Social Mining & Big Data Analytics

SoBigData

RESEARCH INFRASTRUCTURE

Deliverable D3.2

Dissemination and Impact Report and Plan for following period 1

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

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 (60 months)
ENDING DATE	31/12/2024
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	USFD, CNR, UNIPI, IMT, KCL, SNS, PSE, RIE, CEU
DELIVERABLE NUMBER	D3.2
DELIVERABLE TITLE	Dissemination and impact report and plan for following period 1
AUTHOR(S)	Joanna Wright (USFD), Erika Widegren (RIE), Matteo Nicolosi (RIE), Beatrice Rapisarda (CNR), Valerio Grossi (CNR), Roberto Pellungrini (UNIPI)
CONTRIBUTOR(S)	Vittorio Romano (CNR), Marco Braghieri (KCL)
EDITOR(S)	Beatrice Rapisarda (CNR), Valerio Grossi (CNR)
REVIEWER(S)	Marco Braghieri (KCL), Roberto Trasarti (CNR)
CONTRACTUAL DELIVERY DATE	30/06/2021
ACTUAL DELIVERY DATE	07/07/2021
VERSION	V1.3
TYPE	Report
DISSEMINATION LEVEL	Public
TOTAL N. PAGES	68
KEYWORDS	Dissemination activities, communication, events

EXECUTIVE SUMMARY

This Deliverable provides an overview of the dissemination activities that have taken place from M1 to M18 of this project. SoBigData was a highly successful project providing a platform to bring together many experts from the different fields of Data Science and provide an environment for collaborative and fruitful alliances and partnerships. SoBigData++ aims to continue this success by deepening existing relationships and create new opportunities between researchers and industry. The project also aims to reach out further than this to feed in at decision level with policy makers and governments at national and international level. This is a bold objective that the project hopes to achieve during its next four-year phase.

During the latter stages of the SoBigData project, project leaders were looking into the possibility of webinars, video lectures, online assessments etc in order to have resources more readily available to users and also provide additional training materials for a more inclusive platform. As circumstances dictated, Covid-19 has meant that all events from March 2020 onwards have been held virtually. This has meant a huge shift in planning and organisation and a fundamental change to how events are delivered. Being a project with an innovative and flexible foundation has meant that all organisers and host venues have risen to the challenge and put on an impressive range of events that have generally been attended in the expected numbers, if not greater numbers than originally planned for.

Going forward, it is likely, that at least in the short term, events will continue to be held virtually. Depending on the travel recommendations from the World Health Organisation (WHO), some events later on in 2021 might be held face to face. In the meantime, the consortium will continue to assess the success and flexibility of online options. The impact of Covid-19 on a global level has demonstrated that the project can be both flexible and adaptable as well as being innovative and imaginative. The use of technology has not hindered events to the extent it could have, rather it has allowed involvement from those that are unable to travel and has provided a facility to watch interesting talks again, or to watch additional talks that had originally clashed with another. It has also allowed 'break off' sessions to occur in different 'rooms' meaning more involved or specific issues can be discussed by those that wish to.

One of the main challenges that Covid-19 has presented to WP3 is the collection of the data that we require in order to measure the success and impact of the project. As the events are held on a virtual platform, it has been harder to ascertain numbers and in particular the genders of participants as well as their stakeholder type. In order to participate in an event all that is generally required is a code to enter the virtual platform, so registration is not a necessity. Every effort has been made to obtain as much data as possible; however, it is acknowledged that this is an aspect that requires further planning and implementation to ensure any holes in data are minimised.

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.

GLOSSARY

EU	European Union
EC	European Commission
H2020	Horizon 2020 EU Framework Programme for Research and Innovation
WHO	World Health Organisation
OA	Open Access
AI	Artificial Intelligence
RI	Research Infrastructure

TABLE OF CONTENTS

1 Relevance to SoBigData++

The effective dissemination of the SoBigData++ project is of vital importance to its overall success and longevity. The more participants that become involved (whether they be individuals, research groups or companies), share their expertise, collaborate with project members and use the Research Infrastructure, the more successful the project has been.

1.1 Purpose of this document

This document is to provide the consortium with an overview of the dissemination activities carried out in the first period of the project, from M1 to M18, together with a report on the project's impact so far.

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.
- Exploration of initial ideas on generating impact from the project results on the target stakeholders.
- Establishment of the appropriate metrics and indicators in order to assess the success of the dissemination and impact of the project results.

1.2 Relevance to project objectives

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

The number of activities (despite the pandemic) that have already taken place demonstrate the project's commitment to disseminate as widely and as effectively as possible. The variety of activities available also serves to encourage a wide participation from data scientists as well as other stakeholders from diverse backgrounds. Activities also provide a meeting point that can initiate new collaborations, nurture existing partnerships and provide an effective networking platform.

1.3 Relation to other work packages

Dissemination and impact generation activities are horizontal activities that last for the entire duration of the project. Therefore, there is a clear interaction with all work packages. All work packages provide feedback, both for dissemination and impact on stakeholders.

On the other hand, WP3 provides feedback to the rest of the work packages in order to influence their work at multiple levels: assessing activities in the scientific and commercial spheres, drive activities in a correct and timely direction, suggesting changes on different components licensing to ensure the results complementarity, offering ideas on how small enhancements could be better perceived by the stakeholders. In the second half of the project, sustainability will be a particular focus.

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 provides a report of the dissemination activities that have taken place in the reporting period. In more detail, it provides dates in which event took place, partners who organized the event and a short summary of the event including details of the format. It also includes details of events that have been planned but have not yet taken place. The last part of section 2 looks at the results of the project indicators and assesses the progress and success of the dissemination activities.
- Section 3 of this document provides an overview of the project's impact of outreach towards policy makers and the public at large during the reporting period.
- Section 5 gives an update about the stakeholder analysis already published in D3.1.
- Section 5 of the document concludes with consolidated findings so far.

2 Dissemination Activities from M1 to M18

2.1 Introduction and the impact of Covid-19

Covid-19 has obviously had a huge impact on the project and has required many changes to the organisation and delivery of the various activities that have taken place. It has also been the reason for creating several of the events that have been organised promptly in response to the data opportunities afforded by the global pandemic. Many scientific papers have also been written in response to the pandemic.

Since the events have been delivered virtually data collection on participants has proved challenging. With only a link required to join in with the conferences, talks and workshops, age, gender and stakeholder type of the participants has not been collected for every event: therefore, the breakdown of data into these categories is not as comprehensive.

Mandatory forms have been introduced to formalise and simplify the collection of data for every event. The initial teething problems have been tackled and the collection of data will be more comprehensive as the project progresses. However, without a full registration system (which is not always appropriate or viable) data from some events has been ascertained using the information available.

2.2 Timeline of Events

The table below is a comprehensive list of the various events that have taken place from Month 1 to Month 18 of the project together with a brief description of the event.

Future events that have been planned are detailed in Table 11 further on this report.

The first word in the description denotes what type of event took place.

	Event	Date	Partner	Description
Month 2	**VI Scuola Nazionale di Chimica dell'Ambiente e dei Beni culturali**	3-6 Feb 2020	IMT	**School** devoted to students of Environmental Chemistry and Environmental Economics.
	XAI Tutorial 2020 (part of AAAI 2020)	8 Feb 2020	UNIPI	**Tutorial.** The future of AI lies in enabling people to collaborate with machines to solve complex problems. XAI (eXplainable AI) aims at addressing challenges such as trust, communication, and clarity.

	Event	Date	Partner	Description
Month 4	**Realtime Epidemic Datathon**	6 Apr – Jul 2020	ETH Zurich	**Datathon**. A collective open-source real-time forecasting challenge that aims to develop real-time and large-scale epidemic forecasting models.
Month 5	**Data Science Colloquium**	20 May - 8 Jun	SNS, CNR, IMT, UNIPI	**Colloquium** for first year PhD students to provide advice and support for final year projects. Also open to anyone interested.
	SoBigData++ Infrastructure Webinar	21 May 2020	CNR	**Webinar** to introduce the SoBigData RI. It explained how to integrate a new method, how to execute an experiment and how the method engine works. Furthermore, it presented the workspace as an instrument for the sharing of datasets and methods and the promotion of scientific collaborations.
Month 6	**Data Science in Techno-Socio-Economic Systems Online Workshop 2020**	10-11 Jun 2020	ETH Zurich	**Workshop** to enable research exchange and to share insights from different data-intensive disciplines. Aimed at students, academics, and practitioners in the field of data science for techno-socio-economic systems and quantitative finance.
Month 7	**Epidemics and the city: how human mobility and well-being changed during the COVID-19 era**	3 Jul 2020	IMT	**Webinar** discussing how the COVID-19 epidemic changed mobility habits and impacted on individuals' well-being and on the virus transmissibility. First webinar in a series from the perspective of Data Science and Environmental Epidemiology.
	1st SoBigData++ Awareness Panel: **Data Protection for Research and Statistical Purposes: Towards Legally Attentive Datathons**	22 Jul 2020	SSSA	**Awareness Panel** discussing various issues of working within the SBD++ framework and the risks, opportunities, and legal considerations of datathons.
Month 9	**XKDD Workshop** (part of ECML PKDD Conference)	14 Sep 2020	UNIPI, CNR & University of Warsaw	**Workshop** aiming to address important issues related to ethical, fair, explainable and transparent data mining and machine learning from a technical, legal, ethical, or sociological perspective.
	CMF 2020 (part of NetSci)	18 Sep 2020	IMT	**Workshop** covering the challenges facing today's financial and regulatory institutions which require multidisciplinary problem-solving that leverages methodologies from both complex network analysis and data science.

	Event	Date	Partner	Description
	ROME II (part of NetSci)	17 Sep 2020	IMT	**Workshop**. The aim of the **ROME** (**R**educing **O**nline **M**isinformation **E**xposure) satellite NetSci event was to convey state of the art research on the analysis and comprehension of the information system of OSNs (Online Social Networks).
	The 7th Satellite on Quantifying Success (part of NetSci)	17 Sep 2020	CNR	**Workshop** for researchers and scientists to exploit the increasing availability of large-scale datasets which provide an unprecedented opportunity to explore patterns underlying success in the emerging field of the "science of success".
	PSD 2020	23-25 Sept 2020	URV	**Conference**: 'Privacy in Statistical Databases' discusses the issues between the increasing societal and economical demand for accurate information and the legal and ethical obligation to protect the privacy of individuals and enterprises providing the data.
	Can Big Data Bridge Gaps in Migration Statistics?	29 Sep 2020	UNIPI & CNR	**Webinar**. Discussing the limitations of migration data which in turn causes impediments to sound social and economic policies specifically concerning migration and the potential utilisation of Big Data analytics for bridging these gaps.
Month 10	**SocInfo 2020**	6-9 Oct 2020	CNR	**Conference**: SocInfo is an interdisciplinary venue for researchers from Computer Science, Informatics, Social Sciences and Management Sciences to share ideas and opinions, and present original research work on studying the interplay between socially centric platforms and social phenomena.
	Disinfo 2020 (part of SocInfo)	6 Oct 2020	CNR, Tartu	**Workshop** on "Information Disorders: Fake News and Coordinated Inauthentic Behaviours" focuses on the study, modelling, and characterisation of all challenges related to mis- and dis-information, fake news, coordinated inauthentic behaviour and information operations.
	eXplainable Decision-Support Making (within DSAA 2020)	9 Oct 2020	CNR	**Tutorial.** GDELT news (including sentiment and social stability factors) can help in capturing month-to-month peace fluctuations and in revealing significant events. This could bring added value to researchers, policy-makers, and peacekeeping organisations, such as the UN and Red Cross, to facilitate the timely reaction on applying the right policies, prevent detrimental societal effects and contribute effectively to societal progress.
	Building Open Science Gateways to open and linked research outcomes (part of OpenAire Week)	16 Oct 2020	CNR	**Webinar** to present the OpenAIRE services that support research communities, initiatives, and infrastructures at implementing and monitoring the uptake of Open Science principles.

	Event	Date	Partner	Description
	Explaining Explanation Methods	30 Oct 2020	CNR, UNIPI	**Webinar** presenting the existing problems with eXplainable AI (XAI) (particularly in safety-critical or socially sensitive contexts) and the main strategies adopted to solve them.
Month 11	**'The SoBigData Research Infrastructure'** (part of EGI Conference)	2 Nov 2020	EGI	**Workshop** to present an overview of the important aspects a community must take into consideration in order to design and implement a distributed, pan-European, multi-disciplinary research infrastructure for big social data analytics such as SoBigData RI.
	2nd SoBigData++ Awareness Panel: **Research Infrastructure Platforms: Data Protection & IP Issues**	10 Nov 2020	SSSA	**Awareness Panel** to discuss the limits and the opportunities provided by the European data protection and intellectual property framework in respect to research over large datasets and to research platforms' interoperability.
	Machine Learning of Dynamic Processes and Time Series Analysis	26-27 Nov 2020	SNS	**School.** First school on Machine Learning of Dynamic Processes and Time Series Analysis. The purpose of this School is to present recent developments in Machine Learning focusing on dynamical systems. Applications will also be discussed, such as the forecasting of financial time series.
	SoBigData++ Video Presentation by Jesús A. Manjón at European Researchers Night	27-28 Nov 2020	URV	**Video Presentation.** Presented by Jesús A. Manjón to promote the SoBigData++ Project and Research Infrastructure.
Month 12 Month 12	**Explainable Machine Learning for Trustworthy AI** (part of Complex Networks)	2 Dec 2020		**Keynote speech** by Fosca Gianotti providing an insight into eXplainable AI and the limitations caused by the lack of transparency in the decisions made by black box AI systems.
	Big Mobility Data at work: a case of synergy between SoBigData and the Track & Know project	7 Dec 2020	CNR	**Webinar** introducing the context of the EU H2020 project "Track & Know" and, more specifically, three Big Mobility Data Analytics tools developed in T&K, the roots of which lie in the SoBigData project. This webinar will discuss how they evolved, their impact in real applications and the challenges still open.
	Expert Roundtable on Data And Ethics in a Post-Covid World (part of International Forum on Digital and Democracy (IFDad 2020)	10-11 Dec 2020	RIE	**Round Table.** This session was organised with the **SoBigData++ Consortium** and was invitation only. The discussion explored the ethical dilemmas around AI and Big Data in Ethics, Society, and Law, together with technical challenges in data science and Explainable AI, with a special emphasis on Trustworthiness.

	Event	Date	Partner	Description
	Evaluating the significance of network observables with a maximum entropy-based approach	18 Dec 2020	UNIROMA1	**Webinar** introducing the basic concepts around maximum entropy modelling and the proposed methodology. Also, the python package will be presented implementing the methodology under development, called "claude", showing several use cases and examples.
Month 14	**Data in soccer: an athletic trainer's point of view**	12 Feb 2021	UNIPI	**Webinar** focused on describing the advantages of using data science in soccer teams.
	3rd SoBigData++ Awareness Panel: **Medical Device Regulation and Digital Health: Problems and Perspectives**	15 Feb 2021	SSSA	**Awareness Panel** – discussing the issues and considerations surrounding Digital Health, AI, and digital medical devices.
Month 15	**A Scientific Experiment on Privacy and AI Transparency**	11 Mar 2021	BSC	**Workshop/Festival –** conducting an experiment on the state of privacy and AI transparency in Internet. The experiment engaged participants with current consent and privacy forms of popular webpages to assess their accessibility and actual degree of transparency. Discussions were encouraged, to identify general tips (e.g., keywords to look for, how much effort it implies, form traps and how to avoid them, etc.). If the results are of interest, a scientific paper will be produced.
	ECIR 2021	28 Mar - 1 Apr 2021	CNR	**Conference** on Information Retrieval (IR) – searching and organising unstructured information with a focus on encouraging participation from early-career researchers.
	ROMCIR 2021: Reducing Online Misinformation through Credible Information Retrieval (Part of ECIR 2021)	28 Mar – 1 Apr 2021	IMT	**Workshop** covering the importance of providing everyone with access to credible/verified information to mitigate the information disorder phenomenon.
Month 16	**Thematic Workshop on Migration and Big Data**	7 Apr 2021	Not a SBD++ Event	**Webinar:** The first in a series of talks on big data for migration studies with an interdisciplinary audience.
	Incontra Informatica	16 Apr 2021	UNIPI	**Workshop** for high school students who were shown how to read, pre-process, and analyse soccer match event data, showing some basic statistics to describe the performance of players and teams.

