



<i>Project Acronym</i>	<i>SoBigData</i>
<i>Project Title</i>	<i>SoBigData Research Infrastructure Social Mining & Big Data Ecosystem</i>
<i>Project Number</i>	<i>654024</i>
<i>Deliverable Title</i>	<i>Training Programme 2</i>
<i>Deliverable No.</i>	<i>D4.8</i>
<i>Delivery Date</i>	<i>04 2018</i>
<i>Authors</i>	<i>Tobias Blanke (KCL), Mark Coté (KCL), Marco Braghieri (KCL)</i>



DOCUMENT INFORMATION

PROJECT	
Project Acronym	SoBigData
Project Title	SoBigData Research Infrastructure Social Mining & Big Data Ecosystem
Project Start	1st September 2015
Project Duration	48 months
Funding	H2020-INFRAIA-2014-2015
Grant Agreement No.	654024
DOCUMENT	
Deliverable No.	D4.8
Deliverable Title	Training Programme 2
Contractual Delivery Date	31 th August 2017
Actual Delivery Date	30 th April 2018
Author(s)	Tobias Blanke (KCL), Mark Coté (KCL), Marco Braghieri (KCL)
Editor(s)	Roberto Trasarti (CNR), Beatrice Rapisarda (CNR)
Reviewer(s)	Beatrice Rapisarda (CNR), Giulio Rossetti (CNR)
Contributor(s)	Tobias Blanke (KCL), Mark Coté (KCL), Marco Braghieri (KCL), Giulio Rossetti (CNR)
Work Package No.	WP4
Work Package Title	NA3_Training
Work Package Leader	KCL
Work Package Participants	CNR, USFD, UNIPI, FRH, UT, IMT, LUH, KCL, SNS, AALTO, ETHZ, TUDelft
Dissemination	Public
Nature	Report
Version / Revision	V1.1
Draft / Final	Final
Total No. Pages (including cover)	26
Keywords	Training Activities, Summer School, Hackathon, Training Course

DISCLAIMER

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

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

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

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

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

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

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

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

GLOSSARY

NAME	DEFINITION
Summer School	A series of lectures and activities that takes place during higher education summer holidays, generally on specific topics
Hackathon	Event where a number of people convene together in order to engage in computer programming in a limited time frame.
Training Course	Series of lectures and activities aimed at the development of a set of skills
Datathon	Event focused on data analysis and elaboration, where a number of people convene together in order to engage in computer programming in a limited time frame.

TABLE OF CONTENT

DOCUMENT INFORMATION	2
DISCLAIMER	3
GLOSSARY	4
TABLE OF CONTENT	5
DELIVERABLE SUMMARY	7
EXECUTIVE SUMMARY	8
1 Relevance to SoBigData	10
1.1 Purpose of this document.....	10
1.2 Relevance to project objectives	10
1.3 SoBigData project description.....	10
1.4 Relation to other workpackages	11
1.5 Structure of the document	11
2 Initial Planning	12
2.1 Den Haag Partner Meeting on WP4 Planned Training Activities	12
2.1.1 Upcoming SoBigData Training Activities	13
3 Overview of Planned Training Activities per Task	15
3.1 Task T4.1 – Summer Schools	15
3.1.1 Fake news, 25-29 june 2018	15
3.1.2 Lipari school on computational complex and social systems, 19-25 june 2018	15
3.2 Task T4.2 - Training Modules For Stakeholders	16
3.3 Task T4.3 - Datathons	16
3.3.1 ETH crypto datathon 2018, 4-9 june 2018.....	16
3.3.2 CNR soccer data challenge 2018, 12-13 october 2018	17
3.4 Task T4.4 – Addressing Gender and Diversity Issues in Data Science Through Training.....	17
3.5 Other Activities	17
3.5.1 Online Social Networks and Media: Network Proprieties and Dynamics (OSNED) Workshop at The Web Conference, 24 April 2018	17
3.5.2 Alternate track Conference at The Web Conference, 24 April 2018	18
3.5.3 Complexity 72h Workshop, 7-8 May 2018.....	18
3.5.4 SAO 2018 co-located with ASONAM - Social Human Behaviour Analysis through Online Social Networks and Media Workshop, 29 August 2018	18
3.5.5 4 th International Workshop on Dynamics on&of Networks, 28 August 2018	19
3.5.6 Complex systems Society Conference, 23 – 28 September 2018	19
4 Overview of Past Training Activities	20

