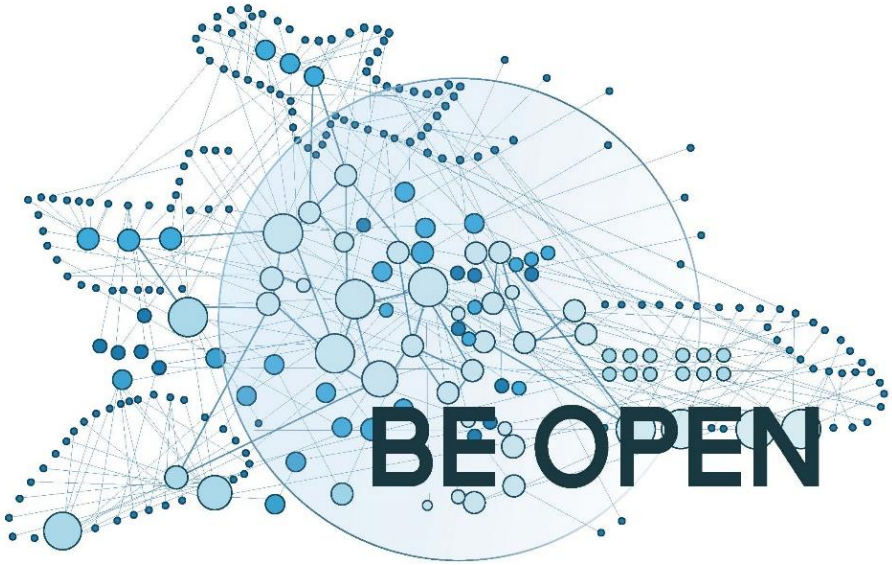
Pending EC/INEA approval

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824323



European forum and oBsErvatory for OPEN science in transport

Project Acronym:	**BE OPEN**
Project Title:	**European forum and oBsErvatory for OPEN science in transport**
Project Number:	**824323**
Topic:	**MG-4-2-2018 – Building Open Science platforms in transport research**
Type of Action:	**Coordination and support action (CSA)**

3.5 TOPOS Sustainability Analysis

Final version

Deliverable Title:	TOPOS Sustainability Analysis
Work Package:	WP 3
Due Date:	June 30th 2021
Submission Date:	June 29th 2021
Start Date of Project:	Jan 2019
Duration of Project:	30 months
Organisation Responsible of Deliverable:	TØI
Version:	Final version
Status:	Final version
Author name(s):	Anja Fleten Nielsen (TØI), Michela Fioretto (FIT), Pavlos Spandis (CERTH), Ove Langeland (TØI).
Reviewer(s):	Kristel Palts (DLR), Jakob Michelmann (VDI)
Nature:	☒ R – Report ☐ P – Prototype ☐ D – Demonstrator ☐ O - Other
Dissemination level:	☒ PU - Public ☐ CO - Confidential, only for members of the consortium (including the Commission) ☐ RE - Restricted to a group specified by the consortium (including the Commission Services)

Document history			
Version	**Date**	**Modified by (author/partner)**	**Comments**
1	9/06/2021	Anja Fleten Nielsen, Michela Fioretto, Pavlos Spandis, Ove Langeland	
2	17/06/2021	Anja Fleten Nielsen, Michela Fioretto, Ove Langeland	
3	24/06/2021	Kristel Palts	Jakob Michelmann had no comments
4	28/06/2021	Anja Fleten Nielsen	

Contents

List of Figures

List of Tables

Abbreviations and Terminology

EOSC	European Open Science Cloud
ESPON	European Observation Network for Territorial Development and Cohesion
GDPR	General data protection regulation
HERMES	Establishing a CompreHEnsive transport Research information Management and Exchange System
HUNT	The HUNT Study - a longitudinal population health study in Norway
IP	Intellectual Property
RDA	Research Data Alliance
TOPOS	Transport fOrum/Observatory for Promoting Open Science
TRIMIS	Transport Research and Innovation Monitoring and Information System
TØI	Institute of Transport Economics

Executive summary

In this document we have identified how to solve specific barriers to make TOPOS sustainable in terms of **1) actual use** and **2) financial funding**. To do this, we have looked on six different types of barriers related to innovation research: resources, behavioural, organizational, information, technology, and legal barriers. The last three were the scope of the GA, but in order to make TOPOS sustainable we also interviewed the organizations about the first three, as these barriers were identified in former WPs as well as in research literature.

The main findings are presented in the table below:

Barrier	Sustainability measure
Resources	There are mainly four different financial models TOPOS possibly could undertake: • Public funding (international, national or regional) • User-payment **1)** Paying for data **2)** Paying for infrastructure • Stakeholders: Annual Organisational membership fees The advantage of international public funding (EU) is that the commission is regarded as a trustworthy organization with a heavy influence on a broad geographical level, possibly making the users of TOPOS more interested in the platform. If wanting to increase Open Science through "obligation", this is more likely to manage if TOPOS is funded by the European Commission.
Organizational	To make TOPOS sustainable an Open Science culture of sharing is also needed. If TOPOS can offer a pool of data available, sharing might also be easier for others, as there is already available data for them to use as well. In order to do this, TOPOS could focus on getting data from public stakeholders first, as they are less sensitive in terms of competitive advantage when sharing data, at least at a lower geographical level. Also, demonstrating that participating in Open Science at an early stage could potentially give benefits in terms of competitive advantage. Some organisations benefit from creating open data/infrastructure in terms of (1) becoming famous for their data/systems etc. which (2) in turns generates more projects and project invitations.
Information	Funding for a quality control organ and continuous dissemination is needed to make TOPOS sustainable. Possibly corporate further with TRIMIS.
Legal	Take advantage of the increased focus on GDPR, offering a two-way log in system for parts of the cloud, making it extremely secure. Making this available to use for sensitive data, if user is paying for storage.
Technological	Semantic Portal to improve the quality of communication between the information provider and the user and supporting self-sustainability

1 Introduction

1.1 Purpose of the document

Interviewing similar initiatives to TOPOS, this document identifies what data/information, technical, legal, organizational, behaviour and financial barriers that needs to be addressed and it is identifying mitigation strategies in order to make TOPOS sustainable.

The initiatives of HUNT, ESPON, TRIMIS and Safety Cube are interviewed by TØI. RDA and Science and Technology Digital Library are interviewed by FIT Consulting, and lastly HERMES is interviewed by CERTH.

1.2 Methodology

First a summary of the barriers presented in former tasks is summarised and categorized based on innovation literature in a table (see table 1) to address what is needed to be solved to make TOPOS sustainable. Based on the summary of already identified barriers we have conducted interviews with stakeholders that gave us useful information on how to solve the specific barriers and make TOPOS sustainable in terms of **1) active use** and **2) financial funding**.

The questions in the interview guide as well as the cases that are chosen for the interviews (organizations, networks or databases) are based upon the identified barriers (see table 1).

After interviews were conducted the information gathered was sorted by type of barrier (see chapter 2), before eventually assessing the essential information that is relevant for TOPOS.

2 Barriers and sustainability

In the former WPs (D 2.4, D 2.1, D 2.2 D 1.2, D 5.1) a lot of barriers related to Open Science were identified. We have made a matrix of the different barriers and sorted them by 'type of barrier' (see figure 1). We here relate to barriers that are widely discussed in the innovation literature, particularly to sustainable innovation and to open innovation (Gohoungodji et al 2020, Gupta et al 2020, de Jesus & Medonca 2018) since they to a large extent resembles barriers to open science. Main barriers of innovation can be divided in the following six categories (figure 1):

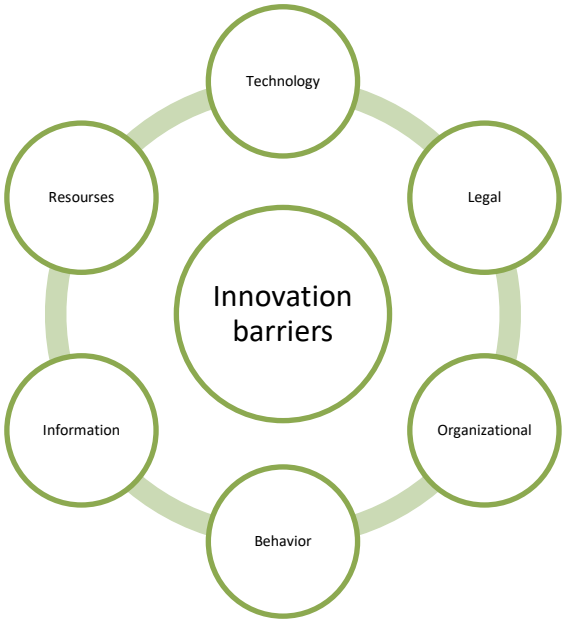


Figure 1 Barriers for innovation

1. **Resources** include lack of human capital/competence, lack of physical resources, time constraints in addition to lack of financial resources
2. **Behaviour** which include attitudes of employees, management, behaviour, habits and routines. Including corporate social responsibility.
3. **Information** which include the ability to comprehend and understand data, absorptive capacity (the organisation's ability to identify, assimilate, transform and use external knowledge/information), lacking knowledge. On a lower level, technical quality of data can be included in this.
4. **Technology** which include lack of technical capacity
5. **Organizational** which include culture, management culture, lacking support systems and commitment.
6. **Legal** which include differences in legal environment, in terms of laws and regulations, between countries /regions

In the BE OPEN Grant Agreement, information, technology and legal (laws and regulations) barriers are mentioned specifically, and in the former WPs resources, behaviour and organizational barriers are also identified as problematic areas for making open science and TOPOS sustainable. We can

classify the barriers mentioned as **necessary conditions** for making TOPOS sustainable. However, they are not **sufficient conditions** for sustainability. Making TOPOS sustainable we need to make sure that it is both **financially sustainable** and sustainable in terms of **used as intended**, i.e.

1. Institutes and researchers upload their projects and data to TOPOS
2. Researchers use TOPOS to get information about projects in their field of interest
3. People (public) reuse data from TOPOS
4. TOPOS is updated to sustain high quality

For further and more detailed information about measuring user participation in Open Science, please see BE OPEN D 5.2

To ensure active usage of TOPOS this can be done in two ways:

- Obligation: All national and EU funded projects has to make their project related material available for EOSC (has to be initiated by EOSC and agreed upon by national funding agencies).
- System of independent choice: TOPOS has to be viewed as a competitive advantage rather than time consuming and disadvantageous.

