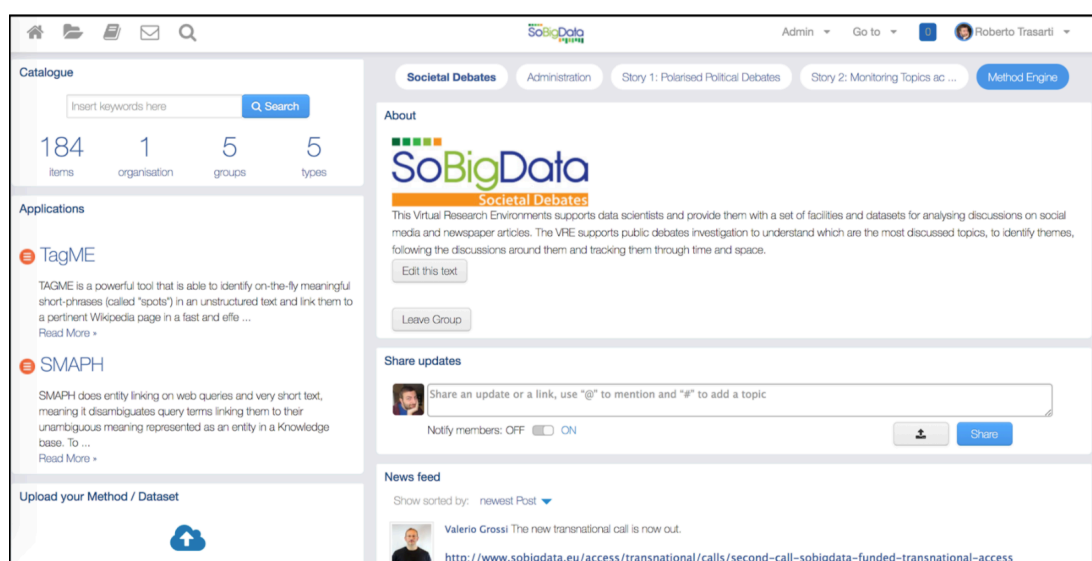


Figure 8. Example of exploratory VRE: Societal Debates.

The main functionality offered by the Exploratories VRE include:

- A dashboard (Fig. 8) for (i) collaborating with other VRE members by posting messages or being informed and reacting to (e.g. commenting, likes) posts, (ii) being informed on the most recent objects added to the VRE workspace, (iii) invite colleagues to participate, (iv) have information on the Exploratory including a description and the list of managers, (v) having some indicators on the user activity in the VRE, and (vi) a direct link to the catalogue and the list of the Applications linked related to the exploratory, e.g. TagMe;
- A set of Story pages related to the Exploratory containing the explanation about the research and the methods used in the exploratory;
- A Method Engine area (Fig. 9), for executing the available methods of interests, if any, directly on the SoBigData e-Infrastructure with the possibility to change the input parameters, thus in line with Open Science principles.