- 5 Reporting Template, Responsibilities and Meeting Schedule 22**
 - 5.1 Reporting Template..... 22**
 - 5.2 Work Package Responsibilities per Partner 23**
 - 5.3 Meeting Schedule..... 24**
- 6 Operative Ethical Board 25**
- 7 Conclusions 26**

DELIVERABLE SUMMARY

The deliverable's objective is to report on the planned training activities organised by SoBigData partners in the period in exam.

The deliverable is organised in five sections, which describe: relevance according to the project's objectives, planning activity, an overview of each future activity according to the WP's tasks, past activities and WP4's work flow.

EXECUTIVE SUMMARY

Deliverable 4.8 is a deliverable that provides insight into future planned activities within WP4 of the SoBigData project. It begins by assessing the work package's main objectives and its relevance among the overall SoBigData project activities.

The deliverable then reports on the SoBigData Winter Meeting which has taken place in Den Haag in February 2018. During this meeting, a specific attention was given to WP4's future activities as each type of event corresponds to a task, such as the organisation of summer schools, the development of training modules for stakeholders, the organisation of datathons and actions which aim to address gender and diversity issues in data science through training.

A calendar for upcoming SoBigData training activities is then provided in order to summarise all future events which will take place in a number of countries, including Italy, Switzerland, France, Greece and Spain. Moreover, all future events have been analysed in light of the different tasks that Work Package 4 has.

Within the work package, the first task (T4.1) regards the organisation of summer schools and in the period in exam there will be two different events which will take place. The first summer school will be held in Paris, France and will regard the issue of fake news and will combine this matter with the approach developed during the GATE Summer Schools regarding the mining of social media and news content. The second summer school which will take place in the period in exam will be held in Lipari, Italy and will be focused on modern data analysis, especially in the big data field, and includes a focus on data mining from Internet of Things. Moreover, it will feature a workshop organised by SoBigData partners, which will focus on digital democracy.

The second task (T4.2) regards training modules for stakeholders. Having collected training materials from different SoBigData partners, WP4 aims to produce a taxonomy and metadata scheme regarding the submissions it has received. This activity will be put in place in order to create content standards and integrate existing training materials.

The third task (T4.3) regards the organisation of datathons. Within this field, the deliverable describes three different events which will take place. The first is the Complexity 72h workshop which will be held in Lucca, Italy, the second will be the ETH Crypto Datathon, which will be held in Zurich, Switzerland and the third will be the Soccer Data Challenge, which will be held in Pisa, Italy.

Moreover, regarding task T4.4, which aims to address the current imbalance in the representation of all gender and minority groups, we describe the planning of an event which will take place in London, United Kingdom in 2018. This event will feature the participation of R-Ladies, a group dedicated to promoting gender diversity in the R community.

The deliverable then features an overview of activities which have taken place between the production of Deliverable 4.1 and the beginning of the period in exam in Deliverable 4.8.

Aside from activities which specifically one of WP4's tasks as the above, a number of other activities have been organised which will feature SoBigData partners. Among these, there are workshops, lectures and conferences.

The final section of the deliverable is dedicated to WP4's workflow, in order to recap all activities which have been put in place in order to assure the best possible cooperation and communication between different partners which are involved in this work package.

1 RELEVANCE TO SOBIGDATA

Work Package 4, entitled Training, aims to establish a joint training and education initiative on social big data within the European Research Area. The Work Package explores and develops both conventional and unconventional training experiences for master students, PhD students and early career post-doctoral researchers as well as an academically interested general public. Likewise, Work Package 4 proposes campaigns aimed at high school students to promote interest in data science with special emphasis on the underrepresentation of minority genders. These experiences include a number of different activities such as summer schools, datathons, training courses and specific high-school oriented courses.

