



<i>Project Acronym</i>	<b><i>SoBigData</i></b>
<i>Project Title</i>	<b><i>SoBigData Research Infrastructure Social Mining &amp; Big Data Ecosystem</i></b>
<i>Project Number</i>	<b><i>654024</i></b>
<i>Deliverable Title</i>	<b><i>SoBigData e-Infrastructure software release 2</i></b>
<i>Deliverable No.</i>	<b><i>D10.6</i></b>
<i>Delivery Date</i>	<b><i>31 July 2017</i></b>
<i>Authors</i>	<b><i>Massimiliano Assante (CNR), Leonardo Candela (CNR), Luca Frosini (CNR), Lucio Lelii (CNR), Francesco Mangiacrapa (CNR), Pasquale Pagano (CNR)</i></b>



## 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.	D10.6
Deliverable Title	SoBigData e-Infrastructure software WP10 release 2
Contractual Delivery Date	31 July 2017
Actual Delivery Date	16 January 2018
Author(s)	Massimiliano Assante (CNR), Leonardo Candela (CNR), Luca Frosini (CNR), Lucio Lelii (CNR), Francesco Mangiacrapa (CNR), Pasquale Pagano (CNR)
Editor(s)	Pasquale Pagano (CNR)
Reviewer(s)	Paolo Manghi (CNR), Valerio Grossi (CNR)
Contributor(s)	
Work Package No.	WP10
Work Package Title	JRA3_SoBigData e-Infrastructure
Work Package Leader	CNR
Work Package Participants	USFD, UNIPI, FRH, UT, IMT, LUH, KCL, SNS, AALTO, ETHZ
Dissemination	Public
Nature	Demonstrator
Version / Revision	V1.0
Draft / Final	Final
Total No. Pages (including cover)	13
Keywords	e-Infrastructure, 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
VRE	Virtual Research Environment
VO	Virtual Organization

# TABLE OF CONTENT

<b>DOCUMENT INFORMATION .....</b>	<b>2</b>
<b>DISCLAIMER .....</b>	<b>3</b>
<b>GLOSSARY .....</b>	<b>5</b>
<b>Table of Content .....</b>	<b>6</b>
<b>EXECUTIVE SUMMARY .....</b>	<b>7</b>
<b>1 Relevance to SoBigData.....</b>	<b>8</b>
<b>1.1 Purpose of this document.....</b>	<b>8</b>
<b>1.2 Relevance to project objectives .....</b>	<b>8</b>
<b>1.3 Relation to other workpackages .....</b>	<b>8</b>
<b>2 Sobigdata e-infrastructure software release.....</b>	<b>9</b>
<b>3 Conclusion .....</b>	<b>12</b>
<b>REFERENCES.....</b>	<b>13</b>

## EXECUTIVE SUMMARY

This document is intended as a follow up of deliverable D10.3: SoBigData e-Infrastructure release plan 2, which detailed the SoBigData infrastructure release plan in terms of the deployment and customization of gCube services on top of the existing and operational D4Science infrastructure. Specifically, the SoBigData e-infrastructure will consist of four main components, i.e. the SoBigData portal, SoBigData Virtual Organization, SoBigData Resource Catalogue, and a number of VREs serving community needs. This deliverable provides the list and pointers to the software packages produced by the project and implementing such components, whose operation today constitutes the SoBigData infrastructure accessible from <http://sobigdata.d4science.org>.

## 1 RELEVANCE TO SOBIGDATA

### 1.1 PURPOSE OF THIS DOCUMENT

This deliverable provides the list and pointers to the software packages produced by the project to implement the functionalities of the SoBigData infrastructure, accessible from <http://sobigdata.d4science.org>. Such functionalities have been described in Deliverables 10.2 and 10.3.

### 1.2 RELEVANCE TO PROJECT OBJECTIVES

One of the main goals of the SoBigData project is to support the so-called Virtual Access. This is ensured by means of an operational e-infrastructure, whose services support scientists at:

- 1) registering (and integrating) their resources in the e-infrastructure (be them datasets, methods, or applications),
- 2) discovering them based on consumption-driven search patterns, and
- 3) combining them in a consistent and uniform manner (e.g. using a method over a dataset and pass the output to an application).

The software listed in this deliverable was produced and deployed in production in the first year of the project and facilitates the above operations.

### 1.3 RELATION TO OTHER WORKPACKAGES

The software was realized following the inputs from WP8 on **Big Data Ecosystem** and WP9 on **Integrating Big Data Analytics Methods and Techniques**. Their inputs provide requirements on how datasets and methods should be described (metadata) to facilitate their discovery and reuse, but also requirements on how datasets and methods should be integrated in the system. Finally, the software has been deployed to serve the needs of WP7 **Virtual Access**, and deliver an operational e-infrastructure for SoBigData.