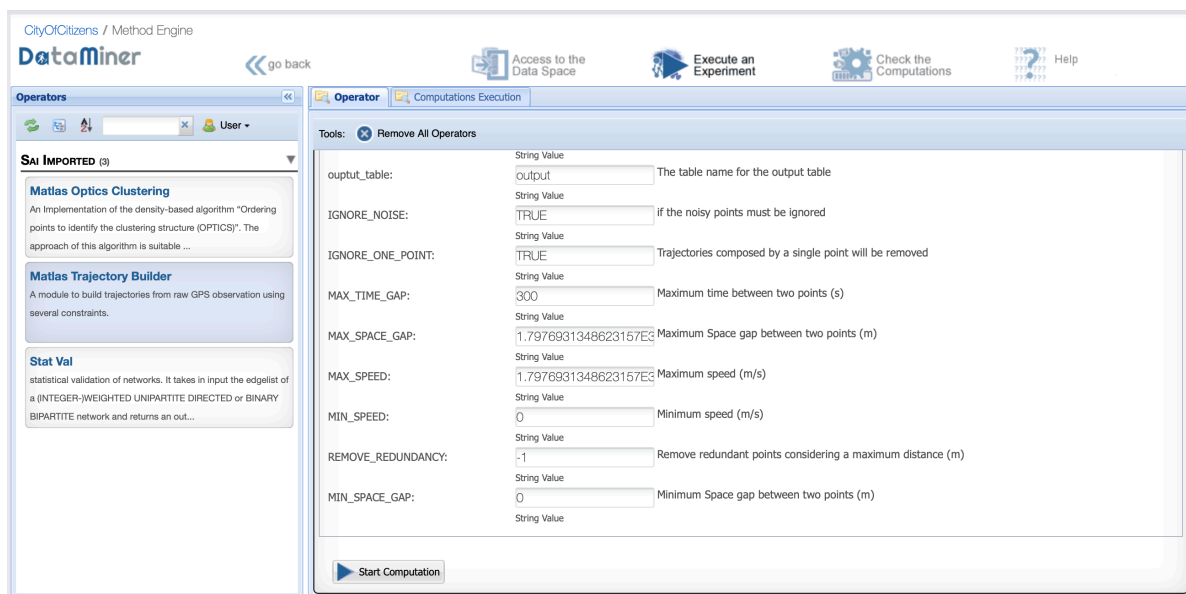


Figure 9. Example of exploratory VRE Method Engine area: City of Citizens.

4.2 APPLICATIONS VRES

These Virtual Research Environments are mainly conceived to provide their members with a VRE-based working environment for exploiting the Applications hosted by the SoBigData e-Infrastructure. It is expected that the majority of users will exploit the Applications VREs in a programmatic way thus the VREs are mainly exploited for managing membership (i.e. authorise users), enable users to acquire a proper token for using the service, and collect statistics. However, they are equipped with the standard set of collaboration-oriented facilities that can be exploited by its members for discussion about the application usage, asking advices and reporting errors as depicted in Figure 4.2 for the TagME VRE.

Figure 10. TagME VRE Home Page

The TagME VRE provides users with access to 3 different applications, described in D10.9 [6]. (TagMe, WAT and SWAT), all under the entity linking “hat”, while the SMAPH VRE and M-Atlas VRE provide users with access to the respective single application also described in D10.9 [6].

The main functionality offered by these Applications VRE include:

- A dashboard (Fig. 10) for (i) collaborating with other VRE members by posting messages or being informed and reacting to (e.g. commenting, favouring) co-workers’ posts, (ii) being informed on the most recent objects added to the VRE workspace, (iii) invite colleagues to participate, (iv) have information on the VRE including a description and the list of managers, (v) having some indicators on the user activity in the VRE, and (vi) acquire an authorization token for exploiting the services offered by the VRE in a programmatic way;
- A shared area in the workspace, for making available objects of interests, e.g. project deliverables, presentations, working notes;
- A Members area, for enabling each VRE member to be informed on the rest of VRE members and acquire details for contacting them;
- A user management area, to enable authorised users (i.e. VRE Managers) to manage other users using or willing to access the VRE. VRE Managers can (i) authorise users in accessing the VRE and its services, (ii) assign or withdraw roles to users, (iii) remove users, and (iv) send a communication to the current users;
- An accounting area, to enable authorised users (e.g. VRE Managers) to analyse usage records pertaining VRE services, e.g. most used service, most active user.

4.3 PROJECT INTERNAL VRES

There exists two project internal Virtual Research Environments, SoBigData.it and SoBigData.eu. The former is conceived to provide the SoBigData.it initiative members with a VRE-based working environment while the latter to provide the SoBigData project members with a VRE-based working environment equipped with tools (conversations, wiki, activity tracker etc.) useful to the project management.

