

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)**

D1.2 Open Science framework, terminology and instruments

Final

Deliverable Title:	D1.2 Open Science framework, terminology and instruments
Work Package:	WP1: Focused objectives of key actors
Due Date:	M7
Submission Date:	M8
Start Date of Project:	01/01/2019
Duration of Project:	30 Months
Organisation Responsible of Deliverable:	German Aerospace Centre (DLR)
Version:	1.0
Status:	Final
Author name(s):	Kristel Palts (DLR), Kolja Kindler (DLR), Elena Campillo Borrás (DLR)
Reviewer(s):	Anja Fleten Nielsen (TOI)
Nature:	<input checked="" type="checkbox"/> R – Report <input type="checkbox"/> P – Prototype <input type="checkbox"/> D – Demonstrator <input type="checkbox"/> O - Other
Dissemination level:	<input checked="" type="checkbox"/> PU - Public <input type="checkbox"/> CO - Confidential, only for members of the consortium (including the Commission) <input type="checkbox"/> RE - Restricted to a group specified by the consortium (including the Commission Services)

Document history			
Version	Date	Modified by (author/partner)	Comments
0.1	02/07/2019	Kristel Palts (DLR) Kolja Kindler (DLR)	Table of contents and first content of layout
0.2	23/07/2019	Kristel Palts (DLR) All task 1.2 partners	Comments for the draft version implemented
0.3	24/07/2019	Kristel Palts (DLR) Elena Campillo Borrás (DLR)	Revised structure and first proposed draft
0.7	31/07/2019	All task 1.2 members Afroditi Anagnostopoulou (CERTH)	Feedback from Partners
0.8	05/08/2019	Kristel Palts (DLR) Elena Campillo Borrás (DLR)	Complete Revised structure and partner comments implemented
0.9	21/08/2019	Anja Flete Nielsen (TOI) Afroditi Anagnostopoulou (CERTH) Kristel Palts (DLR) Elena Campillo Borrás (DLR)	Implementing comments and preparation of final draft
1.0	26/08/2019	Kristel Palts (DLR) Elena Campillo Borrás (DLR)	Final version

Contents

List of Figures.....	5
List of Tables.....	6
Abbreviations	7
Executive summary	9
1 Introduction.....	11
1.1 Purpose of the document.....	11
1.2 The approach.....	12
2 Transport research	14
2.1 Transport modes	15
2.2 Competence areas and main actors.....	16
3 Mapping of current situation	18
3.1 Open Science terminology	18
3.2 Stakeholder-centred analysis	19
3.2.1 Stakeholders experience with Open Science	23
3.2.2 Mode specific study.....	25
3.2.3 Open Innovation.....	32
3.2.4 Derived Challenges concerning Open Science/ Open Data.....	33
3.2.5 Conclusion	36
3.3 Main challenges.....	36
4 Open Science and transport research	39
4.1 Open Science Platforms	42
4.2 Open Software Sources.....	48
4.2.1 Open Internet	49
4.3 Open Science Models.....	50
5 Common understanding of Open Science in Transport sector	52
5.1 Framework	53
5.2 Metrics.....	59
6 Main Outcomes	66
6.1 Potential for implementation in transport research.....	67
7 References.....	69
8 Annex – BE OPEN Survey.....	73

