Social Mining & Big Data Analytics

SoBigData

RESEARCH INFRASTRUCTURE ++

Deliverable D6.1

Periodic Report on TA Activities

Period 1

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	WP6 TA1 Transnational Access
WORK PACKAGE LEADER	USFD
WORK PACKAGE PARTICIPANTS	CNR, UNIPI, UT, IMT, LUH, SNS, AALTO, ETH Zürich, PSE, UNIROMA1, CEU, URV, BSC, UPF, KTH, UvA
DELIVERABLE NUMBER	D6.1
DELIVERABLE TITLE	Periodic Report on TA Activities – Period 1
AUTHOR(S)	Joanna Wright (USFD), Alice Tucker (USFD), Lucy Moffatt (USFD), Beatrice Rapisarda (CNR)
CONTRIBUTOR(S)	
EDITOR(S)	Beatrice Rapisarda (CNR)
REVIEWER(S)	Rajesh Sharma (UT), Valerio Grossi (CNR)
CONTRACTUAL DELIVERY DATE	31/12/2021
ACTUAL DELIVERY DATE	22/12/2021
VERSION	1
TYPE	Report
DISSEMINATION LEVEL	Public
TOTAL N. PAGES	31
KEYWORDS	Transnational access, Short-Term Scientific Missions

EXECUTIVE SUMMARY

This report details the Transnational Access facility within the SoBigData++ project during its first 2 years. It describes the impact of Covid-19 on this aspect of the project, the Call that went out in June 2021, the applications received, the selection process and the visits that went ahead together with visitor feedback and plans for the next 12 months.

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
TNA	Trans National Access
STSM	Short-Term Scientific Mission

TABLE OF CONTENTS

1 Relevance to SoBigData++

1.1 Purpose of this document

This document describes the Transnational Access activities carried out in the first 2 years of SoBigData++. It details the preparation of the first Call (addressing the impact of Covid-19), the application process, initial applicant statistics and outcomes.

1.2 Relevance to project objectives

As detailed below, the Transnational Access facility is integral to the project's objective of growing the SoBigData++ community and widening the reach of the project. TA visits are also a driver to share knowledge and expertise throughout the Data Science community. By benefitting from a SoBigData++ TA visit, the researcher is also guided and supported to become a responsible and ethical Data Scientist; this, in turn, will promote one of the project's primary objectives of creating and maintaining a trusted Research Infrastructure with a highly respected status in the field.

1.3 Relation to other work packages

Transnational Access, by definition, shares and disseminates expertise from many of the project's partners with the newer members of the SoBigData++ community. The expertise on offer covers multiple and wide-ranging skills and knowledge encompassing every WP within the project. The Exploratories focus research into specific areas that pull together the strands of the project across lateral and vertical thematic environments. It is this close relationship between the work packages under the umbrella of the overarching Exploratories that gives the Research Infrastructure its strength and stature within the Data Science community.

1.4 Structure of the document

The report is organized in 3 main sections:

- Section 1 gives the purpose of this document.
- Section 2 provides details of the transnational access call and visits for the first two years of the project.
- Section 3 details the plans and expectations of the next call and the next period of the project.

The report is completed by Appendices as follows:

- Appendix A: the Call text itself from the project website
  http://sobigdata.eu/content/call-2021-22-sobigdata-transnational-access
- Appendix B: shows the Application Form that must be completed by every applicant.

- Appendix C: is the Proposal Evaluation Form displaying the selection and review criteria.
- Appendix D: is the Agreement Template which is essentially the contract between the project and the visitor. This must be signed by a representative of the project as well as the visitor before a TA visit starts.

2 Transnational Access Calls during the Reporting Period

2.1 Introduction and the impact of Covid-19

Covid-19 has obviously had a huge impact on the project resulting in the postponement of all TA visits. With the various local, regional, national and international travel restrictions the Call was held back and all TA activities were put on hold.

A close eye was kept on the ever-changing pandemic situation and when it was deemed viable the Call was released in mid-June 2021. The number of applications has been inhibited due to travel issues and it is expected this will continue to be the case while the pandemic persists.

The project has put a number of processes and checks into place to provide some flexibility to potential visitors. The host and applicant (once their application has been approved) will liaise and discuss the timings of a visit and the travel restrictions of both departure and destination country. The host will provide some flexibility in case of unforeseen or unexpected circumstances and the visitor will be advised to factor in the unforeseen and unexpected into their plans as Covid-19 related issues are outside of the project's control. Furthermore, there is now an option for a successful applicant to postpone their visit for up to 6 months which should provide sufficient flexibility to a visitor.

The project is mindful that the pandemic may still be a feature for some time to come and as such will continue to monitor the situation and offer flexibility due to Covid-19 issues for the foreseeable future.

2.2 First Call

The Call was released in mid-June 2021 offering transnational visits, to complete a short-term scientific mission (STSM) of between 2 weeks to 2 months with up to €5,000 available to cover travel and living expenses. The project can now offer a visit to one of 17 European infrastructures providing wide ranging and varied expertise.