The screenshot shows the SoBigData.eu VRE Home Page. At the top, there is a navigation bar with 'Administration', 'Members', 'Activity Tracker', 'Wiki', and 'Twitter Monitor'. Below this, the page is divided into several sections:

- Statistics:** Shows 'Your Stats in SoBigData.eu' with a profile picture and activity indicators (ACTIVITY, GOT) and counts (1, 0, 1, 2, 2).
- Share updates:** A text input field for sharing updates, with a 'Share' button and a 'Notify members' toggle (OFF/ON).
- News feed:** A section for news, sorted by 'newest Post'. It features a post by Massimiliano Assante dated November 21, 2018, 11:53 AM, with 2 likes and 2 replies. The post discusses the new 'Login with EOSC' functionality. Below the post are two download links: 'Academic Login.png' and 'infrastructure log...'. A reply by Fosca Giannotti eh dated November 21, 2018, 2:39 PM is also visible.
- Top Topics:** A list of hashtags: #refresh2018, #eosc, #orcid, #academicaccounts, and #eudat.
- SoBigData.eu's VRE home:** A table listing workspace items:

Name	Owner	Last modified
Deliverables	LC	06 Oct 17:03 15
Meetings	LC	06 Oct 17:03 15
Milestones	LC	06 Oct 17:03 15
- About:** A section with the SoBigData logo and text describing the project's goal: 'SoBigData proposes to create the Social Mining & Big Data Ecosystem: a research infrastructure (RI) providing an integrated ecosystem for ethnic-sensitive scientific discoveries and advanced applications of social data mining on the various dimensions of social life, as recorded by 'big data'. Building on several established national infrastructures, SoBigData will open...'. It includes a 'See more' link and 'Other options ...'.
- VRE Managers and Groups:** A section with a 'View Managers' button and a list of groups in the VRE: 'Production-Support'.

Figure 11. SoBigData.eu VRE Home Page

The main functionality offered by the VREs include:

- A dashboard (cf. Figure 11) for (i) collaborating with other VRE members by posting messages or being informed and reacting to (e.g. commenting, favouring) co-workers' posts, (ii) being informed on the most recent objects added to the VRE workspace, (iii) invite colleagues to participate, (iv) have information on the VRE including a description and the list of managers, (v) having some indicators on the user activity in the VRE, and (vi) acquire an authorization token for exploiting the services offered by the VRE in a programmatic way;
- A Wiki for collaboratively documenting project related activities;
- An issue tracking system (or ticketing system), to support the planning and monitoring of project related activities including work packages, tasks, deliverables and milestones as well as technology development tasks;
- A shared area in the workspace, for making available objects of interests, e.g. project deliverables, presentations, working notes;
- A Members area, for enabling each VRE member to be informed on the rest of VRE members and acquire details for contacting them;

- A user management area, to enable authorised users (i.e. VRE Managers) to manage other users using or willing to access the VRE. VRE Managers can (i) authorise users in accessing the VRE and its services, (ii) assign or withdraw roles to users, (iii) remove users, and (iv) send a communication to the current users;
- An accounting area, to enable authorised users (e.g. VRE Managers) to analyse usage records pertaining VRE services, e.g. most used service, most active user.

4.4 SOBIGDATA LAB VRE

This Virtual Research Environment is conceived to provide the SoBigData e-infrastructure users with a VRE-based working environment capable of executing and importing methods³ in the SoBigData e-infrastructure.

The main functionality offered by the VREs include:

- A dashboard (for (i) collaborating with other VRE members by posting messages or being informed and reacting to (e.g. commenting, favouring) co-workers' posts, (ii) being informed on the most recent objects added to the VRE workspace, (iii) invite colleagues to participate, (iv) have information on the VRE including a description and the list of managers, (v) having some indicators on the user activity in the VRE, and (vi) acquire an authorization token for exploiting the services offered by the VRE in a programmatic way;
- A Method Importer (cf. Figure 12), a tool enabling the integration of existing methods (e.g. R scripts) in the SoBigData e-infrastructure. Additionally, it allows scientists to update their scripts without following long software re-deploying procedures;
- A Method Importer Documentation page describing how to use the Method Importer tool;
- A Method Engine (cf. Figure 9), for executing the available methods of interests directly on the SoBigData e-Infrastructure with the possibility to change the input parameters, thus in line with Open Science principles.
- A user management area, to enable authorised users (i.e. VRE Managers) to manage other users using or willing to access the VRE. VRE Managers can (i) authorise users in accessing the VRE and its services, (ii) assign or withdraw roles to users, (iii) remove users, and (iv) send a communication to the current users;
- An accounting area, to enable authorised users (e.g. VRE Managers) to analyse usage records pertaining VRE services, e.g. most used service, most active user.