## 2 SOBIGDATA E-INFRASTRUCTURE SOFTWARE RELEASE

In the following the software packages released in the gCube system which contributed by SoBigData and released are divided in five subsystems, namely SoBigData Catalogue, SoBigData Portal, SoBigData VRE Management Framework, SoBigData Accounting Framework and SoBigData VO Management Framework. They are all relevant to the SoBigData e-infrastructure project.

Figure 1 describes how the above-mentioned subsystem are linked together while Table 1 table summarizes the software packages contributed by SoBigData and released in the gCube system

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-system/gcube-releases/tree/v4.6.0>

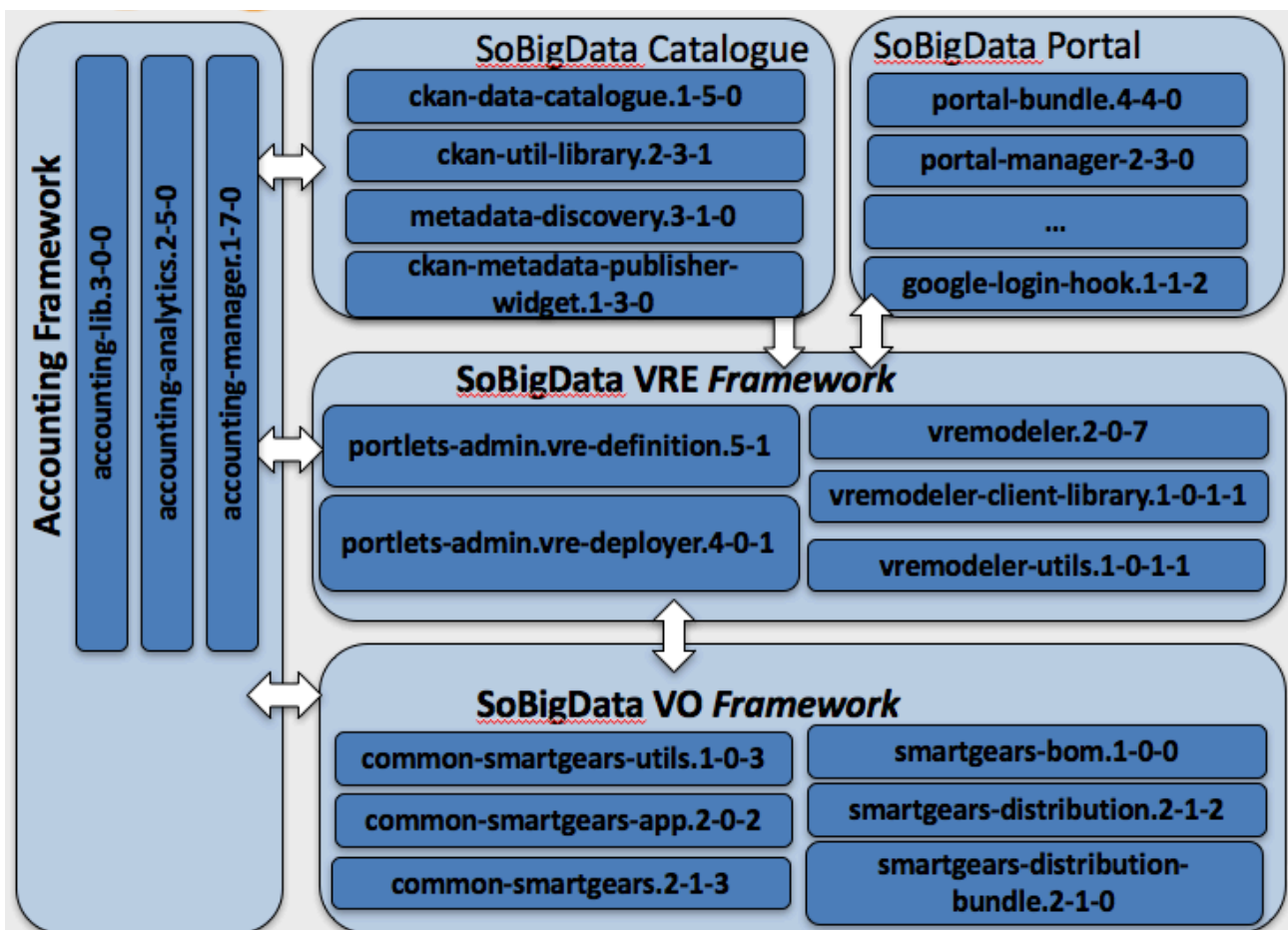


Figure 1. Software Packages in the SoBigData e-infrastructure architecture

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 (<a href="http://ckan.org">http://ckan.org</a>).</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><a href="#">Data-catalogue</a></p> <ul style="list-style-type: none"> <li>ckan-util-library.2-3-1</li> <li>metadata-discovery.3-1-0</li> </ul> <p><a href="#">Portlets-user</a></p> <ul style="list-style-type: none"> <li>ckan-data-catalogue.1-5-0</li> </ul> <p><a href="#">Data-access</a></p> <ul style="list-style-type: none"> <li>ckan-connector.1-1-2</li> </ul> <p><a href="#">Portlets-widgets</a></p> <ul style="list-style-type: none"> <li>ckan-metadata-publisher-widget.1-3-0</li> </ul>
SoBigData Portal	<p>A Liferay based web-portal customized to interface with the D4Science infrastructure and equipped with gCube portlets. This portal 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.</p>	<p><a href="#">Distribution</a></p> <ul style="list-style-type: none"> <li>distribution.portal--distribution.1-2-0</li> </ul> <p><a href="#">Portal</a></p> <ul style="list-style-type: none"> <li>portal.portal-bundle.4-4-0</li> <li>portal-manager-2-3-0</li> <li>portal.linkedin-login-hook.1-1-1</li> <li>portal.google-login-hook.1-1-2</li> <li>portal.login-hook.1-2-0</li> <li>portal.notifications-common-library.1-4-0</li> <li>portal.notifications-portlet.2-4-0</li> <li>portal.social-mail-servlet.2-2-0</li> <li>portal.social-networking-library.1-16-0</li> <li>portal.social-networking-ws.2-0-1</li> </ul> <p><a href="#">Portlets-user</a></p> <ul style="list-style-type: none"> <li>portlets-user.newsfeed.2-4-0</li> <li>portlets-user.myvres.2-3-0</li> <li>portlets-user.notifications.2-4-0</li> <li>portlets-user.shareupdates.2-4-0</li> <li>portlets-user.socialprofile.2-1-0</li> <li>portlets-user.vre-members.2-2-0</li> <li>portlets-user.user-statistics.2-2-0</li> <li>portlets-user.workspace-portlet.6-17-1</li> <li>portlets-user.join-vre.3-3-0</li> </ul>