Table 1 Barriers identified in former tasks for Open Science and TOPOS, sorted by type of barrier and related to problem that needs to be solved by TOPOS

Type of barrier	Barriers and Challenges from former tasks	Problem to solve
Information	Common vocabulary – gap between industry and academia	Linguistic understanding, lack of common terms
Information	Individual researchers don't have knowledge on what data is interesting for others to use	Low awareness may reduce sharing of relevant data.
Information	Low knowledge on Open Science (EOSC, TOPOS) for both researchers in transport industry, and companies outside research	Making it difficult to reach out to users.
Information	Awareness in organisations, attitudes from management, lack of corporate responsibility may affect sharing culture/openness	How can we inform institutes on open science TOPOS? (Marketing)

Information (data quality)	Quality of data = LOW	How do others work with quality of their service? Develop standards for data quality
Behavior	Individual researchers think 'they own their data'	How can we get individual researchers to share their data?
Organizational	Openness vs. local sustainability – restrictive sharing due to risk of *others motives* on using the data	How to increase trust between researcher communities? Develop a sharing culture.
Organizational	Competitiveness – sharing data may give disadvantage for project applications	Getting the institutes interested in sharing data. How can TOPOS increase collaboration and generate competitive advantage instead of disadvantage?
Organizational	Large generation of data outside research (more difficult to get other stakeholders like PTA to share relevant data)	How can we get other stakeholders to share data with TOPOS?
Technology	Large generation of data outside research and variety of stakeholders makes the data fragmented and on different formats	One framework with quality demands. Standardization of data.
Resources + technology	Lack of skilled experts	Skilled experts with miscellaneous experience in the field of analytics together with the understanding of the best use of data sets should be recruited to add value to the research data.
Resources	Huge amount of data needs to be stored for a long time, having reliable archive systems	Requires agreements between countries and research organizations on how to store data safely

Resources	Cost and time for production of sharable data	Develop systems of cost sharing between countries and research organizations
Resources	Funding of TOPOS after June 2021	Agreements on funding schemes
Laws and regulations	Security issues – restrictive sharing due to potential societal risk (e.g. information on number on passengers on different busses in real-time)	Common EU regulations
Laws and regulations	Sharing without discriminations. EU directive.	Common EU regulations
Laws and regulations	Institutes don't want to share data due to GDPR (high fee for making mistakes)	How can we make sure that GDPR issues are avoided?

The identified barriers are on several different levels:

- Individual (i.e. individual researchers thinking 'they own their data')
- Organizational (i.e. institutes not wanting to share data due to competitiveness)
- Industry/sector (i.e. large generation of data outside research)
- Society – both national or international (i.e. security issues related to societal risk or international laws of sharing)

Several of the barriers identified will possibly reduce sharing and use of TOPOS. To avoid this, we have to look at what other similar entities have done to solve the different types of barriers. Below we are presenting the chosen cases and why we found them to be interesting for TOPOS:

1. HUNT research centre in Norway carries out one of the largest health surveys in the world (Health Survey of Nord-Trøndelag). They have approximately health information on 120 000 people, including blood samples on 80 000 people. All these data are available for researchers, but you must apply to use it. As mentioned in task 2.4, the medical fields have more sensitive data than transport research, and a longer tradition on **solving GDPR and ethical issues** and could therefore give valuable input to the project. Also, the **funding** of the data collection is very interesting for TOPOS, as it potentially could give input to possible funding solutions.
2. Safetycube is a former EU funded transport project, with a database totally open to the public. Because this is a former EU funded projects, it is very interesting to look at the **funding scheme** that is in work today and how many people that **use the service regularly** to get input of how TOPOS could solve some of the similar problems in terms of sustainability.
3. ESPON (European Spatial Planning Observation Network) is an applied research programme aimed at supporting the formulation of territorial development policies in Europe. ESPON activities cover all EU Member States, plus Iceland, Liechtenstein, Norway and Switzerland, and involve more than 130 bodies across the continent. They conduct different kinds of studies (thematic, policy impact, cross-theme, scientific networking, capacity-building) and produce territorial data in the form of statistics, analyses and maps. One of the major challenges of

ESPON therefore involves making this material **accessible and understandable** to local policy-makers. At EU level, the results of ESPON research efforts provide a source of comparable information that can be used to improve the Union's competitiveness and its sustainable development. it has to **solve many similar tasks to those of TOPOS**. How ESPON manage and operate its **informational and data-related** activities should therefore be of relevance for TOPOS.

4. RDA (Research Data Alliance) has an **inclusive approach covering all data lifecycle stages**, engaging data producers, users and stewards, addressing data exchange, processing, and storage. It has succeeded in creating the **neutral social platform** where international research data experts meet to exchange views and to agree on topics including social hurdles on **data sharing**, **education and training challenges**, **data management plans** and **certification of data repositories**, disciplinary and interdisciplinary **interoperability**, as well as technological aspects. All these aspects of RDA resulted to be of added value for TOPOS. In particular, the way RDA is economically sustainable through "customised" fees regarding the different type of audience/users could be a valid example for TOPOS since also in our case there are different stakeholders interested in accessing the OS information available. Moreover, both platforms aim at sharing information by fostering a cross-contamination among diverse areas of competence.
5. HERMES was an FP7 EU project focussed on the issues relating to sustainable development, especially in Transportation Research & Development (R&D) and the need for international collaboration and exchange of information and ideas in transport research. To facilitate this, HERMES set out to create a **gateway where transport researchers could share, access and exchange** their research information to an International/Global audience and thereby facilitate international and long-term collaboration. Because this initiative is very similar to what TOPOS is trying to do, HERMES is a highly relevant case.
6. S&TDL – Science and Technology Digital Library has been created by CNR (Italian National Research Council) with the simple but ambitious objective to build a **Digital Library** making science and technology available to everyone, promoting its most widespread use. STDL aims at going beyond the large number of digital libraries rapidly growing mostly at international level offering mainly digitised cultural contents to their users. The S&TDL willingness was to make available at the same time:
   - **research products**
   - **datasets**
   - **data and information on research activities (**projects, R&D actors, their expertise and skills in the respective areas of interest)
   - **digitised** cultural and historical heritage **contents**

   Thanks to Semantic Web advanced tools and techniques, S&TDL would like to provide a constantly expanding **network of relations** and connections among the main parts of the system. These characteristics could be of inspiration for TOPOS since the focus is on making research products available and usable for a broader audience and the promotors are in both cases coming from the research sphere.
7. TRIMIS started in 2017 – at a request from the European Commission as a top-down initiative, however the idea came about as a successor of a previous EU funded project TRIP to improve sharing in transport research. Since TRIMIS is one of few successful EU funded projects that

has survived after the initial funding – and a **possible collaborator or 'competitor' of TOPOS**, we wanted do see if there was anything to learn from this initiative.

3 Interview cases

The organizations interviewed differ in terms of openness. While some have no difference between users, and it's totally open to everyone, others are only open to some users (i.e. you have to be part of a research organization) and also follow a set of rules.

3.1 Safetycube

SafetyCube DSS is the European Road Safety Decision Support System, which has been produced within the European research project SafetyCube, funded within the Horizon 2020 Programme of the European Commission, aiming to support evidence-based policy making. The SafetyCube Decision Support System provides detailed interactive information on a large list of road accident risk factors and related road safety countermeasures. An open access calculator for policy makers is also available. The calculator for Economic Efficiency Evaluation (E3) of road safety counter measures allows to combine information about the effectiveness of a measure (i.e. the percentage of crashes or casualties prevented) with the costs of this measure. The calculator also integrates updated information of crash-costs in the European countries, allowing to express all costs and benefits of a measure in monetary values and conducting cost benefit analysis.

The database is open for everyone to read and use interactively (calculator). It was intended for local decision makers and stakeholders in municipalities and cities so that they could implement the best strategies to reduce traffic accidents, as well as people working with traffic safety in other ways – hence the user group is mixed.

Project webpage with link to SafetyCube DSS: https://www.safetycube-project.eu/

3.2 HUNT

HUNT started with two general practitioners in Verdal that wanted to do a research project on health service in the 80s. They contacted FHI (Norwegian Health Institute) that did tuberculosis screening at the time and they started doing several clinical examinations (diabetes, heart and life quality factors) on the already existing tuberculosis (TB) screening program. Everyone was invited to the TB screening, and the attendance was 90 %. Since everyone that participated was based on birth number the idea to do follow up research came about, and several other examinations and surveys was entered in the later HUNT studies. Since HUNT 1 in 1984-86, the survey has been conducted every 10th year and 230 000 people have participated in HUNT since then. HUNT is now a research center with about 30 employees. They are part of NTNU (University), Public Health medicine and consist of HUNT BIOBANK, HUNT Databank and HUNT cloud that offers services for researchers: storing, analyzing and data handling.

The data is open to everyone that are (1) employed at a research institution, (2) has a project description and (3) has approval of REK (regional ethical committee) and (4) follows other laws and regulations. The whole model is built upon HUNT collecting data that is open to 'everyone'.

They have also built an infrastructure that is open for other researchers to use and store their own data in the bank (i.e. it's used by the hospitals to register number of strokes). This also ensures that other data can be stored somewhere for the future when it can be used by others. They have so much data that they need for a safe, available storage. In the BIOBANK they have biological material from 60.000 people - there is an extreme need for storage. The BIOBANK is also used by other outside HUNT to store their data.

The HUNT Cloud is an even more open solution. Here you can do analysis inside the actual cloud. This they created because of the enormous amount of data. Other can also store their data in this infrastructure. It is also very secure: ISO certified and there is a twostep log in. Also, hospital data can be stored on this server.

3.3 ESPON

"The European Observation Network for Territorial Development and Cohesion, ESPON for short, is a European funded programme under the objective of ""European Territorial Cooperation"" of the Cohesion Policy of the European Union. It is co-funded by the European Regional Development Fund - Interreg. The mission of the programme is to support policy development in relation to the aim of territorial cohesion and a harmonious development of the European territory. Firstly, it provides comparable information, evidence, analyses and scenarios on territorial dynamics and secondly it reveals territorial capital and potentials for the development of regions and larger territories thus contributing to European competitiveness, territorial cooperation and a sustainable and balanced development.

The current ESPON 2020 Programme is carried through by 28 European Union Member States as well as Iceland, Liechtenstein, Norway and Switzerland and the European Commission. The ESPON program follows the EU budgetary periods (7 years periods) the first ESPON programs was ESPON 2006 (1999 – 2006), followed by ESPON 2013 (2007 – 2013) and ESPON 2020 (2014 – 2020) the next program will be ESPON 2027 (2021 – 2027).

ESPON was a bottom-up initiative from a few member states, Luxembourg was particularly active and the Coordination Unit is located there. The initial goal was in 1999, was rooted in the European Spatial Development Perspective (ESDP) which aimed for sustainable development of the territory of the EU. The aim was to collect and harmonise regional data, collection of statistical data was at that time small and disorganized. Also develop typologies in order to be able to compare the development in different European regions, data collection of urban areas (FUA). Gradually Eurostat and OECD also provide better regional data. The ESPON database is a platform for disseminating data - 1) store data from projects, 2) keep data sets ""alive"" by updating. Updating and operability of the database is taken care of by external experts, ESPON Coordination unit design and make specifications and then external experts (universities, research institutions etc.) carry out the work, ESPON launches calls for this work, rather few partners/core teams respond to these call (specialised work).

The database is open to everybody, with a few restrictions. You can go directly to the webpage and collect the data. They are reworking their metadata model to make it completely machine friendly. It is fully discoverable they also allow quires to the website - to link other web applications to their web. For example, external applications that want to use their GIS systems are allowed to connect to database through the webservices.