³ A method in the SoBigData e-infrastructure is a piece of code in Java, Python or R that implements a social mining algorithm / procedure.

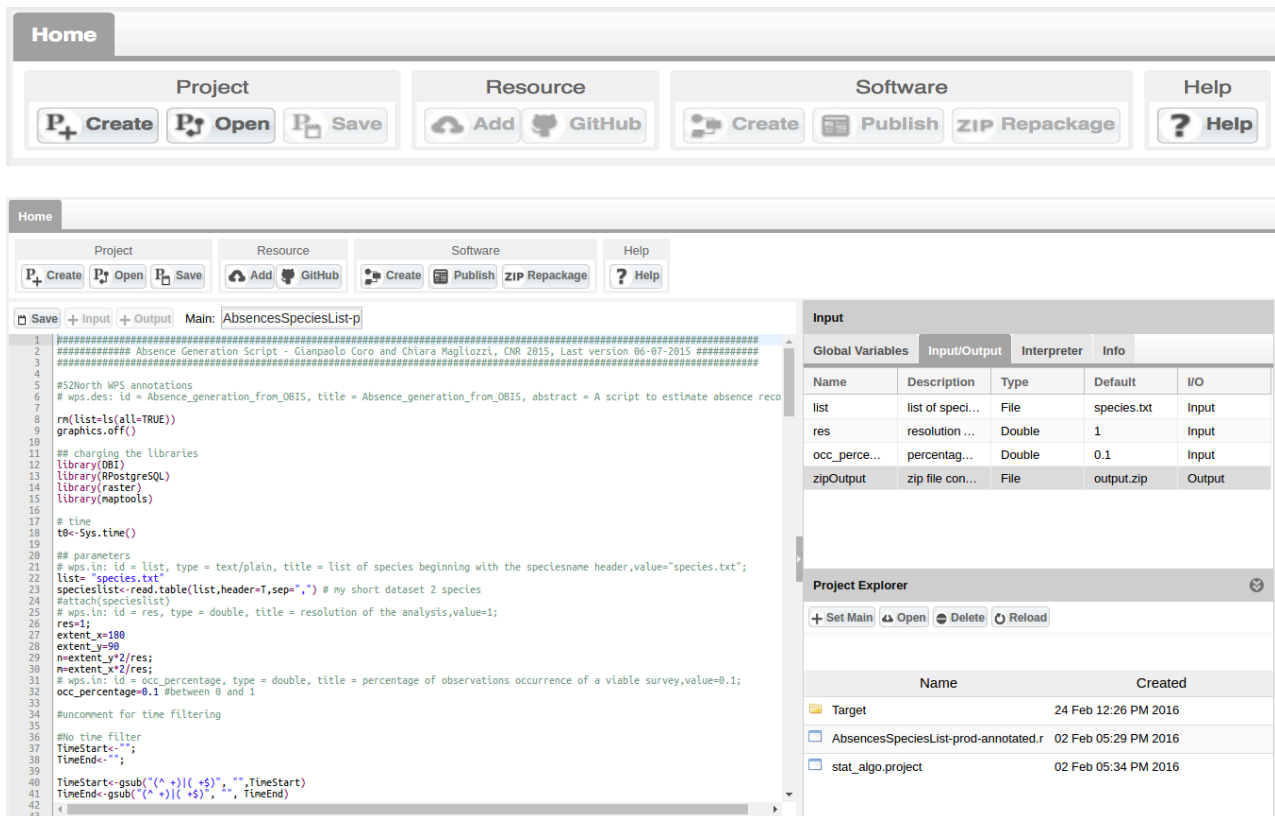


Figure 12. SoBigData Method Importer Tool

4.5 E-LEARNING AREA VRE

The e-Learning Area VRE hosts training materials developed within the SoBigData project. This is the area where all the training course modules provided by SoBigData are categorised and organised.