1.1 PURPOSE OF THIS DOCUMENT

The purpose of this document, entitled ‘Training Programme 2’ is to provide insight on future events within the SoBigData project for the period in exam. As such, Deliverable 4.8 provides an overview of the different type of activities which have been planned by SoBigData partners, likewise providing a reference of past activities and insight on the activities which have been put in place in order to enhance WP4’s workflow.

1.2 RELEVANCE TO PROJECT OBJECTIVES

The training activity within the SoBigData project is aimed at establishing a unique, joint training and education resource centred on social big data. In order to do so, the Work Package explores and develops conventional and unconventional training experience for master and PhD students and post-doctoral trainees.

These experiences include a number of different activities. Project-oriented summer schools and datathons are organised in order to match research (and industrial) needs and people skills. Training courses are developed in order to create new start-up initiatives related to big data. Moreover, actions are put into place in order to illustrate opportunities in the big data field to address gender and diversity issues in data science through training.

1.3 SOBIGDATA PROJECT DESCRIPTION

The SoBigData project’s aim is to create a pan-European research infrastructure. This infrastructure will integrate already established national infrastructures. To this end, the SoBigData project has defined a series of key priorities:

1. Better access to the best national research infrastructures
2. Training a new generation of mobile researchers
3. More effective national research systems
4. Optimal transnational cooperation
5. Accelerating innovation through partnerships with industry
6. Effective data, method, and knowledge sharing

1.4 RELATION TO OTHER WORKPACKAGES

WP2: as training activities fall under the legal and ethical framework of the SoBigData infrastructure

WP3: as training activities have a strong relationship with the dissemination and impact strategies developed for the whole SoBigData project

WP5: as training activities are connected to the innovation activities aimed at industry and other stakeholders

WP10: as training activities will present and educate on the new methodologies and technologies that SoBigData is developing

WP11: as training activities relate to the construction of a benchmarking and evaluation framework for big data analytics and social mining methods.

1.5 STRUCTURE OF THE DOCUMENT

The document describes future events in the period in exam, starting from the planning activity which has taken place during the SoBigData Winter Meeting.

The deliverable details, for each upcoming event, name, date and type of activity.

Moreover, it provides a description for each event organised following the four tasks that Work Package 4 has, which involve the organisation of different activities aimed at different stakeholders.

D4.8 then features a table summarising the activities which have taken place since the production of Deliverable D4.1 and focuses on events ranging from M18 (February 2017) to M32 (April 2018).

Finally, the deliverable recaps Work Package 4's workflow by summarising the reporting template, responsibilities for each partner and meeting schedule, which have all been provided in order to offer an overview of WP4 activities.

2 INITIAL PLANNING

Tasks for Work Package 4 have been discussed and agreed on the last SoBigData Winter Meeting which was held in Den Haag from 19 to 21 February 2018. Jointly with a general overview of the status of all Work Packages, specific attention was given to an overview of Work Package 4, focusing on the SoBigData Training Platform and Materials and planned Training Activities.

2.1 DEN HAAG PARTNER MEETING ON WP4 PLANNED TRAINING ACTIVITIES

During the SoBigData Winter Meeting in Den Haag, a specific focus was dedicated to WP4 planned Training Activities and the following steps were agreed in order to progress on the work package's tasks. The four event types which it was agreed that the WP4 will focus on are:

- Summer Schools
- Datathons/Hackathons
- Workshops
- Training Activities outside academia

Moreover, during the SoBigData Winter Meeting in Den Haag, it was also agreed that WP4 will go forward with a collection activity among SoBigData Partners regarding

- Data Scientist Training Materials

The Work Package, along with the planning activity, has been continuously surveying existing activities on data science education among the consortium partners in order to achieve a comprehensive overview that will help coordinate work within the SoBigData project itself.