In addition to the ESPON database, ESPON also provide several other tools which can be useful for researchers and politicians, such the ESPON Data Navigator, On-line Mapping tool, Mrs ESPON - Monitoring of Macro-regions in Europe, Functional indicators tool, Hyper Atlas, ESPON TIA-tool etc. The different tools can be used for visualisation, comparative analyses of the relative positions of regions, scenarios, decision making etc. The tools could be used even more but many of them are fairly new and not too well known, ESPON informs about these tools on webinars, make training sessions and use the national contact points for marketing etc.

3.4 HERMES

HERMES (Establishing a CompreHEnsive transport Research information Management and Exchange System) project brought together partners from the four transport modes of road; rail; aeronautics and marine. All the mode partners agreed that there was an opportunity to bring together a common area on the internet where researchers could access information concerning transport research organisations; projects and publications via one portal. This would save time and duplication of efforts in searching via multiple search engines.

At the same time, the portal would be a tool for enabling and facilitating the collaboration and cooperation of researchers and the exchange of information outside of the normal working practices within organisations and in European and International projects.

The tool was actually a search engine to help international collaboration in transport. They did not create a new database. It created the “Google” for transport. HERMES (via MoUs) had access to transport research results, data and publications. It also supported the cooperation of researchers (something like a forum between the members).

The search engine had access to open science data where available. We have been connected to USA (TRB) database, Japan official Transportation database, Australia official Transportation database and EU databases.

3.5 Research Data Alliance (RDA)

In 2011 and 2012, many discussions focused on the need for more and better infrastructure to support data-driven research efforts around the world. Both the funding and research communities recognized that the growing gap between research exploration and enabling data infrastructure would slow the advancement of innovation if not better addressed. The concept of establishing an initiative to promote the exchange of data across international boundaries was presented in 2011 to a group of stakeholders. The concept of RDA began to pick up steam. Plans were made for the first RDA Plenary in March of 2013 in Gothenburg, Sweden. Between October 2012 and March 2013, the founding group continued to work towards the development of a real organization. Re A charter, principles, vision, mission, and organizational structure for RDA were drafted. Funders from the U.S. (NSF), the E.U. (the European Commission), and the Australian Government provided support for RDA development in their home regions.

The result, almost eight years later, is a community-driven organization of over 11,000 individual members from 145 countries, 62 organisational members and over 100 working and interest groups dedicated to the development and use of infrastructure that promotes data sharing and data-driven

exploration. The RDA supports the development of both technical infrastructure (code, protocols, tools, models, etc.) and social/community infrastructure (common vocabularies, curricula, pre-standards, etc.) that reduce the difficulties researchers encounter when seeking to access, harmonize, and use data to address research and societal problems.

RDA is based on six fundamental values or Guiding principles:

1. **Openness** – Membership is open to all interested individuals who subscribe to the RDA's Guiding Principles. RDA community meetings and processes are open, and the deliverables of RDA Working Groups will be publicly disseminated.
2. **Consensus** – The RDA moves forward by achieving consensus among its membership. RDA processes and procedures include appropriate mechanisms to resolve conflicts.
3. **Inclusive** – The RDA seeks to promote broad, balanced and inclusive representation of its membership and stakeholder communities.
4. **Harmonization** – The RDA works to achieve harmonization across data standards, policies, technologies, infrastructure, and communities.
5. **Community-driven** – The RDA is a public, community-driven body constituted of volunteer members and organizations, supported by the RDA Secretariat.
6. **Non-profit and technology-neutral** - RDA does not promote, endorse, or sell commercial products, technologies, or services and the development of open and re-usable recommendations and outputs within the RDA is mandatory.

Furthermore, the RDA Vision is that Researchers and innovators openly share and re-use data across technologies, disciplines, and countries to address the grand challenges of society. This is achieved through the practical RDA Mission of building the social and technical bridges that enable open sharing and re-use of data.

So, RDA is completely open and is directly related to the principles and visions of Open Science, Open Research, and FAIR (Findable, Accessible, Interoperable and Reusable).

3.6 S&TDL – Science & Technology Digital Library

The Science & Technology Digital Library Project is the result of an Agreement signed on July 17th, 2012 between the Italian National Research Council (CNR) and the Department for the Digitization of Public Administration and Technological Innovation within the Presidency of the Council of Ministers – later merged into the Agency for Digital Italy (AgID) – as part of a memorandum of understanding between MIUR (Ministry of Education, University and Research) and CNR.

The Project is one of the initiatives of the Italian Digital Agenda to exploit ICTs in order to promote growth, innovation and competitiveness. According to the "Digital Agenda for Europe" it fits in the broader framework of the EU 2020 Strategy. In particular, Established on 1st March 2012 by the Decree of the Minister of Economic Development, in agreement with the Minister for Public Administration and Simplification, the Minister for Territorial Cohesion, the Minister of Education, University and Research and the Minister of Economy and Finance, the Italian Digital Agenda (ADI) transferred the strategies and the principles outlined by the Digital Agenda for Europe to the Italian context, through a coherent plan of concrete initiatives and measures and the effective coordination of public intervention both at central and local level.

The Science & Technology Digital Library is also perfectly aligned with the European and National programmes such as Horizon 2020. In the main documents of the European and National scenario some watchwords are repeated insistently, used as inspiring principles and strengths of the Project:

- research, knowledge, innovation, digital society, digital administration
- access, inclusion, public data, transparency
- ICT, networks, research infrastructures
- sharing, cooperation, community, interoperability, standards
- coordination, integration, harmonisation vs. fragmentation, duplication, dispersion

Hence, in order to systematise knowledge, skills, experiences and high-level resources, the project is committed to:

- promote the emergence of a strong project community that, beginning from the Networking and Information System Unit, which is the holder of the Project within the National Research Council, is open to growing segments of the National Science and Technology sector, establishing intense exchanges and solid relationships with the different communities;
- adequately enhance partners' diverse and complementary expertise.

In particular, S&RDL aims at:

- ensuring an effective, permanent and certified access to information resources, scientific and technical data, expertise, research activities, projects and programs.
- developing methodological approaches and innovative technological solutions that are coherent with the key choices and ongoing modernization within CNR information system;
- creating cutting-edge services tailored n the information needs of different users in order to ensure a large scientific and technological information dissemination and to make research, its processes and its results transparent, open and reusable;
- guaranteeing interoperability with the most important National and International R&D information systems;
- ensuring the harmonization.

3.7 TRIMIS

TRIMIS (Transport Research and Innovation Monitoring and Information system) started in 2017 – at a request from the European Commission as a top-down initiative, however the idea came about as a successor of a previous EU funded project TRIP. TRIMIS is a neutral platform, without any partners, but they interact with a lot of different stakeholders. It is a long-term project, financed by the Horizon 2020 program – it depends of course on the success of it, but in the last four years it has been successful.

TRIMIS is a completely open database. It is composed of many parts, one of them is the database, consisting of 9000 projects and programs. It is supposed to be a one stop shop for anything that has to do with research and innovation in transport and consists of EU founded projects and national

projects. The content of each project consist of who participated in the projects, a summary of the project, the budget, relevant documents, etc. – and if you want to look at the underlying data you can contact the responsible partners of the project. Everything is public.

In addition to giving information through the database, TRIMS also do research analysis on their data. TRIMIS started and is supporting the Strategic Transport Research and Innovation Agenda (STRIA). STRIA is developed around 7 roads maps, which look at the network and transport management, data design etc. Based on this framework TRIMIS analyses and assesses the topics, funding distribution, types of partners, etc. to then see what kind of research and policy gaps exist and what future possible research and policy gaps needed to be addressed.

They are not just working on this approach in relation to STRIA but on different thematic scope and area. I. e. they did a report on waterborne carbonization in Europe, where they assessed all the projects that was related to this. These assessments end up in reports that are also open and published on TRIMIS as well as on the EU Science HUB.

TRIMIS is an interactive tool. It is possible for anyone to submit projects. When they receive a request from a project that is in line and compliance with the scope of TRIMIS, they can **also promote the project.**

4 Analysis

4.1 Resources

The first barrier to innovation is **resources**, which include lack of human capital/competence, lack of physical resources, time constraints in addition to lack of financial resources. In order to make TOPOS sustainable after the project ending the financial resources are one of the most crucial parts. We have gathered information about several different financial solutions that TOPOS could potentially utilize.

Financial resources:

Public funding (international) ESPON is funded on a budget of 50 mill euros - 46 mill by EU, the rest by the four partner states. It originated as an EU funded project, and after it was finished, through a bottom up initiative made by Luxemburg, was working to find a way to finance it. The solution was to create ESPON as a transnational Interreg program.

RDA was also initially funded through grants from U.S. (NSF), the E.U. (the European Commission), and the Australian Government that provided support for RDA development in their home regions.

Public funding (national) 80 % of HUNTs funding is through public funding with a combination of national and local funders: Helse Midt-Norge (local health unit), NTNU (university), Fylkeskommunen (municipality) and Helsedepartementet (national health unit). The whole model is based upon collecting data that benefits 'everyone'.

The Science & Technology Digital Library (S&TDL) has also been created through national public funding and its cornerstones are represented by a strong partnership with key players of the National R&D system – both internal and external to the National Research Council – and their active and proactive participation in project activities. Currently, the Science and Technology Digital Library uses the important contribution of some of the most valuable Italian research institutions.

User-payment 20 % of HUNTs funding is through user-payment for both using data and infrastructure:

1) **Paying for data** Using HUNT data for an article/paper, the organization must pay a fee of 3000 EURO for each article. When a research project applies for funding they therefore calculate in the fee in the application when starting a project based on HUNT data. This way the organization that use HUNT data does not pay for it directly, but through the external project funders.
2) **Paying for infrastructure and service** HUNT also make some money in terms of offering their infrastructure for storing data for others (i.e. hospitals). But first it needs to be quality controlled before entering the HUNT databank. This is a collaboration between HUNT and the user, where they charge for the hour, in addition to an annual fee for storing the data.

Grant funding through regional offices: the RDA authorises and gives validity to the affiliated Regions. It is in fact RDA that provides the forum for the global community to connect and share knowledge that provides the context in which Regions operate. To this end, RDA provides the forum for the international community to connect and share knowledge on all aspects of data sharing. RDA then supports the work of the Regions by various means, including the support of Regional leadership to build the RDA community and create impact by offering small grants (or collaborate on the application for funding) to assist with the creation/development of Regional activities.

