



Deliverable D3.1

Initial Dissemination and Impact Plan



DOCUMENT INFORMATION

PROJECT	
PROJECT ACRONYM	SoBigData-PlusPlus
PROJECT TITLE	SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics
STARTING DATE	01/01/2020 (48 months)
ENDING DATE	31/12/2023
PROJECT WEBSITE	http://www.sobigdata.eu
TOPIC	INFRAIA-01-2018-2019 Integrating Activities for Advanced Communities
GRANT AGREEMENT N.	871042
DELIVERABLE INFORMATION	
WORK PACKAGE	WP3 - NA2 - Dissemination, Impact, and Sustainability
WORK PACKAGE LEADER	USFD
WORK PACKAGE PARTICIPANTS	CNR, UNIPI, IMT, KCL, SNS, PSE, RIE, CEU
DELIVERABLE NUMBER	D. 3.1
DELIVERABLE TITLE	Initial Dissemination and Impact Plan
AUTHOR(S)	Joanna Wright (USFD), Valerio Grossi (CNR), Beatrice Rapisarda (CNR), Erika Widegren (RIE)
CONTRIBUTOR(S)	Paolo Ferragina (UNIPI)
EDITOR(S)	Valerio Grossi (CNR), Beatrice Rapisarda (CNR)
REVIEWER(S)	Paolo Ferragina (UNIPI), Marco Braghieri (KCL)
CONTRACTUAL DELIVERY DATE	31/03/2020
ACTUAL DELIVERY DATE	12 /06/2020
VERSION	1.2
TYPE	Report
DISSEMINATION LEVEL	Public
TOTAL N. PAGES	38
KEYWORDS	Impact plan, stakeholder analysis, dissemination indicators

EXECUTIVE SUMMARY

This deliverable provides an overview of the dissemination and impact plan for the first stage in the SoBigData++ project. It will list the various events that were planned during the first period of activity and describes who their intended audience would have been. However, due to the current Coronavirus pandemic it must be noted that no events or activities will be going ahead until the World Health Organization (WHO) declares it safe to do so. Any events that can be rescheduled will be and details will follow in further deliverables.

Building on the foundation of SoBigData, SoBigData++ will draw on the knowledge and expertise developed over this previous four-year project in order to design bespoke events for appropriate stakeholder audiences and attract participants in greater numbers and from more diverse backgrounds. The events are designed to be numerous, varied and geographically diverse. There will also be a strong emphasis on attracting a higher number of individuals from previously underrepresented areas, such as females, early career researchers and the general public to demonstrate SoBigData++'s commitment to reaching out to a wider audience.

SoBigData++ will also target communication channels more specifically towards policy makers and industry partners to ensure the project will have the greatest possible impact on society at large. Currently, there is a definite imbalance in the demographic of the actors who have chosen to participate in the SoBigData events, heavily biased towards educated individuals with a science background. SoBigData++ aims to open its appeal and attract individuals from more diverse and varied academic and workplace backgrounds. How this will be achieved will be detailed in WP5 and in future deliverables.

At the end of the first year of the project, an intention has been agreed to create a non-profit organisation named SoBigData Association, as the first step towards a legal entity that might take over its tasks at the end of the project. Due to the current pandemic, timescales and targets will need to remain dynamic in line with WHO's guidelines.

DISCLAIMER

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 871042.

SoBigData++ strives to deliver a distributed, Pan-European, multi-disciplinary research infrastructure for big social data analytics, coupled with the consolidation of a cross-disciplinary European research community, aimed at using social mining and big data to understand the complexity of our contemporary, globally-interconnected society. SoBigData++ is set to advance on such ambitious tasks thanks to SoBigData, the predecessor project that started this construction in 2015. Becoming an advanced community, SoBigData++ will strengthen its tools and services to empower researchers and innovators through a platform for the design and execution of large-scale social mining experiments.

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 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.

Copyright © The SoBigData++ Consortium 2020. See <http://www.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 2020.”

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

EU	European Union
EC	European Commission
H2020	Horizon 2020 EU Framework Programme for Research and Innovation
WHO	World Health Organisation
EOSC	European Open Science Cloud
RI	Research Institute
RTD	Research & Technical Development

TABLE OF CONTENTS

1.	Relevance to SoBigData++	7
1.1.	Purpose of this document	7
1.2.	Relevance to project objectives	7
1.3.	Relation to other work packages	8
1.4.	Structure of the document	8
2.	Dissemination Plan	9
2.1.	Dissemination Plan Specification	9
2.1.1.	<i>Collaboration With Other Initiatives</i>	10
2.2.	Dissemination Channels	18
2.2.1.	<i>Social Media Strategy</i>	18
2.2.2.	<i>Branding</i>	19
2.2.3.	<i>Project Website, Portal and Management Platform</i>	19
2.2.4.	<i>Dissemination Material</i>	20
2.2.5.	<i>Addressing the Gender Imbalance in Data Science</i>	21
2.3.	Dissemination Activities and Indicators	21
2.3.1.	<i>Planned Events</i>	21
2.3.2.	<i>Overview of Future Events</i>	25
2.3.3.	<i>Dissemination Indicators</i>	25
3.	Initial Stakeholder Analysis	27
4.	Outreach to Policy Makers and the Public at Large	30
4.1.	About Re-Imagine Europa and its Network	30
4.1.1.	<i>White Papers</i>	31
4.1.2.	<i>Meetings, Expert Round-Tables and Conferences</i>	31
4.2.	Overview of European Priority Areas	31
4.2.1.	<i>European Data Strategy</i>	31
4.2.2.	<i>Conference on the Future of Europe</i>	32
4.2.3.	<i>The European Green Deal</i>	32
4.2.4.	<i>Strong Social Europe for Just Transitions</i>	33
4.3.	Future developments	33
5.	CONCLUSION	34
Appendix A.	List of stakeholder already involved	35

1. Relevance to SoBigData++

One of the aims of SoBigData++ is widening and consolidating the SoBigData community. Therefore, the project aims to establish contacts and networks for data scientists, creating a fertile ground to stimulate ideas and initiatives and also encourage younger minds to consider a future in the field of Data Science. These events are physical and practical components of the broader SoBigData++ project.

SoBigData++ aims to reach the widest possible audience, deepen partnerships with collaborators, establish strong relationships with other linked initiatives and ultimately grow and consolidate the SoBigData community.

1.1. Purpose of this document

The purpose of this document is to provide the consortium with an overview of the dissemination and impact plan of SoBigData++. The objectives can be summarised as follows:

- Communication of project results to the general public, scientific communities, and potential adopters through various dissemination channels.
- Dissemination of the project results to the partners' research and business associates, clients and public through media and existing and future contacts.
- Identification of user communities for potential academic and commercial impact.
- Explore ideas on generating impact from the project results on the target stakeholders.
- Continue to assess and analyse the success of the dissemination and impact of the project results

1.2. Relevance to project objectives

Dissemination of the project results and impact generation are key activities to reach the project objectives.

The initial SoBigData project created a starting community of data scientists from the areas of mathematics and ICT to create and distribute a 'multidisciplinary European infrastructure on Big Data and social data mining'. This next stage of the project will continue to develop the research infrastructure (RI), strengthen the services it offers, provide a range of new tools and services and aims to widen its user base in new key areas, (such as social science, data journalism, digital humanities, political science, demography and data-driven medicine). The project will also address scalability, by collaborating and inter-operating with, among others, the European Open Science Cloud (EOSC) and will expand the availability of supercomputing facilities based on the infrastructure provided at partner sites.

By housing existing research platforms and multiple datasets under the one umbrella, together with substantial human and digital resources, the SoBigData++ RI is able to offer a one stop platform for Data Scientists to come together to train, share, learn from and collaborate with other researchers reaching across disciplines. The platform will provide an inclusive, welcoming and supportive environment to inspire and foster new and innovative experiments in the field of social data mining. The Research Infrastructure takes

care of the legal, ethical, methodological and infrastructural issues arising from working with social data, thus enabling data scientists to focus solely on their research.

To achieve this concept SoBigData++ will follow three synergistic lines of activity:

- Access (improve resource access and usage by users)
- Networking (expand the stakeholder community)
- Joint Research (improve and expand platform functionalities and their provisioning within exploratories – see WP10 for more information)

The dissemination and impact activities, relevant research papers and publications, and reaching as many participants as possible are crucial to the success of the project. By consolidating the project in the realm of data science and establishing itself as one of the foremost and most respected, comprehensive repositories of information and expertise, SoBigData will ensure its long-term sustainability.

1.3. Relation to other work packages

Dissemination and impact generation activities are horizontal activities that last for the entire duration of the project. There is clear interaction between all work packages and their relationship to each other is designed to be reciprocal; in that all work packages work together by sharing information, communicating best practice and providing feedback where required.