	Event	Date	Partner	Description
	Benchmarking and Survey of Explanation Methods for Black Box Models	28 April 2021	SNS	**Webinar** covering black-box models in Artificial Intelligence and the need for explanation methods to reveal how these obscure models reach specific decisions.
Month 18	**4th SoBigData++ Awareness Panel:** **Mobility data sharing: application potential and ethical issues**	10 Jun 2021	URV	**Webinar:** Covering 5 topics: 1. Studying population dynamics without compromising people's privacy. 2. Data Sovereignty in the connected vehicle. 3. The MobiDataLab project. 4. Decentralized anonymization of mobility data. 5. Mobility data reidentification opportunities and… risks.
	HumaneAI - AI & Society Roundtable	30 June 2021	UNIPI	The roundtable is designed as a collective intelligence exercise towards shaping the research questions of Social AI, driven by societal challenges. It is implemented through a structured conversation among inter-disciplinary scientists, looking at the relationship between AI and society from the perspectives of human-AI scientists and social scientists.

Table 1: List of Events that have taken place between M1 and M18

2.3 Periodic Training Planning Report Data

Training activities are an opportunity to disseminate the project, reach new stakeholders and grow the SoBigData community. Therefore, in this deliverable have been listed also the training events.

The collection of data concerning the numbers, ages, stakeholders, and genders of participants has not been as easy to collect due to the virtual nature of the events. Below are tables that show the events that did collect as much data as was practical or possible.

Event	Number of Female Participants	Number of Male Participants	Total number of participants
VI Scuola Nazionale di Chimica dell'Ambiente e dei Beni culturali			120
Realtime Epidemic Datathon	20	30	50
Data Science in Techno-Socio-Economic Systems Online Workshop 2020	60	140	300

Event	Number of Female Participants	Number of Male Participants	Total number of participants
Epidemics and the city: how human mobility and well-being changed during the COVID-19 era			70
1st SoBigData++ Awareness Panel: **Data Protection for Research and Statistical Purposes: Towards Legally Attentive Datathons**			34
XKDD Workshop (part of ECML PKDD Conference)	63	150	213
Complexity Meets Finance: Data, Methods and Policy Implications (part of NetSci)			60
ROME II (part of NetSci)			30
The 7th Satellite on Quantifying Success (part of NetSci)			30
PSD 2020	30	42	72
Can Big Data Bridge Gaps in Migration Statistics?	18	27	45
SocInfo 2020	39	83	127
Building Open Science Gateways to open and linked research outcomes (part of OpenAire Week)	70	39	109
Explaining Explanation Methods	10	38	48
The SoBigData Research Infrastructure (part of EGI Conference)	15	17	32
2nd SoBigData++ Awareness Panel: **Research Infrastructure Platforms: Data Protection & IP Issues**	9	9	18
Machine Learning of Dynamic Processes and Time Series Analysis	60	208	268
Big Mobility Data at work: a case of synergy between SoBigData and the Track & Know project			30
Expert Roundtable on Data and Ethics in a Post-Covid World (part of International Forum on Digital and Democracy (IFDad 2020)			45 (Invite only)
Evaluating the significance of network observables with a maximum entropy-based approach	18	12	30
Data in soccer: an athletic trainer's point of view	5	40	45

Event	Number of Female Participants	Number of Male Participants	Total number of participants
3rd Awareness Panel: **Medical Device Regulation and Digital Health: Problems and Perspectives**	10	15	25
ECIR 2021	346	757	1103
Thematic Workshop on Migration and Big Data	30	25	55
AISC (Aggregate Intellect Science)	20	80	100
A Scientific Experiment on Privacy and AI Transparency	15	15	30
Incontra Informatica	10	17	27
XDSM Tutorial 2020 (part of DSAA 2020)	50	50	100
XAI Tutorial 2020 (part of AAAI 2020)	50	50	100
4th Awareness Panel: **Mobility data sharing: application potential and ethical issues**	10	15	25
HumaneAI - AI & Society Roundtable	35	75	110
Totals	993	1934	3451

Table 2: The number and gender breakdown of participants at each event

The chart below (Figure 1) shows the split between male and female participants in this first period of the project. The split remains largely two thirds male and one third female – this is the general trend for many of the events.

However, gender was impossible to collect for just over 10% of participants (419 out of 3451), and thus Figure 1 (below) provides an accurate statistic of the gender divide between male and female participants.

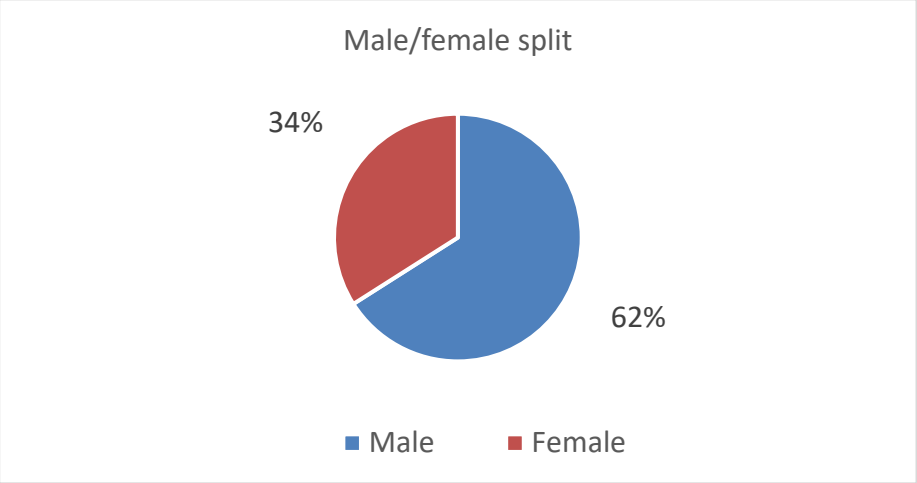


Figure 1: Showing the split between Male and Female participants in M1 to M18 of the Project

The issues regarding the collection of data have already been detailed. The forms that were introduced to collect data on genders, numbers and stakeholders also asked for the age ranges of participants. A range was provided in each of the four categories – detailed below:

- Under 18s
- between 18-30 years of age
- between 30-50 years of age
- over 50 years of age

As there is no requirement for a formal registration, organisers were tasked with collecting (to the best of their ability) estimates of how many people were in each range. Obviously, this method has many flaws, but it is useful as a guide as to which age range the project reaches and appeals to. It is perhaps worth noting here that there were not many under 18s participating in this first period of the project. This is due to the various Covid-19 restrictions that have meant that many events aimed at school children could not take place.

The ranges provided for ascertaining the number of participants in each range are:

- 1-5 participants
- 6-10 participants
- 11-15 participants
- 16-20 participants
- 21-30 participants
- 31-40 participants
- 41-50 participants
- Over 50 participants

The table below gives the estimated number of participants in each age range for the events that were able to collect this data.

Event	Under 18 years of age	Between 18-30 years of age	Between 30-50 years of age	Over 50 years of age
Epidemic Datathon	0	50+	0	0
Data Science in Techno-Socio-Economic Systems Online Workshop 2020	0	41-50	50+	41-50
SocInfo 2020	0	41-50	50+	11-15
Explaining Explanation Methods	1-5	16-20	16-20	1-5
'The SoBigData Research Infrastructure' (part of EGI Conference)	0	6-10	21-30	6-10
2nd SoBigData++ Awareness Panel: Research Infrastructure Platforms: Data Protection & IP Issues	0	0	1-5	11-15
IFDAD 2020	0	11-15	31-40	21-30
Evaluating the significance of network observables with a maximum entropy-based approach	0	11-15	11-15	1-5
Data in soccer: an athletic trainer's point of view	0	11-15	11-15	1-5
3rd SoBigData++ Awareness Panel: Medical Device Regulation and Digital Health: Problems and Perspectives	0	1-5	11-15	6-10
ECIR 2021	0	50+	50+	50+
ROMCIR (part of ECIR 2021)	0	21-30	21-30	21-30
Thematic Workshop on Migration and Big Data	0	16-20	21-30	6-10

Event	Under 18 years of age	Between 18-30 years of age	Between 30-50 years of age	Over 50 years of age
A Scientific Experiment on Privacy and AI Transparency	0	6-10	11-15	6-10
Incontra Informatica	21-30	1-5	0	0
HumaneAI - AI & Society Roundtable	0	21-30	31-40	31-40

Table 3: Breakdown of age groups in selected events

As definitive totals cannot be achieved in each of the 4 age ranges, the lowest number in each range has been added together to provide cursory totals. These totals have been used to create the visual chart below to give as clear a picture as possible of the division of the participants' age ranges.

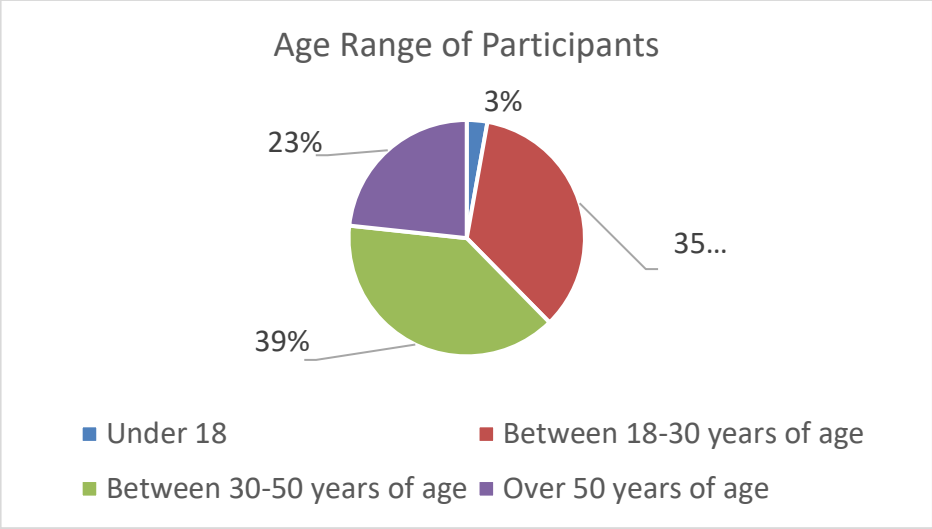


Figure 2: The split of the age ranges from M1 to M18 of the project

For the project, it is relevant to understand who participates in its events and who its message reaches. The table below shows the stakeholder breakdown. Again, the collection of data proved challenging and as such only the events that managed to collect data are included in the table.

Event	Type of participants						
	Teaching/ Academic Institutions			Early-career Researchers	Data Analysts	Industry	Policy Makers
	Undergrads	PhD Researchers	Academics				
Realtime Epidemic Datathon	50+						
Data Science in Techno-Socio-Economic Systems Online Workshop 2020	31-40	41-50	50+	41-50		50+	
SocInfo 2020	1-5	41-50	31-40	31-40		1-5	
Building Open Science Gateways to open and linked research outcomes (part of OpenAire Week)	1-5	6-10	11-15				1-5
Explaining Explanation Methods		11-15	6-10	6-10			
'The SoBigData Research Infrastructure' (part of EGI Conference)	1-5	16-20	16-20	21-30		6-10	6-10
2nd SoBigData++ Awareness Panel: **Research Infrastructure Platforms: Data Protection & IP Issues**			6-10	1-5		1-5	
IFDAD 2020	11-15	16-20	16-20	21-30		16-20	11-15
Evaluating the significance of network observables with a maximum entropy-based approach		6-10	6-10	6-10			
Data in soccer: an athletic trainer's point of view	1-5	6-10	1-5	1-5		6-10	
3rd Awareness Panel: **Medical Device Regulation and Digital Health: Problems and Perspectives**	1-5	1-5	6-10	6-10		1-5	1-5
ROMCIR (part of ECIR 2021)	16-20	16-20	16-20	16-20			
Thematic Workshop on Migration and Big Data	11-15	11-15	16-20	11-15	1-5		

Event	Type of participants						
	Teaching/ Academic Institutions			Early-career Researchers	Data Analysts	Industry	Policy Makers
	Undergrads	PhD Researchers	Academics				
HumaneAI - AI & Society Roundtable	NA	NA	NA	NA	NA	NA	NA

Table 4: Types of participants breakdown of selected events

Figure 3 (below) gives a visual representation of the participants range. The numbers used to produce this chart have been taken from Table 4 above. As previously, the lowest numbers in each range were added together. The chart demonstrates that the spread between the 4 main academic realms – Undergrads, PhD Researchers, Academics and Early Career Researchers is fairly even. The other two stakeholders being individuals working in industry and policy makers show a 12% and a 2% participation respectively. These figures are encouraging and are in line with the project's objectives.

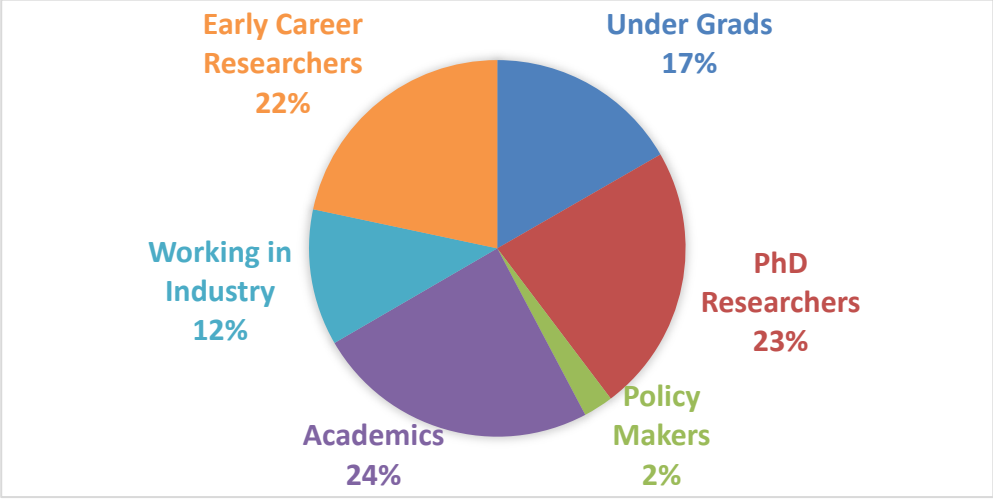


Figure 3: The split of the participants ranges from M1 to M18 of the project

2.4 Detailed Report on Selected Training Activities

The various events that have taken place have been split into the following categories:

- Webinars

- Conferences
- Awareness Panels
- Other Events (Datathons, Colloquiums, Workshops, Festivals)
- SBD++ Flagship Events

Below are short descriptions of the events that have taken place and where applicable there is a link to the event's website or the recording of the event.

2.4.1 *Webinars*

2.4.1.1 EPIDEMICS AND THE CITY: HOW HUMAN MOBILITY AND WELL-BEING CHANGED DURING THE COVID-19 ERA

Objectives: This was a dissemination webinar that was promptly organised to capture the threads of the topical Covid-19 crisis and the impact it has had on society from the perspective of Data Science and Environmental Epidemiology. It took place on 3 July 2020. Experts discussed the effect of Covid-19 on mobility, the impact on people's well-being and on virus transmissibility as well as the quality of life of citizens. It also looked at co-benefits in relation to lockdown.

Participants: The webinar was aimed at experts, stakeholders of the SoBigData++ project and it was also open to the general public. It was advertised via the usual SoBigData++ channels of social media, newsletters and mailing lists and was also promoted through IMT where the event was organised. Approximately 70 people were involved, and the event was well received.

Link: http://sobigdata.eu/events/epidemics-and-city-how-human-mobility-and-well-being-changed-during-covid-19-era

2.4.1.2 CAN BIG DATA BRIDGE GAPS IN MIGRATION STATISTICS?

Objectives: This webinar was organised by UNIPI and CNR and was born from the Horizon2020 HumMingBird project (https://hummingbird-h2020.eu/) which investigates human migration. It took part on 29 September 2020. The event was aimed at experienced and non-expert researchers and academics interested in studying human migration, but coming from different research fields, e.g., sociology, computer science, and demography. The usual methods of advertising were utilised. Various official mailing lists of research institutes and universities, SoBigData ++, HumMingBird were used, and it was also promoted via Twitter (including tags to the SoBigData ++ profile and link to the event).