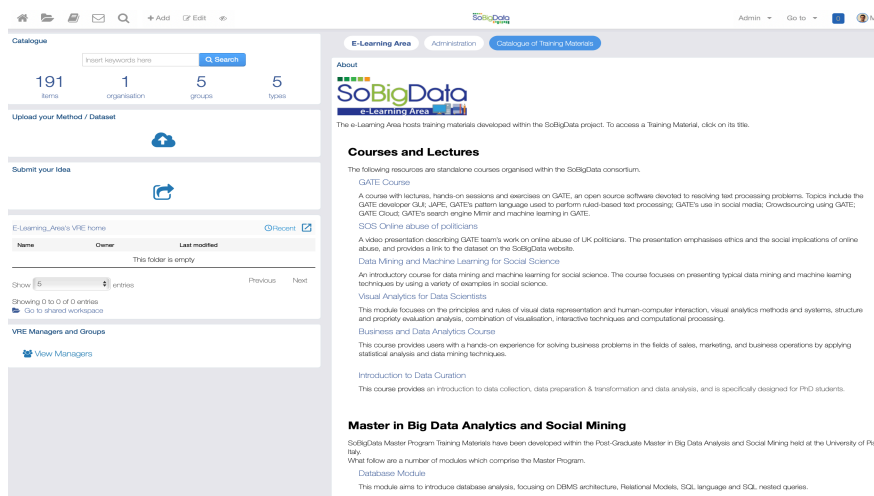


Figure 13. SoBigData E-Learning Area VRE Home Page

The main functionality offered by the VREs include:

- A dashboard listing all the available courses and lectures and for (i) collaborating with other VRE members by posting messages or being informed and reacting to (e.g. commenting, likes) posts, (ii) being informed on the most recent objects added to the VRE workspace, (iii) invite colleagues to join the VRE;
- A direct link to the catalogue and the list of courses via providing possibility to search for catalogue training material in the catalogue as shown in Figure 14.