List of Figures

Figure 1: Workflow of approach.....	12
Figure 2: Main inputs to mapping the situation.....	16
Figure 3: Type of organization.....	20
Figure 4: Competence areas.....	21
Figure 5: Age range of the participants.....	21
Figure 6: Experience level of the participants (years of experience).....	22
Figure 7: Modes of transport participating in the survey	22
Figure 8 Countries of organizations participated in the survey	23
Figure 9: Stakeholders awareness for Open Science	24
Figure 10 Level of usage of Open Science	24
Figure 11: Awareness for Open Science in waterborne transport.....	25
Figure 12: Usage of Open Science in air transport.....	26
Figure 13: Usage of Open Science in waterborne transport mode.....	26
Figure 14: Awareness of Open Data in multimodal transport	27
Figure 15: Usage of Open Data in air transport	27
Figure 16: Usage of Open data in waterborne transport.....	28
Figure 17: Usage of Open Software in waterborne transport mode	28
Figure 18: Awareness of Open Access publication in air transport mode	29
Figure 19: Usage of Open Access Publications in road transport mode	29
Figure 20: Usage of Open Data	30
Figure 21: Usage of data based on mode of transport	30
Figure 22: Result of data analyses in different sectors	31
Figure 23: Result of process data in different sectors.....	31
Figure 24: Open and Closed innovation	33
Figure 25: Main type of Open data used.....	33
Figure 26: Usage of Data platforms.....	34
Figure 27: Possibilities for Open Science.....	34
Figure 28 Stakeholders needs for sharing their experience.....	35
Figure 29: Schematic of framework inputs	39
Figure 30. Research cycle (21).....	40
Figure 31: Difference between current and next generation repository (23)	43
Figure 32: Structural scheme of EOSC.....	45
Figure 33. Framework inputs and requirements.....	52
Figure 34: Proposed framework structure	54
Figure 35. Efficiency measurements in transportation systems (55).....	61
Figure 36: Cost evaluation hierarchy (55)	61
Figure 37: Typical topology of IoT systems (57)	62
Figure 38: Basket of metrics for evaluation of science (46).....	63



List of Tables

Table 1: Stakeholders influenced based on competence areas.....	17
Table 2: Scientific sources in Transport research.....	42
Table 3: Review of diverse Open Science sources by competence areas and stakeholders	48
Table 4: Transport metrics in relation to competence areas (54)	60
Table 5: Open Science usage in Member States	65
Table 6: Classification of Open Science sources according to their potential impact	68

Abbreviations

ACARE	Advisory Council for Aeronautics Research in Europe
API	Application Programming Interface
C-ITS	Cooperative – Intelligent Transport Systems
CAT	Connected and Automated Transport
COAR	Confederation of Open Access Repositories
CORDIS	Community Research and Development Information Service
CPU	Central Processing Unit
CSA	Coordination and Support Action
DOAJ	Directory of Open Access Journals
DORA	Declaration On Research Assessment
EC	European Commission
EOCS	European Open Science Cloud
EPATS	European Personal Air Transportation System
EREA	Association of European Research Establishments in Aeronautics
ERTICO	European Road Transport Telematics Implementation Coordination
EU ODP	EU Open Data Portal
FAIR	Findable, Accessible, Interoperable, Re-usable
FOT	Field Operational Tests
FP	Framework Programme
GHG	Green House Gas
GIPO	Global Internet Policy Observatory.
GPS	Global Positioning System
H2020	Horizon 2020
IETF	Internet Engineering Task Force

IGF	Internet Governance Forum
IoT	Internet of Things
IP	Intellectual Property
IT	Information Technology
ITF	International Transport Forum
KPI	Key Performance Indicator
LeMO	Leveraging Big Data for managing Transport Operations
LOADB	Listing of Open Access DataBases
NDS	Naturalistic Driving Studies
NGO	Non-Governmental Organisation
NTM	Network and Traffic Management
ODP	Open Data Portal
PID	Persistent IDentifiers
RDA	Research Data Alliance
STRIA	Strategic Transport Research Agenda
TOPOS	Transport fOrum/Observatory for Promoting Open Science
TRC	Transport Research Cloud
TRIMIS	Transport Research and Innovation Monitoring and Information System
UDRIVE	European naturalistic Driving and Riding for Infrastructure and Vehicle Safety and the Environment
WoS	Web of Science

Executive summary

The main objectives of the BE OPEN project are to create a framework of common understanding among all stakeholders, to assess the practical impact of Open Science and to identify and put in place the mechanisms to implement it in the field of transport research. The main goal of Task 1.2 is to develop a framework to enable common understanding for Open Science in Transport Research. Task 1.2 is complementary to Tasks 1.1 (1) and 2.1 (2), taking into consideration defined competence areas, stakeholders and scientific sources and evaluating the information in terms of the practical impact of Open Science in transport research.