The project website details the Call and how to apply. The direct URL is detailed here: http://sobigdata.eu/content/call-2021-22-sobigdata-transnational-access, and a copy of the Call text is also available under Appendix A of this document. The Call provides all the relevant details and also the email address of the lead contact person in each infrastructure. There is information included relating to the pandemic and considerations regarding travel restrictions. Below, Table 1 details the 17 Infrastructures and their short names which will be used throughout the remainder of this document.

Short Name of the Infrastructure	Short name of the installation	City	Country
it.SoBigData.eu	CNR-TA	Pisa	IT
uk.SoBigData.eu	GATE-TA	Sheffield	UK

it.SoBigData.eu	UNIPI-TA	Pisa	Italy
ee.SoBigData.eu	UT-TA	Tartu	Estonia
it.SoBigData.eu	IMT-TA	Lucca	Italy
de.SoBigData.eu	L3S-TA	Hannover	Germany
it.SoBigData.eu	SNS-TA	Pisa	Italy
fi.SoBigData.eu	AALTO-TA	Helsinki	Finland
ch.SoBigData.eu	ETHZ-TA	Zurich	Switzerland
fr.sobigdata.eu	PSE-TA	Paris	France
it.SoBigData.eu	SAP-TA	Rome	Italy
sp.SoBigData.eu	URV-TA	Tarragona	Spain
sp.SoBigData.eu	BSC-TA	Barcelona	Spain
sp.SoBigData.eu	UPF-TA	Barcelona	Spain
se.SobigData.eu	KTH-TA	Stockholm	Sweden
nl.SoBigData.eu	UvA-TA	Amsterdam	Netherlands
at.SoBigData.eu	CEU-TA	Vienna	Austria

Table 1 Showing the 17 Infrastructures with their short names and locations

There are currently 7 Exploratories within SoBigData++ and these are listed below in Table 2, together with the infrastructures supporting them:

Exploratory	Infrastructure supporting Exploratory
Sustainable City of Citizens	CNR-TA, IMT-TA, UNIPI-TA
Demography, Economy and Finance 2.0	ETHZ-TA, SNS-TA, UNIPI-TA, UT-TA
Societal Debates and Misinformation	GATE-TA, ETHZ-TA, IMT-TA, CEU-TA, UT-TA
Migration Studies	PSE-TA, CNR-TA, UNIPI-TA
Sports Data Science	UNIPI-TA, CNR-TA
Social Impact and AI and explainable ML	UNIPI-TA, BSC-TA
Network Medicine	SAP-TA

Table 2 Showing the 7 Exploratories and supporting Infrastructures

The project also utilizes Thematic Clusters detailing the expertise within the project. The table 3 details these Thematic Clusters and their acronyms. The applicant will use these Thematic Clusters when planning their project proposal.

Abbreviation	Thematic Cluster
TSMM	Text and Social Media Mining
SNA	Social Network Analysis
HMA	Human Mobility Analytics
WA	Web Analytics
PET	Privacy Enhancing Technology
VA	Visual Analytics

Table 3 Showing the Thematic Clusters and their acronyms

Below, in Table 4, are the 17 Infrastructures together with the Thematic Clusters they have expertise in. Potential visitors will use this information when choosing the most relevant infrastructure for their proposed project.

Infrastructure	Short Name	Thematic Cluster
uk.SoBigData.eu	**GATE-TA**	TSMM, WA
it.SoBigData.eu	**UNIPI-TA**	TSMM, SNA, HMA, WA, PET
it.SoBigData.eu	**SNS-TA**	TSMM, SNA, HMA, WA, PET
it.SoBigData.eu	**SAP-TA**	TSMM, SNA, HMA, WA, PET
it.SoBigData.eu	**CNR-TA**	TSMM, SNA, HMA, WA, PET
it.SoBigData.eu	**IMT-TA**	TSMM, SNA, HMA, WA, PET
de.SoBigData.eu	**L3S-TA**	TSMM, WA
ee.SoBigData.eu	**UT-TA**	TSMM, SNA, WA
fi.SoBigData.eu	**AALTO-TA**	SNA, HMA
se.SobigData.eu	**KTH-TA**	SNA, HMA
ch.SoBigData.eu	**ETHZ-TA**	TSMM SNA WA

sp.SoBigData.eu	URV-TA	TSMM, SNA, HMA, PET
sp.SoBigData.eu	BSC-TA	TSMM, SNA, HMA, PET
sp.SoBigData.eu	UPF-TA	TSMM, SNA, HMA, PET
at.SoBigData.eu	CEU-TA	TSMM, SNA, HMA
nl.SoBigData.eu	UvA-TA	VA
fr.sobigdata.eu	PSE-TA	SNA, HMA

Table 4 Showing the Infrastructures with their short name and Thematic Clusters

2.3 Application Procedure

The Call has a link to the Application Form and the email address to send the completed proposal to. Each application is reviewed within 2 months (and usually within 4 weeks). The proposal is first assessed by the host infrastructure. After it has been accepted by the host, an external reviewer reviews the proposal and, if relevant, provides feedback. The proposal also goes through an Ethical Review process and any deficiencies within the proposal, or possible issues from an ethical perspective are raised and then addressed by the applicant before their project can be accepted. The Application Form is shown in Appendix B.