2.1.1 UPCOMING SOBIGDATA TRAINING ACTIVITIES

The calendar for upcoming SoBigData Training Activities is the following:

Event Type	Event Name	Event Date	Place
Workshop	Online Social Networks and Media: Network Proprieties and Dynamics (OSNED) at The Web Conference	24 April 2018	Lyon, France
Conference	The Web Conference	25 April 2018	Lyon, France
Workshop	Complexity 72h	7 - 11 May 2018	Lucca, Italy
Datathon	Crypto Datathon 2018	04 - 09 June 2018	Zurich, Switzerland
Summer School	Fake News	25 - 29 June 2018	Paris, France
Summer School	Lipari School on Computational Complex and Social Systems	19 - 25 July 2018	Lipari, Italy
Workshop	'From swarm intelligence to digital democracy: new tools for a complex society' within Lipari School on Computational Complex and Social Systems	19 - 25 July 2018	Lipari, Italy
Workshop	SAO 2018 co-located with ASONAM Workshop on Social Human Behaviour Analysis through Online Social Networks and Media	29 August 2018	Barcelona, Spain

Event Type	Event Name	Event Date	Place
Workshop	Dyno 2018 co-located with ASONAM The 4th International Workshop on Dynamics on&of Networks	29 August 2018	Barcelona, Spain
Conference	Complex Systems Society 2018	23 – 28 September 2018	Thessaloniki, Greece
Datathon	Soccer Data Challenge 2018	12 – 13 October 2018	Pisa, Italy

3 OVERVIEW OF PLANNED TRAINING ACTIVITIES PER TASK

The following is an overview of all planned activities divided for each of the four tasks that Work Package 4 has within the SoBigData Project. Since WP4's main focus is Training the four main tasks regard four different types of activities, such as the organisation of summer schools, the development of training modules for stakeholders, the organisation of datathons and actions which aim to address gender and diversity issues in data science through training.

3.1 TASK T4.1 – SUMMER SCHOOLS

The general aim in organising summer schools within the SoBigData project is to prepare the PhD students in science and engineering to work as data scientists. While summer schools provide an introduction to big data analysis, they are also based on a hands-on approach, in order for participants to develop skillsets using the SoBigData Research Infrastructure.

The University of Sheffield has taken a lead in planning and organising the summer schools for the reporting period. The two summer schools which will take place in the timeframe in exam, focusing on digital democracy and fake news, will give the project a perfect mixture of integrating existing excellent work as well as driving new agendas.

3.1.1 FAKE NEWS, 25-29 JUNE 2018

The GATE summer school, originally taking place in Sheffield and organised by SoBigData partner USFD, is taken to Paris to work directly with social scientists at Sciences Po. This summer school will take place from 25 to 29 June 2018. Science Po is based in Paris, France and combines different approaches in social sciences, politics and International Studies with a multidisciplinary approach. Within Science Po, the Médialab was established in 2009 in order to help social sciences and humanities students to take full advantage of big data, focusing on three main fields such as digital methods, media studies and social theory. The Médialab's organisation is based on three main areas such as social sciences, digital data engineering and information design. Hence, within this framework, the SoBigData Summer School on Fake News will combine the approach developed within the GATE Summer School with activities developed within Science Po's Médialab on the subject of fake news.

In 2018, the GATE Summer School will reach its 11th edition, and will focus on mining social media and news content with GATE, an open-source infrastructure developed in order to provide analysis of social media content.

3.1.2 LIPARI SCHOOL ON COMPUTATIONAL COMPLEX AND SOCIAL SYSTEMS, 19-25 JUNE 2018

The Lipari Summer School 2018 edition's is aimed at providing insight into modern data analysis, especially in the big data field, and includes a focus on data mining from Internet of Things. The Summer School hosts a number of speakers from SoBigData partners such as IMT in Lucca, King's College London, SNS in Pisa jointly with speakers from University College London, New York University and Queen Mary University London and University of Catania. Its directors are Alfredo Ferro from University of Catania, Dirk Helbing

from ETH Zurich and Claudio Cioffi-Revilla from George Mason University, United States. The steering committee includes individuals from a wide range of universities including SoBigData partners such as University of Pisa and CNR in Pisa.

Among the Summer School activities there will be a specific workshop entitled 'From swarm intelligence to digital democracy: new tools for a complex society' which will be hosted by SoBigData partners UNIFI, CNR, IMT and ETH Zurich.

All information regarding this event, including lecture abstracts, are available at <https://complex.liparischool.it/>

3.2 TASK T4.2 - TRAINING MODULES FOR STAKEHOLDERS

Task 4.2 is focused on producing open source training modules, which have been developed within training events. These training modules are then integrated in each partner's learning and teaching environments, starting from existing sources. Aside from harmonising existing training modules, T4.2 is centred on the production of new training modules for specific stakeholders that are currently under-served in the big data field, such as social scientists and humanities researchers. All training modules are developed in open educational standards in order to be easily integrated in e-learning environments.