Based on the different modes of transport, their characteristics and the nature of research in the field of transport, a set of challenges will be identified regarding the competence areas and stakeholders defined by the aforementioned tasks. Such challenges will bring out the particularly critical topics that need to be addressed by the framework. The identified main topics will then build up the structure of the proposed framework, which will also need to comply with a set of requirements intrinsic to Open Science. Finally, different Open Science sources will be evaluated to assess their potential for implementation in transport research.

For creating a common understanding of Open Science in the area of transport research, the following steps need to be taken:

1. Evaluate transport research in terms of transport modes, competence areas, stakeholders and their interrelationship
2. Map the current situation, defining the appropriate terminology to ensure common understanding, evaluating the stakeholder experience and needs through a survey and identifying the main challenges
3. Evaluate Open Science sources and their potential impact on transport research
4. Extrapolate the main relevant topics to be addressed by the framework and propose a structure compliant with the key features of Open Science
5. Capture the impact of the implementation of Open Science in transport research with metrics

To evaluate transport research, the categorisation of competence areas and stakeholder identification from Task 1.1 (1) are used.

The current situation of Open Science and transport research is mapped. Furthermore, a unified terminology is defined in order to enable common understanding. By means of a survey, the knowledge of stakeholders from different transport modes is evaluated, analysing their experience and needs for Open Science, and corresponding challenges in agreement with the literature are identified. This information will be used to assess which are the main topics to be considered within the structure of the framework, which should create an optimal research environment, overcoming these challenges.

The following challenges have been highlighted:

- Fragmentation of data and large variety of stakeholders
- Enhancement of data security and privacy principles
- Data quality
- Lack of skilled experts
- Technological issues
- Legal issues
- Funding

Based on the different transport modes, the key competence areas, main actors and challenges identified, a structure is proposed to create a framework featuring collaboration, transparency, univocal referencing, controllability of access and defined analytics to capture the impact. Collaboration and transparency are actively sought to nurture high-quality research and best practice implementation. Univocal referencing enables sharing, yielding to numerous benefits - ranging from accessibility of data to repeatability of studies. Controllability of access is key when dealing with sensitive data, especially when considering Cloud platforms, and can be enforced using data security and privacy principles complementing Open Access. Key Performance Indicators (KPIs) and metrics will finally be used to capture the impact of the implementation of Open Science in transport research.

The main outcomes of Task 2.1 are stated in the final section. These outcomes could be considered as input for the prospective work of Task 3.2 “Set up of the European Open Science in Transport Forum”, Task 3.3 “Set up of Open Science in Transport Observatory”, Task 5.1 “Identification of main challenges, opportunities, constraints and bottlenecks of Open Science in transport research” and 5.2 “KPIs for Open Science in transport evaluation”. Moreover, the outcomes of the survey can be used as a starting point for Task 1.3 “Transport stakeholder needs and objectives”.

1 Introduction

1.1 Purpose of the document

The BE OPEN project is allocated within the Horizon 2020 research and innovation programme as a coordination and support action (CSA). The project strives to assess the practical impact of Open Science in the field of transport research and the key features required for its proper implementation. This project is characterised by a holistic approach to the subject focused on the importance of common understanding. Its main objectives are as follows (3):

- To develop a framework in order to establish a common understanding of operationalizing Open Science in Transport
- To map existing Open Science resources and see how transport research fits in
- To facilitate an evidence-based dialogue to promote and establish Open Science in transport
- To provide the policy framework and guidance for open science implementation in transport
- To engage a broad range of stakeholders in a participatory process for Open Science uptake

Task 1.2 falls therefore within the first target. The main goal of Task 1.2 is to enable common understanding of Open Science in the transport research area. Task 1.2 builds on the previous work of Tasks 1.1 (1) and 2.1 (2) by defining a univocal terminology and assessing its related research and scientific production within the field. Ultimately, processes, practices, infrastructures and technologies are mapped and challenges and opportunities for intervention for Open Science practices identified. This deliverable aims to specifically produce a grid of guidelines that provide a structured definition of general transport context, facilitating the general framing of scientific research and analysing their practices and expectations in implementing Open Science in transport research.