WP3 in particular aims to provide overarching feedback to the other work packages in order to influence the direction of their work. One of the main aims is to ensure the Research Infrastructure of SoBigData++ is relevant and useful to the spheres of academia and industry and to this end, topical discussions, news events and other pertinent subject matter is closely monitored to ensure opportunities are not missed and suitable data is collected and made available to data scientists.

1.4. Structure of the document

The document is categorized into four different sections:

- Section 1 gives a brief introduction, outlines the major purpose of the document and explains the relevance to SoBigData++.
- Section 2 of the document addresses the initial dissemination plan of the SoBigData++ project. It outlines the dissemination channels, details the events planned so far and gives an overview of future activities. The last part of section 2 addresses the dissemination activities indicators, which will be used to measure the progress of the dissemination activities.
- Section 3 of this document outlines the initial stakeholder analysis.
- Section 4 gives an overview of the strategy to maximise the impact of the work developed by SoBigData++ in reaching out to policy makers and the public at large.
- Section 5 of the document concludes with consolidated findings so far.

2. Dissemination Plan

2.1. Dissemination Plan Specification

SoBigData++ will adopt a multi-channel, multi-target dissemination approach for maximum impact. The main dissemination targets will be relevant scientific communities of social data scientists, who currently embrace multiple research communities, such as data mining, pattern discovery, text and social media mining, mobility analytics, social network analysis, visual and Web analytics. We will also complement these with innovation actions and events aimed at companies in our target verticals (especially Small and Medium Enterprises), government organisations, and policy makers.

One of the project objectives is promoting the education of the next generation of data science researchers with training events (see Work Package 4 for more details). Moreover, there will be a full program of scientific papers and presentations at technical and scientific conferences, aimed at sharing the results of SoBigData++ with the scientific community. In accordance with European Commission policy, scientific publications will be made publicly available through open-access.

Moreover, there will be a programme of papers and articles in the information technology and general business literature, as well as presentations at IT and business seminars and conferences. We will also target key stakeholder groups for the project's innovation activities, split by target. Again, this will be directed towards the European community and associated countries.

Targeted events will include but not be limited to:

- **Big Data:** ECIR 2021,
- **General IT, language technologies:** NetSci 2020, CCS2020, Complex Networks, CompleNet 2020,
- **Law & technology:** PSD 2020, Center for the Future of Intelligence,
- **Digital media:** SocInfo 2020,

SoBigData++ will also maintain and continually update a public project website <http://plusplus.sobigdata.eu> It will function both as a project dissemination tool and as the web-front end of the SoBigData++ research infrastructure. The website provides the link to training materials, code examples, API documentation, and online tutorials of the e-infra (<https://sobigdata.d4science.org/>) explaining how the research infrastructure can be used to run new experiments; to integrate in applications; or be used as component services by SMEs and other companies. Access to numerous social data repositories, open-source tools and state of the art methods for near real-time analysis of large social data is also available. SoBigData++ also aims to provide an advanced infrastructure for large scale analysis of social data, which will enable researchers to run experiments seamlessly on computational resources from EOSC and on super-computing facilities.

SoBigData++ will use the social media channels already established in SoBigData (G.A. n. 654024) to promote the project, events and papers. Those channels are in fact the reference point of a large community of data scientists and stakeholders. They are Facebook, LinkedIn, Twitter and YouTube accounts and pages. Furthermore, the website's blog will be used to publish *ad hoc* created posts, in order to widen the audience (see Section 2.2. for details).

Within the limitations imposed by appropriate protection of intellectual property, all scientific results of this project will be made available to the research community and will also be available on the website via open-access.

In addition, a number of big social data analytics tools and resources from SoBigData will be made available as open-source to facilitate take-up. Another vital part of the dissemination and innovation activities will be the support of a community of early adopters of the research infrastructure, through the website, online user collaboration tools, code examples, training events, datathons, etc. It is envisioned that some of these early adopters will become mentors, or support, for future data scientists.

Training materials will be available and used by the academic partners as part of their course materials at postgraduate level, as well as, used by all partners to provide courses to companies and other non-scientific stakeholders. In addition, networking, transnational and virtual, and joint research activities in the project will involve PhD students and early career researchers, thus aiming to enlarge the social data analytics research community.

SoBigData++ will also produce results relevant to other on-going projects, with which it will cooperate (for details see section 2.1.1. In addition, project partners have already established links to complementary projects in all key RTD areas.

The innovation activities (see WP5 for details) will widen the project impact beyond the target scientific communities, to include outreach towards businesses in diverse sectors (especially SMEs), as well as other stakeholders (e.g. government bodies, funders, policy makers). For instance, we will provide consultancy to business and contractual research and development, train data scientists in entrepreneurial skills, contribute to policy making, and participate in relevant industrial and standardisation bodies. To complement these, SoBigData++ will continue to develop an innovation accelerator, which will help measure the impact of the research infrastructure at scientific, societal, economic, policy, and user community levels.

A range of other instruments will also be employed for dissemination including press releases and media engagement, further participation in appropriate networks of excellence, and activities aimed at both the academic and industrial sectors (e.g. delivery of lectures at summer schools, industry-oriented dedicated workshops, etc). These are activities to be carried out as part of WP3, WP4, and WP5.

2.1.1. Collaboration With Other Initiatives

SoBigData++ plans to re-use results and knowledge deriving from a number of many completed or ongoing projects, and to share data, algorithms, knowledge, and dissemination activities with concurrently running projects. Below is a table which details projects with a collaboration already established and where applicable, SoBigData++ consortium members involved are indicated.

PROJECT NAME & URL & Acronym of Consortium Member already involved	PROJECT OVERVIEW
RDA (Research Data Alliance) www.rd-alliance.org CNR	RDA is a worldwide initiative for research data sharing, currently funded by the European Commission, USA, and Australia. RDA is developing policy, standards, and best practice guidelines for research data sharing. CNR, as one of the RDA Europe partners, will build on their guidelines when creating the SoBigData++ research infrastructure. Many SoBigData++ researchers are involved in RDA's activities and interest groups which contribute in synergy to the definition of technology standards for the RI.
ARIADNE www.ariadneinfrastructure.eu/ CNR	ARIADNE is a digital infrastructure for archaeological research and is interesting to us as there is a growing interest in digital archaeological approaches in relation to social media mining. SoBigData++ will cooperate with ARIADNE and sign a memorandum of understanding aimed also to provide access to SoBigData social mining methods.
CESSDA www.cessda.org/	CESSDA is an ESFRI infrastructure to integrate social science data archives across Europe. In the second phase of the project we will expand our efforts to mine social media data sets to other related social data sets and especially work with CESSDA towards a comprehensive preservation strategy of our derived data sets.
E-CLOUD www.pro.europeana.eu/web/european-a-cloud CNR, KCL	Europeana Cloud is a Best Practice Network, submitted under Objective 2.1.a and coordinated by the Europeana Foundation, designed to establish a cloud-based system for Europeana and its aggregators. Europeana Cloud will provide new content, new metadata, a new linked storage system, new tools and services for researchers and a new platform – Europeana Research.
CLARIN www.clarin.eu USFD, UT, KCL	SoBigData++ fits into the vision of the pan-European initiative CLARIN, which is to enable lower thresholds for accessing multicultural and multilingual content. SoBigData++ will show how such shared datasets would be combined with SoBigData++ content analytics services to extract, aggregate and visualize results, in various domains and use cases.
ELG	The European Language Grid project has just commenced and is aiming to build a cloud-based platform for multilingual language technology in Europe. SoBigData++ will benefit from the multilingual text analysis

<p>https://cordis.europa.eu/project/rcn/219378/en</p> <p>USFD</p>	<p>services from ELG, as well as provide itself datasets, services and tools to the ELG.</p>
<p>DARIAH</p> <p>www.dariah.eu</p> <p>KCL, CNR</p>	<p>DARIAH is an ESFRI infrastructure to support transnational research in all arts and humanities disciplines and related social sciences (such as political communication or media research). We will collaborate with DARIAH on social history and media data sets.</p>
<p>RISIS2</p> <p>https://risis.eu/</p> <p>USFD</p>	<p>RISIS2 is an advanced infrastructure for research and innovation policy studies. It will provide an end user community for the tools to be developed by the proposed SoBigData follow-up project</p>
<p>CODATA</p> <p>www.codata.org</p>	<p>The Committee on Data for Science and Technology is working towards better practices for scientific data management and use. It promotes open data and repeatable science through policies and frameworks for data management, also working closely with the Research Data Alliance. As with RDA, SoBigData++ will seek to adopt best practice policies and frameworks arising from CODATA and, in particular, from the task group on information commons for science.</p>
<p>OpenAIRE-Advance H2020</p> <p>https://cordis.europa.eu/project/rcn/212961/factsheet/en</p> <p>OpenAIRE, CNR</p>	<p>OpenAIRE-Advance continues the mission of OpenAIRE to support the Open Access/Open Data mandates in Europe. By sustaining the current successful infrastructure, comprised of a human network and robust technical services, it consolidates its achievements while working to shift the momentum among its communities to Open Science, aiming to be a trusted e-Infrastructure within the realms of the European Open Science Cloud</p>
<p>AI4EU</p> <p>https://www.ai4eu.eu/</p> <p>CNR</p>	<p>AI4EU aims to build a comprehensive European AI-on-demand platform and ecosystem to lower barriers to innovation, to boost technology transfer and enhance the growth of start-ups and SMEs in all sectors through open calls and other actions. The platform will act as a broker, developer and one-stop shop providing and showcasing services, expertise, algorithms, software frameworks, modules, data, computing resources, prototyping functions and access to funding. Synergy with SoBigData++ is already planned and will focus on the Explainable AI exploratory and on dissemination activities.</p>