Since the SoBigData project has started, a lot of additional resources on open education in social data science have become available. Hence, SoBigData has adjusted and now concentrates on harvesting these resources into its open education ecosystem, which is an integral part of its infrastructure.

We continue our call for open e-learning contributions by partners and at the same time evaluate the Jupyter framework for creating a common framework for these resources. We continue a call for digital educational materials directly, together as set of criteria for "nice-to-haves" like the integration of research data, research-led methods, etc. Currently, 22 training resources have been collected within the consortium, ranging from tutorial and short courses to master program modules.

The next step will be to create a taxonomy and metadata scheme as well as content standards to integrate the existing training materials which have been harvested from different SoBigData partners.

3.3 TASK T4.3 - DATATHONS

These events aim to provide theoretical and practical advice to a group of participants in order to collect, analyse and visualise big data in order to address relevant issues. Based on interactivity, datathons aim to bring together computer programmers, storytellers, graphic designers and statisticians who are the grouped in teams and within an event aim to create innovative tools and services on based on big data analysis.

3.3.1 ETH CRYPTO DATATHON 2018, 4-9 JUNE 2018

This event will take place in Zurich, Switzerland between 4-9 June 2018. The Crypto Datathon is aimed at international students which are currently undertaking master or graduate studies and are skilled in big data collection and processing, data mining and Natural Language Processing, machine learning and web systems and web crawling. This event is focused on the scientific analysis of complex systems around

cryptocurrencies and is built around lectures and projects which will be developed during the datathon. Speakers at this event include individuals from several SoBigData partners such as ETH Zurich, SNS in Pisa and the L3S Research Center in Hannover.

All information regarding this event can be found at <https://www.cryptodatathon.com/>

3.3.2 CNR SOCCER DATA CHALLENGE 2018, 12-13 OCTOBER 2018

The Soccer Data Challenge will take place in Pisa, Italy, on 12th and 13th october, 2018. It is aimed at Italian "hackers" skilled in data processing and analysis. The event is focused on the analysis of soccer data, i.e., data describing all the events involving the ball generated by players during all the games of an entire season of the Italian Serie A. A difficult analytical problem on soccer data will be proposed to the hackers, and the solutions provided will be evaluated by a heterogeneous jury composed by people from both sports industry and data science (professional data scientists, researchers in data science, professional soccer coaches and players, sports journalists). Before the award ceremony there will be a panel on the topic "Big Data analytics in sports", involving several individuals from the SoBigData infrastructure.

All information regarding this event can be found at ASK LUCA FOR THE URL (not ready yet)

3.4 TASK T4.4 – ADDRESSING GENDER AND DIVERSITY ISSUES IN DATA SCIENCE THROUGH TRAINING

In order to address the current imbalance in the representation of all gender and minority groups, T4.4 aims to organise specific activities. This task's general objective is to leverage existing networks in order to raise awareness among targeted groups on the possibilities within the data science field.

At the end of 2018, King's College London will organise a datathon specifically dedicated to addressing gender issues in data science. KCL will work with a new partner, R-Ladies, which is dedicated to promoting gender diversity in the R community. R is a de facto standard in statistical computing and visualisation and R-Ladies was founded in 2012 and is organised in several meetups around the world. Its aim is to address underrepresentation of minority genders within the R community. In order to achieve proportionate representation within this field, it encourages and supports minority gender R programmers and their activities through a global network of mentors and developers.

3.5 OTHER ACTIVITIES

Aside from activities which are specific to one of the tasks WP4 has, a number of other activities have been organised which will feature SoBigData partners. Among these, there are workshops, lectures and conferences. The following is an overview of each activity.