In more detail, the application form includes a short biography of the applicant and requests details of their qualifications, experience in the subject area and a personal statement. They must detail their proposed project including a description of the work planned, their expected outcomes and how their project will positively impact the SoBigData++ community. They must also describe how they intend to share their project and its findings more widely.

2.4 Selection Process

A small number of visits can be granted to researchers between SoBigData++ institutions, however the project primarily funds visitors from EU institutions that are not involved in the SoBigData++ research project. In addition, up to 20% of visitors may be accepted from non-EU institutions. The three Fund Managers will monitor these numbers. The three Fund Managers are EGI, UT and USFD.

The applications are judged on 4 basic criteria:

1. The quality of the applicant
2. The quality of the proposed research project
3. Originality of the project
4. Feasibility and Impact of the project on SoBigData++

The project uses a simple 1-5 ranking system for each of the criteria, with 5 being the highest. Each application goes through an external review procedure where a reviewer with no connection to the project

assesses the project's viability and potential against the same 4 criteria above. This is to ensure the application has been judged fairly and objectively from an impartial perspective.

2.5 Legal & Ethical Issues

approval process. If concerns are raised the application and ethical feedback is referred back to the applicant for any issues to be addressed. Revised applications are then reassessed by the Ethical Board before the application can be accepted.

The applicant must show in their application form that they have considered the ethical implications of the data they propose to collect and use for their project. This encourages personal responsibility and having self-awareness of the importance of ethical data handling. This, in turn, ingrains the project's ethos and expectation that becoming an ethical and responsible data scientist is paramount.

2.6 Applicants Summary

The number of applications during the first Call was low which was expected due to the pandemic. The Table 5 shows the applications received and their outcomes.

No.	Name	Gender	EU/Non-EU	Project Title	Host applied for	Fund Manager	Was the application successful?
1	Alex Bravo Serrano	M	EU	Adaptation of online abuse detection algorithms to Spanish	GATE-TA	USFD	Yes
2	Aditya Khanna	M	Non-EU	No details	GATE-TA	USFD	No – due to Covid-19
3	Prashant Khandewal	M	Non-EU	No details	GATE-TA	USFD	No – due to Covid-19
4	Naman Srivastava	M	Non-EU	No details	GATE-TA	USFD	No – due to Covid-19
5	Iknoor Singh	M	Non-EU	Disinformation Analysis	GATE-TA	USFD	Yes
6	Ophir Frieder	M	Non-EU	Conversational IR	CNR-TA	UT	Yes
7	Dino Carpentras	M	EU	Analyzing the attitude space of the Twittersphere	ETHZ-TA	UT	Yes
8	Silvia Gargova	F	EU	Disinformation in social media in Bulgaria	GATE-TA	USFD	Yes

9	Federico Cinus	M	EU	Stance collection and analysis of Twitter data on popular societal debates	KTH-TA	EGI	Yes
10	Koustav Rudra	M	Non-EU	Scalable Interpretable Document Ranking	L3S-TA	USFD	Under review
11	Dijana Tolić	F	EU	Bitcoin bubbles and socio-financial feedback: the digital traces of bubbles in Twitter	ETHZ-TA	UT	Under review
12	Sourangshu Bhattachrya	M	Non-EU	Improving Machine Learning and Trusted-AI systems using Explanations	L3S-TA	USFD	Under review

Table 5 The table below shows the applications received and their outcomes

During this 2-year period (which includes 18 months of Covid-19 restrictions) the project received 12 applications of which 3 could not progress due to Covid-19 restrictions. Of the 9 that progressed, 7 were male applicants and 2 were female. There were 5 EU applicants and 4 non-EU (India and America) applicants.

2.7 Participants Feedback

We have received positive feedback from the visitors who have undertaken access so far in regard to the level of support they received from their host institution, there were some less favorable comments in relation to difficulties encountered due to COVID 19 and more specifically travel restrictions/issues. One example can be found below:

'During this semester-long internship, several people have provided many forms of help and support. Firstly, I would like to thank Dr Kalina Bontcheva & Dr Diana G. Maynard for selecting me to do this project and for continuous guidance throughout the project. Further, I would also like to thank Dr Carolina Scarton and all the people at Natural Language Processing Group, The University of Sheffield who helped me develop the skills necessary to tackle this project.'