Regional in-kind contribution (from the regional offices to RDA):

- Providing in-kind support, in particular:
    - Skills, duration and details will be agreed in collaboration with RDA to ensure the support offered can generate value (e.g. profiles like Communication & Marketing Expert, Outputs & Adoption Expert, Technological Expert, Event Expert, Community Engagement Expert, Operation/Process Expert, Project Manager).
    - Each in-kind staff resource would be provided to the Secretariat at a minimum of 50% of his/her time.
    - Staff may be provided to complete contracted pieces of work (e.g. an analysis of outputs adoption to inform RDA strategy).
- Facilitating the hosting and organizing of Plenaries, like:
    - Organisation and expenses are the Region's responsibility (contribution to Global)
    - Hosting governance meetings before, during, or after plenary.
- Shaping future directions, such as:
    - Interacting with national research funding bodies, ministries and other government officials to influence data policy and digital agendas.
    - Developing robust sustainability plans and business models in collaboration with national funders and governments to ensure continued contribution to RDA.

- Contributing to RDA business and strategy through multiple means including the Regional Assembly.

Regional contributions to the global office to support global operations RDA made the choice to allow regional / national funding to contribute to the regional / national policies and goals in terms of research data management and interoperability. Building regional offices and champions allowed the freedom to apply for funding but at the same time contribute with in-kind staff support and "cash" to the globally coordinated activities. Regional and global grants were obtained to support other activities too, for example, adoption programmes, early career and fellowship awards, bi-annual meeting sponsorship.

Annual Organisational membership fees in addition to the regional contributions, RDA has an annual organisational membership scheme which provides revenue. Such a scheme includes different member profiles, in particular:

- organisational members, including Corporations/Enterprises, Universities, Research and development agencies, Non-profit foundations and/or community organizations, Libraries, Consultant firms, Other types of organisation not already listed that have an interest in furthering the goal of the RDA and have an intent to adopt RDA's infrastructure solutions. The organisational members are asked to contribute with an annual subscription fee based on the number of employees (< than 50 employees: USD1,000 per year | 50 and <250 employees: USD2,000 per year | >250 employees: USD10,000 per year).
- Organisations can become Supporters by contributing to the Business of RDA through an annual financial contribution to be agreed directly with the RDA Foundation. Indicatively contributions of this kind have been between GBP £20,000 to GBP £60,000 per annum.
- Organisations can also become Sponsors of Research Data Alliance Business and / or Work activities through:
    - RDA plenary meeting sponsorship
    - organisation of research data meetings
    - data expert/fellow/early career programmes
    - testing, adoption & implementation projects / pilots
    - RDA plenary meeting travel support
    - Creation of a programme to offer visibility and recognition to data excellence
    - Creation of a training course or series of focused events
- Regional members: each Region will make a yearly financial contribution towards the operating costs of RDA. The amount will be on a sliding scale so that large economies pay more than small economies and will be calculated from the GDP of the region. While every Region may contribute a different amount, a monetary contribution established in a formal arrangement between the Region and RDA is an essential part of being considered as an RDA Region. The source(s) of the funds will be different for the different regions.

In the case of RDA, as a community driven initiative, they had great interest from the global practitioners and experts to get involved and make RDA successful. Hence a conscious decision was made to apply for different funding grants in regions / continents and not to create an intergovernmental organisation. Naturally this allows a great deal of freedom for the regional offices

to design and apply for grants to their national funders. On the other hand, the timing and objectives of the grants are difficult to coordinate across the globe fragmenting the support at times.

Project funding (not funded today) HERMES had an overall budget of € 814.130 where the EU contribution was € 730.150. It was coordinated by the University of Newcastle, UK. Unfortunately, after the end of the project and although the search engine was fully functional, there was not further funding from the EU. Many different ways have been tried to further fund the project but without luck. The search engine ended two years after the project end due to lack of resources for maintenance. The maintenance cost was calculated to be as €6.000 but there was no luck even to this small amount of money.

Safetycube OSS did also not receive any funding after the initial project. However, they have kept up the marketing of their product through other Open Science initiatives – referring to it in newsletters and linking it through other Open Science portals.

4.2 Behaviour

Behavioural barriers include attitudes of employees, management, behaviour, habits and routines. Including corporate social responsibility. Because all the interviewed projects and organizations are built upon open science concepts the people that work in the projects/organizations can't really have other attitudes towards sharing than positive ones – because their funding is based upon open science. So, the problem that 'individual researchers think they own their data' is more relevant other types of projects and organizations where open science it not the main reason for funding, and where the competitiveness for projects is based on other aspects. Hence, we did not ask any questions related to this. The attitudes of the management and social responsibility perspective in the discussion under organizational barriers and benefits of open science confirms this assumption.

4.3 Information

Information barriers include the lacking ability to comprehend and understand data, absorptive capacity (the organisation's ability to identify, assimilate, transform and use external knowledge/information) and lack of knowledge. On a lower level this also includes other types of information – data quality and marketing. In relation to TOPOS we have focused on data quality and outward information spreading i.e. marketing of the service and creating an awareness among users as this is highly relevant for making TOPOS sustainable.

4.3.1 Data quality and maintenance

Depending on who gathers and collect the data the issues related to quality and maintenance differ among the interview cases. SafetyCubes data was collected by different countries, but entered by the project itself, hence they had quite a lot of control on the quality of the end result. For HUNT they both collect and enter most of the data themselves, but also other people outside HUNT use their infrastructure to store their data, which makes the quality control for external entries important. For ESPON all the data is collected by others and the maintenance and update is externalized. TRIMIS upload internally but get information from external partners as well – which makes the quality of the data differ. S&TDL gathers up data internally so that potential issues related to data quality and compatibility can be minimized. In the case of RDA, they do not collect data but they act as a facilitator in promoting the exchange of data across international boundaries. RDA directly and logically tackles

numerous data infrastructure challenges through the work of its Working Groups[1], Interest Groups[2] and Communities of Practice[3].

All data collection and upload done internally

SafetyCubes had some issues with the quality of the data, but because they could communicate directly to those who delivered the data during the project period they had the opportunity to do a lot of internal quality control (programming in R to look for consistency) and communicate back to correct mistakes. Some countries did not have official numbers - then the numbers were based on other sources (i.e., specific projects that had tried to work out numbers). Some part of the information will be outdated after some time. No one volunteered to fund the maintenance of the data. However, a lot of the data will continue to be relevant – especially the research analysis that is done based on different studies. Also, the calculator can be used since you can manually enter and adjust the numbers to fit for your country and year. However newer studies will not be entered, and the individual analysis could be updated when new research is done if maintenance was funded.

In case of the Science&Technology Digital Library data are collected by CNR that plays also the role of intermediator when gathering information coming from the other members of the partnership. This centralized data collection model has been chosen because CNR:

- carries out research in the main areas of knowledge and it is highly *multidisciplinary;*
- its main objective is *to create value through knowledge generated by research*; according to Art. 3 of CNR *Statute*, in particular:
    - It proposes, coordinates and develops strategic research projects of National importance, often appointed directly by the Government, in collaboration with universities and companies, taking into account the research and innovation needs at a regional level.
    - It gives support to International institutions of the European Union, and to National and regional organisations for the evaluation and monitoring of scientific programmes.
    - It collaborates with universities and with the other research institutions in order to promote and spread scientific and technological knowledge and to share research infrastructures.

[1] WGs are short-term (18 months) and come together to develop and implement data infrastructure, which could be tools, policy, practices and products that are adopted and used by projects, organizations, and communities. Embedded within these groups are individuals who will use the infrastructure and help in making it broadly available to the public. Any RDA member can join or initiate a WG.
[2] IGs are open-ended in terms of longevity. They focus on solving a specific data sharing problem and identifying what kind of infrastructure needs to be built. These groups identify specific pieces of work and can start up a WG to tackle those projects. Any RDA member can join or initiate an IG.
[3] CoPs investigate, discuss and provide knowledge and skills within a specific discipline and/or research domain. These groups are committed to directly or indirectly enabling data sharing, exchange and/or interoperability by serving as THE coordination focal point for RDA in specific disciplines/research domains.

- It promotes the enhancement and use of research results and knowledge spread in society.

In particular, the Networking and Information System Unit of CNR has been considered as the *mediating unit*, able to put innovation at the service of the information needs of the different scientific communities in a highly flexible way and without disciplinary bias. Such a Unit, in fact, has gained a vast and solid experience over the years at the service of the CNR scientific network. It has developed innovative technological tools and services – especially platforms and advanced instruments related to information and knowledge management, and effective user-oriented management services.

It is worth noting that the S&RDL has been conceived for being addressed not only to researchers and the different scientific communities, but also to the business sector, students, families, all the citizens that can benefit from information contents and cutting-edge services, profiled exactly to meet their specific needs. Bibliographic, documentary and information resources; scientific, technical and statistical data; research activities and programmes; institutions, competences, expertise; a varied and flexible spectrum of advanced services are provided to everyone, beyond the usual librarian approach, thanks to a unique simplified access.

In line with the S&RDL main objectives fostering the spreading of research products and datasets, information and knowledge about R&D historical contents

- wider usage of information resources in digital format and bibliographic services, purchased through agreements with publishers and commercial aggregators;
- promotion of agreements for cooperative purchase that are useful to foster economies of scale;
- increase in contracts and agreements for the legal deposit (Law n. 106, 2004, April 15th Decree n. 252 of the President of the Republic, 2006 May 3rd).

The purchase activities are currently carried out by many research and academic libraries, that only partially use cooperative strategies in order to maintain and/or develop their collections. Taking advantage from the recent experiences in purchasing policies coordinated within the CNR library system, WP2 goal is to optimize cooperative purchase practices through political, legal and financial proposals developed by partners, that may offer alternative models (ex.gr. open access strategies) and/or integrated with current consortium strategies.

Thanks to cooperation agreements with other academic and National research institutions, with inter-university purchase consortia and with publishers and information content aggregators several benefits will be obtained:

- library collection rationalisation;
- purchase cost savings;
- wider usage of resources and additional services (digital preservation, web services, advanced bibliographic services, etc.).

In order to support collections and other bibliographic documents selection and development activities, technical tools for usage statistics continuous monitoring and for users feedback collection are implemented.

Some data collection internally and some external collectors, upload done internally

TRIMIS upload internally but get information from external partners as well – which makes the quality of the data differ. A challenge that they have had, and still have – is in relation to national funded projects. As they are a part of the EU commission family with CORDIS they have a direct link to all project carried out in the EU, and the quality of the information about the project is also high. National projects are more inconsistent in terms of information about the projects as it is more difficult to follow up on – it is dependent on their network connection, but also on language barriers where a lot of national projects are conducted in the country's own language. This is something they are working on continuously. Data related barriers are the main challenge. TRIMIS could potentially use the national projects to make their analysis on, but the level of detail of data varies quite a lot. This is one aspect they are focusing on to improve, but this is linked to the linguistic barriers. To improve this, they have a tight and regularly contact with their large national networks and assist them in uploading if there is a bunch of projects that needs to be uploaded. The connection with the national contacts (ministry, research institutions, STRIA network) is working well, however it is on voluntary basis – which makes it somewhat vulnerable. For the type of data used for the EU based analysis, the quality of the data is quite high – however they are working internally to add additional sources of information and combine different databases to continue improving this.