The first chapter of this Deliverable defines its goals and ambitions to provide a clear picture of the approach taken to fulfil them. The second section refers to Tasks 1.1 (1) and 2.1 (2) to introduce the different modes of transport, the main stakeholders involved and competence areas identified. This information is used to map the current situation of Open Science in transport research. To ensure common understanding, key concepts are defined within the terminology. Furthermore, the experience and needs of the main actors are assessed through a survey to better define the situation. The main challenges to be addressed by the framework are chosen based on the findings from the literature review and the survey. Bearing in mind the aforementioned challenges, different Open Science platforms, models and software sources are analysed to evaluate their potential impact on the field of European transport research. Using the requirements of the main actors and competence areas as input, a structure for a framework is proposed. Such a framework should too be compliant with the intrinsic features that characterise Open Science, i.e. transparency, collaboration, among others. The assessed platforms are compared to identify their best potential to overcome the challenges and are allocated within the framework following that criterion. To estimate the effectiveness of Open Science in the field of transport research a set of Key Performance Indicators (KPIs) also needs to be defined. The use of metrics will finally capture the potential for implementation.

1.2 The approach

The implementation of our approach can be well defined quoting the words of the European Commission: *“As open as possible, as closed as necessary”*. (4)

A framework of common understanding could improve communication between the different stakeholders, facilitating the use of Open Science in the transport sector. To achieve this, it is important to identify the main objectives and the requirements of the key categories of actors. It will compile common principles, definitions and technical data made available for research. Using adequate methodologies, the effectiveness of Open Science will be optimised through the implementation of best practices and modern technologies.

A framework can be considered as a workspace within which projects or research can be undertaken. The characteristics of such a workspace would then depend on the users and their needs and expectations (5). Therefore, in order to create a suitable framework for the application of Open Science in transport research, the different modes of transport should be assessed to identify who the stakeholders are and which are the competence areas within which their research is comprised. The work in that area builds on the previous work of Tasks 1.1 (1) and 2.1 (2). During that assessment, different challenges will be identified based on the stakeholders’ needs and expectations extrapolated from the survey, the different competence areas and the characteristics of Open Science. The framework should provide tools to overcome such challenges. These tools will be grouped into topics that need to be addressed and compared to existing platforms and models to assess the potential for implementation. The resulting framework should be collaborative, transparent, allow controlled access, create univocal referencing and include analytics to capture the impact.

Figure 1 shows a scheme of the approach that has been devised for this Deliverable 2.1.