3.5.1 ONLINE SOCIAL NETWORKS AND MEDIA: NETWORK PROPRIETIES AND DYNAMICS (OSNED) WORKSHOP AT THE WEB CONFERENCE, 24 APRIL 2018

The first International Workshop on Online Social Networks and Media: Network Proprieties and Dynamics will be co-located with The Web Conference in Lyon, France and will be held on 24 April 2018. The activity

will focus on the analysis of network properties and dynamics at all scales within online social networks and media. There will be a special focus on emerging phenomena dynamics, such as diffusion of information, the acquisition of the role of influencer, trust among users, polarization of opinions, diffusion of 'fake news'. The program committee chairs include researchers from SoBigData partner CNR.

All information regarding this event can be found at <https://osned.netsons.org/>

3.5.2 ALTERNATE TRACK CONFERENCE AT THE WEB CONFERENCE, 24 APRIL 2018

Within The Web Conference which will be held in Lyon between 23 – 27 April 2018, among the alternate and other track sessions, there will be a specific focus on international projects. Among these, the SoBigData project will be featured with a focus on social mining and big data ecosystems. The session will be hosted by SoBigData partners CNR and USFD.

All information regarding this event can be found at <https://www2018.thewebconf.org/alternate-sessions-details/?track=international-projects#international-projects>

3.5.3 COMPLEXITY 72H WORKSHOP, 7-8 MAY 2018

The Complexity 72h Workshop is organised and hosted by SoBigData partner IMT School for Advanced Studies and will take place in Lucca, Italy between 7 and 8 May 2018. Its organisers are Alberto Antonioni from University College London, Eugenio Valdano from URV in Tarragona Spain and Tiziano Squartini, assistant professor at the IMT School for Advanced Studies in Lucca, which is a SoBigData partner. This event aims to bring together young researchers which are currently involved in different complex systems fields. Participants will form working groups which will work on carrying out a project over the span of three days (72 hours). Each group's objective is to upload a report of their activity within the workshop on arXiv by the end of the event. Participants will be assisted by a team of tutors which will guide each group in the development of their activity. Teamwork projects will be based on a wide range of subjects, such as link prediction on social networks, network sensitivity, reaching brain networks beyond correlation analysis, altered modularity in psychiatric disorders, bitcoin analysis and human contact networks and multi-pathogen interaction. Moreover, individuals will attend lectures from scientists which focus on different fields of complex systems. Lectures will be held by professors and researchers from SoBigData partners such as CNR in Pisa and IMT in Lucca jointly with professors and researchers from INSERM, the French National Institute of Health and Medicine Research and the University of Zurich.

All information regarding this event can be found at <https://complexity72h.weebly.com/>

3.5.4 SAO 2018 CO-LOCATED WITH ASONAM - SOCIAL HUMAN BEHAVIOUR ANALYSIS THROUGH ONLINE SOCIAL NETWORKS AND MEDIA WORKSHOP, 29 AUGUST 2018

SAO is a workshop which will be co-located with the ASONAM (Advances in Social Network Analysis and Mining) Conference which will be held in Barcelona, Spain on 28 – 31 August 2018. Online social networks and media can be viewed both as challenging technical environment for computer scientists but also as a source of individual behaviour analysis through the digital footprints, regarding mobility patterns, opinions and interest spreading. Within this frame, SAO aims to provide an interdisciplinary reference forum focusing

on the use of online social networks and media as a ‘big data microscope’ which can be used to investigate human social behaviour. Workshop chairs include researchers from SoBigData partner CNR.

All information regarding this event can be found at <http://cnd.iit.cnr.it/sao2018/index.html>

3.5.5 4TH INTERNATIONAL WORKSHOP ON DYNAMICS ON&OF NETWORKS, 28 AUGUST 2018

The 4th International Workshop on Dynamics on&of Networks (DYNO) will be co-located with the ASONAM (Advances in Social Network Analysis and Mining) Conference which will be held in Barcelona, Spain on 28 – 31 August 2018. The aim of the DYNO workshop is to provide a place for researchers to share their latest results on both dynamics on the network, such as information propagation and viral diffusion and dynamics of the network, which revolves around the new relations that are created as new entities join the network and old ones vanish. The general purpose of this workshop is to encourage principled research that will lead to the advancement of the social science in time-evolving networks. The organising committee and program committee feature researchers from SoBigData partners such as CNR and UNIPI.