Below are the links to the blog posts we have received to date:

- http://sobigdata.eu/blog/tna-experience-adaptation-online-abuse-detection-algorithms-spanish
- http://sobigdata.eu/blog/disinformation-analysis-tna-experience

2.8 Costs so far

To date the total costs incurred for visits are €3,943.65, these costs have only been incurred by partner USFD. There are no further costs to report for this period at present.

3   Plans for the next 12 months

The plan for the next 12 months of the project is to continue monitoring the Covid-19 situation while providing administrative support to TA applicants and visitors. Host contacts will be encouraged to promote the TA visits – particularly to female researchers.

The call will remain an Open Call and will be promoted at regular intervals via the project website, Twitter and through the usual email channels.

Feedback will continue to be collected and analyzed to make any relevant improvements to the process and visitor experience.

4   Conclusions

Covid-19 has had a huge impact on the TA Work Package and this is reflected in the vastly reduced number of applications the project has received. It is accepted that as long as the pandemic persists, a return to the usual number of applications is unlikely.

The visits that have taken place have been successful and well received with the visitor gaining valuable experience and support for their projects and establishing meaningful academic connections and working relationships within the community. Some of the visitor's blogs can also be seen in the periodical SoBigData++ Magazine.

SoBigData++ will continue to be as flexible and as accommodating as possible to TA applicants and visitors and it is hoped that the next period of the project will receive an increased number of applications.

Appendix A. Call for SoBigData++ - Funded Transnational Access

Publish Call for Proposal: Tuesday 15 June 2021 8:30 CEST

Submission deadline: Sunday 31 July 2022 23:00 CEST

Application: 📄 Application form - SoBigData++ TNA 2021.doc

Applicant Notification: within 2 months of submission

Visits period: 1 September 2021 / 31 July 2022

The SoBigData++ project invites researchers and professionals to apply to participate in Short-Term Scientific Missions (STSMs) to carry forward their own big data projects. These opportunities are offered as part of SoBigData++'s Transnational Access (TNA) activities and applications can be submitted at any time.

You should submit your project application by the end of each calendar month to ta-admin@sobigdata.eu.

We welcome applications from individuals with a scientific interest, professionals, startups and innovators that may benefit from training in data science and social media analytics. In order to apply you have to fill the Project Application Form.

Applications will be reviewed within 2 months of submission, subject to COVID-19 restrictions. (Please note all applications received between 15 July 2021 and 31 August will not be processed until after 1 September 2021). Each application will be assessed not only on its scientific merit but will take into consideration the current Covid-19 regulations of both the visitor's departure country as well as the destination country. Please plan your visit to start at least 4 months from the end of the month when you submitted your application and expect some adjustments to your planned dates to take into account Covid-19 travel restrictions which are outside of our control.

Travel, international, national and local restrictions due to Covid-19 are beyond the control of either the project or the visitor and as such we would ask that you are flexible and understanding of unexpected and fast changing situations. All regulations that the welcoming infrastructure put in place are for your safety and the safety of others and must be adhered to. We will continue to monitor the situation constantly and will adapt and be flexible where we can.

Please be aware SoBigData++ can only fund the visit at the agreed rate for the agreed placement dates and will not cover any incurred expenses during periods of imposed quarantine or other travel or working restrictions placed on the visitor (in either the departure or arrival country) as a result of Covid-19. Please factor this information into your plans and be prepared for re-scheduling of your TNA visit, should sudden changes in local, national and international policies make this necessary.   Please note Covid-19 tests will not be paid for by SoBigData++.

Under no circumstances should you book or plan to start your visit without an email confirming your application has been successful, that an ethical approval has been granted and a contract has been signed.

Funding for a **short-term scientific mission** (3 weeks to 2 months) is available **up to 5000 euros** per participant (to cover the cost of daily subsistence, accommodation, and economy flights/train).

STSM bursaries are awarded on a competitive basis, according to the procedure described in the application pack and eligibility criteria below, and based upon the quality of the applicant, the scientific merit of the proposed project, and their personal statement.

Success rates from previous calls have been very high (65-75%). Applications from female scientists are particularly encouraged.

Visitors are welcome subject to the host country's and **host institution's Covid-19 regulations**. We will consider offering up to a 6-month postponement of an accepted application if travel restrictions are imposed due to Covid-19.

Pre-requisites for projects to carry out hosted research:

- Good understanding of social data and, ideally, track record of prior social data analysis projects
- Experience with using at least one of machine learning, natural language processing, and/or complex networks algorithms

Pre-requisites for projects to integrate new tools/datasets/services:

- An already existing open-source tool for social media mining to be integrated

  OR

- An already created openly licensed dataset of relevance to SoBigData++, that can be integrated within the infrastructure

Applications will be hosted at one of the following centers (infrastructures).