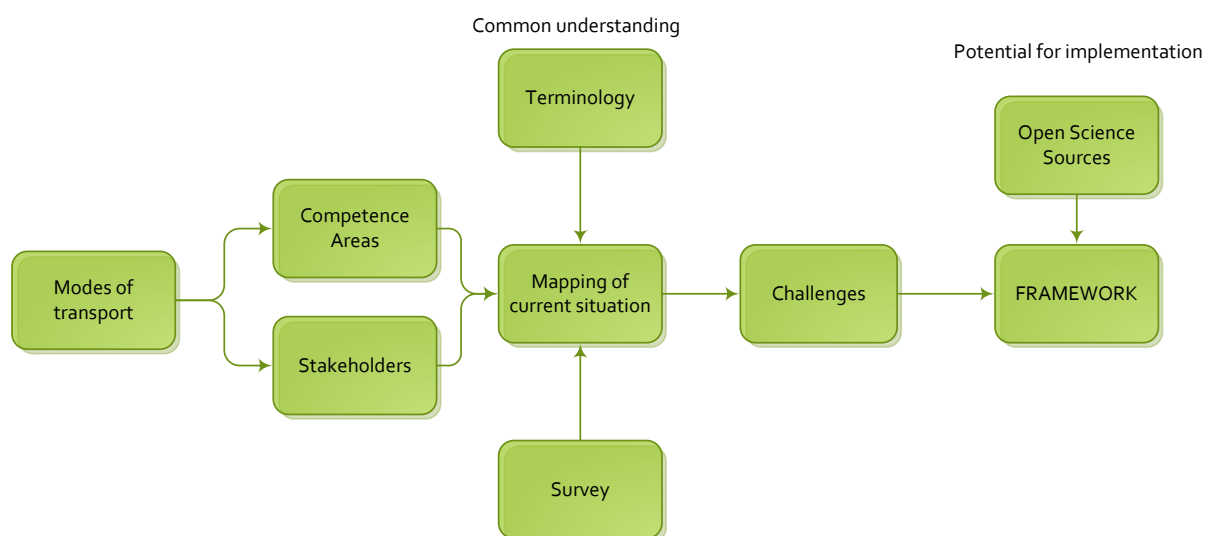


Figure 1: Workflow of approach

To this end, the following steps need to be taken:

1. Evaluate transport research in terms of transport modes, competence areas, stakeholders and their interrelationship
2. Map the current situation, defining the appropriate terminology to ensure common understanding, evaluating the stakeholder experience and needs through a survey and identifying the main challenges
3. Evaluate Open Science sources and their potential impact on transport research
4. Extrapolate the main relevant topics to be addressed by the framework and propose a structure compliant with the key features of Open Science
5. Capture the impact of the implementation of Open Science in transport research with metrics

In the research field, the BE OPEN project mainly focuses on the major influential research organisations such as ECTRI, EURNEX, EATEO, WEGEMT, EUROCONTROL, UITP and VTT. Based on the analysis provided by Deliverable 1.1 (1), stakeholders of transport research could be further classified as primary and secondary; primary referring to those groups directly affected and secondary to those indirectly affected. Representatives of primary stakeholders participated in a survey conducted within Task 1.2 in an attempt to identify their needs, motivators and requirements of using Open Science in transport research.

Moreover, the following competence areas (defined in Deliverable 1.1 (1)) were considered within this survey in order to map the practical impact of Open Science and define how it is used in the field of transport research:

- Business modelling
- Environmental
- Legal/ Regulatory
- Socio-economic
- Technological
- Transport planning

The derived results of the survey are presented and analysed in Section 3 of this report and provide the basis of developing the framework of common understanding of Open Science in transport research described in Section 5.

2 Transport research

Technological innovation, political decisions and new mobility concepts as well as social and economic trends lead to a system transformation. Within the transport sector, research is being undertaken on, for example, CO₂ reductions, efficiency optimisation and propulsion electrification, but the main goal is to develop a sustainable mobility system. The transport sector is a fast growing area where future transport systems will differ from those of today. That is can already be noted, for example, in the field of Urban Air mobility and data atomization, which means new technologies and big datasets are being created. The usage of Big Data and innovative tools for data management become more and more important. (6)

The European Commission project Mobility 4 EU – European Action Plan for user-centric and cross-modal transport in 2030, has listed six main action areas: Low/zero emission mobility; automation and connected transport; Safety and (cyber)security in transport; mobility planning and cross-modal/cross-border transport and integration of novel mobility services in public transport and inclusion- putting the user in the centre. For each area roadmap, recommendations for researchers, innovation and implementation in transport were proposed, of technical as well as non-technical nature. With innovative transport solutions across all modes for reducing the environmental footprint and making better use of existing transport network, stakeholders – industry, operators, policy makers, users etc. from different modes of transport have to work tightly together to make better use of the transport data produced and shared. To ease the societal challenges for future transport demand and supply automation, disruptive technologies and models need to be implemented. Huge amounts of datasets are being created through the use of new technologies and automation - connected and cooperative driving, smart infrastructures (c-ITS) and smart traffic management systems.

The transport research field has to be adapted to the upcoming requirements. That also means that the Research Institutions and Industry together with Regulatory units from all transport modes have to work tightly together for innovation in the field of transport. Using Open Science also includes Open Reproducible Research, where the main idea is to practice Open Science to offer users free access to experimental elements for research reproduction (7). Transport data is faced with some challenges to use Open Reproducible Research. With automation and use of different IT solutions there is a growing importance of cyber security (8). A legal framework at national as well as international level needs to be agreed on. Furthermore, the lack of stakeholders' commitment to share the data and research and the need for agreement on the specification and availability of quality data need to be addressed.