All information regarding this event can be found at <http://dyno.sci-web.net/>

3.5.6 COMPLEX SYSTEMS SOCIETY CONFERENCE, 23 – 28 SEPTEMBER 2018

The Complex Systems Society Conference will be held in Thessaloniki, Greece between 23 – 28 September 2018. Since 2004, the European Conference on Complex Systems organized by the CSS, is the most important annual meeting for the complex systems research community, and previous editions have been held in Lucca, Italy (2014); Tempe, Arizona, United States (2015); Amsterdam, Netherlands (2016); Cancun, Mexico (2017). The conference is focused on the field of Complex Systems and in 2018 will feature a lecture as plenary speaker by Fosca Giannotti, from SoBigData partner CNR.

All information regarding this event can be found at <http://ccs2018.web.auth.gr/>

4 OVERVIEW OF PAST TRAINING ACTIVITIES

A number of Training Activities has taken place among various SoBigData Partners since the production of Deliverable D4.1, which detailed activities organised between M6 (February 2016) and M17 (January 2017) of the SoBigData Project. Hence, the period in exam ranges from M18 (February 2017) to M32 (April 2018).

	Event Name	Event Type	Date
M17	SoBigData Master Program	Master Program	January – December 2017
M20	PhD+ program	Lectures and coaching and mentoring activities	19 April – 25 May 2017
M22	Computational Fake News Analysis: A Practical Workshop	Workshop	15 – 16 June 2017
M22	10 th GATE Summer School	Summer School	26 – 30 June 2017
M23	Lipari Summer School on Computational Complex and Social Systems	Summer School	16 – 22 July 2017
M23	2nd SoBigData Workshop City of Citizens @ Lipari Summer School	Workshop	19 July 2017
M25	SoBigData Conference: Data Science for Society	Conference	8 September 2017
M25	European Conference on Machine Learning & principles and practices of knowledge discovery in databases	Workshops	18 – 22 September 2017
M25	Nowcasting Well Being in Societies	Workshop	29 September 2017

	Event Name	Event Type	Date
M26	Data Science in Societal Debates Workshop at the 4 th IEEE International Conference on Data Science and Advanced Analytics	Workshop	19 October 2017
M27	Social Impact Data Hack	Datathon	10 – 12 November 2017
M27	GARR Conference – The data way to science	Conference	15 – 17 November 2017
M27	Big data and alternative data sources on Migration: from case studies to policy support	Workshop	30 November 2017
M30	PhD+ program	Lectures and coaching and mentoring activities	6 February – 8 March 2018
M30	SoBigData Winter Meeting	Meeting	20 February 2018

Regarding the activities that have taken place in the period in exam, SoBigData scientists from different project partners have taken part in a wide range of different activities, constantly expanding the project's horizon and its connections with relevant entities in the field. A complete overview of activities, including activities which have been organised externally from the SoBigData project but to which SoBigData partner have taken part in can be found at <http://www.sobigdata.eu/events/past>.

Within the events organised by SoBigData partners, we can provide specific focus on the 'Social Impact Data Hack' which took place in Tartu, Estonia between 10 – 12 November 2017. All project partners contributed to the immense success of the activity. The topic of the datathon was the analysis of the impact of public investment subsidies on enterprise performance and their network structure in Estonia during 2011–2016. There were 220 participants, 12 teams and 18 mentors from all around Europe. The analysis used micro-level data on enterprise VAT and payroll declarations from Estonian Tax and Customs Board, including unique data on sales between enterprises. The tax data have been combined with data on investment subsidies from Estonian Agricultural Registers and Information Board.

5 REPORTING TEMPLATE, RESPONSIBILITIES AND MEETING SCHEDULE

In order to assess the organisation, participation and outcome of each training activity, a reporting template has been produced. For each training activity within the SoBigData project, partners are required to provide insight on the event type, its organisation and data on participants and outcome.

In this section, we are also detailing the responsibilities for each partner involved in Work Package 4 and the general meeting schedule that has been developed in order to enhance the workflow between WP partners.

5.1 REPORTING TEMPLATE

The following is the reporting template that partners are asked to provide after each event they have organised in order to produce a homogeneous dataset on organisation, participants and outcome.