<p>Humane AI Flagship http://www.humaneai.eu/ CNR, CEU, UNIPI</p>	<p>The Humane AI Flagship will develop the scientific foundations and technological breakthroughs needed to shape the ongoing artificial intelligence (AI) revolution. The goal is to design and deploy AI systems that enhance human capabilities and empower people -- both as individuals and society as a whole -- to develop AI that extends rather than replaces human intelligence. Synergy with SoBigData++ is already planned and will focus on the Ethical Framework and the Explainable AI.</p>
<p>WeVerify https://weverify.eu/ USFD</p>	<p>WeVerify - Wider and Enhanced Verification for You. An EU co-funded horizon 2020 project that deals with algorithm-supported verification of digital content. Its aim is to address the advanced content verification challenges through a participatory verification approach, open source algorithms, low-overhead human-in-the-loop machine learning and intuitive visualizations. Synergy with SoBigData++ is already planned and will focus on Societal Debate and online misinformation exploratory.</p>
<p>PERICLES http://periclesproject.eu/ KCL</p>	<p>PERICLES is addressing the challenge of ensuring that digital information (including research data) remains accessible in an environment that is undergoing continual change. SoBigData++ will seek cooperation with PERICLES and sign a memorandum of understanding. SoBigData++ will also benefit directly from the knowledge and tools developed by PERICLES as regards digital preservation of the research data generated by this project.</p>
<p>ABIDA https://www.abida.de/en LS3-IRI</p>	<p>ABIDA (Assessing Big Data), funded by the German Federal Ministry of Education and Research, explores social opportunities and risks of the generation, linking and analysis of huge amounts of data and develops options for political action, research and development. The project examines the societal impacts associated with Big Data by using the methods of technology assessment oriented to dialogue and participation. Synergies with SoBigData++ are already planned with the activities of Critical Data Literacy</p>
<p>DADD https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/R033188/1</p>	<p>DADD (Discovering and Attesting Digital Discrimination) is a *novel cross-disciplinary collaboration* to address open research questions on bias and discrimination following a continuously-running co-creation process with academic (Computer Science, Digital Humanities, Law and Ethics) and non-academic partners (Google, AI Club), and the general public, including technical and non-technical users. This three-year project (2018-21) is funded through the ICT Cross-Disciplinarity and Co-Creation of the Engineering and Physical Sciences Research Council (UK). DADD will design ground-breaking machine learning methods to</p>

	certify whether or not datasets and algorithms discriminate by automatically verifying computational non-discrimination norms, which will in turn be formalised based on socio-economic, cultural, legal, and ethical dimensions, creating the new *transdisciplinary field of digital discrimination certification*.
Zones of Data Translation https://gtr.ukri.org/projects?ref=AH%2FR008477%2F1 KCL	Understanding how social and cultural data are transformed into value: This critical data literacy research project will collaborate with the civically minded technology and data advocacy NGO, the Berlin-based Tactical Tech Collective to co-design workshops and tools which will allow humanities and social, science researchers, and the general public to explore mobile operating systems, apps, and analyse permission coding which articulates the flow of social and cultural data. The workshops will be held in a Mozilla-sponsored event in San Francisco, and it, along with the tools, be made available through open access, as will an Impact Report on best practices. Synergies with SoBigData++ are already planned, with particular reference to Data Literacy dissemination and raining activities.
ARA http://l3s.de/~anand/interpretability/ L3S	With this grant we plan to develop algorithms for post-hoc explanations of black box rankers. In particular we focus on text-based neural network rankers that learn feature representations which are hard to understand for developers and end users alike. Synergies with SoBigData++ are already planned, with focus to Explainable AI Exploratory
PRORES http://prores-project.eu/ CNR	PRO-RES project aims to produce a guidance framework regarding the delivery of Responsible Research and Innovation (RRI), which is required from researchers and research funding and performing organizations (RFPO), in order to balance political, institutional and professional contradictions and constraints. Synergies with SoBigData++ dissemination events are already planned within SoBigData Legal and Ethical framework.
Track & Know https://trackandknowproject.eu CNR	Big Mobility Data Analytics Toolbox and the Predictive Complex Event Recognition Toolbox. The project is aimed at producing Analytics Toolboxes for real-time detection and forecasting of the mobility related events. Synergies with SoBigData++ are already planned and will focus on the City of Citizens exploratory.
HarmonicSS https://www.harmonicss.eu	HarmonicSS is a medical research project working to provide a platform with open standards and tools, designed to enable secure storage, governance, analytics, access control and controlled sharing of medical information at multiple levels along with methods to make outcomes

LS3-IRI	comparable across centers. HarmonicSS engages methods of network medicine and big data mining services. Synergies with SoBigData++ are already planned with the activities of new special interest group on Network Medicine
AUTOWARE http://www.autowareeu.org/ CNR	AUTOWARE will establish and push forward an open Cyber Physical Production ecosystem, allowing SMEs to access all the different components in order to develop digital automation cognitive solutions for their manufacturing processes. AUTOWARE will connect several initiatives for strengthening the European SME offer on cognitive autonomous products and leveraging cognitive autonomous production processes and equipment towards manufacturing SMEs. Synergies with SoBigData++ are already planned, within the Innovation WP.
REPLICATE https://replicateproject.eu/ CNR	The cities of San Sebastian, Bristol and Florence are deploying smart city solutions which integrate energy, mobility, and information & communication technologies in city districts. Their objectives are to address urban complexities, specific to these cities, and to develop a model which will be replicated in other cities throughout Europe after the project. Specific problems tackled in the project are: renewable power generation and energy storage, 'intelligent lighting', integrated with the electricity grid; electric mobility, such as e-cars, or e-bikes, but also solutions for deliveries. Synergies with SoBigData++ are already planned, within the exploratory City of Citizens
ESPRIT http://www.esprittransport-system.eu/ CNR	The ESPRIT project aims to develop a purpose-built, light weight L category electric vehicle that can be stacked together to gain space. Thanks to pioneering coupling systems, up to eight ESPRIT vehicles can be nested together in a road train, seven being towed, for an efficient redistribution of fleets and a smartly-balanced and cost efficient transport system. To prove the ESPRIT concept, the project includes also a suite of data analytics, modelling, and simulation tools to predict, once ESPRIT vehicles are deployed, the economic, social and environmental benefits as well as key operating strategies. Synergies with SoBigData++ are already planned, within the exploratory City of Citizens.
AMA - Applied Machine Learning Academy https://ama-academy.eu/ LS3	The aim of the AMA project is s to build up the Applied Machine Learning Academy (AMA) based on a sustainable, flexible and applied qualification concept. AMA will offer a rich portfolio of ML learning units - covering courses as well as labs for hands-on experiences - and an expertise sharing network. The focus will be on foundational ML competences on ML competences for the application domain intelligent production (Industry 4.0). An active dialog with Industry and

	SMEs will ensure a demand-driven training portfolio. Synergies with SoBigData++ are already planned, within the WP3 training and WP5 Innovation
<p>H2020-MSCA-RISE MASTER Multiple ASpects Trajectory management and analysis</p> <p>http://www.masterproject-h2020.eu/</p> <p>CNR</p>	<p>MASTER's objective is forming an international and inter-sectoral network of organisations working on a joint research programme to define new methods to build, manage and analyse rich mobility data. The scientific concept of MASTER is driven by research challenges around the definition, management and analysis of holistic trajectories and potential market opportunities benefitting the European society in the field of tourism, sea monitoring and public transportation. We propose methods to analyse and infer knowledge from holistic trajectories, considering as vital issues the privacy and Big Data dimensions.</p>
<p>Simple-ML</p> <p>https://www.l3s.de/en/projects/simple-ml</p> <p>LS3</p>	<p>The efficient application of current Machine Learning (ML) procedures requires a very high level of expert knowledge, which stands in the way of a widespread use of ML approaches, especially by small and medium-sized enterprises. The goal of the Simple-ML project is to significantly improve the usability of ML processes in order to make them more accessible to a broad user group. The definition of a domain-specific language (DSL) is a central contribution of the project: The DSL provides an integrated description of the ML workflows and their components, and can be specified by textual and graphical editors. Synergies with SoBigData++ are already planned, with focus to Explainable AI Exploratory and the design of WorkFlow Language of WP9</p>
<p>Active knowledge discovery in graphs (AGRA)</p> <p>Academy of Finland ICT2023 project</p> <p>Aalto</p>	<p>Graph mining is the discipline of analyzing data represented as graphs; it is applied to extract knowledge and discover hidden structure in graph datasets. In this project we develop active methods for a broader range of graph-mining tasks. Our hypothesis is that ground-truth knowledge can help guiding the mining algorithm, select among equally good solutions, improve computational efficiency, and overall improve significantly the quality of obtained results. We are developing novel and solid theoretical foundations, carefully analyzing the algorithmic performance both in theory and practice. Synergies with SoBigData++ are already planned, with focus to Filling the Gaps task of WP8</p>
<p>AIDA</p> <p>Academy of Finland ICT2023 project</p> <p>Aalto</p>	<p>In Adaptive and intelligent data aims at applying data-driven methods to optimize the performance of networking systems remains poorly investigated and largely unexplored and proposes a radical paradigm shift by making the data in the network to be the key entity. The objective is to enhance networks with capabilities of large-scale data analytics, so as to enable novel network-management policies that anticipate the network behaviour and make intelligent choices. We</p>