HUNT carry out most of the data collection themselves, which ensures high quality when entering it into their different systems. They are also very concerned about harmonizing data – in terms of using identical surveys to be able to generate comparable data. When you extract a file from the database you automatically get a generation of a metadata connected to the file. We get paid for the quality of our data through the fee of publishing with the HUNT data. As mentioned earlier, this fee is covered by the funder of the project (so you have to add it into the budget when you apply for money). However, they also have some data that is collected by others, and this also needs to be reusable. "Sometimes data is collected by others who are using HUNT data in return. Then HUNT control that the data externals deliver to them is of proper quality before they gain access to any of the HUNT data." As mentioned earlier, externals also utilize HUNTs infrastructure for storing their data, and also here they do a quality control. This is a collaboration between HUNT and the user, where they charge for the hour, in addition to an annual fee for storing the data.

All data collected externally, uploaded by an external service

ESPON collects data on regional indicators on different geographical levels all over Europe. "Our projects are constantly struggling with data. Especially on regional data on lower geographical levels (NUTS3)." Eurostat has a hard time catching up and are moving quite slow, and they must rely on the commitment of the statistical institutes in different member states. It is quite difficult for them to tell member states to implement new indicators and, they don't collect it the same way – which means it's not comparable. Another issue is that in many topics there are not enough relevant data, especially newer topics like circular economy. It's also difficult to go beyond NUTS3 as there are lots

of gaps on city data. Municipalities are constantly merging and changing boundaries which adds another level of difficulties – as there are changing codes and over 100.000 units to keep track of.

In terms of data quality, spatial problems are another battlefield: "To connect statistical data to spatial data is very difficult because of the changing nature of NUTSs regions (changes in codes)". Project are struggling to understand why the spatial data is not joining up with the statistical data. The researcher needs to have some background knowledge to understand the mechanisms behind why this occurs and how to connect historical data to the boundaries. The same dataset offers several variations of the code, but the data might be in the previous version – which is often misunderstood. ESPON try to assist the users with clearer metadata and to give the spatial data with the statistical data whenever possible. "We have regional boundaries that correspond to the NUTS region for the relevant year. Now our database enables when you upload the data you define the version for the right level. The system pics it up." Before the projects just dumped their data, which was not useful for the end user as they did not know what the data consisted of. Also, they have tried to reduce duplications by complementing the existing uploads, instead of updating twice.

The uploading and updating (maintenance) are externalized to the original service that built the database. This is the same team that work on the Eurostat database which is an advantage in terms of making sure the data can be harmonized. This work requires quality checks and proper strategy behind it. We have different channels for uploading the data. We also need to synthesize the data that adds most value - key datasets. There is a process before the upload on what needs to be uploaded where and how. The metadata is defined, and the systems then give them a file on how to make the template that makes it easier for machine learning to read.

RDA deals with all the problems and challenges related to data, but they do not actually host or store any data. RDA looks at the data issues and seeks to identify pragmatic and implementable solutions to them. That's why there are several RDA activities around Data Quality and many of the RDA groups work on interoperability aspects[4].

RDA is an open platform and anyone can access information and re-use it / download it in a completely open way. As a community, RDA has individual membership which is free and only requires acceptance of their guiding principles and code of conduct. Being a member (currently there are almost 11,500 from 145 countries) allows individuals to set up, join, contribute to the working and interest groups via the web platform, as well as provide open comments and community feedback to the outputs and activities conducted via RDA.

HERMES faced databases' compatibility problems, which was solved by APIs and 3rd party software. The databases had to have common fields to store and share the same information and not to lose any. For the USA and Japan databases HERMES have used API links to interconnect these databases. On the other hand, for the Australian database interconnection Google has been asked to create a middleware application to connect the two databases. It was a paid service from Google. There were not any quality issues. The connected organizations (i.e. TRB and OECD) secured the quality of the data. The used data were updated by the connected organizations and not from the HERMES project.

[4] https://www.rd-alliance.org/rda-for-you

After all, the objective of the HERMES project was the creation of this search engine. They were very careful on cyber attached not to lose any vital information of the registered users.

4.3.2 Awareness of users / Marketing of service

There are different ways of *measuring user awareness* among the projects/organisations:

Google views and website visits SafetyCube measures user awareness through views on Google. The first period, October 2017-2019, they had 100.000 separate views on Google. 48 % outside of EU. The second period, 2019-2021, they also had 100.000 separate views on Google – so in four years they had 200.000 views on Google. HERMES used to measure the views using Google views. Exact numbers could not be recalled but the figures were very positive. TRIMIS also have had a successful visit on their website, and the total number of unique TRIMIS website visitors in 2020 was more than 100 000.

Projects based on data Another way of measuring user awareness can be through actual use of data. HUNT have had 215 doctoral thesis and 400 projects based on their data.

Webinar attendance a last way of measuring awareness could be participation rates on event or webinars related to the organisation. ESPON have been marketing through webinars and trainings and they have held 10 workshops with 25-35 people in each workshop.

Members Measuring the increase of members is also a way of measuring awareness. The growth of RDA over the eight years it has been active has been quite considerable, in terms of members (individual and organisational), in terms of the actual working and interest groups (over 140 since inception with over 90 currently active). The community has produced 50 flagship outputs to date, all open and available for reuse. They have over 150 documented cases of adoption of these outputs in different organisational and disciplinary scenarios.

Several *marketing strategies* are used to get people aware of their projects, and to market their service:

Targeted Google searches: SafetyCube states that they end up on the first page on a lot of Google searches in relation to traffic safety due to good keywords, lots of content and a strategic name of the project.

External linking: All the partners in the project have a direct link to SafetyCube on their webpage, reaching a larger audience. This strategy was also used by HERMES. HERMES got in touch with research centers, companies and universities that were related with transportation. Most of them liked to host the HERMES logo on their main page of the website to introduce the HERMES services to their clients, students, etc. The HERMES logo redirected the user to the search engine of HERMES.

Newsletters: During the project SafetyCube did a lot of marketing in the beginning with newsletters. The project is still mentioned in the newsletters from the Road Safety Observatory which gives the project continuous promotion even if the project is finalized. TRIMIS also does a lot of their promotion via newsletters to their subscribers. RDA regularly also update their members with monthly newsletters, summary PowerPoints (that they can adapt and reuse), blogs, events,

webinars. They encourage their members to use the events, news and blog submissions (moderated for spam purposes) facilities on the web site.

Webinars, seminars, workshops and other events.: There are different types of events that can be used for promoting, and all interviewed cases have used this strategy. SafetyCube held a huge launching seminar at the end of the project and did presentations during the project period and HERMES used presentations of the project on workshops, conferences, etc. to help marketing the search engine. ESPON use webinars to promote their work and hold workshops for politicians, researchers and students. Sometimes they have targeted workshops with slightly different training fitting to the participants joining. TRIMIS is also quite active as a dissemination actor and participate as much as possible in events and conferences to promote TRIMIS. TRIMIS does not have their own TRIMIS event, but are part of European Commission events, where they are promoting their projects. It has therefore turned into well-established network in the transport sector. This could maybe also be addressed to the fact that the network of the stakeholders was very supporting from the beginning of the process. One of the most important impacts of RDA has been the value its Plenaries bring to RDA members, stakeholders, and other groups. The semi-annual RDA Plenaries (typically 2.5-3 days) are dynamic, high-energy, productive, and inclusive. The Plenaries are primarily working meetings and although each Plenary has high-profile keynote speakers and excellent panels, the majority of the time is spent doing, rather than listening. There are often a dozen "officially scheduled" RDA group meetings going on at the same time and many "unofficial" group meetings, allowing RDA members to make face-to-face progress on their work more rapidly than possible when they interact remotely (often in vastly different time zones). An additional value of RDA Plenaries for members is whom they attract. Many different stakeholders (Program Officers from a variety of national R&D agencies and non-profits, policy makers, representatives from other organizations, publishers, journalists, and others) attend the RDA Plenary to collaborate, get the "lay of the land" for a broad scope of problems and disciplines, and interact both with the RDA community and each other. Because RDA provides a neutral and collaborative environment, many groups co-locate meetings with RDA, stretching out the three-day Plenary to a crowded week of synergistic activities.

User feedback during the project: SafetyCube made sure that their users could give feedback directly via user surveys and presentations throughout the project.

Proactive media participation: is a strategy used a lot by HUNT. There are several reasons for being proactive in the media. It is important to update both politicians, participants and the research environment.

Paper promotion: HUNT has made sure that people that publish using HUNT data must include HUNT in the title of the article. This way they will get marketing every time someone publish using their data.

Merch: HERMES beermats were designed and produced from a marketing company. The mat had the HERMES logo on it together with the search engine link. They gave these mats to everyone on any occasion (e.g., post it together with an information letter, share them on a workshop or conference).

Social Media: TRIMIS does a lot of their promotion through social media, where they have their own LinkedIN account, and use DG MOVEs Twitter account. RDA also invest considerable time on their

social media accounts. Particularly Twitter and LinkedIn as the have proved to be the most interesting and valuable.

Utilizing national networks: RDA identifies national champions in countries and leveraging on their networks, experience and connections within their country. The value for them is to receive and provide updates on a global level. This is particularly important where the working language is not English as it offers direct channels to the end users.

Domain / discipline specific ambassadors: RDA also has experts on data from specific science and research communities that facilitate an open dialogue with data practitioners and organisations working in domain specific areas[5].

Even though there are lots of different ways of promoting, as stated by SafetyCube – the most important thing is that the content is good and useful to the users – "so that everyone knows that this is the place to go for the most important information". In terms of marketing there are also some pitfalls that is mentioned by ESPON. Sometimes their tools are not always up to date, and therefore they have been changing their strategies more towards a framework approach to avoid the problem of not fronting the latest tools.

4.4 Technology

Technology which include lack of technical capacity in terms of available competence and storage capacity. Some of the interviewed companies does not have technical issues that is relatable to TOPOS.

Communication to the people entering data into the database. A problem mentioned by SafetyCube was that there were some communicational issues related to uploading into the database, as there was a lot of studies of different types. Due to the large difference between types of studies they could not make detail descriptions on what should be entered into the system, also they had no system to identify missing data. There is a need to explain concepts in detail - sometimes people who does not know what something means just make a guess.

Search functions and keywords are important to think about for the end user. If you are looking for 'pedestrians' - will you find the study mentioning pedestrian and cyclist or pedestrians only? Synonyms: entering keywords - if someone is using a small/capital letter, using a short wording or a long wording of the same. Some synonyms in different disciplines means different things.

Presentation of study weaknesses. For this project different studies had different weaknesses. How should the study weaknesses be presented? Should it be open field or predetermined answer options? Some weak studies have not mentioned that it has weaknesses - if they have not filled it out it is not necessarily a good study. Should this be evaluated by someone else or just use what the

[5] https://www.rd-alliance.org/rda-disciplines/rda-europe-ambassadors

study points out itself? Sometimes missing info is the reason behind a study appears to not have weaknesses - this is an issue when someone outside this project use the data/studies.