Please, note that the user group leader and the majority of the users must work in a country other than the country(ies) where the installation is located. This rule does not apply:

- If the applicant is from an international organization, the Joint Research Centre (JRC), an ERIC or similar legal entities;
- in case of remote access to a set of installations located in different countries offering the same type of service

The goal is to provide researchers and professionals with access to big data computing platforms, big social data resources, and cutting-edge computational methods. STSM visitors will be able to:
- Interact with the local experts
- Discuss research questions
- Run experiments on non-public big social datasets and algorithms
- Present results at workshops/seminars

The STSM visits will enable multi-disciplinary social mining experiments with the SoBigData++ Research Infrastructure assets: big data sets, analytical tools, services and skills.

Description of the TNA offered
The SoBigData++ RI manages vertical, thematic environments, called exploratories, on top of the SoBigData++ infrastructures, for performing cross-disciplinary social mining research. The Transnational Activities offered in this call will be for Short-Term Scientific Missions (STSM), between 3 weeks and 2 months.

Under this call, there will be two kinds of proposals funded: STSM research proposals and STSM tool/data integration proposals. Each kind is described in more details next.

Each STSM research proposal needs to be aligned with **one of the following exploratories** and also specify which of the centres listed above they wish to access:

- **Sustainable Cities for Citizens**
    - Infrastructures involved:
        - CNR, IMT, or UNIPI in **it.SoBigData.eu**
- **Demography, Economy and Finance 2.0**
    - Infrastructures involved:
        - ETH Zürich - **ch.SoBigData.eu**
        - SNS, CNR, or UNIPI in **it.SoBigData.eu**
        - UT - **ee.SoBigData.eu**
- **Societal Debates and Misinformation**
    - Infrastructures involved:
        - USFD - **uk.SoBigData.eu**
        - ETH Zürich - **ch.SoBigData.eu**
        - IMT - **it.SoBigData.eu**
        - CEU - **at.SoBigData.eu**
        - UT - **ee.SoBigData.eu**
- **Migration Studies**
    - Infrastructures involved:
        - PSE - **fr.SoBigData.eu**
        - CNR or UNIPI in **it.SoBigData.eu**
- **Sports Data Science**
    - Infrastructures involved:
        - UNIPI, or CNR in **it.SoBigData.eu**
- **Social Impact of AI and explainable ML**
    - Infrastructures involved:
        - UNIPI - **it.SoBigData.eu**
        - BSC - **sp.SoBigData.eu**
- **Network Medicine**
    - Infrastructures involved:
        - UNIROMA - **it.SoBigData.eu**
- **Other Infrastrucure available**
    - LUH - **de.SoBigData.eu**
    - AALTO - **fi.SoBigData.eu**
    - URV or UPF in **sp.SoBigData.eu**
    - KTH - **se.SobigData.eu**
    - UvA - **nl.SoBigData.eu**

Each STSM integration proposal can focus on the integration of already existing open-source tools for social media mining or social datasets within the SoBigData++ infrastructure as a whole or via integration within one of the national facilities listed above. Applications focused on interoperability and integration with other European Research Infrastructures are also strongly encouraged. Applicants for integration proposals are strongly encouraged to contact informally their target host institution, to discuss and ensure technical feasibility of the proposal prior to submitting the application.

Applications are invited at the following centres (infrastructures):

- **uk.SoBigData.eu, USFD, United Kingdom**
    - **Thematic Cluster:** Text and Social Media Mining (TSMM), Web Analytics (WA)
    - **Contact:** Kalina Bontcheva k.bontcheva@sheffield.ac.uk
    - **Location:** *Sheffield, United Kingdom*
- **it.SoBigData.eu, CNR, UNIPI, IMT, SNS, UNIROMA, Italy:**