	achieve this by introducing artificial intelligence into the network and exploiting the learning and adaptation possibilities that it offers. Synergies with SoBigData++ are planned, with focus to SNA task of WP8
NESTOR Academy of Finland project Aalto	In NESTOR: Network structure from group response we develop methods to address certain families of network-inference problems. In our setting we assume a set of entities, and data that record the observed behavior of different groups (subsets) of these entities. We refer to such observed behavior by the term group response. The objective is to infer the underlying network structure over the entities. We will explore different variants of the problems of increasing complexity, richness, and ability to model real-world application scenarios. We will apply the resulting methods on real-world datasets. Synergies with SoBigData++ are already planned, with focus to SNA task of WP8
Social Exclusion, City of Espoo funded project Aalto	Text analytics has been used to understand, detect and predict human social and human mental health. Machine learning techniques and Natural Language Processing tools can detect behavioural patterns and social trends from human-generated data, like surveys and social data banks. In this project we will analyse datasets from the City of Espoo Municipality, in Finland, to detect and understand the process leading to social exclusion. Synergies with SoBigData++ , are already planned, with particular reference to Well Being Exploratory
Territori Aperti UAQ	This project will focus on the unique “disaster recovery” data regarding the devastating earthquake occurred in 2009 in the province of L’Aquila, Italy, and the subsequent reconstruction of the region in the following years. Among the aims this study is a data and model driven deeper understanding of the resurrection of the socio-economic structure of a territory after a catastrophic event.

Table 2. 1 Projects with a collaboration already established

The leading centers from different research communities (i.e. economics, network medicine, digital humanities, political studies, economy and finance) will act as a community hubs and will be tasked with enhancing SoBigData++’s userbase and platform usage across their respective communities, thus expanding and solidifying the SoBigData community.

2.2. Dissemination Channels

2.2.1. Social Media Strategy

Accordingly with the *Social media guide for EU funded R&I projects*¹, social media allow to reach an extremely wide and targeted audience, thus to maximise the impact and successful exploitation of research results. Therefore, SoBigData++ will adopt a multi-channel, multi-target communication and dissemination approach for maximum impact. Diverse media engagement will be, in fact, an essential component of the project.

In our social media strategy, have been considered Facebook, Twitter, and LinkedIn as the most appropriate social media in order to communicate and disseminate the project SoBigData++, excluding Instagram for the lack of images to share.

The SoBigData++ social media channels are the ones already set up for the project SoBigData (G.A. 654024) in order to grow the starting community already created: Twitter (@SoBigData), Facebook (www.facebook.com/SoBigData) and LinkedIn (<https://it.linkedin.com/company/sobigdata>). Furthermore, we created also a YouTube channel to share with our community online videos (<https://www.youtube.com/channel/UCRMtvc9ovcISG4EymRuczHQ>).

These communication and dissemination means are promoted through the more traditional channels (flyers, leaflets website, presentations, posters) and all partners will ensure that the project has active social media engagement. Conversely, social media will be used to actively promote the other communication and dissemination tools, i.e. website, publications, deliverables, software releases, press releases, blog posts, and other news.

The regular newsletter/magazine is another effective tool for dissemination. There is already a substantial mailing list of subscribers to this and it is well received within the SoBigData++ community (See section 2.2.5 for further details of the newsletter/magazine).

SoBigData++ will place a stronger focus on communicating its output with the general public and policy makers via high-quality, popularized material in newspapers and magazines as well as dedicated whitepapers, leveraging the partnership with a network of journalists referring to the new project partner Re-Imagine Europe (RIE). Scientific Publications will be available to all and their existence will be promoted and advertised in various communication formats – the project website, newsletters, social media and blogs.

Although all beneficiaries contribute to the project's social media activities, we designated a Communication Manager to oversee all of this, as suggested by the *Social media guide for EU funded R&I projects*. Our Communication Manager will set up and manage social media accounts, centralise the information to be shared and communicate with the audience, including replying to messages. The Communication Manager will also be responsible to adapt the message to the audience, customising the language and content, as well as retweeting, replying and starting online discussion when needed.

¹ https://ec.europa.eu/research/participants/data/ref/h2020/other/grants_manual/amga/soc-med-guide_en.pdf

2.2.2. Branding

The project logo is a restyling of the SoBigData logo. It includes the ++ symbols as part of the project name and the words 'Social Mining & Big Data Analytics' and 'Research Infrastructure' as a sort of pay-off, describing exactly what the project is about. The choice of maintaining the style of the previous logo is due to the need to maintain continuity with the visual identity of the SoBigData Research Infrastructure. The project logo will appear on all project presentations, online videos, web pages, training materials, etc.



Figure 2. 1 The SoBigData++ Logo

The visual identity of the project will be showcased on all SoBigData++ literature from the project posters, leaflets, brochures and training materials to the presentations at SoBigData++ events and especially at other events where SoBigData++ will have a presence.

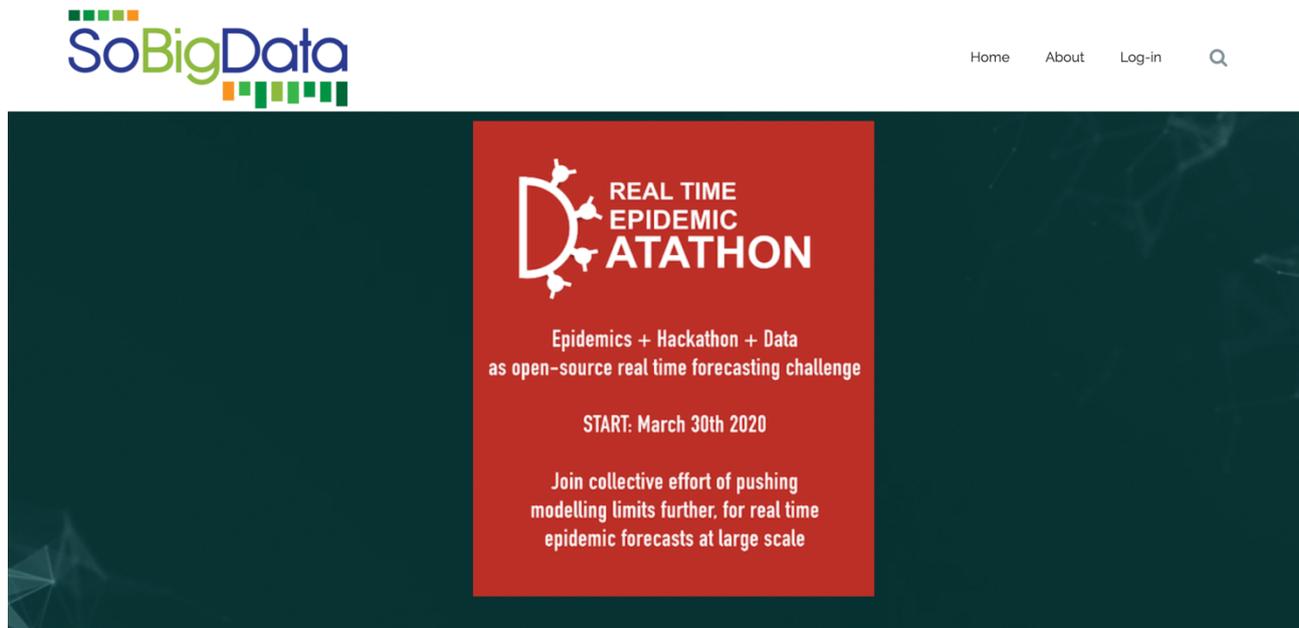
All partners will utilize these branded templates when creating presentations, training resources, posters and all other dissemination material. These templates contain the relevant funding acknowledgements as well as the project identity and will serve to promote the SoBigData++ RI and disseminate its message at all events and through all the relevant dissemination channels.