Technical competence HUNT has had issues with getting people with the correct IT-competence to work on the extended development of the HUNT databank. They are struggling to get people with the correct competence to apply because it's very specialized. This is also mentioned by ESPON that all the time ensure their staff have plenty of knowledge for maintaining and developing the databases as this affects the quality of the database. "We are about 23 people. Mostly 3 people that have the capacity. This is a risk factor with changes in the staff.". TRIMS have a separate IT support that has built the website from scratch, but there are no technical barriers relevant for TOPOS. They have 6 people working on the content in addition to the IT support.

Public procurement. According to ESPON it is important to think about technicalities when using external resources. Any public procurement process is determined on how well you are able to define the expectations. The main barriers are interoperability when working with open standards and open source: "To link the different tools together with the database. Not duplicating, so other sources can get the data. You don't have to update all systems all the time, but general you have one only principle when you deal with data - so you work on standards for disseminations - build the architecture so you can link other systems into it.

Semantic Portal. With reference to S&TDL, it has been conceived to be a Semantic Portal (SEAL) based on the use of Semantic Web, which:

- enhances and facilitates the access, sharing and exchange of information contents and services between the different communities and their reuse, through a user-centred approach;
- is able to ensure the integration of the Portal and its interoperability functionalities with the project components and within the wider National and International R&D information system.

The use of Semantic Web solutions and services allowed the different communities to actively contribute to the Portal, to its conception, design, planning and evolutionary development. According to the enhancement of the social dimension of the Web, a Social Semantic Web Portal (SSWP) has been selected since it can promote knowledge sharing, widespread participation, better profiling and advanced customization of services among the project partners and the user communities, combining the current social web needs with the semantic web ones. Semantic web, in fact, results to be a very useful tool for making information more meaningful and understandable to both people and computers. This because system developers are enabled to find relationships between tagged information using ontology, which allows for a common logic and structure for web pages. In particular, there are small program portions, called software agents, that can easily locate and combine information from different sources.

As defined in the framework of the "SWAD-Europe – W3C Semantic Web Advanced Development for Europe" project[6], a Semantic Portal then refers to an information portal in which the information is acquired and published in semantic web format and in which the structure and domain model is made

[6] https://cordis.europa.eu/project/id/IST-2001-34732

explicit (e.g. in the form of published ontologies). Such a project carried out an interesting analysis on the possible advantages features of the Semantic Portal approach can have, as illustrated in the following table, that can be of help also in the case of TOPOS.

Table 2: Traditional design approach vs. Semantic Portal approach ***[Source: "SWAD-Europe deliverable 12.1.7: Semantic Portals Demonstrator- Lessons Learnt" Report]***

Traditional design approach	**Semantic Portal**
Search by free text and stable classification hierarchy.	Multidimensional search by means of rich domain ontology.
Information organized by structured records, encourages top-down design and centralized maintenance.	Information semi-structured and extensible, allows for bottom-up evolution and decentralized updates.
Community can add information and annotations within the defined portal structure.	Communities can add new classification and organizational schemas and extend the information structure.
Portal content is stored and managed centrally.	Portal content is stored and managed by a decentralized web of supplying organizations and individuals. Multiple aggregations and views of the same data is possible.
Providers supply data to each portal separately through portal-specific forms. Each copy has to be maintained separately.	Providers publish data in reusable form that can be incorporated into multiple portals but updates remain under their control.
Portal aimed purely at human access. Separate mechanisms are needed when content is to be shared with a partner organization.	Information structure is directly machine accessible to facilitate cross-portal integration.

The SWAD-Europe project has so identified a set of advantages to using semantic web standards for information portal design as reported in the box below.

Ontologies

The use of an explicit, shared domain ontology enables both data sharing and richer site structure and navigation including multidimensional classification and browsing schemes. Use of the Semantic Web standards for encoding these ontologies also enables the ontologies themselves to be shared and reused across portals. Several projects have already derived benefits from ontology-driven portal designs [SEAL][WEB-PORTALS].

Evolution

Requirements change over time leading to extensions to the information model. The semantic web helps in two ways. Firstly, the user interface and submission tools can be generated from the declarative ontology. Secondly, the semi-structured data representation of RDF permits new data properties and types to be incrementally added without invalidating existing data, in such a way that both original and extended formats can be used interchangeably. This suggests an alternative approach to information portal design. Instead a long top-down design cycle, we start from a seed ontology and information structure that we extend incrementally.

Community extensions

Whilst many portals support constrained community annotations, such as comments and ratings, the semantic web approach allows more extensive community customization. For example, during work on a portal for wildlife multimedia it became clear that many user communities would like specialized navigation of the data (based on formal species taxonomy or behavior depicted), which was unfeasible for the centralized portal provider. Using the decentralized approach it is possible for communities to develop these specialist navigation structures as a set of external RDF annotations on the portal data. The central site can then aggregate the community-provided enrichments.

Aggregation and decentralization

One problem with traditional information portals is that they are often dependent on the responsiveness of the central maintainers, so that if funding disappears, so may the data. In the semantic web approach supplying groups host their own data and the portal becomes an aggregating service. Central organization is still needed (for example, to provide the initial impetus and ensure that appropriate ontologies and controlled vocabularies are adopted). However, once the system reaches a critical mass it can more easily be self-sustaining - anyone can run an aggregator service and ensure continued access to the data or a new supplier can add data to the pool without a central organization being a bottleneck.

[Source: "SWAD-Europe deliverable 12.1.7: Semantic Portals Demonstrator- Lessons Learnt" Report]

4.5 Organizational

Organizational barriers include culture, management culture, lack of support systems and commitment. In terms of culture and commitment, there are two main stands on Open Science – those who see sharing as problematic in terms of RoI (BEOPEN, 2019, D 2.4), and those who see open science as an advantage and even plan their business idea around open science. Advantage from sharing vs. the competitive advantage from not sharing was therefore an interesting aspect to address in the interviews in terms of selling TOPOS and open science in general to users. All the projects interviewed were/are intended to be open science projects from the beginning and the funding was/is based on that. Therefore, there has not been any issues in terms of competitiveness as they have based their financial plan on the fact that they deliver open data. However, because they are producing a lot of open data, we also asked what kind of benefit they see from working with

open science, in terms of shedding light on the culture and attitudes these types of organization have.

"In terms of competitiveness there is rather a lack of heads to use the data than a lack of data". As HUNT states in the interview there is a much larger issue to get enough heads to utilize all their data than there is competition on using it. **Genetic research has been a driver** internationally in terms of the sharing culture in health research because of the need for a large dataset for analysis. "We have a capacity problem in terms of being active in all the research. The results are many and it is used a lot. I will often get invited for joining other people's research within my field". Hence the data sharing gives HUNT **more opportunities than missed competitive advantage.** The data is so rich there is no limits on what to do with the data. The thought behind the sharing is quite **altruistic:** "How can we make better public health?". Another driver for sharing is that the **international spread and use of the data is considered quite high status.** The culture in medical papers and projects is also to include as many writers as possible to show that you have the best people on your project – this is different from some other disciplines where its considered higher status to be the only writer, showing you can do everything yourself. This has probably made the sharing culture stronger.

The SafetyCube project talks about the different levels of benefits gained from their open science project. First of all, it's a benefit for road safety – creating a high-quality tool to support the system – which **benefits the society as a whole**. But also, it is a benefit for the experts which are becoming better experts. They themselves also gets benefits as (1) they have the tool to use for their own job, (2) they **get status and marketing benefits** of creating the system (3) potentially get a new client. The university behind the SafetyCube system believe in the model of open science – and has done several projects related to open science. **"The benefit of becoming known, that's a bigger benefit than selling small pieces of the information."** They have made a name for themselves and is now invited to projects similar to SafetyCube. This is in line with what HUNT mentioned as well about the benefits of open data and the marketing around it. A last benefit of creating an open tools is that after the project is finished someone is needed to work on the tool – which potentially generate more work for the inventor of the system/tool that often get the job of updating the technical parts. A last input from Safetycube is the fact that having the data open was used as a **selling point to get funded** from the EU – turning it into a competitive advantage compared to other projects.

The positive effects of TRIMIS – or what they are working towards - is that they are recognized as a single point, one stop shop for research and innovation. It can be a neutral tool for innovation for a broad audience of stakeholders: policy makers, academia, transport industry, planning. If TRIMIS can be a reference point – or an initial point where stakeholders look for information, TRIMIS see this as added value for them self as they are **supporting research and identifying research gaps**. This will possibly also **stimulate collaboration among different stakeholders** – and **reduce duplication of work** if people see that something is already studied in the database. TRIMIS is therefore supporting and going towards the Open Science paradigm. The HERMES project also saw similar advantages of open data to **avoid repetition of work**, get **awareness of past work** – past projects and the quicker development of the applications.

There are many benefits to data sharing and reuse according to RDA. They have produced two policy documents on this, one in 2010 and one in 2016 and the recommendations are as valuable today as they were then:

- Riding the Wave: How Europe can gain from the rising tide of scientific data[7]
- The Data Harvest Report – sharing data for knowledge, jobs and growth[8]

In particular, the first document well represents the fact that data should be considered a sort of infrastructure, which means that the intangible knowledge, represented by different data formats, values and uses, may have a **huge value if made seamlessly available to use, reprocess and cross-fertilize to promote the creation of new knowledge.** Given the variety of scientific information – in terms of data formats and types, but also of people and communities creating and utilizing the data – even within the same scientific community, there are different points of view, different ways of analysing, sharing and handling data. Additionally, there is also diversity in how the data are stored, categorised and mapped, and in who can access what kinds of data, and how. Hence, **aachieving an interoperable data infrastructure is a significant challenge but on the other hand also a great enabler** for easily accessing relevant data files anywhere in the world, in any field.

The second report highlights how Open Science could provide benefits to different stakeholders. As shown in the table below, additional value can derive from open data **not just in science but across the economy** by, for instance, **increasing the efficiency and effectiveness of businesses** or allowing for new products and services identification.

Table 3: The benefits of open data ***[Source: "The Data Harvest Report – sharing data for knowledge, jobs and growth" RDA European Report, December 2014]***

Stakeholder category	Potential benefits of open data
Citizens	All people will benefit from the products and services that are developed around open data and sharing – directly or indirectly. More accountable, efficient and effective businesses and government result. Most importantly, citizens are empowered, have the information they need to make decisions in all spheres of life; they are engaged.
Entrepreneurs	Open data is a source of inspiration for entrepreneurs and provides the raw material for new products and services. Examples include the opening up of weather data, leading to the private sector provision of information services, making global positioning data available to a mass market in satellite navigation systems, and making Human Genome data freely available to the genomics sector. No one organisation has the money or the expertise to extract the full value from its data. Opening it up to entrepreneurial imagination will foster unthought-of innovation.
Scientists	Freely exchanging data will transform the nature of what it means to be researchers. It will make their work easier and faster, as more data and tools are put within reach. It will open new

[7] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=707
[8] https://www.rd-alliance.org/data-harvest-report-sharing-data-knowledge-jobs-and-growth.html

	research avenues, crossing old boundaries of discipline, institution or country. It will create new career opportunities, and get more researchers crossing borders. And, through greater engagement with fellow citizens, it will enhance their status and relevance in society at large.