The Open Science concept is gaining popularity and importance in the whole world. **Open Science** is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society. Europe has the culture and ability to share research activities across national boundaries, which along with its remarkable research and knowledge base place it in a leading position in the world to promote and expedite the new Open Science way of working. It includes Open Access, Open Data, Open Reproducible Research, Open Science Policies and Open Science tools. (7) In fact, nearly 10% of the budget allocated for the Horizon 2020 Work Programme 2018-2020 (9) is channelled to direct or indirect support of Open Science and towards this direction, the European

Open Science Cloud (EOSC) (10) has been developed to enable sharing and re-using research data across disciplines and borders, taking into account relevant legal, security and privacy issues.

In the context of transport research, the BE OPEN project aims to create a common understanding for Open Science by analysing work methods, tools and available information.

2.1 Transport modes

Transport research includes all modes of transport for passenger as well as freight transport. For harmonized definitions it is important to agree on the categorisation of transport modes. In the International Transport Forum (ITF), the Glossary for Transport statistics, the following classification of transport modes is used: Railway Transport, Road Transport, Inland Waterway Transport, Pipelines (Oil and Gas) Transport, Maritime Transport, Air Transport and Intermodal Freight Transport. (11)

Key Research Innovation pathways, as defined in STRIA roadmap, are: Active management of CAT technologies, user and societal acceptance, socio economic impacts, environmental and climate impacts, human machine interface, innovative hybrid technologies, cybersecurity and data protection, ICT infrastructures, optimised use of the Internet of Things, data and governance.

- Connected and automated transport (CAT) is an important area of digital technology that promises a number of benefits for the individual, the society and the economy in the various fields of transportation. In road transport, it can provide more safety, better social inclusion and higher efficiency; in railways, it enhances the performance of the overall system, including train operations, traffic management, and maintenance. It creates opportunities for new mobility services; in waterborne, it can improve the safety of shipping and the efficiency of transport and logistics as well as benefit the environment; while in aviation, where it has been implemented for long time with the autopilot being the most prominent example, even higher levels of automation are aimed for. CAT supports the competitiveness of the European transport manufacturing, telecom and IT industries on worldwide markets, and enables potentially disruptive innovation, which may lead to new services concerning the transportation of both people and freight. There are benefits for all transport modes and is therefore important for all the transport modes.

Within TRIMIS the following general challenges towards CAT have been identified: development of technology for hard- and software, vehicle, infrastructure, data communication and decision-making levels, including validation for the technologies and the real-world testing (12). Mode-specific challenges include:

- a) Road: drivers, passenger, and other road users have to be well understood and appropriately addressed. Proper business and operational models need to be derived.
- b) Railways: long-life expectancy of rolling stock and infrastructures, the difference in legacy systems and diversity of operational rules in various European countries could slow down the connected – automated system.

- c) Waterborne: the vessels are large, with high inertia in a comparatively slow-moving environment. Recently there have been introduced new automated systems but it is moving in slow pace.
- d) Air: the different sets of national rules and regulations, funding issues and also the state sovereignty make it difficult to harmonise the air traffic networks across Europe. Future implementation of Unmanned Aerial Systems yields further issues concerning the airspace management. The Single European Sky initiative is already underway.

Network and traffic management systems (NTM) are used for optimisation and management of transport networks operations. The defined bottlenecks across the modes of transport could cause traffic jams, increased pollutant emissions and have an environmental impact. The aim, described in the STRIA roadmap, is to develop and advance multimodal transport system. Making data publicly available could encourage off-peaks traffic, alternative routes through intelligent applications and user information services. Integrated urban traffic management and mobility information systems can therefore contribute to optimising transport flows through cities as well as rural areas.

The markets for different modes of transport are different, the waterborne sector, for example, is a wait-and-see philosophy in road transportation new technologies could be implemented faster. Lots of new technologies are also being developed for Air transport but the deployment is relatively slow due to strict regulations and certification processes.