Participants: It was expected that this event would attract participation of around 30 people. However, 45 individuals accessed this webinar; of those 27 were male and 18 were female. The number of females is over a third which shows an improving ratio for encouraging more women into the realm of Data Science. The number of participants exceeded the expected number by 50%, which is a convincing result and

demonstrates the success of the project's outreach. The organisers were very satisfied with the level of participation and furthermore, various interesting questions and discussions arose at the end of the presentation for further investigation.

Link: https://www.youtube.com/watch?v=ROQzO33KSnA

2.4.1.3 BUILDING OPEN SCIENCE GATEWAYS TO OPEN AND LINKED RESEARCH OUTCOMES (PART OF OPENAIRE WEEK)

Objectives: These Webinars were created to showcase what OpenAire can offer research communities and infrastructures and took place from 12 – 16 October 2020. One of the Webinars presented how OpenAire are helping to build a system for impact monitoring for SoBigData++.

Participants: This event was well attended by approximately 109 participants, the majority of whom were female (about 70%). The event was aimed at Teaching/Academic Institutions, researchers, Law & Policy makers and Public Administration. The majority of the participants were academics and researchers but the event also attracted policy makers.

2.4.1.4 EXPLAINING EXPLANATION METHODS

Objectives: This webinar which took place virtually on 30 October 2020. The lack of transparency on how AI systems make decisions is a clear limitation in their adoption in safety-critical and socially sensitive contexts. Consequently, research in eXplainable AI (XAI) has recently caught much attention, with specific distinct requirements for different types of explanations for different users. In this webinar, the existing explanation problems were presented together with the main strategies adopted to solve them.

Participants: This event attracted 48 participants, of which 10 were female. This equates to just over 20% female participation. Most participants were between 18 and 50 years of age and came from a teaching/academic and research background.

2.4.1.5 BIG MOBILITY DATA AT WORK: A CASE OF SYNERGY BETWEEN SOBIGDATA AND THE TRACK & KNOW PROJECT

Objectives: This webinar took place on 7 December 2020 and was created to discuss the EU H2020 project "Track & Know (T&K)" and its collaboration with the SoBigData++ project. More specifically, the aim was to look at three Big Mobility Data Analytics tools that had been developed during the T&K project and to learn about their roots in the SoBigData project; how they evolved thanks to the contact with concrete project pilots; their impact in real applications and the challenges still open.

Participants: The Webinar had 2 speakers from ISTI-CNR - Luca Pappalardo and Mirco Nanni. It attracted 30 participants – although no gender/age or stakeholder breakdown could be recorded. The Webinar was aimed at Teaching/Academic Institutions, Research, Data Analysts, and Industry. This is an example of

SoBigData++'s collaborations with other projects which are born from a mutually beneficial objective to support and advance the Data Science industry as a whole and to promote the ethical and responsible use of data (being is one of the main objectives of the SoBigData++ project).

2.4.1.6 EVALUATING THE SIGNIFICANCE OF NETWORK OBSERVABLES WITH A MAXIMUM ENTROPY-BASED APPROACH

Objectives: This Webinar took place on 18 December 2020 and was presented by Enrico Maiorino, from Brigham Women's Hospital, Harvard University, Boston. It was about how to statistically validate observed results on Networks. Maiorino introduced the basic concepts around maximum entropy modelling and the proposed methodology. He also presented the python package implementing the methodology which he is developing, called "claude", showing several use cases and examples.

Participants: This workshop was an hour long based on a 50-minute presentation and a 10-minute Q&A session. There were 30 participants, 18 female and 12 males. Approximately half the participants were between 18-30 years of age, between 11 and 15 were between 30 and 50 years old and between 1 and 5 was over 50. The webinar was aimed at the Teaching/Academic and Research sector and Data Analysts. All participants were either PhD Researchers, Early Career Researchers or Academics.

Link: https://www.youtube.com/watch?v=MevkCj2H1Qg

2.4.1.7 DATA IN SOCCER: AN ATHLETIC TRAINER'S POINT OF VIEW

Objectives: This Webinar took place on 12 February 2021 and introduced the concept of using data to tailor training programmes for athletes and maximise the effectiveness of the training schedule. Cristoforo Filetti (athletic trainer from Paris Saint-Germain) spoke for 30 minutes about how he uses data during workdays in order to schedule a training program. After the presentation, there was a 15-minute Q&A session where the moderators of the event asked questions about his work. The Webinar was advertised via SoBigData++ contacts and the Twitter account, the LinkedIn account of one of the organisers and was also promoted via involved soccer experts in order to widen the outreach to the relevant audience. The recording can be seen via the SoBigData++ YouTube channel posted on both the SoBigData++ website and Twitter account.

Participants: The webinar itself had 40 participants and 88 people have watched the recording on the SoBigData++ YouTube channel. During the live event there were 35 male and 5 female participants. They were mostly between 18 and 50 years old and either in the academic world, or sports coaching realm. The SoBigData++ Infrastructure was referenced during the event.

Link: https://youtu.be/40PAAPum9V4 | https://www.youtube.com/watch?v=40PAAPum9V4&t=4s

2.4.1.8 AISC (AGGREGATE INTELLECT SCIENCE) - BENCHMARKING AND SURVEY OF EXPLANATION METHODS FOR BLACK BOX MODELS

Objective: Organised by SNS, this webinar took place online of 28 April 2021. It was a 45-minute event with a 30-minute presentation of recent work on artificial intelligence followed by a Q&A session of approximately 15 minutes. The widespread adoption of black-box models in Artificial Intelligence has enhanced the need for explanation methods to reveal how these obscure models reach specific decisions. Retrieving explanations is fundamental to unveil possible biases and to resolve practical or ethical issues. Nowadays, the literature is full of methods with different explanations. This study shows a visual comparison among explanations and a quantitative benchmarking of various explainers. Participants were very interested in the topic asking multiple questions both during the Q&A session and after the event by email. Requests for follow ups and details about future works were also received.

Participants: There were approximately 100 participants, of which 80 were male and 20 were female. No further breakdown of ages, stakeholders etc. was collected. The event was aimed at teaching and academic institutions, researchers and from industry.

Link: https://aisc.ai.science/

2.4.2 *Conferences*

2.4.2.1 NETSCI 2020

NetSci 2020 is an annual conference of the Network Science Society, which aims to bring together leading researchers and practitioners working in the emerging research area of network science. The NetSci conference fosters multi-disciplinary communication and collaboration in network science research across computer and information sciences, physics, mathematics, statistics, the life sciences, medicine, food science, neuroscience, environmental sciences, social sciences, finance and business, arts and design. Within the NetSci Conference were 3 Satellites organised by SoBigData++; *CMF 2020, ROMEII* & *The 7th Satellite on Quantifying Success.*

2.4.2.2 CMF 2020 - COMPLEXITY MEETS FINANCE: DATA, METHODS AND POLICY IMPLICATIONS

Objectives: This satellite aimed to bridge the gap between the fields of complex networks theory and finance by bringing together young scholars, experienced researchers and policy makers from central banks interested in interdisciplinary research to discuss state-of-the-art work, share knowledge and create opportunities for novel and fruitful collaborations. The event also intended to bring cutting-edge academic research in contact with industry experience and impact. Therefore CMF 2020 was open to researchers, scholars, industry stakeholders and policy makers demonstrating SoBigData++'s aim to reach out and mesh the academic realm with policy makers and the economic world.

Participants: There were 7 invited speakers, of which 2 were female. There were also 4 contributed speakers, of which 1 was female. The event attracted just under 60 participants, although being an online event we do not have a breakdown of the participants into gender. The usual channels of social media and the SoBigData++ newsletters were used to advertise the event and it was also promoted via the main NetSci conference.

Website: https://sites.google.com/view/cmf20/home

2.4.2.3 ROME II (REDUCING ONLINE MISINFORMATION EXPOSURE)

Objectives: The aim of the **ROME II** satellite is to convey state of the art research on the analysis and comprehension of the information system of OSNs (Online Social Networks). OSN's role in shaping the political debate is crucial: misinformation in OSNs can distort and divert public discourse. The spread of hoaxes, propaganda, and rumours impacts different areas of social interest, such as political elections, public health, and social tensions. It is therefore necessary to develop the proper tools in order to detect the various facets with which the false information spreads on the web and to measure its effect and pervasiveness. The speakers were asked to pre-record their presentations and be online during the satellite to answer questions and take part in discussions. The event was advertised through the usual social media channels and via promotion from the main conference.

Participants: There were 4 invited speakers, of which 3 were female. There were also 4 contributed speakers, of which 1 was female. The Satellite was aimed primarily at academia – in particular - PhD Students and researchers and Data Analysts. The event attracted approximately 30 participants. The satellite progressed smoothly, and the participants were able to enjoy an informative and instructive event.

Website: https://sites.google.com/view/cmf20/home

2.4.2.4 THE 7TH SATELLITE ON QUANTIFYING SUCCESS

Objectives: This satellite event took place on 17 September 2020. It was aimed at computational social scientists interested in understanding the relationships between performance and success in several contexts, from scientific publication to sports, cinema, and writing. It was advertised mainly via social media posts on Twitter and LinkedIn, and it was also emailed to various mailing lists of European projects and universities.

Participants: there were five invited talks and five contributed talks. Of the invited talks, 3 were male and 2 were female speakers. There was also a good mix of male and female speakers for the contributed talks. The speakers were all asked to record their presentation in advance and answer questions 'live' from the audience during the event. The event attracted approximately 30 participants.

Website: http://www.onurvarol.com/netsci20-qs7/

2.4.2.5 THE SOBIGDATA++ INFRASTRUCTURE (PART OF THE EGI CONFERENCE)

Objectives: The workshop presented an overview of the important aspects a community must take into consideration in order to design and implement a distributed, pan-European, multi-disciplinary research infrastructure for big social data analytics such as SoBigData RI. This workshop was a dissemination event organised into 4 talks:

1) Data Science, Multidisciplinary & AI
2) Ethics & Privacy
3) Training the next generation of Data Scientists for Social Good
4) TagMe: A success story of how the integration boosted a research result

Participants: There were 32 participants, of which 15 were female and 17 were male demonstrating an equal split. Most participants were between the ages of 30 and 50 with a small number being under 18 or over 50. Most of the participants (between 21-30 individuals) were early career researchers and most of the remaining participants were from an Academic background being either in teaching positions or PhD Researchers. However, this event also attracted a significant number of individuals who were classed as policy makers (between 6-10) and a number from industry (between 6 -10). This is yet another event where SoBigData++'s reach is demonstrably beyond the realm of Academia and is pushing into industry and attracting the attention of policy makers. This is where the project will find its most influential bearing and where the most impact can be attained.

Website: https://indico.egi.eu/event/5000/overview

2.4.2.6 DATA AND ETHICS IN A POST-COVID WORLD (PART OF IFDAD 2020 - INTERNATIONAL FORUM ON DIGITAL AND DEMOCRACY)

Objectives & Participants: This was an expert roundtable discussion which took place on 10 December 2020 and was invitation only. The discussion involved 45 people including the speakers. Among the experts speaking were Fosca Giannotti and Dino Pedreschi of SoBigData++ and Erika Widegren (the chief executive of RIE - one of SoBigData++'s partners). The expert panel included Fosca Giannotti, Jeroen Van den Hoven, Dino Pedreschi, Jospeh Domingo, Caterina Sganga, Juan M Duran and Kalina Bontcheva from SoBigData++. This session was organised with the SoBigData++ Consortium to explore the ethical dilemmas around AI and Big Data in Ethics, Society, and Law and the technical challenges in data science and Explainable AI, with a special emphasis on Trustworthiness. The Roundtable's aim was to bring together the SoBigData++ experts and some of the policymakers that are involved in the parliamentary commissions (e.g., AIDA) that are shaping the new European regulations on data and artificial intelligence. The outcomes of the discussions will help develop a white paper on the topic to be published in early 2021. This is a clear example of how the SoBigData++ project is involved with policy making at its highest level and sharing its expertise to help shape the new European regulations on data and artificial intelligence. The discussion was both productive and meaningful.

2.4.2.7 ECIR 2021 (EUROPEAN CONFERENCE ON INFORMATION RETRIEVAL)

Objective: This 5-day event took place virtually between 28 March – 1 April 2021. It was organised by CNR and was aimed at teaching and academic institutions, researchers, and industry. Although the conference had to convert to a virtual format it maintained the same structure of previous ECIR conferences. The first day was devoted to tutorials and the Doctorial Consortium. The next 3 days were occupied by the main conference and the last day consisted of satellite workshops and the industry day. On the first day, two full-day (FD) tutorials and six half-day (HD) tutorials were offered. The three days of the main conference featured 3 keynote talks, 50 full paper presentations (out of 211 submissions, for a 23.7% Acceptance Rate(AR)), 11 presentations from the Reproducibility Track (with a 47.8% AR), 7 presentations of papers recently published on the Information Retrieval Journal, 12 short-paper presentations of the CLEF 2021 Labs, 39 poster presentations of short papers (with a 28.5% AR), 15 demos (with a 48.4% AR), and a panel on the theme of "OpenAccess and IR Literature". Finally, the last day of the conference hosted the Industry Day and five satellite workshops.

Participants: This conference was aimed at researchers from academia and industry and the anticipated number of participants was approximately 200. Converting the conference to a virtual event meant the fee structure needed to be reconsidered. It was decided to charge a flat nominal fee of €150 to authors only and to grant a free registration to all other attendees. It was hoped this decision would maximise participation – especially from developing countries. This strategy proved effective with over 1,100 registrations from 63 countries. Of this number, 1,028 participants logged into the conference at least once. This jump in numbers represents a five-fold increase on expected numbers meaning the conference reached a far larger number of individuals. Of the number of 1103 registrations, 346 were females, 757 were male. This is approximately 31% female and 69% male following a similar trend of other SoBigData++ events.

Link: https://www.ecir2021.eu/

Figure 4: Screenshot of the OpenAccess and IR Literature panel

2.4.3 *Awareness panel*

2.4.3.1 1ST SOBIGDATA++ AWARENESS PANEL: DATA PROTECTION FOR RESEARCH AND STATISTICAL PURPOSES: TOWARDS LEGALLY ATTENTIVE DATATHONS

Objectives: this Webinar took place on 22 July 2020 and was organised by Scuola Superiore Sant'Anna (SSSA) in Pisa, Italy as an open event. Details of the event were circulated via email to the SoBigData++ mailing list and the link to the event was freely accessible. The program included 3 short talks from experts in the field followed by a Q&A session. The topics covered were:

1. 'Data Processing for Scientific Research and Statistics and the SoBigData++ Framework'
2. 'Datathons: Risks & Opportunities'
3. 'Legally Attentive Datathons: Ready for the Check list'

Participants: Two of the 3 speakers were female demonstrating one of the main aims of the project - to showcase females in Data Science with the purpose of inspiring and encouraging more females into this research sector. The event was attended by 34 participants, although no breakdown into gender was available.

2.4.3.2 2ND SOBIGDATA++ AWARENESS PANEL: RESEARCH INFRASTRUCTURE PLATFORMS: DATA PROTECTION & IP ISSUES

Objectives: This awareness panel took place on 10 November 2020 and was an opportunity to discuss the European data protection and intellectual property framework in respect of research over large datasets and research platforms' interoperability and the principles under article 5 of the GDPR law. Also, the barriers and limitations that exist under Intellectual property law (especially resulting from hybrid or overlapping IP regimes) will be highlighted. The panel was led by experts drawing attention to and discussing the considerations that researchers need to give to the ethical processing of data and the research exceptions that relate to text and data mining. The 4 topics discussed are as follows:

1. Program Prof. Dr. Giovanni Comandè (SSSA): 'Data Protection and Research: Market Innovation and Protection of Data Subjects' Fundamental Rights'
2. Dr. Giulia Schneider (SSSA): 'The GDPR's Research Exception and the incidence of data protection principles on data sharing practices'
3. Prof. Dr. Caterina Sganga (SSSA): 'Overview of IP-relevant aspects of research infrastructures and online practices'
4. Dr. Giulia Priora (SSSA): 'A closer look at recent European developments: TDM exception and Open Access national policies'

Participants: there were 18 participants who joined for the training session – 9 male and 9 female. The session was aimed at Teaching/Academic institutions and Industry. There were a few early career researchers (between 1 & 5), a few from industry (between 1 & 5) with the majority being Academics (between 6 &10).