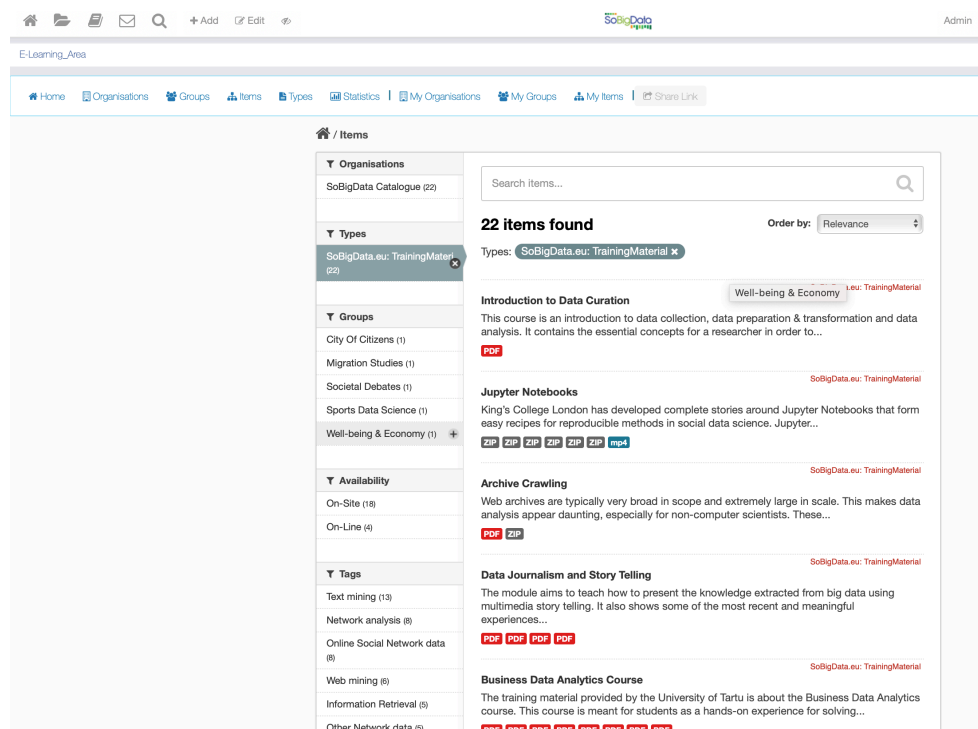


Figure 14. SoBigData E-Learning Area VRE Catalogue page

4.6 SOBIGDATA VIRTUAL RESEARCH ENVIRONMENTS SOFTWARE RELEASE

This section contains the list and pointers to the software packages produced by the project and implementing the components described in the specification above, whose operation today constitutes the SoBigData.eu Virtual Research Environments of the SoBigData infrastructure and accessible via the following links:

EXPLORATORY VRES

- The CityOfCitizens VRE <https://sobigdata.d4science.org/web/cityofcitizens>
- The MigrationStudies VRE <https://sobigdata.d4science.org/group/migrationstudies>
- The SocietalDebates VRE <https://sobigdata.d4science.org/group/societaldebates>
- The SportsDataScience VRE <https://sobigdata.d4science.org/group/sportsdatascience>
- The WellBeingAndEconomy VRE <https://sobigdata.d4science.org/group/WellBeingAndEconomy>

APPLICATION VRES

- The TagMe VRE <https://sobigdata.d4science.org/web/tagme>
- The SMAPH VRE <https://sobigdata.d4science.org/web/smaph>
- The M-Atlas VRE <https://sobigdata.d4science.org/web/m-atlas>

PROJECT INTERNAL VRES

- The SoBigData.eu VRE <https://sobigdata.d4science.org/group/sobigdata.eu>
- The SoBigData.it VRE <https://sobigdata.d4science.org/group/sobigdata.it>

SOBIGDATA LAB AND TRAINING VRES

- The SoBigdata Lab VRE <https://sobigdata.d4science.org/web/sobigdatalab>
- The E-Learning Area VRE https://sobigdata.d4science.org/group/e-learning_area

All the software released are open source, EUPL licensed and available and published on

- Subversion Repository publicly accessible at <http://svn.research-infrastructures.eu/public/d4science/gcube/trunk/>
- GitHub platform accessible at <https://github.com/gcube-team/gcube-releases/tree/v4.13.0>

Subsystem	Description	Packages
SoBigData VRE Management Framework	Services supporting the creation and operation of VREs, i.e. a rich array of gCube-based services enacting the creation and operation of Virtual Research Environments by relying on the available resources (those appearing in the Catalogue);	<p>VRE-Management</p> <ul style="list-style-type: none"> • vre-management.vremodeler.2-0-8 • vre-management.vremodeler-client-library.1-0-1 • vre-management.vremodeler-utils.1-0-1 <p>Portlets-admin</p> <ul style="list-style-type: none"> • portlets-admin.vre-definition.5-1-0 • portlets-admin.vre-deployer.4-2-0
SoBigData Data Analytics Framework	Services supporting the execution and integration of the SoBigData Methods in the SoBigData Infrastructure	<p>Portlets-user</p> <ul style="list-style-type: none"> • portlets-user.data-miner-manager.1-9-0 • portlets.user.statistical-algorithms-importer.1-13-0 <p>Data-analysis</p> <ul style="list-style-type: none"> • data-analysis.dataminer.1.5.5
SoBigData Accounting Framework	Services supporting the collection of resources metrics,	<p>Accounting</p>