2.2 Competence areas and main actors

In order to build up a suitable framework for the implementation of Open Science in transport research, the needs and expectations of the different stakeholders need to be observed and the competence areas within the field considered. Using this information as input towards building the structure of the framework will grant its functionality for the application.

Figure 2 shows a schematic of the main inputs considered to map the current situation.

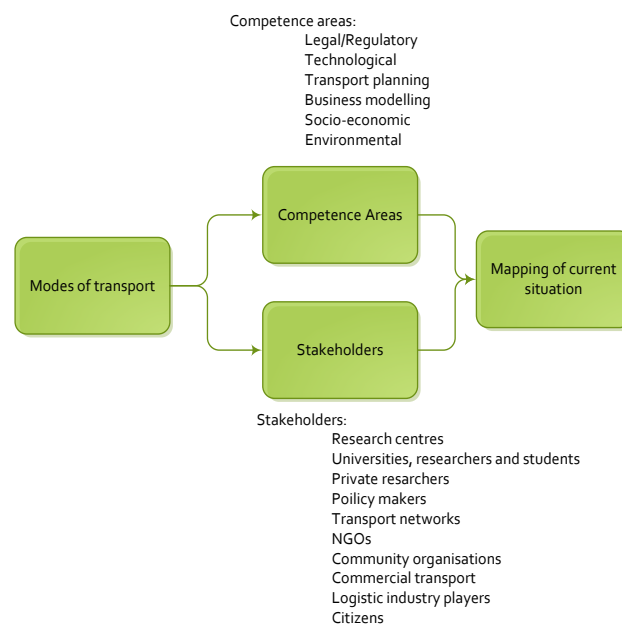


Figure 2: Main inputs to mapping the situation

The transport sector appears to be formed of six main competence areas, defined in the BE OPEN project, Deliverable 1.1 (1):

1. **Legal/ Regulatory:** regulations, action plans, policy instruments and incentives
2. **Technological:** new technologies and innovation applied to transport for technological improvements
3. **Transport planning:** sustainable transport vision for reducing impact to the environment, economy and society
4. **Business modelling:** description of means and methods a company applies to earn revenues projected in its plans
5. **Socio-economic:** accessibility of transport and basic services
6. **Environmental:** delivering substantial socioeconomic benefits and at the same time lowering impacts on the environmental system.

Furthermore, the following stakeholders were identified: research centres, universities, researcher and students, private researcher, policy maker, transport networks, NGOs, community organisations, commercial transport, logistic industry players and citizens. Task 1.1 (1) provides an analysis of primary stakeholders (i.e. stakeholders who are directly influenced) and secondary stakeholders (i.e. stakeholders who are indirectly influenced by the topic/ project/ strategy/ regulation in question in a specific competent area). In light of these aspects, Table 1 presents which competence area of the transport research sector affects either primary or secondary stakeholders.

Competence area	Primary stakeholder	Secondary stakeholders
Legal/ Regulatory	Policy makers and public authorities, with the participation of transport networks and commercial and logistics industry players.	Research centres and universities together with researchers and students,
Technological	Research centres and universities together with commercial transport and logistics industry players followed by transport network and policy makers	Transport network and policy makers
Transport planning	Public authorities, transport networks and policy makers	Commercial transport and logistics industry players and research centres and universities.
Business modelling	Policy maker, Public authority, Transport networks, Commercial and logistics transport players"	Research centres and universities
Socio-economic	public authorities, commercial transport and logistics industry players and transport network	Transport network
Environmental	Research centres and universities, public authorities, commercial transport and logistics industry players and policy makers	NGOs and community organizations together with citizens

Table 1: Stakeholders influenced based on competence areas

3 Mapping of current situation

As previously discussed, in order to apply such a holistic approach to the creation of a framework of common understanding, it is necessary to properly define the current status of the topic. To this end, 3 main topics have been deemed important: Open Science terminology, stakeholder experience and needs and main challenges.