2.2.3. Project Website, Portal and Management Platform

The SoBigData++ website is the visual and online face of the whole project and serves not only as the welcoming hub for the early adopters of the project and its consortium members but also as the welcoming bridge between external data scientists and the SoBigData++ project.

The website includes information about the project, access to high level publications, training materials, advertisements for forthcoming events and details on how to become involved and sign up for events. Provide access to numerous social data repositories and open-source tools, code examples, API documentation, online tutorials. There is now also the added benefit of access to super-computing facilities which are provided through our partnership with, among others, the EOSC. These facilities provide state of the art methods for near real-time analysis of large social data which is particularly pertinent in this current global pandemic. Indeed, ETH Zurich are currently running a real-time forecasting challenge that aims to develop real-time and large-scale epidemic forecasting models. The website also provides an attractive entry

point to the SoBigData virtual research infrastructure, as described in WP7 and WP9. The Research Infrastructure is highly important to the longevity of the SoBigData++ project and therefore having an up to date, user friendly, trust worthy and innovative website to promote both its existence and its capabilities is vital to reaching out to data scientists, SMEs, policy makers as well as the general public.



SoBigData is the European Research Infrastructure for **Big Data** and Social Mining. From data to knowledge, investigating stories **ethically**, paying attention to citizens privacy.

ta.eu/events/epidemic-datathon

Figure 2. 2 Snapshot of website home page

2.2.4. Dissemination Material

Positive practices developed during the SoBigData project will be continued in SoBigData++, such as the active presence on social media and the lively website forums. There will also be stimulating vlogs and blogs and online training videos to further promote the project online.

The half-yearly newsletter/magazine (www.sobigdata.eu/newsletter) will continue to showcase the best results from the community, with a specific focus on the TNA visits and other relevant content, such as the training activities. There is a public email address newsletter@sobigdata.eu which can be utilised to receive relevant content for the newsletter and indeed its dissemination to subscribers.

The project will ensure the relevant presence of senior scientists as ‘evangelists’ in key events and conferences to attract and inspire participants. High-quality flyers and other communication material will be disseminated in a targeted and appropriate manner to advertise and promote the project and events.

High-level scientific publications will be produced and shared, and creative, thematic workshops aimed at diverse multi-disciplinary audiences will be organised across various European locations.

Press releases are published by the coordinator at key milestones and/or events. In SoBigData++ two major press releases are planned: on the creation of the SoBigData Association and the Foundation. Moreover, other press releases will be published when important collaboration agreements are signed or major international initiatives are started. Media coverage will be tracked and reported on the project website, social media channels, and dissemination reports.

2.2.5. Addressing the Gender Imbalance in Data Science

In line with SoBigData++'s commitment to addressing the gender imbalance in data science, the project has in place travel grants to encourage female participants (and minority group researchers) to attend data science conferences and events. These grants will be advertised and promoted throughout the project on all the communication channels.

To further inspire females into the realm of data science SoBigData++ aims to leverage existing networks (e.g. the Women Network within the European Association on Data Mining and Machine Learning, AI club for Gender Minority), in order to raise awareness regarding the opportunities provided by employment in the field of data science. SoBigData++ has planned specific actions for pursuing the promotion of gender balance in all dissemination and training events both on the speaker's and audience's side. Success Indicators will also be disentangled by gender where possible. SoBigData++ will continue to promote the field of data science to the under-represented research communities and will actively deliver this message throughout the project.

2.3. Dissemination Activities and Indicators

SoBigData++ will develop a set of tools and techniques that will be used to carry out diverse and novel activities and create compelling events and workshops that live-up to the wide-ranging challenges arising in big social data analytics. The material's content and design will be adapted to the needs, expertise and focus of the different target groups and dissemination channels as described. Our dissemination and communication tools will include a balanced mix of traditional (workshops, summer schools, conferences and scientific publications) and innovative methods (datathons, online videos, social media, specialist web front end) to promote take-up and reach widest possible impact.

2.3.1. Planned Events

Due to the current global pandemic of Coronavirus – all these events will be postponed until it is safe to reschedule them.

Tables below is to show the events that would have gone ahead if circumstances were different. As reported in deliverable D1.2, the consortium is continuously monitoring the evolution of the emergency and its impact on the project. At the moment (May 2020), we are trying to reschedule dissemination and training events, also considering the use of on-line tools.

Conferences:

Conference	Date	Location	Description	Target Audience
ECIR 2021, 43rd European Conference on Information Retrieval	March/ April 2020	Lucca, Italy	Conference on searching and organizing unstructured information	Researchers from academia and industry, about 200
PSD 2020 – Privacy in Statistical Databases	23-25 September 2020	Tarragona, Italy	Conference about finding trade-offs to the tension between the increasing societal and economical demand for accurate information and the legal and ethical obligation to protect the privacy of individuals and companies	Researchers, statistical agencies / 40-50
NetSci2020	September 2020	Rome, Italy	Conference about Network Science	TBC
SocInfo 2020 – Social Informatics	October 2020	Pisa, Italy	International conference on social informatics	Researchers 80-120
CCS2020	October 2020	Palma de Mallorca, Spain	Conference on Complex Systems	TBC
Complex Networks	December 2020	Madrid, Spain	Conference about Network Science	TBC
Center for the Future of Intelligence	TBA	Cambridge, UK	Issues in XAI2: Understanding and explaining in healthcare	TBC
CompleNet2020	TBA	Exeter, UK	Conference about Network Science	TBC
PhD School in Data Science	September 2021	Rome, Italy	Summer school for Ph.D. students and young researchers on data-science topics	TBC

Table 2.2 List of Conferences planned before the COVID19 emergency

SoBigData++ Events

Event	Date	Location	Description	Target Audience
VI Scuola Nazionale di Chimica dell'Ambiente e dei Beni culturali	February 2020	Italy	School devoted to students of Environmental Chemistry and Environmental Economics	TBC
Issues in XAI 3: Ethics and Epistemology of AI in healthcare	November 2020	Delft, The Netherlands	Workshop with ethicists, philosophers and practitioners using XAI in healthcare.	TBC
EGI Conference	November 2020	Amsterdam, The Netherlands	Data-centric computing, focusing on the challenges of communities moving to petabyte and exabyte-scale computing	TBC
Realtime Epidemic Datathon	April – July 2020	Online – organised by ETHZ	A collective open source real-time forecasting challenge that aims at joining forces to develop real-time and large-scale epidemic forecasting models	70-100 open to everyone

Table 2.3 List of Events planned before the COVID19 emergency

Lectures & Workshops:

Event	Date	Location	Description	Target Audience
VI Scuola Nazionale di Chimica dell'Ambiente e dei Beni culturali	February 2020	Italy	'Cities and climate change'.	TBC

NetSci2020	September 2020	Rome, Italy	Satellite Workshop on Complexity meets Finance: Data, Methods and Policy Implications	Researchers/ Industry stakeholders/ Policy Makers
NetSci2020	September 2020	Rome, Italy	Workshop - ROME II – Reducing Online Misinformation Exposure	TBC
ECML PKDD Workshop	September 2020	Ghent, Belgium	AI and the risks involved concerning data protection, ethical and legal issues etc	TBC
Data Science in Techno-Socio-Economic Systems Online Workshop 2020	June 10-11 th 2020	Online – organised by ETHZ	To enable research exchange and share insights from different data-intensive disciplines. An intensive two-day immersion with world-class experts from industry and academia.	Approx. 300 Students, academics, practitioners of data science

Table 2.4 List of planned Lectures and Workshops

Scientific Publications in 2020:

Journal	Authors	Title	Published/ Submitted
Journal of Ecological Indicators	S. Maranghi, ML Parisi, A. Facchini*, et al	Integrating urban metabolism and life cycle assessment to analyse urban sustainability	TBC
Artificial Intelligence	Juan M. Durán	Explaining scientific explanations in medical AI	Submitted
BMC Medical Informatics and Decision Making	Juan M. Durán - Karin Jongsma	Who is afraid of black box algorithms? Epistemic opacity and moral trust in medical AI	Submitted

Proceedings of VLDB	P. Ferragina, G. Vinciguerra	The PGM-index: a fully-dynamic compressed learned index with provable worst-case bounds	Published
New Journal of Physics	Federica Parisi, Tiziano Squartini, Diego Garlaschelli	A faster horse on a safer trail: generalized inference for the efficient reconstruction of weighted networks	Published

Table 2.5 List of Scientific Publications

2.3.2. Overview of Future Events

USFD are planning a Workshop on Misinformation – RDSM 2020 in September 2020, in Spain. This will be the 3rd edition of the RDSM workshop and will particularly focus on online information disorder and its interplay with public opinion formation. The aim is to put a great emphasis on the usefulness and trust aspects of automated solutions tackling the misinformation and fake news among other themes.

The conference SocInfo 2020 – Social Informatics, organized by CNR, is planning to switch his format to an online conference.

UAQ are discussing a possible collocation for a workshop at the European conference on Software Architecture.