Link: https://www.youtube.com/watch?v=NOQVBzsmLeo

2.4.3.3 3RD SOBIGDATA++ AWARENESS PANEL: MEDICAL DEVICE REGULATION AND DIGITAL HEALTH: PROBLEMS AND PERSPECTIVES

Objectives: The 3rd Awareness panel organised by SSSA was held on 15 February 2021 and concerned the issues and considerations that arise regarding the use of digital devices for health purposes. Digital Humanities is an emerging field that brings together social scientists and computer scientists to encourage collaboration and allow research on a far larger scale than ever before. However, alongside advancements in technology there needs to be a parallel advancement in the regulations, ethical decision making and a robust legal framework to protect patients' privacy. This panel raised these problems and perspectives and opened up a discussion that proposed interesting questions.

Participants: There were approximately 15 participants, mostly from a teaching/academic background including PhD Researchers, early career researchers and academics. The event also attracted individuals working in industry and policy makers. This is another example of where the SoBigData++ project crosses over from the academic realm to the real world; share its knowledge, expertise and be at the forefront of providing the academic background of policies and regulations.

Link: https://www.lider-lab.it/2021/02/12/third-sobigdata-awareness-panel-medical-device-regulation-and-digital-health-problems-and-perspectives-3/

2.4.3.4 4TH SOBIGDATA++ AWARENESS PANEL: MOBILITY DATA SHARING: APPLICATION POTENTIAL AND ETHICAL ISSUES

Objective: The 4th SoBigData++ Awareness Panel was about the application potential and ethical issues of mobility data sharing. Four experts exchanged their thoughts on this topic and a new European project about mobility data sharing was also presented. The topics covered were:

1. Decentralized anonymization of mobility data.
2. Data Sovereignty in the connected vehicle.
3. The MobiDataLab project.
4. Studying population dynamics without compromising people's privacy.
5. Mobility data reidentification opportunities and… risks.

The presentations covered different aspects of mobility data sharing and drove to an interesting discussion about the need to protect mobility data – even though this data is not currently considered to be particularly sensitive.

Participants: There were 25 participants in this Awareness Panel - with 10 female and 15 males. The event was aimed at Research and Data Analysts. It was advertised on Twitter, via the email contacts and on the SoBigData++ website.

Link: https://crises-deim.urv.cat/SBD_awareness_panel/

2.4.4 *Other Events (Datathons, Colloquiums, Workshops, Tutorials, Festivals)*

2.4.4.1 VI SCUOLA NAZIONALE DI CHIMICA DELL'AMBIENTE E DEI BENI CULTURALI

Objectives: This event was one of the few this year that was a face-to-face event as it occurred in February 2020 before the international lockdowns. It was organised by IMT and was held at the University of Siena, Italy. This school was dedicated to the analysis of general knowledge and chemical-environmental knowledge relating to the study of the climate system, the role of anthropic contribution to the greenhouse effect and its consequences on the distribution of chemicals, ecosystems and the techno sphere.

The link between environmental chemistry and other disciplines was addressed in order to identify the survey methodologies and techniques that can be adopted both in the context of mitigation and adaptation to climate change. This is the first edition of the school dedicated to climate change, which today represents the main challenge factor for the sustainability of human activities on planet Earth.

Participants: This event was aimed at masters and PhD students and post doc researchers and was advertised through the official channels of the University of Siena, IMT and Società Chimica Italiana. It attracted approximately 120 participants who were from different Italian and EU Universities. It was a 5-day programme encompassing around 20 hours of lectures from experts in the fields of environmental chemistry, ecological economics, and sustainability studies. The event was very highly successful.

Link: http://www.scuolacabc.it/

2.4.4.2 XAI TUTORIAL 2020 AT AAAI 2020

Objective: this Tutorial was organised by UNIPI and took place on 8 February 2020. The objective of the tutorial was to provide answers to the following questions: What is XAI? Why should we care? How is it critical? How does it work? What did we learn? What is next?

The future of AI lies in enabling people to collaborate with machines to solve complex problems. Like any efficient collaboration, this requires good communication, trust, clarity and understanding. XAI (eXplainable AI) aims at addressing such challenges by combining the best of symbolic AI and traditional Machine Learning. This topic has been studied for years by many different communities of AI, with different definitions, evaluation metrics, motivations, and results. This tutorial provided a snapshot of the work of XAI to date and surveys the work achieved by the AI community with a focus on machine learning and symbolic AI related approaches. The needs of XAI in real-world and large-scale applications was explained and state-of-the-art techniques and best practices were presented.

Participants: No specific data was collected; however, it was estimated that there were approximately 100 participants with an even split between male and female. The tutorial was aimed at Teaching/ Academic Institutions, Research, and Industry.

Link: https://xaitutorial2020.github.io/

2.4.4.3 REAL-TIME EPIDEMIC DATATHON.

Objectives: ETH Zurich seized the opportunity to work with the real time data of the unfolding Covid-19 pandemic and provide an opportunity for data science researchers to work on a collective open-source real-time forecasting challenge. The Datathon was open to everyone (either as an individual, or a team) and ran from April 2020 to July 2020. Publicly available data was used, and data scientists were encouraged to contribute to the global open-source scene by releasing real-time epidemic forecasting models.

Participants: the event had 37 participants and although no formal registration was taken to record genders and ages, there were approximately 8 females and 29 male participants.

Link: https://www.epidemicdatathon.com

2.4.4.4 DATA SCIENCE IN TECHNO-SOCIO-ECONOMIC SYSTEMS ONLINE WORKSHOP 2020

Objectives: ETH Zurich also ran a Workshop between 10 and 11 June 2020. This Workshop was open to students, academics, and practitioners in the field of data science for techno-socio-economic systems and quantitative finance. Data science has emerged as a universal tool to study different systems and domains such as finance, economics, social sciences, and complex systems. The aim of this workshop was to enable research exchange and to share insights from different data-intensive disciplines. For this purpose, the workshop brought together a unique blend of world-class experts from industry and academia for an intensive two-day immersion into their domains.

Participants: The workshop was planned for approximately 300 participants. There were actually 348 registered participants, and the event was well received. The organisers did adjust the schedule for the online format from full days to half days to ensure the participants had a more positive experience. Approximately two thirds of the participants came from America (224) and the other third from Europe (101). The event also attracted a small number from Asia (13), the Middle East (4) and South America (4). Most of the participants were from academia: 209 masters or post-doc students, 17 professors and 5 lecturers. However, there were approximately 89 participants from industry and 33 who registered as CEOs or Company Directors. This demonstrates how the project is reaching out further than academia to individuals who have influence in the financial sector. Another of the project's aims is to encourage more participation from female data scientists. Participation in this event was approximately two thirds male (approximately 230+) and one third female (approximately 100).

Link: https://www.eth-courant-workshop.com

2.4.4.5 EXPLAINABLE AI XKDD WORKSHOP

Objectives: This workshop was the last event of a series of initiatives that the research group has organised in the last 5 years. It took place on 14 September 2020. The purpose of XKDD, eXplaining Knowledge Discovery in Data Mining, is to encourage principled research that will lead to the advancement of explainable, transparent, ethical, and fair data mining and machine learning.

Participants: This event was open to the research community on Machine Learning and Data Mining. It was advertised through the usual channels - mailing lists, Twitter and via other relevant projects as well as SoBigData++, namely AI4EU and ERC XAI. Also linking with the main conference ECML/PKDD meant that the XKDD workshop had a wide visibility within the community and the event attracted 213 participants – far more than had been originally planned for. Although there is no specific data on the gender breakdown it has been estimated that the participants were 30% female and 70% male. This ratio is a common thread throughout this project and shows a solid starting point for pushing the number of females up even further. There were 2 invited speakers, 7 accepted papers involving numerous authors and the event encompassed a four-hour program. The event went smoothly and generated useful questions and proposals for future research. As a direct result of the workshop there was a commitment to set up a communication group to work with the research issues which were raised and community building actions that are going to be implemented.

Link: https://kdd.isti.cnr.it/xkdd2020/

Figure 5: A screenshot of the speakers at the XKDD Workshop

2.4.4.6 XDSM TUTORIAL 2020

(Part of DSAA 2020 (The 7th IEEE International Conference on Data Science and Advanced Analytics))

Objective: This event was organised by UNIPI and was held on 9 October 2020. The purpose of the XDSM Tutorial was to illustrate state-of-the-art approaches for explainable data mining and interpretable

machine learning. To establish and consider the problems, issues, and current challenges and to encourage principled research that will lead to the advancement of explainable, transparent, ethical, and fair data mining and machine learning.

Participants: No specific data was collected on numbers and stakeholders, but approximately 100 individuals participated with an estimated even split between male and female. The event was aimed at Teaching/Academic Institutions, Research, Data Analysts, and Industry.

Link: http://dsaa2020.dsaa.co/tutorials/

2.4.4.7 SCHOOL: MACHINE LEARNING OF DYNAMIC PROCESSES AND TIME SERIES ANALYSIS

Objectives: This event took place on 26-27 November 2020 and was the first School on Machine Learning of Dynamic Processes and Time Series Analysis. The purpose of the school was to present recent developments in Machine Learning focusing on dynamical systems.

The school had 4 plenary speakers each giving a 2.5-hour lecture. Of the invited speakers, 2 were male and 2 were female; again, demonstrating SoBigData++'s commitment to promote the work of females in the industry and encourage more female participation. In addition to the plenary speakers, there were 7 contributed talks of 20 minutes each.

Participants: The event was aimed at Teaching/Academic Institutions, Researchers, Data Analysts and Industry. The school attracted 268 participants, 60 females and 208 males. The number of participants varied between the 4 plenary talks; however, the typical number of participants was approximately 120.

2.4.4.8 EUROPEAN RESEARCHERS' NIGHT 27 – 28 NOVEMBER 2020

Objectives, Participants and Schedule, website

The European Research Night is a public event dedicated to the dissemination of science that takes place every year in 300 cities in 30 European countries at once and, for the third time, was organized in the counties of Tarragona. It aims to bring research and its protagonists closer to the public of all ages and make research and innovation known in a smooth and fun way, although this year it will be different due to the covid-19.

Link: http://tarragona.nitdelarecerca.cat/ca/

Jesús Manjón (URV) recorded a short 2-minute presentation to promote the SoBigData++ infrastructure within this event. As this event is organised for the general public this was an opportunity to disseminate the existence of the infrastructure outside of the Data Scientist community. The reach of this event is extensive throughout Europe and the intended audience includes both children and adults.

The picture below is taken from the event's website.

Figure 6: A screenshot of the SBD++ video presentation at the European Researchers Night taken from the event website

2.4.4.9 DATA SCIENCE COLLOQUIUM

Objectives: The Colloquium commenced 20 May and ran to 8 June. There were 3 sessions per week, each session was for 2 hours. The seminars were held by both professors and 2nd and 3rd year PhD students. The objective of the Colloquium was to help first year students – but they were also open to anyone interested.

The focus of the Colloquium is primarily on helping first-year PhD students to broaden their views towards the finalisation of their research projects, also being inspired by the work of their 2nd and 3rd-year colleague students. But the colloquium is also a great opportunity for anybody interested to know more about the scientific community and the project's research activities.

2.4.4.10 A SCIENTIFIC EXPERIMENT ON PRIVACY AND AI TRANSPARENCY

Objective: This session was organised by BSC and took place on 11 March 2021.

The objective was to conduct an experiment on the state of privacy and AI transparency on the Internet. A set of visual signs designed to replace current consent and privacy online forms (see https://www.linkedin.com/pulse/how-inform-users-intuitively-non-intrusively-rights-garcia-gasulla/) were introduced to the participants. The experiment engaged the participants with current consent and privacy forms of popular webpages to assess their accessibility and actual degree of transparency.

The participants were assigned the proposed visual signs to popular websites and discussions were encouraged to identify general tips (e.g., keywords to look for, how much effort it implies, form traps and how to avoid them, etc.). A scientific paper may be produced to discuss the results of the experiment. All participants would be acknowledged in the paper and any participants who actively contribute to the manuscript will be invited as co-authors.

Participants: There were 30 participants with an equal split of 50:50 male/female. There were between 6-10 people aged between 18-30 years. There were between 11-15 people aged between 30 -50 and between 6-10 were over 50 years of age.

The experiment was open to all – including Teaching/Academic Institutions, Research, Data Analysts, Industry, Public Administration, and the General Public, however no stakeholder data was obtained.

Link: https://www.mozillafestival.org/en/

2.4.4.11 ROMCIR 2021. REDUCING ONLINE MISINFORMATION THROUGH CREDIBLE INFORMATION RETRIEVAL

Objectives: This workshop was organised by IMT as part of ECIR 2021. It took place on 1 April 2021.

The central topic of ROMCIR concerns providing access to users to credible and/or verified information, to mitigate the information disorder phenomenon. In this context, “information disorder” means all forms of communication pollution, from misinformation due to ignorance, to the intentional sharing of false content. This topic is so broad as it concerns different contents (e.g., web pages, news, reviews, medical information, online accounts, etc.), different web and social media platforms (e.g., microblogging platforms, social networking services, social question-answering systems, etc.), and different purposes (e.g., identifying false information, accessing information based on its credibility, retrieving credible information, etc.).

The event consisted of two keynotes speakers and six users presenting their contributions. The contributions were published on CEUR at http://ceur-ws.org/Vol-2838/

Participants: Due to virtual nature of this event and the fact that registration was not taken, it is estimated that there were approximately 80 participants. The split between male and female was estimated to be equal.

The event was directed at Teaching/Academic Institutions, Researchers and Data Analysts. It is estimated that there were between 21-30 individuals in each of the following categories: 18-30 years, 30-50 years and over 50 years. The split was estimated to be fairly equal between each of following categories: Undergrad/Post Grad, PhD Researcher, Early Careers, and Academics.

Link: https://romcir2021.disco.unimib.it/

2.4.4.12 THEMATIC WORKSHOP ON MIGRATION AND BIG DATA

Objectives: This event was organised by Institut Convergences Migrations Paris and although was not a SoBigData++ event, it involved researchers who delivered talks which acknowledged SoBigData++. It took place on 7 April 2021.

It was the inaugural event of a series of 3 workshops and a training course dedicated to the subject of migration. The objective was to introduce the audience to the general topic and present state of the art research using big data for the study of migration and to offer a critical overview of the potential, as well as the challenges, posed by this type of data. The series of workshops will also identify the training needs in the methods used in this type of research for both early career researchers and experienced researchers.

This first workshop was a series of four talks, Alina Sîrbu's, 'Human Migration: the big data perspective' and Simone Bertoli's "Tell me what you eat, I will tell you who you are" both acknowledge the SoBigData++ project.

Participants: The audience was interdisciplinary and included both students and researchers. There were 55 participants: 30 female and 25 males. There were between 16 and 20 people aged 18-30 years, between 21 and 30 aged 30-50 years and between 6 and 10 people aged over 50 years. The participants were Teaching/Academic Institutions, Researchers and Data Analysts. There were between 11 and 15 undergraduates and postgraduates, between 11 and 15 PhD Researchers, between 11 and 15 Early Career Researchers, between 16 and 20 Academics and between 1 and 5 were Data Analysts.

Link:https://www.icmigrations.cnrs.fr/2021/03/10/conf-thematic-workshop-big-data-and-migration-wednesday-7-april-2021-9h30-12h30-online/

2.4.4.13 INCONTRA INFORMATICA

Objective: This event was designed to encourage High School students to consider a degree in Computer Science. This event was organised by UNIPI and was held online on 15 and 16 April 2021. High school students were shown how to read, pre-process and analyse soccer match event data. They were shown some basic statistics to describe the performance of players and teams. The event went well and the engagement from the students was strong.

Participants: There were 27 participants, of which 10 were female and 17 were male. Being High School students, the majority were under 18 and there were between 1-5 students who were over 18.

Link: https://didattica.di.unipi.it/incontra-informatica/

2.4.4.14 HUMANEAI - AI & SOCIETY ROUNDTABLE

Objective: This event took place virtually on 30 June 2021 having been organised by the University of Pisa. It was designed as a collective intelligence exercise towards shaping the research questions of Social AI, driven by societal challenges. The format of the two-and-a-half-hour conversation was four fire-start speeches purposefully setting the stage for the 4 breakout sessions. The topics of these sessions were bias, inequality, polarization and social good. To end the event there was a final restitution session.

Participants: There were 110 participants, 75 male and 35 female. There were between 21 and 30 participants who were aged between 18 and 30 and there were between 31-40 participants in each of the 30-50 age range and in the over 50 years age range. The event was aimed at Teaching, Research, Data Analysts, and Industry, although no data was collected regarding stakeholder status.