		<ul style="list-style-type: none"> <li>• portlets-user.data-miner-manager.1-5-0</li> <li>• portlets-user-statistical-algorithms-importer.1-6-0</li> </ul>
SoBigData Accounting Framework	Services supporting the collection of resources metrics, 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.	<a href="#">Accounting</a> <ul style="list-style-type: none"> <li>• accounting-lib.3-0-0</li> <li>• accounting-analytics.2-5-0</li> <li>• accounting-manager.1-7-0</li> </ul>
SoBigData VO Management Framework	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;	<a href="#">Common</a> <ul style="list-style-type: none"> <li>• common-smartgears-utils.1-0-3</li> <li>• common-smartgears.2-1-3</li> <li>• common-smartgears-app.2-0-2</li> </ul> <a href="#">Distribution</a> <ul style="list-style-type: none"> <li>• smartgears-bom.1-0-0</li> <li>• smartgears-distribution.2-1-2</li> <li>• smartgears-distribution-bundle.2-1-0</li> </ul>
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);	<a href="#">VRE-Management</a> <ul style="list-style-type: none"> <li>• vre-management.vremodeler.2-0-7</li> <li>• vre-management.vremodeler-client-library.1-0-1-1</li> <li>• vre-management.vremodeler-utils.1-0-1-1</li> </ul> <a href="#">Portlets-admin</a> <ul style="list-style-type: none"> <li>• portlets-admin.vre-definition.5-1-0</li> <li>• portlets-admin.vre-deployer.4-0-1</li> </ul>

Table 1.

### **3 CONCLUSION**

SoBigData e-Infrastructure software release 2 can be deployed as a production service due to the adoption of stable and supported software packages listed in this deliverable and past experience of the supporting development teams. The gCube Framework is evolving according to service-computing and loosely coupled components principles. The SoBigData e-Infrastructure system will be operated by keeping maintenance, refinement, and extension costs at minimum.

## REFERENCES

- [1] gCube Framework dissemination site: [www.gcube-system.org](http://www.gcube-system.org)
- [2] gCube documentation: [wiki.gcube-system.org](http://wiki.gcube-system.org)
- [3] gCube Source Code Repository [svn.research-infrastructures.eu/gcube/trunk/](http://svn.research-infrastructures.eu/gcube/trunk/)
- [4] gCube Artifact Repository on GitHub [github.com/gcube-system/](https://github.com/gcube-system/)
- [5] D4Science Hybrid Data Infrastructure dissemination site: [www.d4science.org](http://www.d4science.org)
- [6] D4Science Hybrid Data Infrastructure service site: [services.d4science.org](http://services.d4science.org)
- [7] SoBigData e-Infrastructure service site: [sobigdata.d4science.org](http://sobigdata.d4science.org)