These two reports highlight the importance of data sharing and foster an international effort to stimulate and coordinate work on data sharing by proposing the following actions:

- Do require a data plan, and show it is being implemented.
- Do promote data literacy across society, from researcher to citizen.
- Do develop incentives and grants for data sharing.
- Do develop tools and policies to build trust and data-sharing.
- Do support international collaboration.
- Don't regulate what we don't yet understand.
- Don't stop what has begun well.

For what concerns the S&TDL, the need of including within the digital library a great variety of data induced the adoption of a sort of "evolutionary" data model following an agile approach aiming at standardizing as much as possible the entry of information and on the other side managing any unforeseen issues that could come out. To this end, a particular effort has been required to the ICT people for data integration, homogenization and standardization. Another critical aspect resulted to be the level of data accessibility and usability since the digital infrastructure can play the role of facilitator, but data owners are the ones who decide on the degree of data access.

4.6 Legal

Laws and regulations

In terms of GDPR and other legal issues a lot of the interviewed companies does not have this type of barriers. HERMES did not face any issues with GPRD, as the project existed only before the GDPR laws. Also, some organizations have data that is not afflicted by the GDPR laws, like ESPON or in TRIMIS case where all data is already publicly available reports and articles.

GDPR HUNT was already focusing on GDPR before the GDPR law came, so it was unproblematic when the GDPR laws were established. However, it had quite large consequences for health research in terms of international collaboration. The different laws in different countries has made it problematic to pool data from different countries. Especially with countries outside of Europe and the US. A Scandinavian project they are working on have they solved the restrictions of data sharing with **pooling the results (meta-data analysis) instead of the initial data.** There has been a lot of fear in the beginning that one was doing research that was not allowed and that you would get really high financial fines. HUNT has a **two-way log in system for the cloud, making it extremely secure.**

The Health research law has something called **broad and extended consent** - where you agree that the data can be used for other research in the future than what it specifically intended to be used for

when collected. HUNT is based on this type of consent, otherwise it would not be possible to use the data. The regional ethical committee then approve all use of the data, that its within the consent. There has been some instances where they have gotten passive concent when needing to use a lot of blood samples for genetic research. When the data collected, genetic research was not a field. This has been extreamly valuable and have been done 2-3 times.

HUNT **data is stored nationally, but you can get access through a central server**, and the data will therefore not be exported. This way you **can shut down the access when you want** as well. You can do analysis of several countries through a central server. Still we do also have some projects where data is exported. An **advantage of GDPR** is that similar security systems implemented in several countries can possibly **make it easier to collaborate** as its the same laws that are in use.

"Another issue in Norway has been to link datasets from different registers in Norway. We have a unique possibility to link data because of the birth number of all the citizens (few countries have birth numbers). The infrastructure and quality of data in Norway is amazing, but if you are going to link the data it will take more than 1 year to get it done. This is now trying to be solved by a health-data-service where you get reply much sooner than today. They are making **a cloud system where you have available all the tools to make analysis in the cloud system**. It is supposed to be avaialable for governments, researchers and industry that want to do research, and also to the participants: Health analysis platform: E-direktoratet.

Ownership of the platform. 'Think about having a national (public) platform with data vs. commercial platform. It's important to own the infrastructure nationally to be able to have better control over the data, and then possible to buy infrastructure from commercial actors' (HUNT).

Intellectual property rights In the SafetyCube project some of the work was not original, so there were some issues in terms of intellectual property rights - this was solved with links to the original data. Also, in the case of S&TDL a key issue resulted often to be the intellectual property rights when asking researchers to share not only metadata but data.

Licence In terms of legal and regulation barriers, for ESPON they are mostly related to that some of their projects must pay a licence to use the data. Also, some discussion how to deal with aggregated data where there is different legislation in different countries.

Grant Agreement issue: HERMES was not allowed to commercialize the service in any way even just to keep the service alive, to get the basic funding to maintain the service. There was a term like the following on the Grant Agreement: "Access to data was granted on a not-for-profit use".

Licence and IP policy: In the case of RDA, the legal and ethical aspects around datasets and personal data are complex, irrespective whether the data set is small and relatively simple or very large with a great deal of sensitive information. In terms of licensing and IP policy they have a very open ones as reported in the following box.

RDA Output status, review, and approval.

Discussion Documents need no review. Anything an RDA member has rights to and feels advances the work of RDA may be posted on the open RDA web site (e.g. within an WG or IG space). Discussion Documents are available to the entire membership under a general Creative Commons Attribution

Only 4.0 license (CC-BY). Discussion Documents published on the RDA Web site need to have a value for and be maintained by a WG, IG, or other RDA group. Groups may define their own publishing criteria.

RDA Polices, Case Statements, and Charters are available under the default CC-BY 4.0 license. Formal processes are addressed in the Document Publication Policy.

Implementations are the responsibility of the person(s) or organization(s) hosting them. RDA does not maintain or formally endorse these products. They are, however, reviewed in the sense that they may have been proposed as part of a WG Case Statement, and TAB, Council, or the membership may have made suggestions to improve how the Implementations advance the work of RDA. Further, to foster open interoperability and wide-spread adoption, RDA strongly encourages that these Implementations be made available as openly as possible. RDA promotes the use and broader adoption of products that adhere to RDA principles. RDA generally recommends open source licenses from the BSD family or similarly permissive licenses where possible.

RDA Recommendations undergo formal review as endorsed Recommendations of RDA as defined in the Process and Criteria for RDA Recommendations . They are produced by WGs and IGs where robust consensus and transparency are basic principles. The Technical Advisory Board (TAB) shall have a chance to comment on draft versions and give advice on whether the content fits with the technical intentions of RDA. Finally, the Council needs to assert whether the document was developed and shared in accordance with the vision, principles, and processes of RDA. Recommendations must be open for public use and adoption. The default license should be the Creative Commons Attribution Only 4.0 license (CC-BY) or the Creative Commons CC0 1.0 Public Domain Waiver (CC0). Authors may chose either the CC-BY license or CC0 waiver, but in some cases other licenses or waivers may apply as defined in the Process and Criteria for RDA Recommendations. RDA strongly encourages users and contributors of RDA Recommendations to adhere to Norms for contributing to and using RDA products.

RDA does not hold patents. Anyone contributing to an RDA Recommendation must disclose any known patent or any known pending patent application they hold that may restrict the open use of the RDA Recommendation. If the patent holder does not allow unrestricted use of the patented material, the material may not be part of an RDA Recommendation.

5 Main findings and discussion

Based on the interviews we have gathered the information that is most important in order to make TOPOS sustainable. The main findings are presented below according to the same topics as in the

rest of the document: barriers related to resources, organization, technology, legal and information. Behaviour barriers are excluded for this section as this was not addressed in the interviews.

Resources. As described in chapter 4, there are several different financial solutions to choose from in order to make TOPOS sustainable:

- Public funding (international, national or regional)
- User-payment
  3) Paying for data
  4) Paying for infrastructure
- Stakeholders: Annual Organisational membership fees

Like TRIMIS and ESPON which originally steam from former EU project initiatives, BEOPEN could aim at getting continuously support from the European Commission. The advantage of having the European Commission as a funder is that it will be easier to sell in the concept of Open Science to the users. The commission is regarded as a trustworthy organization with a heavy influence on a broad geographical level, possibly making the users of TOPOS more interested in the platform as well. As stated in chapter 2, if desired to increase Open Science through "Obligation", this is more likely to happen if TOPOS is funded by the Commission. Where they could suggest that all data and relevant information from EU funded projects (and national projects if agreed upon by national funding agencies) must be made available for EOSC uploaded to their respective EOSC cloud system – TOPOS being the one for transport.

Having several funding partners on a national and or regional level like HUNT and S&TDL is also an option. However this will take quite a lot of work to get in place, especially since national or regional funders might not be as mature in terms of implementing Open Science as the European Commission, and therefore it might be more difficult to sell in the concept to potential funders. It is also more demanding to work towards several funders instead of one. TOPOS might also be unequally utilized in different countries and depend more on which nations/regions that are funding on the initiative.

Another possible strategy is to utilize HUNTs user payment. Where i.e. users of TOPOS data could use the data for free, unless they are making publications with the data in papers. However, this could potentially cause less use of TOPOS data, as not all countries or institutions would necessarily have the financial aids to pay for such a fee. This would then again lead to less equality among researchers – working against one of the main advantages of Open Science. On the other hand, if the data is available for a lower price than what it cost to do own data collection, this could be a win-win for TOPOS and the institutions using the data. However, you would still miss out on other potential individual users – and not taking advantage of citizen science where individuals make new solutions based on free available data. Therefore, this model might not be the best suited for increasing Open Science.

The other solution with paying for infrastructure for storing your own data safely and securely could be possible idea for TOPOS. At the current state there is no secure log-on system at TOPOS as there is at the HUNT Cloud. However, as GDPR issues are becoming a more and more important for all research organizations – the need for safe storage is something TOPOS could potentially gain a competitive advantage if offering this kind of system in the future. This way research organizations

could store their data in TOPOS securely, and therefor also increase the chance for them to share data that is possible to share with other users of TOPOS. In order to have as much as possible open, this system for sensitive data could be a separate part of TOPOS.

A last possible financial solution is the organisational fees from users, as the RDA system:

- *organisational members* contributing with an annual subscription fee based on the number of employees – for TOPOS this could potentially be members of the BEOPEN project, universities, industry stakeholders etc.
- *supporters* by contributing through an annual financial contribution to be agreed directly with TOPOS.
- *sponsorships* Sponsoring activities like meetings, projects, training courses etc.
- *regional members* As the RDA also has regional funding their regional members also pay a yearly contribution. For TOPOS this could also be other types of funders, depending on what funding plan is chosen.

This system might be less reliable in terms of cashflow, as it is dependent on the number of members and yearly sponsorship. Also, if the data is available to all, but just a couple of members pay for it, the unequal playing field that is mentioned as a negative consequence of Open Science (D5.3) might come into fruition.

Organizational benefits of open science and sharing of data is mentioned by several of the interviewed. The sharing culture is based upon:

- Status: Benefits for the organization creating open data/infrastructure in terms of (1) becoming famous for their data/systems etc. which (2) in turns generates more projects and project invitations
- Based their financial model on doing open science – using it as a benefit when applying for funding
- Societal benefit and altruistic motives:
    - better road safety
    - better public health
    - supporting research and identifying research gaps
    - stimulate collaboration among different stakeholders
    - reduce duplication of work
    - value if made seamlessly available to use to create new knowledge
    - achieving an interoperable data infrastructure is a significant challenge but on the other hand also a great enabler
    - benefits not just in science but across the economy by, for instance, increasing the efficiency and effectiveness of businesses or allowing for new products and services identification
- More data available than its possible to do research on
- Some fields are drivers for open science (genetic research mentioned in medicine, climate research mentioned in WP 2).