Link: https://www.humane-ai.eu/event/ai-society-roundtable/

2.4.5 *SBD++ Flagship Events – PSD2020 and SocInfo*

2.4.5.1 PSD 2020 (PRIVACY IN STATISTICAL DATABASES)

Objectives: The SoBigData++ sponsored event, PSD 2020 went ahead between 23 and 25 September 2020 in a virtual format. This event takes place every 2 years, and this year was originally due to be held in Arezzo, Italy - which was then changed to Tarragona, Catalonia; however due to the continued travel restrictions of the global pandemic it was converted to an online conference.

Participants: The event was aimed at researchers and statistical agencies and although the conference originated in Europe, it is open to all and contributions and attendees from across the globe are welcomed. This year's event was planned for approximately 40-50 attendees and contrary to expectations, the event attracted 72 attendees. Of these, 42 were male and 30 were female – demonstrating a fairly even spread. The speakers were also a good mix of male and female, and event went very well.

The organisers were able to maintain the structure of the event, albeit losing the social aspects, and every speaker had the opportunity to present their paper and answer questions from the audience. The organisers were also able to utilise a feature of Zoom that enabled them to set up parallel rooms for some attendees so that they could engage in more detailed discussions about a specific presentation. This facility turned out to be an advantageous benefit of holding the conference online.

The organisers are very pleased with how the event went and stated:

"More than 70 attendees of 17 countries, representing national statistical offices, universities, research centres and companies, shared their last contributions on privacy models, microdata protection, protection of statistical tables, protection of interactive and mobility databases, record linkage and alternative methods, synthetic data, data quality, and case studies. The conference has been carried out without incidences and

given the circumstances, we are very glad of the high level of achieved interaction between the attendees, that is one of the main goals of organizing a conference. Despite the organization success, we hope the next edition PSD2022 will be hold face-to-face."

Figure 7: A screenshot of the PSD 2020 event

To promote the project, the organisers sent every attendee an email presenting the SoBigData++ platform encouraging them to take a look.

Link: https://unescoprivacychair.urv.cat/psd2020/

2.4.5.2 SOCINFO 2020

Objectives: This event was the 12th International Conference on Social Informatics, and this year was organised by CNR and sponsored by SoBigData++ among others. It took place from 6-9 October.

The conference is designed to bridge the gap between Computer Sciences, Informatics, Management Sciences and Social Sciences and provide a forum for researchers from all related communities to collaborate, share knowledge and opinions and to present original research work about the interplay between socially centric platforms and social phenomena. The goal of Social Informatics is to create better understanding of socially centric platforms not just as a technology, but also as a set of social phenomena. In this edition, presenters faced a variety of topics, from human migration to gender equality, from misinformation to fact-checking, from polarisation to ideological bias of opinions, and from unemployment and health.

The ethos of the SocInfo conference is to foster strong working relationships and encourage networking between the various disciplines. The talks are chosen to reflect and highlight the cross over and interaction between the different synergies of research. The workshops are designed to be interactive involving discussion, activities, brainstorming and networking. Rather than long talks, organisers are encouraged to structure the workshops as interactive events allowing plenty of time for interdisciplinary discussion, the exchange of ideas, sharing relevant research and making meaningful connections with data scientists from other geographical or scientific backgrounds. To achieve this objective, the numbers are kept relatively small, at around 35-40 participants.

Participants: This year the conference was transformed into a 4-day online event with a full schedule including 4 keynote speakers, 4 workshops and 3 tutorials. There were also 10 sessions organised for 33 presented papers and 14 posters. The best paper was voted for by the regular attendees. There were 127 attendees of which 39 were female – approximately a third. Most participants were aged between 18 and 50 being PhD researchers, early career researchers and academics. There were a few undergraduate and postgraduate students and a few individuals were classified as working in industry.

Link: https://kdd.isti.cnr.it/socinfo2020/

2.4.5.3 DISINFO 2020

Objectives: Social networking platforms are a crucial component of the public sphere, fostering discussions and influencing the public perception for a myriad of topics including politics, health, climate change, economics, migration, to name but a few. On the one hand, this represents an unprecedented opportunity to discuss and propose new ideas, democratizing information and giving voice to the crowds. On the other hand, however, new socio-technical issues arise.

Among the most pressing issues is the spread of fictitious and low-quality information (e.g., fake news, rumours, hoaxes). These questionable means are often used to influence the opposing side about controversial and polarizing topics, or simply to sow discord and erode trust in governments, institutions, and societies. The spread of low-quality information is sometimes carried out by groups of coordinated or automated accounts that pollute and tamper with our social environments by injecting and resharing a large number of targeted messages. These issues are currently exacerbated by the recent advances in AI that have made it easy and convenient to fabricate plausible texts, to create high-quality images of non-existing people,

and to impersonate public characters in videos (e.g., deepfakes), at large. All these aspects jointly contribute to making our online social ecosystems the ideal landscape for deceit and manipulation. Therefore, prompt responses are expected from decision makers, scholars, and practitioners in order to limit the spread and impact of these ailments.

The International Workshop on "Information Disorders: Fake News and Coordinated Inauthentic Behaviours (DisInfo'20)" focuses on the study, modelling, and characterisation of all challenges related to mis- and dis-information, fake news, coordinated inauthentic behaviour and information operations.

Participants: This workshop was organised by CNR and took place on the 6 October as part of the SocInfo 2020 conference. It was a half-day event split into 2 sessions. Each session had a keynote speech from a leading world researcher and 3 talks discussing contributed articles. The number of attendees planned for was 15 – appealing mostly to Academics, Researchers and Data Analysts. The actual number of attendees doubled expectations with 30 participants with an equal split of male and female participants. The participants were PhD Researchers and Early career researchers. The equal split between the genders is another demonstration of how SoBigData++ events are successfully appealing to females as well as males.

Link: https://kdd.isti.cnr.it/socinfo2020/workshops.html#workshop3, http://ci.iit.cnr.it/disinfo20

2.5 Scientific Production

Many scientific papers have been written during the first period of this project. Table 5 lists the papers acknowledged by the project as reported by the publications tab in the EU Sygma portal. The main categories are Journal Articles, Publication in Conference proceedings/Workshop, and Book Chapters. All the publications are open access, either green or gold.

No.	Type	Title	Authors
2	Article in Journal	UTLDR: an agent-based framework for modeling infectious diseases and public interventions	Giulio Rossetti, Letizia Milli, Salvatore Citraro, Virginia Morini
3	Chapter in a Book	A "Learned" Approach to Quicken and Compress Rank/Select Dictionaries	Antonio Boffa, Paolo Ferragina, Giorgio Vinciguerra
4	Chapter in a Book	Learned Data Structures	Paolo Ferragina, Giorgio Vinciguerra
5	Article in Journal	The PGM-index	Paolo Ferragina, Giorgio Vinciguerra
6	Article in Journal	Evaluating community detection algorithms for progressively evolving graphs	Remy Cazabet, Souâad Boudebza, Giulio Rossetti
7	Article in Journal	Conformity: A Path-Aware Homophily Measure for Node-Attributed Networks	Giulio Rossetti, Salvatore Citraro, Letizia Milli
8	Publication in Conference proceedings/Workshop	Fair Detection of Poisoning Attacks in Federated Learning	Ashneet Khandpur Singh, Alberto Blanco-Justicia, Josep Domingo-Ferrer, David Sanchez, David Rebollo-Monedero
9	Publication in Conference proceedings/Workshop	An individual-level ground truth dataset for home location detection	Luca Pappalardo, Leo Ferres, Manuel Sacasa, Ciro Cattuto, Loreto Bravo

10	Chapter in a Book	Pattern detection in large temporal graphs using algebraic fingerprints	Suhas Thejaswi, Aristides Gionis
11	Article in Journal	General Confidentiality and Utility Metrics for Privacy-Preserving Data Publishing Based on the Permutation Model	Josep Domingo-Ferrer, Krishnamurty Muralidhar, Maria Bras-Amoros
12	Publication in Conference proceedings/Workshop	Deep Learning for Human Mobility: a Survey on Data and Models	Luca M.; Barlacchi G.; Lepri B.; Pappalardo L.
13	Publication in Conference proceedings/Workshop	Achieving Security and Privacy in Federated Learning Systems: Survey, Research Challenges and Future Directions	Blanco-Justicia, Alberto; Domingo-Ferrer, Josep; Martínez, Sergio; Sánchez, David; Flanagan, Adrian; Tan, Kuan Eeik
14	Article in Journal	Attention dynamics on the Chinese social media Sina Weibo during the COVID-19 pandemic	Hao Cui, János Kertész
15	Article in Journal	Cross-Lingual Sentiment Quantification	Andrea Esuli, Alejandro Moreo, Fabrizio Sebastiani, Erik Cambria
16	Publication in Conference proceedings/Workshop	Boilerplate Removal using a Neural Sequence Labeling Model	Jurek Leonhardt, Avishek Anand, Megha Khosla
18	Publication in Conference proceedings/Workshop	Estimating countries' peace index through the lens of the world news as monitored by GDELT	Vasiliki Voukelatou, Luca Pappalardo, Ioanna Miliou, Lorenzo Gabrielli, Fosca Giannotti
19	Publication in Conference proceedings/Workshop	Explaining the difference between men's and women's football	Pappalardo, Luca; Rossi, Alessio; Pontillo, Giuseppe; Natilli, Michela; Cintia, Paolo
20	Article in Journal	Secure monitoring in IoT-based services via fog orchestration	Alexandre Viejo, David Sánchez
21	Article in Journal	The limits of differential privacy (and its misuse in data release and machine learning)	Josep Domingo-Ferrer, David Sánchez, Alberto Blanco-Justicia
22	Article in Journal	Identifying and exploiting homogeneous communities in labeled networks	Salvatore Citraro, Giulio Rossetti
25	Chapter in a Book	Maximizing diversity over clustered data	Guangyi Zhang, Aristides Gionis
26	Publication in Conference proceedings/Workshop	Modelling Human Mobility considering Spatial,Temporal and Social Dimensions	Cornacchia, Giuliano; Rossetti, Giulio; Pappalardo, Luca
27	Publication in Conference proceedings/Workshop	Are Social Networks Watermarking Us or Are We (Unawarely) Watermarking Ourself?	Bertini, Flavio; Sharma, Rajesh; Montesi, Danilo
29	Article in Journal	Automatic Assessment of Privacy Policies under the GDPR	David Sánchez, Alexandre Viejo, Montserrat Batet
30	Article in Journal	Multi-Dimensional Randomized Response	Josep Domingo-Ferrer, Jordi Soria-Comas
31	Chapter in a Book	Explaining Sentiment Classification with Synthetic Exemplars and Counter-Exemplars	Orestis Lampridis, Riccardo Guidotti, Salvatore Ruggieri
33	Article in Journal	Utility-Preserving Privacy Protection of Textual Documents via Word Embeddings	Fadi Hassan, David Sanchez, Josep Domingo-Ferrer
34	Article in Journal	Towards better social crisis data with HERMES: Hybrid sensing for EmeRgency ManagEment System	Marco Avvenuti, Salvatore Bellomo, Stefano Cresci, Leonardo Nizzoli, Maurizio Tesconi
35	Chapter in a Book	Dynamics of Scientific Collaboration Networks Due to Academic Migrations	Pavlos Paraskevopoulos, Chiara Boldrini, Andrea Passarella, Marco Conti

36	Publication in Conference proceedings/Workshop	Firms' Challenges and Social Responsibilities during Covid-19: a Twitter Analysis	Patuelli, Alessia; Caldarelli, Guido; Lattanzi, Nicola; Saracco, Fabio
37	Publication in Conference proceedings/Workshop	Time-Varying Volatility in Bitcoin Market and Information Flow at Minute-Level Frequency	Barjašić, Irena; Antulov-Fantulin, Nino
39	Publication in Conference proceedings/Workshop	Explain and Predict, and then Predict Again	Zijian Zhang, Koustav Rudra, Avishek Anand
43	Publication in Conference proceedings/Workshop	Characterising different communities of Twitter users: Migrants and natives	Kim, Jisu; Sîrbu, Alina; Giannotti, Fosca; Rossetti, Giulio
44	Article in Journal	The academic wanderer: structure of collaboration network and relation with research performance	Pavlos Paraskevopoulos, Chiara Boldrini, Andrea Passarella, Marco Conti
45	Publication in Conference proceedings/Workshop	Tweet Sentiment Quantification: An Experimental Re-Evaluation	Moreo, Alejandro; Sebastiani, Fabrizio
47	Publication in Conference proceedings/Workshop	The relationship between human mobility and viral transmissibility during the COVID-19 epidemics in Italy	Cintia, Paolo; Pappalardo, Luca; Rinzivillo, Salvatore; Fadda, Daniele; Boschi, Tobia; Giannotti, Fosca; Chiaromonte, Francesca; Bonato, Pietro; Fabbri, Francesco; Penone, Francesco; Savarese, Marcello; Calabrese, Francesco; Guzzetta, Giorgio; Riccardo, Flavia; Marziano, Valentina; Poletti, Piero; Trentini, Filippo; Bella, Antonino; Andrianou, Xanthi; Del Manso, Martina; Fabiani, Massimo; Bellino, Stefania; Boros, Stefano; Urdiales, Alberto Mateo; Vescio, Maria Fenicia; Brusaferro, Silvio; Rezza, Giovanni; Pezzotti, Patrizio; Ajelli, Marco; Merler, Stefano; Vineis, Paolo; Pedreschi, Dino
48	Publication in Conference proceedings/Workshop	MP Twitter Engagement and Abuse Post-first COVID-19 Lockdown in the UK: White Paper	Farrell, Tracie; Bakir, Mehmet; Bontcheva, Kalina
49	Chapter in a Book	Structural Invariants in Individuals Language Use: The “Ego Network” of Words	Kilian Ollivier, Chiara Boldrini, Andrea Passarella, Marco Conti
51	Article in Journal	Internal migration and mobile communication patterns among pairs with strong ties	Mikaela Irene D. Fudolig, Daniel Monsivais, Kunal Bhattacharya, Hang-Hyun Jo, Kimmo Kaski
52	Publication in Conference proceedings/Workshop	Re-Assessing the "Classify and Count" Quantification Method	Alejandro Moreo; Fabrizio Sebastiani
53	Article in Journal	A Critical Reassessment of the Saerens-Latinne-Decaestecker Algorithm for Posterior Probability Adjustment	Andrea Esuli, Alessio Molinari, Fabrizio Sebastiani
54	Publication in Conference proceedings/Workshop	Heterogeneous document embeddings for cross-lingual text classification	Alejandro Moreo, Andrea Pedrotti, Fabrizio Sebastiani

55	Article in Journal	Finding Path Motifs in Large Temporal Graphs Using Algebraic Fingerprints	Suhas Thejaswi, Aristides Gionis, Juho Lauri
56	Publication in Conference proceedings/Workshop	Sharp Thresholds for a SIR Model on One-Dimensional Small-World Networks	Becchetti, Luca; Clementi, Andrea; Denni, Riccardo; Pasquale, Francesco; Trevisan, Luca; Ziccardi, Isabella
59	Article in Journal	Predicting seasonal influenza using supermarket retail records	Ioanna Miliou, Xinyue Xiong, Salvatore Rinzivillo, Qian Zhang, Giulio Rossetti, Fosca Giannotti, Dino Pedreschi, Alessandro Vespignani
60	Publication in Conference proceedings/Workshop	Benchmarking and Survey of Explanation Methods for Black Box Models	Bodria, Francesco; Giannotti, Fosca; Guidotti, Riccardo; Naretto, Francesca; Pedreschi, Dino; Rinzivillo, Salvatore
61	Publication in Conference proceedings/Workshop	Expansion and Flooding in Dynamic Random Networks with Node Churn	Becchetti, Luca; Clementi, Andrea; Pasquale, Francesco; Trevisan, Luca; Ziccardi, Isabella
62	Publication in Conference proceedings/Workshop	Testing properties of signed graphs	Adriaens, Florian; Apers, Simon
63	Chapter in a Book	Improved mixing time for k -subgraph sampling	Ryuta Matsuno, Aristides Gionis
64	Article in Journal	Evaluating local explanation methods on ground truth	Riccardo Guidotti
65	Publication in Conference proceedings/Workshop	Searching for polarization in signed graphs: a local spectral approach	Han Xiao, Bruno Ordozgoiti, Aristides Gionis
66	Article in Journal	GLocalX - From Local to Global Explanations of Black Box AI Models	Mattia Setzu, Riccardo Guidotti, Anna Monreale, Franco Turini, Dino Pedreschi, Fosca Giannotti
67	Chapter in a Book	Discovering Dense Correlated Subgraphs in Dynamic Networks	Giulia Preti, Polina Rozenshtein, Aristides Gionis, Yannis Velegrakis
69	Chapter in a Book	Mining Dense Subgraphs with Similar Edges	Polina Rozenshtein, Giulia Preti, Aristides Gionis, Yannis Velegrakis
70	Publication in Conference proceedings/Workshop	Finding large balanced subgraphs in signed networks	Bruno Ordozgoiti, Antonis Matakos, Aristides Gionis
72	Publication in Conference proceedings/Workshop	Explainable Classification of Brain Networks via Contrast Subgraphs	Tommaso Lanciano, Francesco Bonchi, Aristides Gionis
74	Publication in Conference proceedings/Workshop	Analysis of bank leverage via dynamical systems and deep neural networks	Fabrizio Lillo, Giulia Livieri, Stefano Marmi, Anton Solomko, Sandro Vaienti
77	Article in Journal	ANGEL: efficient, and effective, node-centric community discovery in static and dynamic networks	Giulio Rossetti
79	Article in Journal	Error Estimation of Ultra-Short Heart Rate Variability Parameters: Effect of Missing Data Caused by Motion Artifacts	Alessio Rossi, Dino Pedreschi, David A. Clifton, Davide Morelli
80	Article in Journal	Change of direction asymmetry across different age categories in youth soccer	Athos Trecroci, Alessio Rossi, Thomas Dos'Santos, Damiano Formenti, Luca Cavaggioni, Stefano Longo, F. Marcello Iaia, Giampietro Alberti
81	Article in Journal	A Public Dataset of 24-h Multi-Levels Psycho-Physiological Responses in Young Healthy Adults	Alessio Rossi, Eleonora Da Pozzo, Dario Menicagli, Chiara Tremolanti, Corrado Priami, Alina Sîrbu, David A.