- o **Thematic Cluster:** Text and Social Media Mining (TSMM), Social Network Analysis (SNA), Human Mobility Analytics (HMA), Web Analytics (WA), Privacy Enhancing Technology (PET)
  - o **Contact CNR:** Roberto Trasarti roberto.trasarti@isti.cnr.it
  - o **Contact SNS:** Fabrizio Lillo fabrizio.lillo@sns.it
  - o **Contact UNIPI**: Anna Monreale anna.monreale@unipi.it
  - o **Contact IMT:** Angelo Facchini angelo.facchini@imtlucca.it
  - o **Contact UNIROMA:** Aris Anagnostopoulos aris@diag.uniroma1.it
  - o **Locations:** *Pisa, Lucca, Rome, Italy*
- **ee.SoBigData.eu, UT, Estonia:**
  - o **Thematic Cluster:** Text and Social Media Mining (TSMM), Social Network Analysis (SNA), Web Analytics (WA)
  - o **Contact:** Rajesh Sharma rajesh.sharma@ut.ee
  - o **Location:** *Tartu, Estonia*
- **de.SoBigData.eu, LUH, Germany:**
  - o **Thematic Cluster:** Text and Social Media Mining (TSMM), Web Analytics (WA)
  - o **Contact:** Avishek Anand anand@l3s.de
  - o **Location:** *Hannover, Germany*
- **fi.SoBigData.eu, AALTO, Finland:**
  - o **Thematic Cluster:** Social Network Analysis (SNA), Human Mobility Analytics (HMA)
  - o **Contact:** Kimmo Kasky kimmo.kaski@aalto.fi
  - o **Location:** *Aalto, Finland*
- **ch.SoBigData.eu, ETH Zurich, Switzerland:**
  - o **Thematic Cluster:** Text and Social Media Mining (TSMM), Social Network Analysis (SNA), Web Analytics (WA)
  - o **Contact:** Nino Antulov-Fantulin nino.antulov@gess.ethz.ch
  - o **Location:** *Zurich, Switzerland*
- **fr.sobigdata.eu, PSE Paris, France:**
  - o **Thematic Cluster:** Social Network Analysis (SNA), Human Mobility Analytics (HMA)
  - o **Contact:** Hillel Rapoport hillel.rapoport@univ-paris1.fr
  - o **Location:** Paris, France
- **sp.SoBigData.eu, URV, BSC, UPF, Spain:**
  - o **Thematic Cluster:** Text and Social Media Mining (TSMM), Social Network Analysis (SNA), Human Mobility Analytics (HMA), Privacy Enhancing Technology (PET)
  - o **Contact URV:** Jesus Manjon jesus.manjon@urv.cat
  - o **Contact BSC:** Dario Garcia dario.garcia@bsc.es
  - o **Contact UPF:** Carlos Castillo carlos.castillo@upf.edu
  - o **Location:** Barcelona, Tarragona, Spain
- **se.SobigData.eu, KTH, Sweden:**
  - o **Thematic Cluster:** Social Network Analysis (SNA), Human Mobility Analytics (HMA)
  - o **Contact:** Aristides Gionis argioni@kth.se
  - o **Location:** Stockholm, Sweden
- **nl.SoBigData.eu, UvA, Amsterdam:**
  - o **Thematic Cluster:** Visual Analytics (VA)
  - o **Contact:** Richard Rogers r.a.rogers@uva.nl
  - o **Location:** Amsterdam, Netherlands
- **at.SoBigData.eu, CEU, Austria:**
  - o **Thematic Cluster:** Text and Social Media Mining (TSMM), Social Network Analysis (SNA), Human Mobility Analytics (HMA)
  - o **Contact:** Janos Kertesz kerteszj@ceu.edu
  - o **Location**: Vienna, Austria

User selection panel: TBA

Appendix B.   Application form

SoBigData++ Transnational Access Application form

Application form to be submitted to ta-admin@sobigdata.eu , accompanied by a 2 page CV as an appendix. Please make it all 1 PDF file.

Family Name		Nationality	
First Name		Birth year	
Gender (*tick the appropriated item*):　　　❑ female　　　❑ male			
Home Institution:			
Legal Status of Home Institution Code[1]		Home Institution Country Code[2]:	
Function / Job / Title:		Position Code[3]:	
Mailing Address:			
Phone (office):		Phone (mobile):	
Fax:		E-mail:	

---

[1] **UNI**=University, **RES**=Public Research Organisation, **SME**=Small or Medium Enterprise, **PRV**=Other and/or profit or not profit Private Organisation, **OTH**= Other Organisation

[2] **AL**=Albania, **AT**=Austria, **BE**=Belgium, **BG**=Bulgaria, **CH**=Switzerland, **CY**=Cyprus, **CZ**=Czech Republic, **DK**=Denmark, **EE**=Estonia, **FI**=Finland, **FR**=France, **DE**=Germany, **GR**=Greece, **HR**=Croatia, **HU**=Hungary, **IS**=Iceland, **IE**=Ireland, **IL**=Israel, **IT**=Italy, **LV**=Latvia, **LI**=Liechtenstein, **LT**=Lithuania, **LU**=Luxembourg, **MT**=Malta, **MK**=Macedonia, **ME**=Montenegro **NL**=Netherlands, **NO**=Norway, **PL**=Poland, **PT**=Portugal, **RO**=Romania, **SK**=Slovakia, **SI**=Slovenia, **ES**=Spain, **SE**=Sweden, **SR**=Serbia, **TR**=Turkey, **GB**=United Kingdom

[3] **UND**=Undergraduate, **PGR**=Post graduate (student with a first University degree or equivalent), **PDOC**=Post-doc researcher, **TEC**=Technician, **EXP**=Experienced researcher (professional researcher).

Short Biography *(max 500 words)*
Please provide details of your academic qualifications and experience in your research area
Note: Please supplement this with a 2-page CV, as an appendix to your application.

Recent Publications (please list 5)

Web page and open code:

Please provide pointers to your web page and any open-source code you have on github or elsewhere.

SoBigData++ Infrastructure access being applied for:
Name of the host institution:
Start date: Location:
Length of Visit Requested (in weeks):
Specific requests for preparation or training *(optional)*

NOTE: Project title, Project objectives and *Description of the planned work, and References* are intended to provide information on the project that the applicants are proposing to work on during the trans-national access at the above-named host institution.