BSC and UT are in discussions about the possibility of submitting a proposal for a workshop with the support of USFD. No further details are available at this time.

2.3.3. Dissemination Indicators

The success of the communication activities will be closely monitored and reported in the periodic activity deliverables. We will be following the dissemination map to ensure we reach the intended stakeholder targets and keep a comprehensive record of the categories and numbers of stakeholders approached.

There are set targets for the different channels of communication and expectations will be detailed for a high level of social media presence and outreach to both Industry stakeholders and Policy Makers. We are in the process of implementing tighter controls on monitoring the participants in the various events, from pre-registration information to ‘on the day’ registration. We will be collecting feedback where possible and requesting contact details for the dissemination of future events and news.

In order to monitor the activities on social media and website, we will use Google Analytics, Twitter Analytics and Facebook Insight, as suggested by the *Social media guide for EU funded R&I projects*².

² https://ec.europa.eu/research/participants/data/ref/h2020/other/grants_manual/amga/soc-med-guide_en.pdf

The table below details the targets SoBigData++ aims to reach during the three reporting periods.

Success Indicators	Target Reporting Periods		
	1	2	3
Extent of audience reached	2,500	5,000	6,500
Number of SMEs, big companies and other institutional stakeholders engaged through all dissemination activities (Companies & Institutions)	100	150	200
Papers published on OpenNAIRE portal	30	60	90

Table 2.6 Success Indicators for Reporting Period 1, 2 and 3

3. Initial Stakeholder Analysis

One of the key points for moving forward from a starting research community of big-data pioneers to a wide multi-disciplinary and well-established community of big data researchers, developers and innovators is to create a community of the stakeholders explicitly or implicitly involved in the consortium, and then study their needs. The main goal of this study is to define a strategy that will engage these stakeholders, and more than these ones in the future, in order to guarantee the long-term sustainability of project's results. This will necessarily need a strong coordination with the innovation activities of WP5, in order to maximize project's impact, and a valorisation of the free virtual access (VA) offered by the SoBigData++ infrastructure, as described in WP6 and WP7.

As a result, for the first time in big-data social mining research, a variegated community of researchers, developers, innovators, and other stakeholders will have the opportunity to access a rich social data ecosystem of datasets and software tools, in a unified manner, and to conduct large-scale, multi-disciplinary experiments on extracting social knowledge. In this context, Task 3.1 has the objective of enlarging significantly the stakeholder community in three directions: *i)* experts in data analysis from outside computer science, thus empowering their tools with the ones developed and made available in SoBigData++ infrastructure; *ii)* innovators that use big data analytics for their business, thus expanding the features and functionalities of their analytics platforms; and *iii)* users from other disciplines, thus facilitating cross-disciplinary research on social data (e.g. Digital Humanities, Political Science, Medicine, etc.).

All project's partners have an extensive experience in acquiring, managing and analysing real-world Big Data from heterogeneous sources and have the capacity to transform their research findings into tools and technologies usable by the scientific community at large. They have also shown in the past to be also able to transfer to the industrial and innovator communities expertise and methodologies. So, we have developed a first comprehensive stakeholder analysis in order to understand who are the stakeholders already involved with SoBigData++ partners either in some project's tasks or that collaborate with them via formal agreements in other projects still concerning with topics related to Social Big Data.

The analysis is based on three main dimensions:

1. **type of Stakeholder:** this dimension represents our potential stakeholders. Currently, it includes as potential stakeholders: Industries, Policy and law makers, Researchers, Journalists, Teaching institutions, Public Administration, Data analysts, Journalists, Non-profit organizations, Public.
2. **application fields:** this dimension represents the actual sectors where the research developed in this project can be applied. Currently, it includes: Finance, Energy, Economy, Humanities, Laws, Statistics, Retail, Telecommunication, Computer Science, Sport Analytics, Transport, Social science, Insurance, Medicine, Social policies, Ethics.
3. **type of collaboration:** this dimension refers to Stakeholders already involved in SoBigData++. For these ones we have specified the type of current collaboration: e.g. if the stakeholder is using the SoBigData E-infra, or it has participated (or is participating) to a SoBigData++ event, or if it collaborates in some other way, e.g. the stakeholder provides data to SoBigData++.

In order to include in this analysis the industrial world, task “T5.1 - Partnerships with industry and contributions to policy making” takes concrete steps towards this particular category of stakeholders and will measure the impact of the SoBigData++ research infrastructure at societal, policy, and community levels. It will monitor, analyze, and enhance the social processes around data science and innovation. In this context, we notice that also the Exploratories in WP10 will amplify new experiments achievements towards the vertical scientific communities and potential stakeholders by activating appropriate dissemination channels. We recall that SoBigData++ continues to use social media as a prime dissemination channel to promote the research infrastructure and its services to the large number of stakeholders listed below. The channels include also relevant groups on LinkedIn, Facebook, and blogs in the fields of data mining and pattern discovery, social network analysis, text and social media analysis, semantic web, business intelligence, new media. We have to mention that also WP4 with task “T4.1 - Online Training Modules” will create training modules adapted to stakeholder requirements in collaboration to WP2, WP5, WP7 and WP10.

The following two figures report a summary of the results obtained from this analysis. In particular, Figure 3.1 shows the macro-areas of the activities of the stakeholders already involved with one/some SoBigData++ activities; whereas Figure 3.2 shows the distribution of their application fields.

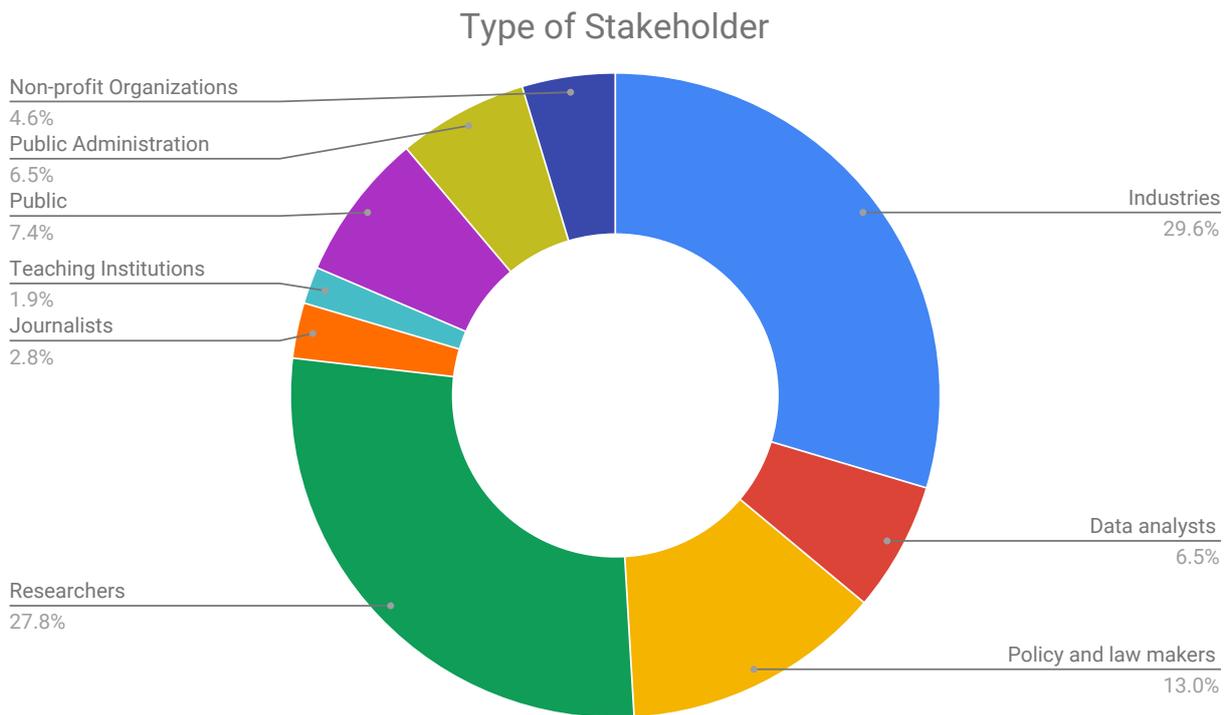


Figure 3.1 The macro-areas of the activities of the stakeholders already involved

Being SoBigData++ a research project, it is not surprising that most stakeholders come from the research field. Therefore, one of the aims of SoBigData++ is to positively impact onto the society, by accelerating innovation within industries and law makers, as primary fields of application of our scientific results, tools and methodologies.