			Clifton, Claudia Martini, Davide Morelli
85	Publication in Conference proceedings/Workshop	Dynamic Hard Pruning of Neural Networks at the Edge of the Internet	Lorenzo Valerio, Franco Maria Nardini, Andrea Passarella, Raffaele Perego
86	Publication in Conference proceedings/Workshop	Optimising cost vs accuracy of decentralised analytics in fog computing environments	Lorenzo Valerio, Andrea Passarella, Marco Conti
87	Chapter in a Book	Explaining Misclassification and Attacks in Deep Learning via Random Forests	Rami Haffar, Josep Domingo-Ferrer, David Sánchez
88	Chapter in a Book	Efficient Detection of Byzantine Attacks in Federated Learning Using Last Layer Biases	Najeeb Jebreel, Alberto Blanco-Justicia, David Sánchez, Josep Domingo-Ferrer
90	Chapter in a Book	Characterizing Social Bots Spreading Financial Disinformation	Serena Tardelli, Marco Avvenuti, Maurizio Tesconi, Stefano Cresci
91	Article in Journal	Interaction Strength Analysis to Model Retweet Cascade Graphs	Paola Zola, Guglielmo Cola, Michele Mazza, Maurizio Tesconi
92	Publication in Conference proceedings/Workshop	Measuring Immigrants Adoption of Natives Shopping Consumption with Machine Learning.	Riccardo Guidotti, Mirco Nanni, Fosca Giannotti, Dino Pedreschi, Simone Bertoli, Biagio Speciale, Hillel Rapoport
93	Article in Journal	A decade of social bot detection	Stefano Cresci
94	Article in Journal	On the performance of learned data structures	Paolo Ferragina, Fabrizio Lillo, Giorgio Vinciguerra
96	Publication in Conference proceedings/Workshop	Temporal mixture ensemble models for intraday volume forecasting in cryptocurrency exchange markets	Nino Antulov-Fantulin, Tian Guo, Fabrizio Lillo
101	Article in Journal	An ethico-legal framework for social data science	Nikolaus Forgó, Stefanie Hänold, Jeroen van den Hoven, Tina Krügel, Iryna Lishchuk, René Mahieu, Anna Monreale, Dino Pedreschi, Francesca Pratesi, David van Putten
102	Publication in Conference proceedings/Workshop	Private Sources of Mobility Data Under COVID-19	Raquel Pérez Arnal, David Conesa, Sergio Alvarez-Napagao, Toyotaro Suzumura, Martí Català, Enric Alvarez, Dario Garcia-Gasulla
108	Article in Journal	Evaluation of home detection algorithms on mobile phone data using individual-level ground truth	Luca Pappalardo, Leo Ferres, Manuel Sacasa, Ciro Cattuto, Loreto Bravo
110	Publication in Conference proceedings/Workshop	Ecology in the digital world of Wikipedia	Fumiko Ogushi, Janos Kertesz, Kimmo Kaski, Takashi Shimada
111	Article in Journal	Human-agent coordination in a group formation game	Tuomas Takko, Kunal Bhattacharya, Daniel Monsivais, Kimmo Kaski
112	Publication in Conference proceedings/Workshop	Morningness-eveningness assessment from mobile phone communication analysis	Chandreyee Roy, Daniel Monsivais, Kunal Bhattacharya, Robin I.M. Dunbar, Kimmo Kaski
114	Publication in Conference proceedings/Workshop	Some Challenges in Monitoring Epidemics	Vaiva Vasiliauskaite, Nino Antulov-Fantulin, Dirk Helbing
118	Article in Journal	Uncovering the mesoscale structure of the credit default swap market to improve portfolio risk modelling	I. Anagnostou, T. Squartini, D. Kandhai, D. Garlaschelli

120	Publication in Conference proceedings/Workshop	Time-varying volatility in Bitcoin market and information flow at minute-level frequency	Irena Barjašić, Nino Antulov-Fantulin
122	Publication in Conference proceedings/Workshop	Low-dimensional statistical manifold embedding of directed graphs	Thorben Funke, Tian Guo, Alen Lancic, Nino Antulov-Fantulin
127	Publication in Conference proceedings/Workshop	Definition of an enriched GIS network for evacuation planning	Evans Etrue Howard, Lorenza Pasquini, Antinisca Di Marco, d Eliseo Clementini
129	Article in Journal	Data science: a game changer for science and innovation	Valerio Grossi, Fosca Giannotti, Dino Pedreschi, Paolo Manghi, Pasquale Pagano, Massimiliano Assante
131	Publication in Conference proceedings/Workshop	The effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions	Antonio F. Peralta, Matteo Neri, János Kertész, Gerardo Iñiguez

Table 5: List of all publication from M1 to M18 as listed in the EU portal Sygma

2.5.5.1 SPECIAL ISSUE - INTERNATIONAL JOURNAL OF DATA SCIENCE AND ANALYTICS

Data Science is rapidly changing the way we do business, socialize, and govern society, and the way we make scientific research. In this context, we introduce the resulting collection of papers in the special issue entitled "Social Mining and Big Data Ecosystem for Open, Responsible Data Science", published in the International Journal of Data Science and Analytics (JDSA) - Volume 11, issue 4 published May 2021 (https://link.springer.com/journal/41060/volumes-and-issues/11-4). The special issue provides an account of a global scientific trend that is profoundly transforming science, providing it better means to foster social good. We ended accepting six papers, covering a broad spectrum of challenging issues:

- ***Data Science: a game-changer for science and innovation.*** This paper introduces how data science impacts science and society, including ethical and governance issues connected with managing data that touch upon aspects of human behaviour. https://link.springer.com/article/10.1007/s41060-020-00240-2
- ***Measuring objective and subjective well-being: dimensions and data sources.*** This work illustrates the approaches for measuring well-being. The authors distinguish between objective and subjective well-being and surveys the theoretical background, the relevant dimensions of well-being, the new data sources for measurement, and relevant recent studies. https://link.springer.com/article/10.1007/s41060-020-00224-2
- ***(So) Big Data and the transformation of the city.*** This paper discusses the main issues of urban data analytics, focusing on privacy issues, algorithms, applications, and georeferenced data from social media. As concrete case studies of urban data science tools, the authors leverage the results obtained in the "City of Citizens" thematic area of the SoBigData initiative, which includes a virtual research environment with mobility datasets and urban analytics methods by several institutions around Europe. https://link.springer.com/article/10.1007/s41060-020-00207-3

- ***Human migration: the big data perspective.*** In this paper, the authors answer the question "How can big data help to understand the migration phenomenon?" through an analysis of various phases of migration, comparing traditional and novel data sources and models at each phase. They focus on three phases of migration - the journey, the stay, and the return - at each phase describing state of the art and recent developments and ideas. https://link.springer.com/article/10.1007/s41060-020-00213-5.
- ***A Workflow Language for Research e-Infrastructures.*** This work outlines the HyWare language and platform. The language is an extension of the traditional workflow languages enabling the definition of workflows, including automatic and manual analytical steps to replicate and build large-scale data-driven experiments. https://link.springer.com/article/10.1007/s41060-020-00237-x
- ***An ethical and legal framework for social data science.*** This paper provides a framework for research infrastructures that enable ethically sensitive and legally compliant data science, helping data scientists frame the appropriate self-assessment questions to ensure an ethical, responsible design, implementation, and deployment of data science projects. https://link.springer.com/article/10.1007/s41060-020-00211-7

SoBigData RI is proud to offer this collection to the attention and scrutiny of the scientific community, and we recall that all the papers are published under Open Access rules.

2.6 Dissemination Impact

In Deliverable 3.1 - "Initial Dissemination and Impact Plan" we have defined a set of indicators in order to monitor the success of the communication activities. The success of the communication activities is 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.

2.6.1 *Project dissemination indicators*

Table 6 reports the indicators defined in Deliverable 3.1. The following indicators represent some metrics in order to evaluate the dissemination activities.

The Extend of audience reached includes all the audience reached through the activities done within the various communication channels of the project, in particular social media, website, press coverage and events, as shown in Table 7.

The Number of SMEs, big companies and other institutional stakeholders includes all the various stakeholders engaged through the dissemination activities of the project. A detailed description of this success indicator can be found in Section 4 "Stakeholder Analysis".

The Papers published on OpenAIRE portal includes all the paper on the portal, distinguished between Open access and Not Open Access, as shown in Table 8.

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 OpenAIRE portal	30	60	90

Table 6: Success Indicators for Reporting Period 1, 2 and 3

Extent of audience reached	Metrics	Results
Web site	**Visits**	19868
Twitter	-Tweets: 86 -**Followers: 1508**	1508
Facebook	-Posts: 368 -**Followers: 495** -Likes: 132	495
YouTube	-YouTube channel subscribers: 37 **-video views: 917**	917
Press coverage	**-number of traditional press articles**	13
Events	**-total number of participants: 3416** -male participants: 1894 -female participants: 998	3416
Total		**26217**

Table 7: Extent of audience reached for reporting period 1

As we can see from the tables, we have largely achieved the Success Indicators established in D3.1. The results achieved show that the SoBigData community is increasingly growing, thanks to the variety of activities done and a very active consortium on the dissemination activities, despite the difficulties due to the lockdown for Covid-19 and the impossibility of organize events in presence and meet in person.

Papers published on OpenAIRE portal	Results
Open Access	67
Restricted	6

Closed Access	2
Total	**75**

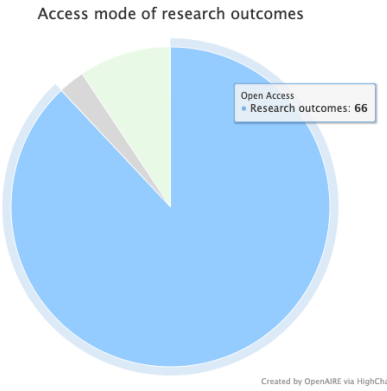
Table 8: Papers published on OpenAIRE portal for reporting period 1

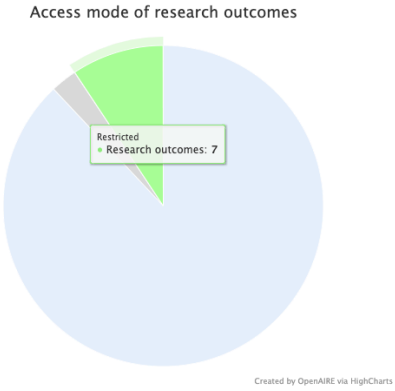
As we can appreciate form the Table 8, the Success Indicator is largely reached also for the papers published on the OpenAire portal. Furthermore, the majority of the publications within the project is in Open Access, showing that the commission's indications to move towards open science are increasingly accepted by the consortium.

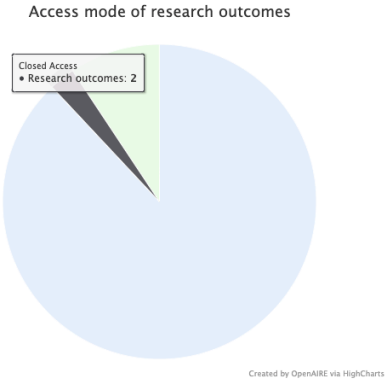
It is worth saying that, as this is a first release of the SoBigData++ page on the OpenAire portal, a check is needed in order to see if no data are missed. In fact, the numbers of publications in OpenAIRE are different from the ones collected among the researchers and reported on Table 5 and on EU Sygma portal. The SoBigData++ page on the OpenAIRE portal is available for consultation at:

https://explore.openaire.eu/search/project?projectId=corda__h2020::a45280b6e42e263da0c4e85e5bf6845b

In the graphs below, we can see at glance the numbers of Open Access publications.







2.6.2 *Impact of Press Coverage and social media*

Promoting the project on the internet and social media is one of the most immediate and simplest methods of dissemination. The project website is updated regularly with events and blogs keeping the data science

community up to date with current news stories and opinion pieces. Recordings of events are often posted so that they can be viewed again, or by an additional audience (again widening the availability of events and opening up the limits of how far SoBigData++ talks and events reach). However, the website will only be accessed by the community who know of its existence or individuals who opportunely discover the site through their search criteria.

Twitter, however, has a different method of disseminating messages and this can be exploited by the SoBigData++ Project. Tweets are a much quicker and more dynamic way for SoBigData++ to disseminate notice of future events, links to live events and talks or recordings of past events. The short and precise nature of Tweets means that any followers can read the post and make quick decisions on which reports, events, talks or webinars they are interested in. This is a much more direct manner to reach the data science community as it embraces and exploits the short attention span users give to each Tweet.

Furthermore, the hashtag feature of Twitter allows individual data scientists to promote SoBigData++ through their own personal accounts by 'hashtagging' '@SoBigData' in their post. These posts therefore reach an audience who are not currently 'followers' of the SoBigData++ Twitter account. This is a facility that could be encouraged within the SoBigData++ infrastructure as a simple and effective way to disseminate the project's existence and message.

In Table 9 there is a selection of the Tweets posted on the SoBigData Twitter account in November 2020 (month chosen at random) to provide a sample of the Tweets. Tweets regarding links to publications acknowledging SoBigData++ have not been included, however this is another example of disseminating papers and articles in a more immediate way to the largest possible audience. The sample of posts in the table only relate to Tweets posted on the SoBigData account and not posts that have tagged '@SoBigData' by individuals or posts that been retweeted by SoBigData.

It is also worth noting that some projects or events that operate within the SoBigData++ Infrastructure have their own Twitter account to further promote their events and blogs. An example of this would be SocInfo 2021 '@socinfo21'.