Project title:
Project objectives (max 500 words)

Description of the planned work (max 2 pages, figures and tables included)

Personal statement: expected achievements and impact (max 500 words)

Outline the reasons why you are interested in the TNA, what you wish to gain from it, and how you expect it to contribute to your own research in the future.

Dissemination plans (max 500 words)

Describe how you plan to bring lessons learned in the TNA to the attention of members of your project team, academic department or organization, or the research community at large. This could be in the form an informal talk, a brief report, a blog post, or a publication.

Ethical self-reflection

Please provide some reflection on the ethical integrity and implications of your project. Describe the dataset(s) you are planning to use in your project, with particular emphasis on the ethical implication. Does it contain personal data? Is it public? Do you think that real identities of data subjects can be revealed by the access of this dataset or sensitive information can be inferred by the publication of your results?

References

Main scientific field of the project

Give the particular field of research in your project, e.g., "Social network analysis", "Natural Language Processing", or "Machine Learning".

1. **Request for SoBigData++ sponsorship?**   ❏ Yes   ❏ No

2. **Any other EU project related to this proposal?** ❏ Yes   ❏ No

In case of positive response, specify:

_____
_____
_____
_____
_____

3. **How did you hear about SoBigData++?**
    ❑ personal contact   ❑ web page   ❑ conference   ❑ twitter
    ❑ article   ❑ other (*specify*)

I confirm that, I will send a Summary Report no later than 2 months after the end of the access visit, and that the results of the proposed project will be published specifying that the project leading to the publication has received funding from the European Union's Horizon 2020 programme: "*The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871042 (SoBigData++)*"

Your application will be sent to a partner for review as well as an External Reviewer and the SBD++ Ethics Board. By signing this application form you are agreeing to it being shared (together with your CV and any supporting documents) in confidence with all relevant parties.

Date:

_____

Signature:

_____

Appendix C. Proposal evaluation form

Proposal Evaluation Form **SoBigData++ Transnational Access: Proposal Evaluation form**

Completed evaluation forms to be sent to ta-admin@sobigdata.eu AND joanna.wright@sheffield.ac.uk

Reviewer Name		Nationality	
Gender (*tick the appropriated item*): □ female □ male □ prefer not to say			
Affiliation:			
Home Institution Country Code		Job Title:	
E-mail:			

Project title:
Applicant Name:
Applicant's Institution:
Quality of the applicant: Please provide a brief summary on the quality of the applicant, based on their CV.   **Applicant score (delete as appropriate): 1 – unsatisfactory; 2 – adequate; 3 – good; 4 – very good; 5 – excellent**
Quality of the proposed research project: *Please comment.*   **Quality score (delete as appropriate): 1 – unsatisfactory; 2 – adequate; 3 – good; 4 – very good; 5 – excellent**

Originality of the proposed research project: *Please comment.* **Originality score (delete as appropriate): 1 – unsatisfactory; 2 – adequate; 3 – good; 4 – very good; 5 – excellent**	
Feasibility and Impact on SoBigData++: *Discuss whether the proposal is feasible given the short duration of the visit. Also comment on how the proposed project would help enhance SoBigData++'s impact through joint publications, outreach to new countries/research groups, open source software, or any other planned contributions that can **realistically** be expected to arise.* **Feasibility Score (delete as appropriate): 1 – unsatisfactory; 2 – adequate; 3 – good; 4 – very good; 5 – excellent**	
Overall Score	**(Delete as appropriate) 1 – unsatisfactory; 2 – adequate; 3 – good; 4 – very good; 5 – excellent**

In the event of equal scoring, weighting will be given to users who have not previously used the SoBigData++ infrastructure, to young researchers and to those working in countries where no research infrastructure is available. The selection panel shall also pay regard to issues of gender equality.

Appendix D. Agreement Template

Department
Of
Computer
Science.

17 November 2021

Head of Department
Professor Guy Brown

Joanna Wright
Regent Court
211 Portobello
Sheffield
S1 4DP

Tel: +44 (0) 114 222 1942
Email: joanna.wright@sheffield.ac.uk

Dear **USER NAME**

<div align="center">

TITLE

</div>

The SoBigData++ Consortium are pleased to inform you that your project **TITLE** (annex 1) has been accepted by the selection panel.

The duration of your visit to **LOCATION** is for **DURATION**, starting **START DATE** and ending **END DATE**.

In line with University financial regulations please refer to annex 2 for information with regards to expenditure and allowable costs. Please ensure that these guidelines are adhered to so we can ensure that your funding claim is eligible in line with EC and University legislation. Please note, the **last date for receiving expenses claims is (DAY/DATE – suggest 6 weeks after visit)**. After this date expense claims will NOT be processed. Please ensure you have returned your blog post and end of visit results report so as not to delay the process.