- 1) *What were the training objectives?*
- 2) *How many people/participants attended (please break down by gender and age (youth, mature, old), and other relevant groupings (such as community, private professional, civil society, government professional)).*
- 3) *How did the training sessions attempt to address the training objectives?*
- 4) *What was the overall cost of holding the training?*
- 5) *What are the immediate and longer-term benefits resulting from the event?*
- 6) *How will you monitor or evaluate the outcome of this event into the future?*
- 7) *Was there anything you would do differently next time?*
- 8) *If you have any pictures of the training event, please use to illustrate your report.*

5.2 WORK PACKAGE RESPONSIBILITIES PER PARTNER

The following tables detail person-months per participant for each of the SoBigData partners involved in Work Package 4 and their main responsibilities.

Participants short name	KCL	CNR	USFD	UNIFI	FRH	UT	LUH	AALTO	ETHZ	TU Delft
Person-months per participant	10	6	4	4	4	4	4	2	6	4

Partner	Responsibilities
KCL	Overall management; Task lead for T 4.2 (including reporting)
CNR	Summer schools; training modules (with particular focus on integration with RI tools) and datathons; widening participation
USFD	Task lead for T 4.1 (including reporting); workshop on widening participation
UNIFI	Task lead for T 4.4 (including reporting); workshop on widening participation; datathon
FRH	Training materials on visualisation; support widening participation
UT	Datathon
LUH	Summer School and widening participation
AALTO	Training material
ETHZ	Task lead for T 4.3 (including reporting); summer school; datathon
TU Delft	Training material on data science ethics; support on widening participation

5.3 MEETING SCHEDULE

The following is the meeting schedule among WP4 partners, which has been developed and followed in order to enhance communication between SoBigData partners in this work package.

Date & Place	Intra WP	Multi-WPs WPs number	Telecom	In Person	Objective / Deliverable / Activities
Kick-Off Meeting in Pisa					Kick-off of activities: agree upon general work plan and responsibilities; detailed planning for next months
Approximately every 3-6 months; Virtual	X		X		Regular telecom meeting to discuss progress, issues and planning for next period
Various Workshops	X	X	X	X	ad-hoc telecoms when required

6 OPERATIVE ETHICAL BOARD

The General Assembly has also decided to create an additional Board in the project called Operative Ethical Board, which is responsible to be a contact point for students and visitors. This board will be used for example to validate the future TNA applications or to help the students if they have doubts for their projects (avoiding to use the local ethical boards which are usually difficult to reach). Moreover this board will be responsible also for validate the experiments, datasets and methods which will be published by the e-infrastructure platform. The researcher in this new board are: **Francesca Pratesi** which will be also the chair of the board, **Maria Francesca Donati**, **Stefanie Hänold**, **David van Putten**. The project will co-found the board covering the PM of the members.

7 CONCLUSIONS

The current deliverable reported the decisions as well as planned actions, regarding WP4, as discussed within the SoBigData consortium during the Winter Meeting that took place in February 2018.

Within such work package, the planned actions regarding the first task (T4.1) involve the organisation of two summer schools: a first one, on fake news, will be held in Paris, France, the second will be held in Lipari, Italy and will be focused on big data analysis.

As second action, the collection and categorization of training materials available within the consortium were planned (T4.2). The extraction of metadata and description of a taxonomy will allow to create content standards and integrate existing training materials.

The third action, (T4.3), regards the organisation of three datathons: the first one is the Complexity 72h workshop which will be held in Lucca, Italy; the second will be the ETH Crypto Datathon, which will be held in Zurich, Switzerland; finally, the third will be the Soccer Data Challenge, which will be held in Pisa, Italy.

Finally, regarding task T4.4, the consortium planned an event, featured by the R-Ladies, which will take place in London, United Kingdom in 2018. Such event will aim to address the current imbalance in the representation of all gender and minority groups.

The deliverable also lists other training activities that involve members of the consortium such as the organisation conferences and workshops. For all the training activities was designed a common reporting template to produce a homogeneous synthetic information on organisation, participants and outcome of the activity.

Finally the deliverable include the creation of the operative ethical board to help students and visitors in their projects or to evaluate new content for the e-infrastructure.