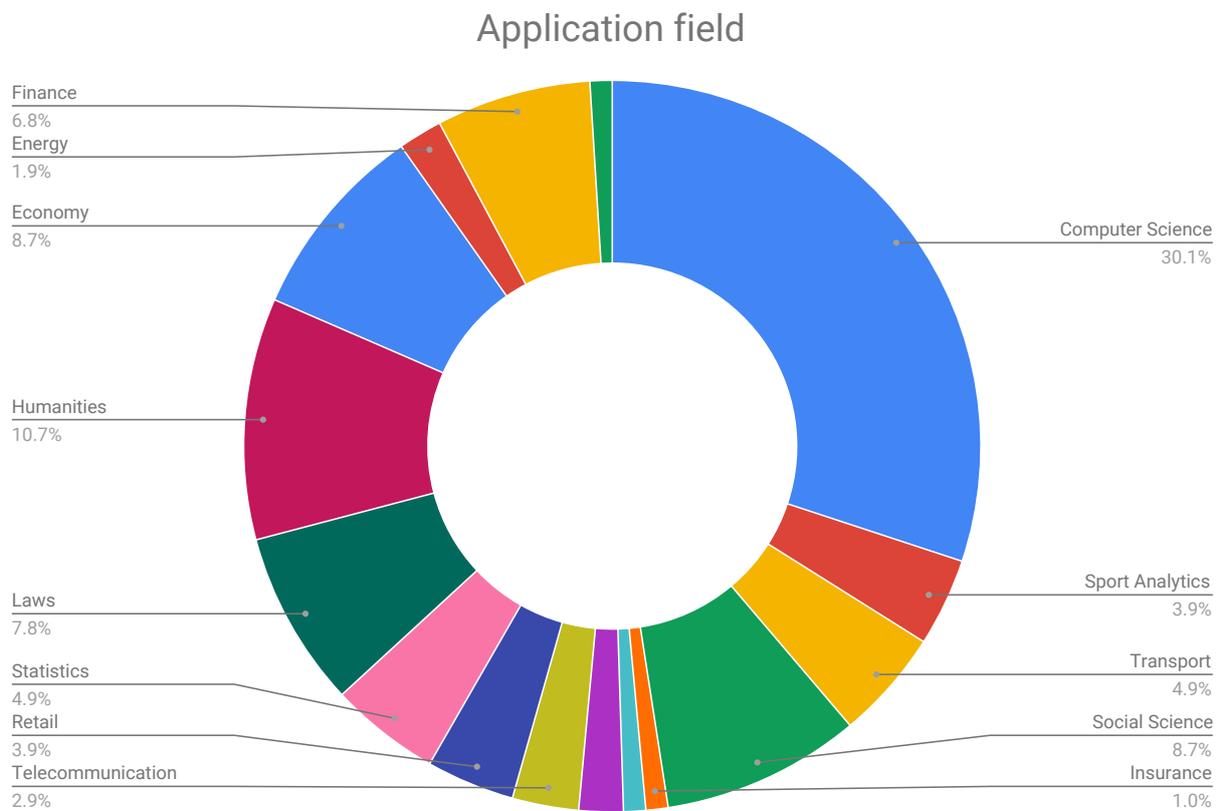


Figure 3.2 Distribution of the application fields of the stakeholders already involved

Due to the nature of the consortium it is not a surprise that Computer Science is the major application field, with 30.1% of stakeholders. More interesting is the fact that a significant percentage is shared by economy, humanities, social science or law stakeholders.

Appendix A shows the complete list of stakeholder already involved in the Institutions of the SoBigData++ Consortium.

4. Outreach to Policy Makers and the Public at Large

The last decade has seen digital technologies changing the economy and society and affecting the daily lives of European citizens. Data is at the centre of this transformation, and the consequent innovation will bring benefits for citizens. At the same time, the increasing volume of non-personal industrial data and public data in Europe, combined with technological change in how the data is stored and processed, will constitute a potential source of growth and innovation.

It is therefore essential that a project like SoBigData++ is widely known and understood, not just within the community of potential users, but also with the public at large and policy-makers.

The European policy agenda includes a number of initiatives where the work developed by SoBigData++ will be relevant. This will relate not only to the European Data Strategy but also all those policy priorities addressed by the exploratories including: “Media Action Plan”, “European Industrial Strategy”, The “European Green Deal”, “Conference on the Future of Europe”. The COVID-19 crisis has further underlined the need for deeper societal debate about many of these issues and it will be imperative that the consortium be engaged in these debates.

This section will give an overview of the strategy to maximise the impact of the work developed by SoBigData++ in reaching out to policy makers and the public at large.

4.1. About Re-Imagine Europa and its Network

Founded by President Valéry Giscard d'Estaing and based in Brussels, Re-Imagine Europa was created as an “incubator” for new political ideas. It works with members and partners to deliver world-class interdisciplinary and intersectoral research, the facilitation of dialogue and engagement with national and international political processes.

RIE acts as a convener, catalyst and communicator. The aim is to be both innovative and strategic in its approach as well as pragmatic and concrete. Re-Imagine Europa works on the principles of evidence-based policy making, inclusive and collaborative approaches and working at an interdisciplinary and intersectoral level to build shared narratives and new ideas on topics of strategic importance. Re-Imagine Europa is nonpartisan and founded in honour of the friendship and partnership between President Giscard d'Estaing and Chancellor Helmut Schmidt and their efforts to define a future that went beyond personal, national or partisan interests.

Re-Imagine Europa works closely with the European Commission and European Parliament in order foster new thinking and innovative solutions to some of the most challenging problems of our time. With a network of over one hundred knowledge partners across Europe, including the Organisation for Economic Co-operation and Development (OECD), the Joint Research Centre of the European Commission, the European Parliament Research Service, the Foundation for European Progressive Studies (FEPS), European Liberal Forum, the Wilfried Martens Centre for European Studies, the European Youth Forum, the European Trade Union Confederation, "la Caixa" Foundation, Cariplo Foundation, Calouste Gulbenkian Foundation, ALL

European Academies, to name a few, RIE develops inclusive methodologies for policy discussion engaging the public at large and key stakeholders in the debate to find innovative solutions.

4.1.1. White Papers

Throughout the project, Re-Imagine Europa will produce a number of white-papers based on the reports from the workshops developed T3.3 – Multidisciplinary Community Building and Outreach. These white papers will be decided together with the consortium and reflect the policy and public discourse agenda. In order to maximise the impact of the white-papers the production will be staggered and planned in relation to European conferences and timelines.

Dissemination of the white-paper will be organised through specific events, presentation at ongoing events, dissemination across the network of Re-Imagine Europa and through our media partners.

4.1.2. Meetings, Expert Round-Tables and Conferences

Another key strategy in order to ensure the relevance of the results of SoBigData++ for policy makers and the public at large is to organise a number of strategy expert round-tables and conferences as well as engage in key policy events and events with the public at large.

Re-Imagine Europa will be organising a number of strategic events at the European Parliament in order to present results of the work of the SoBigData++ consortium.

This chapter will give an overview of the strategy to maximise the impact of the work developed by SoBigData++ in reaching out to policy makers and the public at large.

4.2. Overview of European Priority Areas

There are a number of key priority areas that will be shaped in the coming years where input from the SoBigData++ consortium could be very important. Please find below an overview of key policy areas that should be targeted.

4.2.1. European Data Strategy

The digital transformation has brought data at the core of our activities. Any action generates increasing amounts of data, so European policymakers have become increasingly aware of the importance of data ownership and the use that can be made of data. How data is collected and used must place the interests of the individual first, in accordance with European values, fundamental rights and rules.

Citizens will trust and embrace data-driven innovations only if they are confident that any personal data sharing in the EU will be subject to full compliance with the European Unions' strict data protection rules. For this reason, the commission is steadily working to create a single market for data, which respects the rules of privacy and ethics that have always characterized the European lifestyle.

The construction of fair, reliable and trustworthy laws and practices will allow the free circulation of data within the union, creating the conditions for increasing the innovation of many sectors. Some examples may be transportation, healthcare and agriculture. The benefits of such a setting, even related to the development of artificial intelligence and other digital technologies, are significant and undeniable.

In this context, a project such as SoBigData++ must be the forerunner both for good data use practices and to improve citizens' awareness of the use that is made of their data, the responsibilities connected to this and the benefits they can obtain by transparent and careful management of this process.

4.2.2. Conference on the Future of Europe

The democratic system of the European Union is unique. It encompasses 500 million people and transcends borders. To make it even more vibrant, interactive and relevant to our citizens, we need to use new methods. President Ursula von der Leyen has pledged to do this by giving Europeans a greater say on what the Union does and how it works for them. This is the central premise behind the idea of a Conference on the Future of Europe.

As a major pan-European democratic exercise, the Conference will be a new public forum for an open, inclusive, transparent and structured debate with citizens around a number of key priorities and challenges. It will be a bottom-up forum, accessible to all citizens, from all walks of life, and from all corners of the Union, and should reflect Europe's diversity. It will be open to civil society, the European institutions and other European bodies, including the Committee of the Regions, the European Economic and Social Committee, as well as national, regional and local authorities, parliaments and other stakeholders – all contributing as equal partners.

It would be very relevant for the Exploratory on Societal Debates and Misinformation Analysis be engaged and involved in this exercise.