No.	What was posted	Date Posted
1.	Falling Walls Circle Table: Understanding the Scientific Method in the 21st Century	5 Nov 2020
2.	SocInfo 2020: a point of view from organizers	27 Nov 2020
3.	Responsible AI: from principles to action	27 Nov 2020
4.	A contribution on basic questions regarding AI in law	27 Nov 2020
5.	Dynamics of Scientific Collaboration Networks in Academic Migrations	27 Nov 2020
6.	A report of the "Human Migration – Potential areas for combinations of Big Data" workshop at SOCINFO 2020	27 Nov 2020
7.	Network Medicine: Disease Genes Prioritization Problem	27 Nov 2020
8.	FairLens: Auditing Black-box Clinical Decision Support Systems	27 Nov 2020

9.	Relationship between External and Internal Workloads in Elite Soccer Players	27 Nov 2020
10.	"Girls Just Wanna Have Fun?" Participation Trends and Motivational profile in women during sky events	27 Nov 2020
11.	Does immigration make Europeans less supportive of redistribution?	27 Nov 2020

Table 9: A sample of Tweets on the SoBigData++ Twitter account in the month of November 2020

Traditional press coverage is used much less by the project as the reach of the press is not as prevalent as the internet or social media and exists on a more national or regional basis rather than the international framework of the internet and social media. Posting directly on the internet and social media also enables direct authorship and removes the barrier between data scientist and audience. To date there have been 13 articles published which are detailed in Table 10:

Country	Publication	Title	Date	Link:
Italy	Quotidiano di Sicilia	**Big data, via a "Challenge Us": la ricerca supporta le imprese** (Big data, go to "Challenge Us": research supports businesses)	30/03/2021	https://mail.google.com/mail/u/0?ui=2&ik=17eeed4bb4&attid=0.1&permmsgid=msg-f:1698094678443534932&th=1790d81cf79baa54&view=att&disp=inline
Italy	Il Tirreno	**Analisi dei dati gratis Ecco la sfida lanciata alle imprese toscane** (Free data analysis Here is the challenge launched to Tuscan companies)	30/03/2021	https://mail.google.com/mail/u/0?ui=2&ik=17eeed4bb4&attid=0.2&permmsgid=msg-f:1698094678443534932&th=1790d81cf79baa54&view=att&disp=inline
Italy	University of Pisa Website	**Big data: da gruppo internazionale di ricercatori consulenza gratuita alle aziende** (Big data: from an international group of researchers free advice to companies)	29/03/2021	https://www.unipi.it/index.php/news/item/20548-big-data-da-gruppo-internazionale-di-ricercatori-consulenza-gratuita-alle-aziende
Italy	La Nazione Pisa	**Big Data, i ricercatori diventano consulenti per migliorare il business delle aziende** (Big Data, researchers become consultants to improve companies' business)	30/03/2021	https://www.lanazione.it/pisa/cronaca/big-data-i-ricercatori-diventano-consulenti-per-migliorare-il-business-delle-aziende-1.6190197
Italy	PISA Today	**'Challenge us': consulenza gratuita alle aziende che vogliono 'immergersi' nei big data** ('Challenge us': free advice to companies that want to 'dive' into big data")	29/03/2021	https://www.pisatoday.it/cronaca/challenge-us-programma-big-data.html

Italy	Gonews.it	**Challenge Us, consulenza gratuita alle aziende sui big data** (Challenge Us, free advice to companies on big data)	29/03/2021	https://www.gonews.it/2021/03/29/challenge-us-consulenza-gratuita-alle-aziende-sui-big-data/
Italy	Stamp Toscana	**Big data: da ricercatori consulenza gratuita ad aziende** (Big data: from researchers free advice to companies)	29/03/2021	https://www.stamptoscana.it/big-data-da-ricercatori-consulenza-gratuita-ad-aziende/
Italy	Sant'Anna Scuola Universitaria Superiore Pisa	**Big data: da gruppo internazionale di ricercatori consulenza gratuita ad aziende che vogliono diventare più competitive. Opportunità del programma "Challenge Us", organizzato dal Sant'Anna con CNR, Università di Pisa, Università la Sapienza di Roma** (Big data: from international group of researchers, free consultancy to companies that want to become more competitive. Opportunities of the "Challenge Us" program, organized by Sant'Anna with CNR, University of Pisa, la Sapienza University of Rome)	25/03/2021	https://www.santannapisa.it/it/news/big-data-da-gruppo-internazionale-di-ricercatori-consulenza-gratuita-ad-aziende-che-vogliono
Italy	in Toscana	**Dalla Sant'Anna consulenze gratis alle aziende sulle potenzialità dei Big data** (From Sant'Anna free advice to companies on the potential of Big data)	29/03/2021	https://www.intoscana.it/it/articolo/big-data-consulenze-sant-anna/
Italy	Scuola Sant'Anna Twitter Account	#ChallengeUs #BigData consulenza gratuita alle imprese per diventare più competitive grazie al programma coordinato da @ScuolaSantAnna #Pisa con @CNRsocial_ @Unipisa @SapienzaRoma @unitartu @sheffielduni @ETH	29/03/2021	https://twitter.com/ScuolaSantAnna/status/1376508705018146820?s=20

		Center for The #Study of #Democracy. Domande entro il 15 maggio 2021 https://twitter.com/intoscana/status/1376476475952570370		
Estonia	University of Tartu, Institute of Computer Science	**Join the Challenge Us programme to explore the potential of big data in your business**	09/04/2021	https://www.cs.ut.ee/en/news/join-challenge-us-programme-explore-potential-big-data-your-business
Switzerland	ETH Zurich	**An initiative to help small and medium sized companies to manage their (big) data problems**	12/03/2021	https://coss.ethz.ch/news/news/2021/03/challenge-us.html

Table 10: Press coverage recorded from month 1 to month 18 of the project

2.7 Overview of Future Events/Plan for Period 2

Table 11 details planned events from month 19 onwards.

	Event	Date	Partner	Description
Month 21	**PhD School in Data Science**	Sept 2021		No further details
Month 22	**Ethics and Privacy of Big Data Use for Migration Research**	7-8 Oct 2021	UNIPI	**Workshop** that will discuss the theoretical, philosophical, legal, and ethical aspects of Big Data with a focus on migration. It will look at Big Data and new technologies for migration studies; real-life applications and socio-political consequences of Big Data and AI employment for migration.
	DSAA 2021	6-9 Oct 2021		**Conference:** The IEEE International Conference on Data Science and Advanced Analytics (DSAA) features its strong interdisciplinary synergy between statistics (sponsored by ASA), computing, and information/intelligence sciences (by IEEE and ACM), and cross-domain interactions between academia and business/industry for data science and analytics. https://dsaa2021.dcc.fc.up.pt/

	SciDataCon 2021	10 Oct / 30 Nov 2021		**Conference:** SciDataCon is a conference for peer-reviewed research and practice presentations, which covers all aspects of the role of data in research, society, and policy. http://sobigdata.eu/events/sobigdata-ri-scidatacon-2021

Table 11: Detailing planned future events for the next period of the project

All events planned during the current time are considering both virtual and 'in person' events so that when travel restrictions allow, organisers are prepared for either eventuality. Many of the benefits of the virtual events may be incorporated into 'in person' events to make them more inclusive and open to wider participation.

2.8 Magazine

The project is also publishing a six-month Magazine, to reach the vast community of SoBigData.

The Magazine is accessible both in virtual and printable version and it is published on the social media channels and website of the project in pdf format. The Magazine is also disseminated through specific mailing lists. A page on the project website has been dedicated to subscribe/unsubscribe to the newsletter, to make registration management faster and more effective (see http://www.sobigdata.eu/newsletter).

The magazine has an Editorial Board composed by Fosca Giannotti, Roberto Trasarti, Beatrice Rapisarda, Marco Braghieri and Valerio Grossi and an Editorial Secretariat composed by Beatrice Rapisarda and Marco Braghieri to manage the contents and the publication of the magazine. The magazine is published every 6 months and contains a summary of the latest news about the project, as well as the latest and most interesting publications of project partners.

3 Impact of outreach towards policy makers and the public at large

3.1 Introduction

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 policymakers. It is equally important that the community is well-informed about the concerns of the public at large as well as what is happening in the policy agenda.

The European policy agenda includes several 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 initial work of the project SoBigData++ to maximise the impact in reaching out to policy makers and the public at large.

The importance of increasing awareness with policymakers and the public at large became acute during the pandemic where governments across the world were battling with the multifaceted and complex issues around using big data and AI to help slow down the spread of the pandemic.

In order to be able to swiftly react to policymakers and public concerns relating to big data and AI as well as work with policymaker on the more long-term issues and understanding of the impact, opportunities, and challenges that these technologies, for SoBigData++ the projected teamed-up with Re-Imagine Europa.

Through bilateral meetings, workshops, events, and contribution to debates, SoBigData++ managed to give initial input into several of the most pressing debates around these issues to the European Commission (in particular with Vice President Vera Jourova ahead of the presentation of the European Democracy Action Plan and the Cabinet of Commissioner Vestager following the presentation of the Digital Services Act), the European Parliament and key European government representatives (more below).

While the inability to hold events in person made it more complex to present the project and its outcomes in some contexts, the increasing attitude to participation in virtual events opened new possibilities of

interaction and collaboration, as seen from the large number of workshops and events mentioned in the previous chapter. It is important to note that that the virtual setup created some problems in collecting data due to conflicting privacy concerns: for instance, those on gender, age, occupation, and level of education of the events' participants are not available for some events.

3.2 Outreach to Policymakers

SoBigData++, through its work and priorities, is developing a wealth of knowledge and expertise that are of great value to European policy makers especially as “a Europe Fit for the Digital Age” has been placed as one of the key priorities of the current European Commission.

Within this broader context, SoBigData++ has relevant input for a number of different action plans and priority areas, as well as input on specific topics treated by the Exploratories of WP10.

During this first phase of the SoBigData++, as the work developed in the exploratories is in its initial stages, outreach work focused on the crosscutting issued dealt with under the broader “Europe fit for the Digital Age” umbrella, specifically looking at:

3.2.1 *Artificial Intelligence*

Artificial intelligence (AI) can help find solutions to many of society’s problems. This can only be achieved if the technology is of high quality and developed and used in ways that earns peoples’ trust. Therefore, an EU strategic framework based on EU values will give citizens the confidence to accept AI-based solutions, while encouraging businesses to develop and deploy them.

This is why the European Commission has proposed a set of actions to boost excellence in AI, and rules to ensure that the technology is trustworthy.

The Regulation on a European Approach for Artificial Intelligence and the update of the Coordinated Plan on AI will guarantee the safety and fundamental rights of people and businesses, while strengthening investment and innovation across EU countries.

AI has also been highlighted as the key priority for the Slovenian Presidency of the Council of the EU (July 2021-December 2021).

SoBigData++ has participated in several events and meetings on this issue ahead of the European Commission communication on a European approach to Artificial intelligence. Initial interactions with key policymakers on this issue include MEPs and initial conversation with DG CONNECT (a full list of meetings is provided below). As the Slovenian semester starts SoBigData++ will increase this work and is planning a targeted event in the second half of 2021.

3.2.2 *European Data Act*

The European data strategy aims to make the EU a leader in a data-driven society. Creating a single market for data will allow it to flow freely within the EU and across sectors for the benefit of businesses, researchers, and public administrations.

At the end of 2020, European Commission published a Proposal for a regulation on European data governance (Data Governance Act). With the Data Governance Act, the EU Commission wants to create a framework that will allow Europe to become a leading data economy, especially for industrial data. The Data Governance Act is part of a wider European Data Strategy, which will be complemented by a Data Act that will govern access to data in business-to-business (B2B) relationships.

SoBigData++ has contributed to the discussions around European Data Strategy both through bilateral meetings with key representatives within the European Commission as well as with key MEPs. In December 2020, SoBigData++ hosted an Expert Roundtable on Data and Ethics in a Post-COVID World" (a full list of meetings is provided below). A white paper is currently being developed to be presented in the autumn of 2021.

3.2.3 *Digital Services Act*

The Digital Services Act and Digital Markets Act encompass a single set of new rules applicable across the whole EU to create a safer and more open digital space. Published in December 2020, these acts aim to provide a modern legal framework that ensures the safety of users online, establishes governance with the protection of fundamental rights at its forefront, and maintains fair and open online platform environment.

The rapid and widespread development of digital services has been at the heart of the digital changes that impact our lives. While there is a broad consensus on the benefits of this transformation, the problems arising have numerous consequences for our society and economy. A core concern is the trade and exchange of illegal goods, services, and content online. Online services are also being misused by manipulative algorithmic systems to amplify the spread of disinformation, and for other harmful purposes. These new challenges and the way platforms address them have a significant impact on fundamental rights online. The accelerating digitalisation of society and the economy has created a situation where a few large platforms control important ecosystems in the digital economy. They have emerged as gatekeepers in digital markets, with the power to act as private rule-makers. These rules sometimes result in unfair conditions for businesses using these platforms and less choice for consumers.

SoBigData++ has been following the developments of the DSA with key meetings with the Cabinet of Executive Vice President Vestager held in January 2021 to create awareness of the work of the consortium and assess how it could contribute to this discussion.

3.2.4 *European Democracy Action Plan*

In December 2020, the European Commission published the European Democracy Action Plan to empower citizens and build more resilient democracies across the EU. Standing up to challenges to our democratic systems from rising extremism and perceived distance between people and politicians, the Action Plan sets out measures to promote free and fair elections, strengthen media freedom and counter disinformation.

It is particularly with this last issue in mind that SoBigData++ has been engaging in this debate with numerous stakeholders and through the work developed by Re-Imagine Europa's Task Force on Democracy in a Digital Society, chaired by Professor Manuel Castells and with the participation of leading stakeholders in the field.

In fact, SoBigData co-organised and participated to the International Forum on Digital and Democracy on the 10-11 December 2020. The event was opened by Vice President Věra Jourová (Vice President for Values and Transparency of the European Commission) and brought together hundreds of key thinkers and players in the field for 2 days of lively discussions.

3.2.5 *Digital Skills*

All Europeans need digital skills to study, work, communicate, access online public services and find trustworthy information. However, many Europeans do not have adequate digital skills. The Digital Economy and Society Index (DESI) shows that 4 out of 10 adults and every third person who works in Europe lack basic digital skills. There is also low representation of women in tech-related professions and studies, with only 1 in 6 ICT specialists and 1 in 3 science, technology, engineering, and mathematics (STEM) graduates being women. The European Commission has set targets in the European skills agenda and the digital education action plan to ensure that 70% of adults have basic digital skills by 2025. These initiatives aim to reduce the level of 13–14-year-olds who underperform in computing and digital literacy from 30% (2019) to 15% in 2030. This is an area where the expertise and work of the consortium can help achieve the European targets. Activities towards these goals have so far been focused on creating awareness with key stakeholders about the opportunities created by the consortium.

3.2.6 *Other key issues*

Other key issues followed by SoBigData++ include:

3.2.6.1 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. SoBigData++ is now exploring the possibility for the Exploratory on Societal Debates and Misinformation Analysis be engaged and involved in this exercise.

3.2.6.2 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. The exact measures will be be explored during the second stage of the project.

3.2.6.3 STRONG SOCIAL EUROPE FOR JUST TRANSITION

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 "String Social Europe for Just Transitions" could benefit from input from several of the exploratories in particular "Demography, Economy and Finance". During the first 18 months of SoBigData++ the outreach activities have focused mainly on policymakers, as set out in the initial proposals, to prepare for a broader outreach to the public. This as a more robust effort can be developed later in the project when outcomes from the Exploratories become more substantial. In the meantime, outreach toward the public at large has been through participation at events, online presence, and articles.

3.3 Research, Strategic meetings, Workshops and Events

3.3.1 *Research and Monitoring*

During the first 18 months of activity, SoBigData++ has followed and reported on major European policy initiatives to explore how best to create synergies between the work developed under the project and policy

proposals. This has been complemented by following media and public discourse across Europe to see how the public discourse around big data, AI and privacy are being framed and addressed.

3.3.2 *Strategic Meetings*

Initial meetings have been held with a range of stakeholders to create awareness about SoBigData++ and to discuss how it could contribute to the policy objectives both at a European and national level.