	<p>i.e. an array of services automatically collecting per-resource usage metrics, integrated into the gCube SmartGears container that enables automatic accounting of user calls.</p>	<ul style="list-style-type: none"> • accounting-lib.3-4-0 • accounting-analytics.2-8-0
<p>SoBigData VO Management Framework</p>	<p>A set of basic services created and operated in the context of D4Science to serve the needs of SoBigData. This Virtual Organisation (VO) realizes the actual operational context for realizing and operating the SoBigData e-Infrastructure and its resources in autonomy with respect to the other communities and initiatives supported by D4Science;</p>	<p>Common</p> <ul style="list-style-type: none"> • common-smartgears-utils.1-0-3 • common-smartgears.2-1-3 • common-smartgears-app.2-0-2 <p>Distribution</p> <ul style="list-style-type: none"> • gcube-smartgears-bom.1-0-0 • smartgears-distribution.2-2-0

5 CONCLUSION

This deliverable described the software deployed to serve the needs of the SoBigData community, delivering the platform and the VREs planned in “D10.4 SoBigData e-Infrastructure release plan 3” [7]. It provided software specifications for the (i) SoBigData portal, available at <https://sobigdata.d4science.org/> (ii) the SoBigData Catalogue, available at <https://sobigdata.d4science.org/catalogue>, and (iii) the thirteen (13) Virtual Research Environments that have been deployed by the project, namely:

- The SoBigData.eu VRE <https://sobigdata.d4science.org/group/sobigdata.eu>
- The SoBigData.it VRE <https://sobigdata.d4science.org/group/sobigdata.it>
- The TagMe VRE <https://sobigdata.d4science.org/web/tagme>
- The SMAPH VRE <https://sobigdata.d4science.org/web/smaph>
- The CityOfCitizens VRE <https://sobigdata.d4science.org/web/cityofcitizens>
- The MigrationStudies VRE <https://sobigdata.d4science.org/group/migrationstudies>
- The SocietalDebates VRE <https://sobigdata.d4science.org/group/societaldebates>
- The SportsDataScience VRE <https://sobigdata.d4science.org/group/sportsdatascience>
- The WellBeingAndEconomy VRE <https://sobigdata.d4science.org/group/WellBeingAndEconomy>
- The SoBigdata Lab VRE <https://sobigdata.d4science.org/web/sobigdatalab>
- The E-Learning Area VRE https://sobigdata.d4science.org/group/e-learning_area

For each of the above-mentioned software specifications the deliverable also provided the reader with a comprehensive list and pointers to the software packages produced by the project and published on the GitHub platform.

REFERENCES

- [1] Assante, M., Candela, L., Frosini, L., Lelii, L., Mangiacrapa, F., Pagano, P. (2016) SoBigData e-Infrastructure software release 1. SoBigData Project Deliverable D10.5, July 2016
- [2] Candela, L., Castelli, D., Manzi, A., Pagano, P. (2014) Realising Virtual Research Environments by Hybrid Data Infrastructures: The D4Science Experience. International Symposium on Grids and Clouds (ISGC) 2014, Proceedings of Science PoS(ISGC2014)022
- [3] Candela, L., Castelli, D., Pagano, P. (2013) Virtual Research Environments: An Overview and a Research Agenda. Data Science Journal, Vol. 12, pp. GRDI75-GRDI81, DOI [10.2481/dsj.GRDI-013](https://doi.org/10.2481/dsj.GRDI-013)
- [4] Candela, L., Manghi, P., Pagano, P. (2016) SoBigData e-Infrastructure Release plan 1. SoBigData Project Deliverable D10.2, March 2016
- [5] Cornolti, M., Ferragina, P., Ciaramita, M. (2013) A framework for benchmarking entity- annotation systems. In *Proceedings of the 22nd international conference on World Wide Web (WWW '13)*. ACM, New York, NY, USA, 249-260. DOI 10.1145/2488388.2488411
- [6] Assante, M., Candela, L., Manghi P. and Pagano, P. (2018) SoBigData e-Infrastructure Resource adaptation to register to the e-infrastructure 2. SoBigData Project Deliverable D10.9, September 2018
- [7] Assante, M., Candela, L., Manghi, P., Pagano, P. (2018) SoBigData e-Infrastructure Release plan 3. SoBigData Project Deliverable D10.4, November 2018