The fact that Open Science is such a new paradigm, participating in Open Science at an early stage could potentially give benefits in terms of competitive advantage – especially if this turns out to be the new normal of science. In relation to TOPOS sustainability, this should be used as a selling point to get different stakeholders to use and potentially fund TOPOS. The altruistic motives behind supporting Open Science mention by the interviewed cases is something that is recognisable from the impact study in BE OPEN Deliverable 5.3 where reducing duplication of work, creating better solutions, increase collaboration and increased efficacy and economic gains was also mentioned as the main beneficiary impacts open science would have on the future.

Depending on how willing stakeholders outside research (public transport operators, governments, teleoperators etc.) are on sharing data – the transport research community could also end up in the situation mentioned by HUNT – the problem is not lack of data, but lack of researchers utilizing the data. This way the loss of competitive advantage might not be of such a great concern for researchers. However, as also stated in BE OPEN Deliverable 5.3 – if there is an unequal playing field in terms of who shares their data, there is potentially conflicts related to sharing. If TOPOS is able to offer a pool of data available, sharing might also be easier for others, as there is already available data for them to use as well. In order to do this, TOPOS could focus on getting data from public stakeholders first, as they have less to lose in terms of competitive advantage when sharing data. At least at a lower geographical level.

Information focused on both data quality and external information spreading.

Data quality

When it comes to assuring the quality of the data, none of the interviewed candidates had the same system as TOPOS in terms of who is responsible for uploading. In TOPOS the individual researchers upload their data, while the other interviewed organizations have either an external service that uploads data, or most is done internally to ensure the quality. But basically, all the interviewed organizations do extensive quality control, one way or the other. Even though TOPOS works differently, there is still a need, and probably also a larger risk that the quality of data might not be good if it is not going through a quality control organ. HUNT also have some data in their system that is uploaded from external sources, however, they are not allowed to upload data without it going through a quality check first. To ensure that researchers will use TOPOS as their main source of transport related data/information, the quality of the data is crucial. Therefore, funding for a quality control organ should also be put in place to make TOPOS sustainable.

Both ESPON and HUNT talk a lot about harmonizing data and having proper metadata that follows the data. This part will definitely be important to follow up from a quality control organ, as individual researcher will have quite different knowledge about this. TRIMIS also mentions language barriers in terms of gaining access to all types of projects that are being executed. TOPOS should also consider how to solve these types of issue. HUNT actually charge for doing the quality control before external enter their data in the database – but this is possible because the paying organization also get a safe storage for their sensitive data.

Awareness and marketing

As mentioned by the projects there are several different ways of *measuring user awareness* among the projects/organisations:

- Google views and website visits
- Projects based on data
- Webinar attendance
- Members

In terms of *increasing the awareness, the marketing strategies* the projects/organisations have chosen are also quite broad:

- Targeted Google searches
- External linking form other established organisations
- Newsletters
- Webinars, seminars, workshops and other events
- User feedback during the project
- Proactive media participation
- Paper promotion – obligation to use the name of the organization on the paper heading
- Merch
- Social Media: Particularly Twitter and LinkedIn
- Utilizing national networks:
- Domain / discipline specific ambassadors

A lot of these dissemination strategies has also been used by TOPOS/the BEOPEN project through the project period – however to build TOPOS to a sustainable, functional cloud system for sharing data – there is definitely a need to continue the dissemination activity. However financial support is needed for this, as the only free dissemination strategy from above is *external linking form other established organisations*. A possible solution could be to establish a collaboration with TRIMIS, as they are working with similar activities, and already have a successful establishment.

Laws and regulations

Legal barriers for TOPOS, depends on what type of data that will be shared. If it will only share already public data like TRIMIS, GDPR will not be a concern. However, TOPOS could actually take advantage of the increasing GDPR focus. As mentioned, HUNT has a two-way log in system for the cloud, making it extremely secure. This system they are making available for others to use, if paying for the storage. They are also making a cloud system where you have available all the tools to make analysis in the cloud system. If TOPOS is able to offer such a system to users, they could also get financial income from making their system available for other institutions. It would be much more expensive for each individual institution to create such a system, so potentially this could be a win-win situation for both TOPOS and the organizations using the system.

Technology

Based on the input on technology it is important for TOPOS to consider the following:

- Making sure to have clear explanations for the meta-data that is entered into the cloud system
- Making sure to have clear explanations to external people working on technical solutions
- Make sure the search functions and keywords are relevant for the end user
- It can be difficult to find the right resources for ITs

An additional feature that TOPOS could take advantage of is the use of a Semantic Portal to improve the quality of communication between the information provider and the user. This translates in the possibility to make a multidimensional search so allowing for multiple views of the same data. This could also support the self-sustainability of the portal since in a semantic web approach supplying groups host their data and the portal becomes an aggregating service so ensuring continued access to the data without the potential bottleneck deriving from a central organization.

6 Conclusions

In this document we have identified how to solve specific barriers to make TOPOS sustainable in terms of **1) actual use** and **2) financial funding**. To do this, we have looked on six different types of barriers related to innovation research: resources, behavioural barriers, organizational barriers, information barriers, technology barriers, and legal barriers. The last three was the scope of the GA, but in order to make TOPOS sustainable we also interviewed the organizations about the first three, as these barriers were identified in former WPs as well as in research literature.

The main findings are presented in the table below:

Barrier	Sustainability measure
Resources	There are mainly four different financial models TOPOS possibly could undertake: • Public funding (international, national or regional) • User-payment 1) Paying for data 2) Paying for infrastructure • Stakeholders: Annual Organisational membership fees The advantage of international public funding (EU) is that the commission is regarded as a trustworthy organization with a heavy influence on a broad geographical level, possibly making the users of TOPOS more interested in the platform. If wanting to increase Open Science through “obligation”, this is more likely to manage if TOPOS is funded by the European Commission.
Organizational	To make TOPOS sustainable a sharing culture is also needed. If TOPOS can offer a pool of data available, sharing might also be easier for others, as there is already available data for them to use as well. In order to do this, TOPOS could focus on getting data from

	public stakeholders first, as they have less to lose in terms of competitive advantage when sharing data. At least at a lower geographical level. Also, demonstrating that participating in Open Science at an early stage could potentially give benefits in terms of competitive advantage. Some organisations benefit from creating open data/infrastructure in terms of (1) becoming famous for their data/systems etc. which (2) in turns generates more projects and project invitations.
Information	Funding for a quality control organ and continuous dissemination is needed to make TOPOS sustainable. Possibly corporate further with TRIMIS.
Legal	Take advantage of the increased focus on GDPR, TOPOS could be offering a two-way log in system for parts of the cloud, making it extremely secure. Making this available to use for sensitive data, if user is paying for storage.
Technological	TOPOS could take advantage of the use of a Semantic Portal to improve the quality of communication between the information provider and the user and supporting self-sustainability

7 REFERENCES

Adesiyun, Adewole, Tromaras, Akiviadis, Almeras, Caroline, Antunes, António, Antunes, Maria de Lurdes, Barateiro José, Erdelean, Isabella, Jasiūnienė, Vilma, Manola, Natalia, Mavropoulou, Eleni, Milenkovic, Milos, Nielsen, Anja Fleten, Olsen, Silvia, Paliukaitė, Migle, Pereira, Ana, Prändl Zika, Veronika and Kornelija Ratkevičiūtė. (2019). Open/FAIR data, software and infrastructure in European transport research. BEOPEN project D 2.2. https://beopen-project.eu/storage/files/beopen-d22-open-fair-data-software-and-infrastructure-ineuropean-transport-research.pdf

de Jesus, A & Mendonça, S. (2018); "Lost in Transition? Drivers and Barriers in the Eco-innovation Road to the Circular Economy", *Ecological Economics*, 145: 75-89.

Gohoungodji, P., N'Dri, A. B., Latulippe, J-M., Barreto Matos, A. L. (2020); "What is stopping the automotive industry from going green? A systematic review of barriers to green innovation in the automotive industry". *Journal of Cleaner Production* 277 (2020) 123524

Himanshu Guptaa, H., Kusi-Sarpongb, S. & Rezaei, J. (2020); "Barriers and overcoming strategies to supply chain sustainability innovation", Resources, Conservation & Recycling 161 (2020) 104819

Nielsen, A. F., Anagnostopoulou A., Tozzi M., Angelopoulos N., Tveit A.K., Tromaras A., Lindberg G. (2020). Governance and operational models. BEOPEN D2.4 https://beopen-project.eu/storage/files/beopen-d24-governance-and-operational-models.pdf

Reynolds, D., P. Shabajee, S. Cayzer and D. Steer. “2.1.7 Semantic Portals Demonstrator -lessons Learnt Swad-europe Deliverable 12.1.7: Semantic Portals Demonstrator-lessons Learnt Contents.”

Yannis, George and Katerina Folla. 2019. European forum and oBsErvatory for OPEN science in transport Project. Open access publications and the performance of the European transport research. BEOPEN project D 2.1. Available at: https://beopen-project.eu/storage/files/beopen-d21-open-access-publications-and-the-performance-of-the-european-transport-research.pdf

8 ANNEX I

Interview guide:

Introduction

1. Tell me in brief how your [database, webpage, network etc.] started and which is the main aim of it. What was the initiative behind it? (tell me when it started, who took initiative to it (project, government etc.))
2. Is your [database, webpage, network etc.] open – to what extent, and how is it related to open science?
3. What were your main barriers in the start-up phase? Have these barriers been overcome?

Information & data barriers:

4. What type of data-related barriers have you had to consider during the process? (Quality, others?)
    a. Have you had any issues related to quality of data?
    b. How do you ensure the quality of the data?
    c. How do you solve issues related to maintenance of data etc. (continuous updating – automatic updates, resources)?
    d. Have you had any issues related to data interoperability (dealing with different formats)?

5. Have you had issues related to awareness/knowledge of users?
    e. How many people use your service (individuals or organizations)?
    f. Any problems related to getting people to use the service, how is this solved? (problem at the organization or problem on user-side?)
    g. And have you done any marketing of the service? (How?)

Organizational barriers:

6. Have you encountered any issues regarding a lack of dialogue between data providers and re-users?
    a. The lack of dialogue with the users
    b. The lack of information about the updates of already opened datasets
    c. The lack of information about the future datasets to be opened

7. Have you had any issues in terms of competitiveness (related to sharing data)?
    a. If not: why?
    b. If: how do you solve this?
8. Have you had any advantages when sharing data open? What positive effects are generated?
9. Have you any differences of access depending on organizations using your data (researches vs. others)?

Technical barriers

10. What type of technical issues/barriers/obstacles have you had to consider during the process?

Law and regulations

11. What type of legal issues/barriers/obstacles have you had to consider during the process?
    a. Have you had any issues related to GDPR? How is this solved? If not: why is it not a problem?
    b. Any other legal barriers?

Recourses

12. How was it initially funded?
13. How did you choose your funding plan?
14. How is it funded today?
15. Do you get any funding by the government or do you have members that fund the initiative?
16. Or other solutions? (pay-by-use)