You may be required to receive Ethical Approval from the Host Institution. If so full details will be provided to you before your arrival.

In addition you may be required to sign a non-disclosure agreement which will be provided and signed by you on the day of arrival for your visit.

I should be grateful if you would confirm that you are happy to accept this offer on the above terms by signing and returning a copy of this letter to me at the above address.

Accepted on behalf of

SoBigData++ Consortium   Accepted by

Signature:

                                                             Signature:

Name: **Joanna Wright**   Name:

Date: Date:

Appendix E. Expenses Guide

Trans-National Access visits 2020-2024

Travel Expense Guide

Researchers should be advised that travel costs incurred while undertaking a Trans-National Activity must be made by the most economical method available.

· **Air Fares**

   o   Should be Economy Class and be by the cheapest route where possible (whilst considering reasonable travel durations).

· **Rail Fares**

   o Standard Class Rail Fares or discounted rail should be used where possible and bookings should be made at least 7 days in advance of travel. First class travel will not be reimbursed.

· **Public transport**

   o Should be used where possible. Car mileage will normally be reimbursed to the equivalent value of a standard class rail fare.

· **Subsistence**

   o Will only be reimbursed for participants who require overnight accommodation (i.e. not for the host organisation). We will provide a daily rate up to a total of €45. Please provide original receipts where possible.

· **Accommodation**

   o   Should be a reasonable price for the location of the meeting and ideally have self-catering facilities.

   o   Accommodation (ideally in a short stay apartment) can be arranged by the University/Infrastructure on your behalf. Should you wish to arrange your own accommodation, please let us know as soon as possible.

●   Please note, the Project cannot approve accommodation arranged through the company AirBnB. If you choose to book accommodation through this company you do so at your own risk and must satisfy yourself that the accommodation is appropriate for your requirements. SoBigData++ will not be liable for any issues arising from accommodation booked through AirBnB.

Partner	Standard self-catering apartment accommodation cost per night	Standard hotel accommodation cost per night
USFD	€70	€95
UNIPI	€60	€85
SNS	€60	€85
UNIROMA1	€70	€100

CNR	€60	€85
IMT	€60	€100
LUH	€60	€100
UT	€50	€80
AALTO	€100	€150
KTH	€100	€120
ETHZ	€150-200	€150-200
URV	€40	€100
BSC	€60	€110
UPF	€60	€95
CEU	€60	€100
UvA	€120	€130
PSE	TBA	€70

Please note that SoBigData++ cannot reimburse for accommodation from the organisation AirBnB.

If you have any queries regarding costs of travel then please contact com-resadmin@sheffield.ac.uk, ta-admin@sobigdata.eu and your host contact.

The maximum you can be refunded for your whole stay is €5,000. If you incur more than this then the funding provided by the TNA SoBigData++ Consortium will be capped.
Trans-National Access (TNA) is granted by 17 national infrastructures:

Infrastructure	Country	Lead Name	Lead Email
USFD (Gate)	UK	Kalina Bontcheva	k.bontcheva@sheffield.ac.uk
UNIPI	Italy	Anna Monreale Riccardo Guidotti	anna.monreale@unipi.it riccardo.guidotti@isti.cnr.it
SNS	Italy	Fabrzio Lillo Michele Benzi	fabrizio.lillo@sns.it michele.benzi@sns.it
UNIROMA 1	Italy	Aris Anagnostopoulos Fazzone Adriano	aris@diag.uniroma1.it fazzone@diag.uniroma1.it
IMT	Italy	Angelo Facchini	angelo.facchini@imtlucca.it
LUH	Germany	Avishek Anand	anand@l3s.de

UT	Estonia	Rajesh Sharma Marlon Dumas	rajesh.sharma@ut.ee marlon.dumas@ut.ee
KTH	Sweden	Aristides Gionis	argioni@kth.se
Aalto	Finland	Kimmo Kaski Kunal Bhattacharya Noora Suominenderios	kimmo.kaski@aalto.fi kunal.bhattacharya@aalto.fi noora.suominenderios@aalto.fi
ETHZ (Nervousnet)	Switzerland	Vaiva Vasiliauskaite Nino Antulov-Fantulin	vaiva.vasiliauskaite@gess.ethz.ch nino.antulov@gess.ethz.ch
URV	Spain	Jesus Manjon Paniagua Josep Domingo	jesus.manjon@urv.cat josep.domingo@urv.cat
BSC	Spain	Dario Garcia	dario.garcia@bsc.es
UPF	Spain	Carlos Castillo	carlos.castillo@upf.edu
CEU	Vienna	Olga Peredi Márton Karsai Janos Kertesz	peredio@ceu.edu karsaim@ceu.edu kerteszj@ceu.edu
UvA	Netherlands	Richard Rogers	r.a.rogers@uva.nl
PSE	France	Hillel Rapoport	hillel.rapoport@psemail.eu