4.2.3. The European Green Deal

Climate change and environmental degradation are an existential threat to Europe and the world. To overcome these challenges, Europe needs a new growth strategy that transforms the Union into a modern, resource-efficient and competitive economy where there are no net emissions of greenhouse gases by 2050, economic growth is decoupled from resource use and no person and no place is left behind.

The European Green Deal is the European Commission's roadmap for making the EU's economy sustainable. This will happen by turning climate and environmental challenges into opportunities across all policy areas and making the transition just and inclusive for all.

The results if several of the exploratories would be useful to be developed in line with some of the priorities of the European Green Deal.

4.2.4. Strong Social Europe for Just Transitions

Individuals and businesses in the EU can only thrive if the economy works for them. The EU's unique social market economy allows economies to grow and to reduce poverty and inequality. With Europe on a stable footing, the economy can fully respond to the needs of the EU's citizens. The European Commission's focus on "Strong Social Europe for Just Transitions" could benefit from input from several of the exploratories in particular "Demography, Economy and Finance".

4.3. Future developments

We will continue to ensure awareness of the political processes in Europe and how the work of the consortium might be useful for various priority areas. Continuous work with stakeholders from across Europe will ensure that we can increase the impact and usefulness of the work done as well as create awareness about the capacity of big data to support evidence-based policymaking. Working with social and traditional media we will also be able to engage with the public at large.

5. CONCLUSION

This deliverable is an initial dissemination and impact plan of SoBigData++. It aims to detail the dissemination strategy to be adopted throughout the project lifetime and the first version of the impact plan. It outlines also a first stakeholder analysis and the outreach to policy makers and the public at large.

Appendix A. List of stakeholder already involved

Below the complete list of stakeholder already involved in the Institutions of the SoBigData++ Consortium.

Partner	Stakeholder already involved in projects with the Consortium's partner	Type of Stakeholder	Application field of the stakeholder
1 - CNR	Tiscali	Industries	Computer Science
1 - CNR	Wyscout spa	Industries	Sport Analytics
1 - CNR	FC Barcelona	Data analysts	Sport Analytics
1 - CNR	Comune di Firenze - Mobility Agency	Policy and law makers	Transport
1 - CNR	Comune di Pisa - Mobility Agency	Policy and law makers	Transport
1 - CNR	Comune di Roma	Researchers	Social Science
1 - CNR	Comune di Roma - Mobility Agency	Policy and law makers	Transport
1 - CNR	Generali Group	Industries	Insurance
1 - CNR	FC Internazionale Milano	Data analysts	Sport Analytics
1 - CNR	IRPET	Researchers	Social Policies
1 - CNR	ISTAT	Researchers	
1 - CNR	Nova 24 Ore	Journalists	Medicine
1 - CNR	Octo Telematics	Industries	Telecommunication
1 - CNR	Privacy Authority	Policy and law makers	
1 - CNR	Regione Toscana	Policy and law makers	Transport
1 - CNR	Roma	Data analysts	Sport Analytics
1 - CNR	Tecnalia	Industries	Computer Science
1 - CNR	Telecom Italia	Industries	Telecommunication
1 - CNR	UNICOOP	Industries	Retail
1 - CNR	University of Firenze	Researchers	Statistics
1 - CNR	University of Pisa	Teaching Institutions	Computer Science
1 - CNR	WIND	Public	Telecommunication
1 - CNR	Presidenza del Consiglio dei Ministri	Public Administration	Computer Science
1 - CNR	Hamad Bin Khalifa University (HBKU) – College of Science and Engineering (CSE)	Researchers	Computer Science
1 - CNR	University of Pisa	Researchers	Computer Science
2 - USFD	DCMS	Policy and law makers	Laws
2 - USFD	First Draft	Journalists	Social Science
2 - USFD	Ontotext	Data analysts	Computer Science

2 - USFD	AFP	Journalists	Humanities
2 - USFD	Text mining solutions	Data analysts	Computer Science
3 - UNIFI	Comune di Venezia	Policy and law makers	Economy
3 - UNIFI	ENEL	Industries	Energy
4 - FRH	Euractive	Policy and law makers	Laws
5 - UT	STACC	Data analysts	Computer Science
5 - UT	Swedbank	Industries	Finance
9 - SNS	Fineco Asset Management	Industries	Finance
9 - SNS	Fondazione Cassa di Risparmio di Firenze	Industries	Finance
9 - SNS	Unicredit	Industries	Economy
3 - UNIFI	Bloomberg (London)	Industries	Finance
3 - UNIFI	CityPost	Industries	Retail
3 - UNIFI	PlayeRank	Industries	Computer Science
3 - UNIFI	SADASDB	Industries	Computer Science
3 - UNIFI	MIT	Researchers	Computer Science
3 - UNIFI	Harvard	Researchers	Computer Science
8 - KCL	YRS	Public	Ethics
8 - KCL	British Library	Public Administration	Humanities
8 - KCL	TNA	Public Administration	Humanities
8 - KCL	Cultural Institute	Public Administration	Humanities
8 - KCL	PyData London	Data analysts	Computer Science
8 - KCL	United Nations High Commissioner for Refugees	Policy and law makers	
8 - KCL	Médecins Sans Frontières	Public	
8 - KCL	Science Gallery London	Public	Humanities
8 - KCL	Tactical Tech	Public	Humanities
8 - KCL	Hospify	Public	Humanities
8 - KCL	Mycroft	Public	Humanities
8 - KCL	Microsoft	Public	Humanities
6 - IMT	KPN	Industries	Social Science
6 - IMT	ING	Policy and law makers	Economy
6 - IMT	ABN-Amro	Policy and law makers	Economy
6 - IMT	ISC-CNR	Researchers	Economy
6 - IMT	University of Zurich	Researchers	Economy

6 - IMT	University of Leiden	Researchers	Economy
6 - IMT	ENEL FOUNDATION	Industries	Social Science
7 - LUH	Internet Archive	Researchers	Computer Science
7 - LUH	German National Library	Public Administration	Computer Science
1 - CNR	Innovalia	Industries	Computer Science
1 - CNR	3D KUMO	Industries	Retail
1 - CNR	Oxford University	Researchers	Social Science
25 - UAQ	Comune L'Aquila	Policy and law makers	Computer Science
25 - UAQ	CGIL, CISL, UIL	Policy and law makers	Computer Science
23 - UPF	European and Global Flood Awareness Systems (EFAS/GloFAS)	Public Administration	Computer Science
23 - UPF	University of Barcelona Group of Advanced Studies on Violence	Researchers	Laws
25 - UAQ	Regione Abruzzo	Policy and law makers	Statistics
25 - UAQ	ASL Avezzano Sulmona L'Aquila	Public Administration	Medicine
7 - LUH	Amazon	Industries	Computer Science
7 - LUH	Schufa	Industries	Computer Science
11 - ETHZ	NYU Courant	Researchers	Finance
11 - ETHZ	SIX Group Ltd.	Industries	Finance
11 - ETHZ	DS3lab ETH	Researchers	Computer Science
1 - CNR	Amazon	Industries	Retail
11 - ETHZ	University of Bologna & Scuola Normale Superiore	Researchers	Finance
31 - SSSA	Knauf	Industries	Economy
31 - SSSA	Loccioni	Industries	Economy
31 - SSSA	Enel	Industries	Energy
31 - SSSA	Fincantieri	Industries	Transport
19 - CEU	Renyi Institute Budapest	Researchers	Statistics
19 - CEU	Charles University Prague	Researchers	Statistics
19 - CEU	RWTU Aachen	Researchers	Statistics
19 - CEU	Grandata	Industries	Computer Science
19 - CEU	Idylle	Industries	Computer Science
19 - CEU	UNICEF Innovation	Non-profit Organizations	Computer Science
19 - CEU	ISI Foundation	Researchers	Computer Science
19 - CEU	IXXI Complex System Institute	Researchers	Computer Science

17 - CNRS	Internet Policy Review	Researchers	Laws
17 - CNRS	CSI Centre de Sociologie de l'Innovation	Researchers	Social Science
17 - CNRS	Internet Governance Lab	Researchers	Laws
17 - CNRS	Internet Society France	Non-profit Organizations	Laws
17 - CNRS	International Association for Media and Communication Research	Non-profit Organizations	Social Science
17 - CNRS	Citizen Lab https://citizenlab.ca	Researchers	Laws
17 - CNRS	Delta Chat	Industries	
17 - CNRS	TRACES	Non-profit Organizations	Humanities
17 - CNRS	Center for Research and Innovation	Teaching Institutions	Humanities
17 - CNRS	Fondation Internet Nouvelle Génération	Non-profit Organizations	Laws
17 - CNRS	Network of centers for Internet	Researchers	Social Science
1 - CNR	DFKI	Researchers	Computer Science
1 - CNR	Engineering	Industries	Computer Science
1 - CNR	University of Warsaw	Researchers	Social Science
1 - CNR	University of Patras	Researchers	Computer Science