Key meetings include:

NAME	POSITION	ORGANISATION
Věra Jourová	Vice President for Values and Transparency of the European Commission	European Commission
Roberto Viola	Director-General of DG CONNECT	European Commission
Romano Prodi	Former President of the European Commission, Former Prime Minister of Italy	
Werner Stengg	Member of Cabinet of Executive Vice President Margrethe Vestager	European Commission
Federico D'Incà	Italian MP and former Minister for Parliamentary Relations	National Government
Professor Manuel Castells	Minister of Universities of Spain	National Government
Iban Garcia del Blanco	Member of the European Parliament (Vice Chair of JURI/AIDA - S&D)	European Parliament
Magdalena Adamowicz	Member of the European Parliament	European Parliament
Karen Melchior	Member of the European Parliament (JURI/FEMM - Renew)	European Parliament
Maria da Graça Carvalho	Member of the European Parliament	European Parliament
Anna-Misel Asimakopoulou	Member of the European Parliament (Vice Chair of INTA/AIDA - EPP)	European Parliament
Kim Van Sparrentak	Member of the European Parliament (IMCO/AIDA - Greens)	European Parliament
Isabel Wiseler-Lima	Member of the European Parliament (AFET/DROI - EPP)	European Parliament
Svetla Tanova	EUROPEAN SCIENCE-MEDIA HUB	European Parliament
Anni Hellman	Deputy Head of Unit - Media Convergence and Social Media - DG CONNECT	European Commission

Table 12: Key meetings with Stakeholders

3.3.3 *Workshops and Events*

The first big event and parallel workshops were organized at the end of 2020, following initial preparation with key stakeholders and policymakers. The events organized or co-organised by SoBigData++ include:

3.3.3.1 INTERNATIONAL FORUM ON DIGITAL AND DEMOCRACY

Co-hosted by Re-Imagine Europa and Associazione Copernicani, in partnership with SoBigData++, the International Forum on Digital and Democracy took place on the 10th and 11th of December 2020. The two-day conference aimed to be a meeting place for politics and academics and a moment of confrontation,

promoting international collaboration through the exchange of information, ideas and best practices. The forum saw the participation, among others of (in alphabetical order):

- Prof. Ricardo Abramovay, Sociologist, political scientist and professor
- Ms Magdalena Adamowicz, Member of the European Parliament
- Mr Richard Barbrook, Author of several influential essays on information society and winner of the 2008 Marshall McLuhan Prize for Outstanding Book in the field of Media Ecology
- Prof. Manuel Castells, Chairman of the Re-Imagine Europa Taskforce on Democracy in a Digital Society, Wallis Annenberg Chair Professor of Communication Technology and Society at the Annenberg School of Communication, University of Southern California, Spanish Minister for Universities
- Min. Federico D'Incà, Italian Member of Parliament and former Minister for Parliamentary Relations
- Ms Věra Jourová, Vice President for Values and Transparency of the European Commission
- Mr Gianluca Misuraca, Strategic policy advisor on Digital Governance, Technology Diplomacy and Human-Centric Artificial Intelligence and a Research Fellow at the Department of eGovernance and Public Administration of the Danube University –
- Ms Nanjala Nyabola, Writer and political analyst, author of "Digital Democracy, Analogue Politics: How the Internet Era is Transforming Politics in Kenya."
- Mr Marco Pierani, Director of Public Affairs and Media Relations and Executive Board Member at Euroconsumers
- Mr Romano Prodi, Former President of the European Commission, Former Prime Minister of Italy
- Prof. Stefano Quintarelli President of Steering Committee AGID (Government agency for digital Italy) and member of Scientific Committee of Associazione Copernicani
- Prof. Jeffrey Sachs, Economist, academic, public policy analyst
- Prof. Stephen Stedman, Senior Fellow at the Freeman Spogli Institute for International Studies, Professor of Political Science at Stanford University and Secretary-General of the Kofi Annan Commission on Elections and Democracy in the Digital Age
- Mons Marcelo Sánchez Sorondo, Argentine Catholic bishop and Chancellor of the Pontifical Academy of Sciences and the Pontifical Academy of Social Sciences
- Mr Roberto Viola, Director-General of DG CONNECT – European Commission

For the full programme please see:

https://ifdad.org/wp-content/uploads/2020/12/20201209-IFDaD_10-11-program.pdf

Although the Forum had to be moved online due to the ongoing pandemic, the event saw the participation of over 1000 participants from across the world during the 2-day of debates.

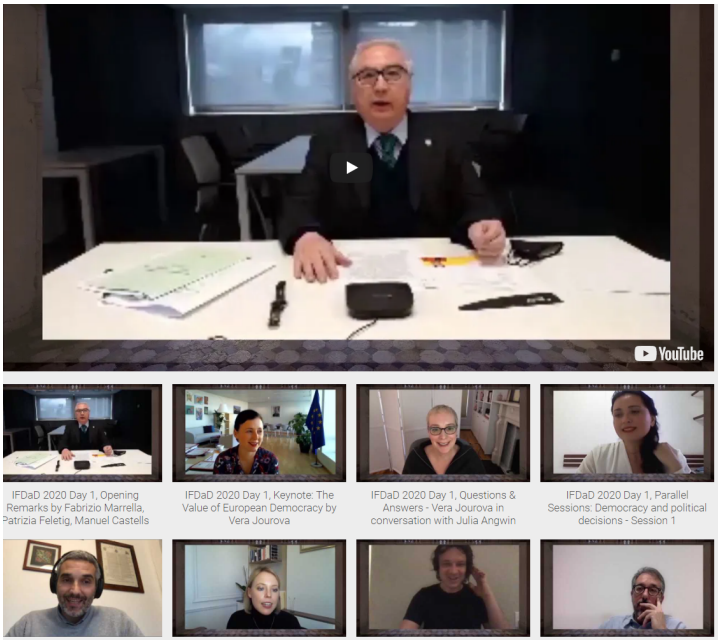


Figure 8: A screenshot of the International Forum on Digital and democracy

3.3.3.2 EXPERT ROUNDTABLE ON DISINFORMATION, NARRATIVES AND THE MANIPULATION OF REALITY

The role of narratives in shaping people's minds has become an important area of research and debate. An increasingly influential stream of research demonstrates the importance of cognition, emotion and values in political decision-making, and that political cognition is emotionally shaped by certain narratives. The session explored the ethical dilemmas around disinformation and the use of narratives and emotions in manipulating reality, in particular in recent years when entire societies have made choices that seem "rationally" counterintuitive.

The COVID-19 pandemic and the accompanying 'infodemic' described by the WHO has placed science disinformation at the centre of global debates. The outcomes of this debate will used as a starting basis for RIE's work in establishing a European Observatory on Narratives.

3.3.3.3 EXPERT ROUNDTABLE ON DATA AND ETHICS IN A POST-COVID WORLD

This session explored the ethical dilemmas around Artificial Intelligence (AI) and Big Data in Ethics, Society, and Law, and related technical challenges in Data Science and Explainable AI, with a special emphasis on trustworthiness. The outcomes of the discussions will be used in developing a White Paper, that will be published in 2021.

The keynote speech, “Matching Ethics and Law in AI: policy and practical implications of the Legal Legs of a “Trustworthy” SoBigData++”, by Professor Giovanni Comandè, presented the perspective of SoBigData++ on this timely issue.

The roundtable saw the participation of key policymakers including:

- Alain Strowel, Professor of Intellectual Property and Media Law at UCLouvain
- Iban Garcia del Blanco, Member of the European Parliament (Vice Chair of JURI/AIDA - S&D)
- Isabel Wiseler-Lima, Member of the European Parliament (AFET/DROI - EPP)
- Karen Melchior, Member of the European Parliament (JURI/FEMM - Renew)
- Anna-Misel Asimakopoulou, Member of the European Parliament (Vice Chair of INTA/AIDA - EPP)
- Kim Van Sparrentak, Member of the European Parliament (IMCO/AIDA - Greens)
- Alessandro Annoni, President of the International Society for Digital Earth

The outcomes of this debate are being developed into a white paper to be presented in the second half of 2021.

Satellite session

Organised as a high-level Round-Table. this session has limited seating and will be allocated on a first-come first-served basis

11:00 - 12:30 **EXPERT ROUNDTABLE ON DATA AND ETHICS in a Post-COVID WORLD**

This session is organised with SoBigData++ Consortium and will explore the ethical dilemmas around AI and Big Data in Ethics, Society, and Law, together with technical challenges in data science and Explainable AI, with a special emphasis on Trustworthiness

Keynote: **Matching Ethics and Law in AI: policy and practical implications of the Legal Legs of a "Trustworthy" AI**

Giovanni Comandé, *Full Professor of Private Comparative Law at Scuola Superiore S. Anna Pisa, Italy. PhD. SSSA, LLM Harvard Law School, Founder and Director of the LIDER-LAB (www.li-der-lab.eu). Attorney at law (Pisa since 1995). (New York Bar since 1997). Mediator mediation trainer. He has taught and researched in numerous universities in the world. Coordinator of LEADs Legality Attentive Data Scientists H2020 project*

PANELLISTS:

Alain Strowel, *Professor of Intellectual Property and Media Law at UCLouvain*
Iban Garcia del Blanco, *Member of the European Parliament (Vice Chair of JURI/AIDA - S&D)*
Isabel Wiseler-Lima, *Member of the European Parliament (AFET/DROI - EPP)*
Karen Melchior, *Member of the European Parliament (JURI/FEMM - Renew)*
Anna-Misel Asimakopoulou, *Member of the European Parliament (Vice Chair of INTA/AIDA - EPP)*
Kim Van Sparrentak, *Member of the European Parliament (IMCO/AIDA - Greens)*
Alessandro Annoni, *President of the International Society for Digital Earth*

EXPERT PANEL:

Fosca Giannotti, *Director of research of computer science at the Information Science and Technology Institute "A. Faedo" of the National Research Council, Italy. Coordinator of SoBigData++. ERC Advanced Grant entitled XAI – Science and technology for the explanation of AI decision making.*
Jeroen Van den Hoven, *Professor of Ethics and Technology. Delft University of Technology. Editor in Chief of Ethics and Information Technology and Member of the European Group on Ethics in Science and New Technologies (EGE).*
Dino Pedreschi, *Professor of Computer Science at the University of Pisa. "Scientific director of the line "Social AI" in the EU network Humane-AI-Net.*
Josep Domingo, *Distinguished Professor of Computer Science and an ICREA-Acadèmia Researcher at Universitat Rovira i Virgili, Tarragona, Catalonia. UNESCO Chair in Data Privacy.*
Caterina Sganga, *Associate Professor in Comparative Private Law. Coordinator of the H2020 project reCreating Europe.*
Juan M. Durán, *Assistant Professor of Epistemology, Ethics and Technology. Delft University of Technology. Herbert A. Simon awardee for his research on Philosophy and Computing*
Kalina Bontcheva, *Research Professor at the University of Sheffield. Deputy Coordinator of the SoBigData++ project. Scientific director of the WeVerify project.*

The outcomes of the discussions will help us develop a white paper on this topic to be published in early 2021.

SoBigData

Figure 9: The program of the Round Table on Data and Ethics in a post-covid world

3.4 Final Assessment

The impact of initial outreach activities towards policymakers and the public at large have been very successful allowing the project SoBigData++ to be presented to a number of key stakeholders, impact key discussions in the policy arena and better understand the policy landscape. For these first 18-months priority was given to the outreach towards policymakers seeing the wealth of issues of relevance on the agenda and to allow different explanatories and other activities of SoBigData++ to strengthen before planning key outreach activities. This decision was further encouraged by the ongoing pandemic that changes the timelines and workings of the project slightly. With a clearer overview of priorities for the coming years, SoBigData++ is continuing to develop its outreach strategy towards policymakers and the public at large. In order to facilitate this internally, a monthly newsletter will be published and circulated within the consortia to create stronger awareness of what is happening in the policy sphere and ideas for action or meeting. This will make it easier for the different parts of the consortium to be aware of the opportunities of engaging with policymakers directly.

4 Stakeholder Analysis Updates

Stakeholder involvement and analysis represent key activities of the SoBigData RI. The main aims are to promote the use of RI serviced and accelerating the technology transfer. During last year, it has been difficult to acquire new Stakeholders for the consortium. This is mainly due to the COVID-19 pandemic: travel restrictions caused by COVID severely disrupted the ability of researchers to travel, meet, and engage with potential new stakeholders. This means that, with respect to the stakeholder analysis performed for deliverable D3.1 "Initial Dissemination and Impact Plan", there have been little changes.

Overall, 9 of the partners of the consortium indicated a total of 20 new stakeholders, while the other partners reported no new stakeholder. In Figure 10, we report the macro-areas of activities of the new stakeholders and in Figure 11 we report the application fields.

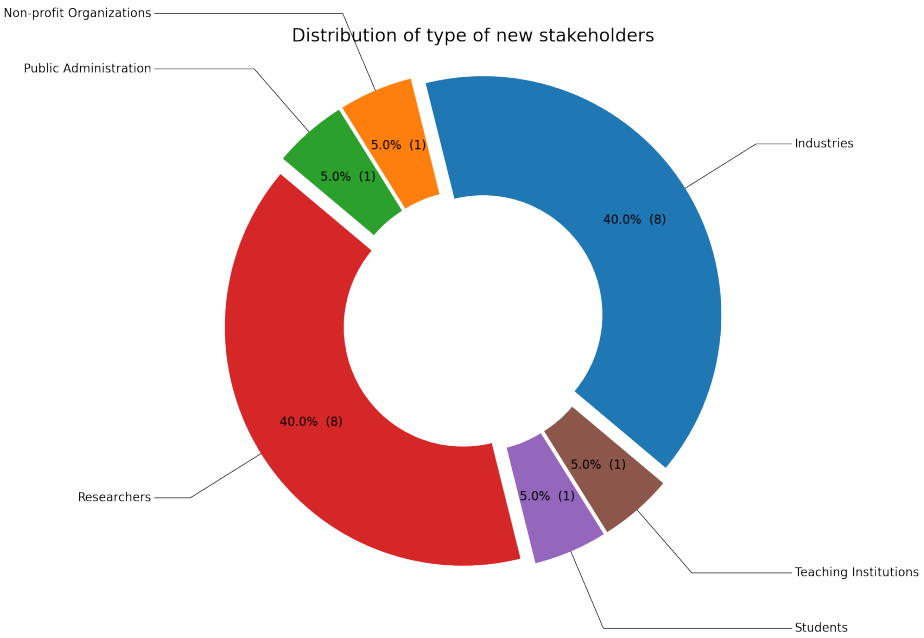


Figure 10: Distribution of macro-areas of activities of newly acquired stakeholders

Both for the active stakeholders and the ones who still have to begin their collaboration, the main contribution to consortium partners is the sharing of data for analysis and the co-development of applications. This is somewhat reflected in the macro-areas analysis, as industries and researchers can both benefit from the sharing and analysis of new datasets. The overall distribution of the application fields and macro-areas for the new stakeholders is similar to what has been found in deliverable D3.1.

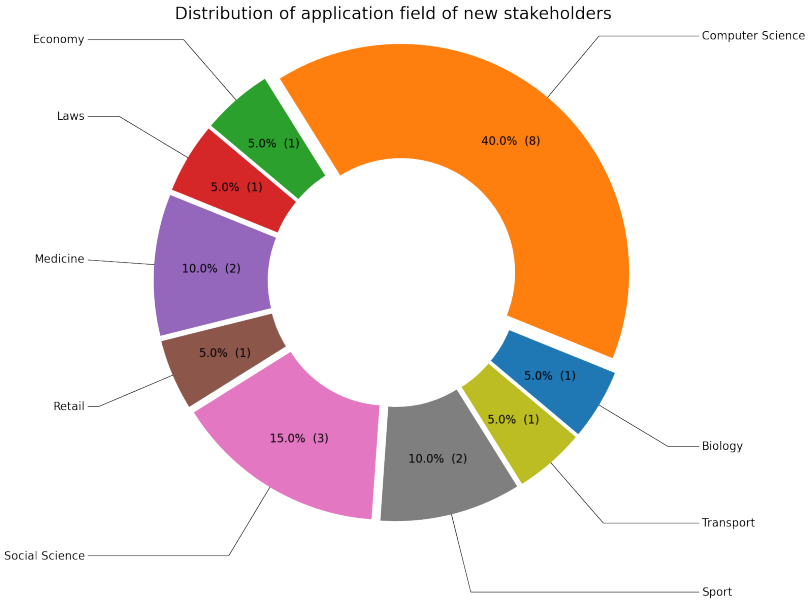


Figure 11: Distribution of application fields of newly acquired stakeholders

As the world keeps battling the COVID-19 pandemic, the hope is that with new countermeasures, researchers and personnel of the consortium may begin to travel again, thus granting them more reach in the coming future.

5 Conclusions

This deliverable provided an overview of the dissemination activities carried out in the first period of the project (01 January 2021 - 30 June 2021), together with a report on the project's impact so far.

As we can appreciate, reading the document, all the actions done to reach general public, scientific communities, and potential adopters through various dissemination channels have reached the objectives and KPI established in D3.1. Furthermore, we provided a comprehensive list and short description of all the events performed. Finally, a short update of the stakeholders' analyses is provided.

Furthermore, we can consider successful also the project's impact of outreach towards policy makers, despite the COVID-19 crisis.