A global definition of Open Science terminology is a key to ensure the common understanding of the framework. Considering the major differences between modes of transport and stakeholders already described, it is important to develop a language for common discussion and exchange that provides clarity and univocal terms. A review of stakeholder experiences and needs will pinpoint important aspects regarding the implementation of Open Science in transport research on a first-hand experience basis, allowing extrapolating important subjects to be addressed by the framework and enhancing a user-centred approach. Finally, the main challenges will be summarised as a combined output of the literature review and the survey, bringing out the major topics to be considered as input when building the structure of the proposed framework.

3.1 Open Science terminology

First it is important to identify the basic terminology used to create a harmonised approach for Open Science in transport research. It has been noted that the use of data is in a central position when referring to Open Science. In the field of transport a lot of data is created, which constitutes a need for big datasets. In order to share quality data and to create a common understanding in terminology within the “*data sharing of transport research data*”, the following definitions are proposed (13):

Data – “*any piece of information whose value might be used during analysis and impact its result.*” This means that participants, characteristics, weather, traffic and driving conditions could be considered as data and part of a dataset.

Open Data – Open Data are online, free of cost, accessible data that can be used, reused and distributed provided that the data source is attributed.

Open Big Data - Big Data refers to data sets that are too large, or too complex, or too fast-moving or too weakly structured to be evaluated by manual and conventional methods of data processing. Previously (In Task 2.1 (2)), the following projects related to Big Data were identified:

- *LeMO – Leveraging Big Data for managing Transport Operations - targets at developing a strategy that defines the necessary research efforts for the realisation of the big data economy in the transport sector.*
- *AutoMat - project with the objective to innovate an open ecosystem for Vehicle Big Data. The need to make mined and anonymous vehicle data, while building upon current trends in Big Data, is highlighted.*

Metadata: any piece of information necessary to use or properly interpret data. It could be divided into following categories (13):

- *Descriptive Metadata* - describes precisely each component of the dataset, including information about its origin and quality.
- *Structural Metadata*- describes how the data is organized.
- *Administrative Metadata* - sets how the data can be accessed and implemented.

Dataset – One definition for dataset is “*any piece of information necessary to use or properly interpret data*”. euroFOT and UDRIVE are two important EU funded projects that collect large datasets. Datasets are great for data research while providing answers for research questions and the interest in data re-use is increasing.

- *euroFOT* – identified and coordinated an on-the-field testing of new Intelligent Vehicle Systems with potential for improving the quality of European road traffic. Project ended in June 2012. (14)
- *UDRIVE* - first large-scale European Naturalistic Driving (meaning that the behaviour of road users is observed unobtrusively in a natural setting) Study on cars, trucks and powered-two-wheelers. (15)

3.2 Stakeholder-centred analysis

To evaluate the usage, knowledge and stakeholders needs for Open Science in the field of transport research, a survey was organized. More specifically, the aim was to identify user experience in different transport modes for: Open Data, Open Science, Open Software and tools and Open Access. In addition, the knowledge of Open Innovation is mapped. In order to create a clear view it is important to have common understanding of the terminology.

- **Open Data** - Open Data are online, free of cost, accessible data that can be used, reused and distributed provided that the data source is attributed.
- **Open Science** – Organisation for Economic Cooperation and Development (OECD), defines Open Science as: “*to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction*”. Furthermore, Open Science is about extending the principles of openness to the whole research cycle fostering sharing and collaboration as early as possible thus entailing a systemic change to the way science and research is done. (16)
- **Open (Source) Software and tools** - Software for which source code is public and can be freely copied, shared and modified. (17)
- **Open Access** - possibility of unrestricted online access to scientific articles (18)

The survey was distributed among different stakeholders, including different modes of transport, mainly from research institutes of CERTH, DLR, WEGMET, EATEO, HUMANIST, and UITP. In addition the survey was distributed to ACARE and EREA members. Altogether approximately 200 surveys were sent out from which 50 completed applications were received and analyzed. Since the BE OPEN project is focused on the transport research data, the input from Universities and research establishment is very important. As shown in the Figure 3, altogether 88% replies were received from

the named organizations and 8 % replies are coming from commercial transport and logistic industry players.

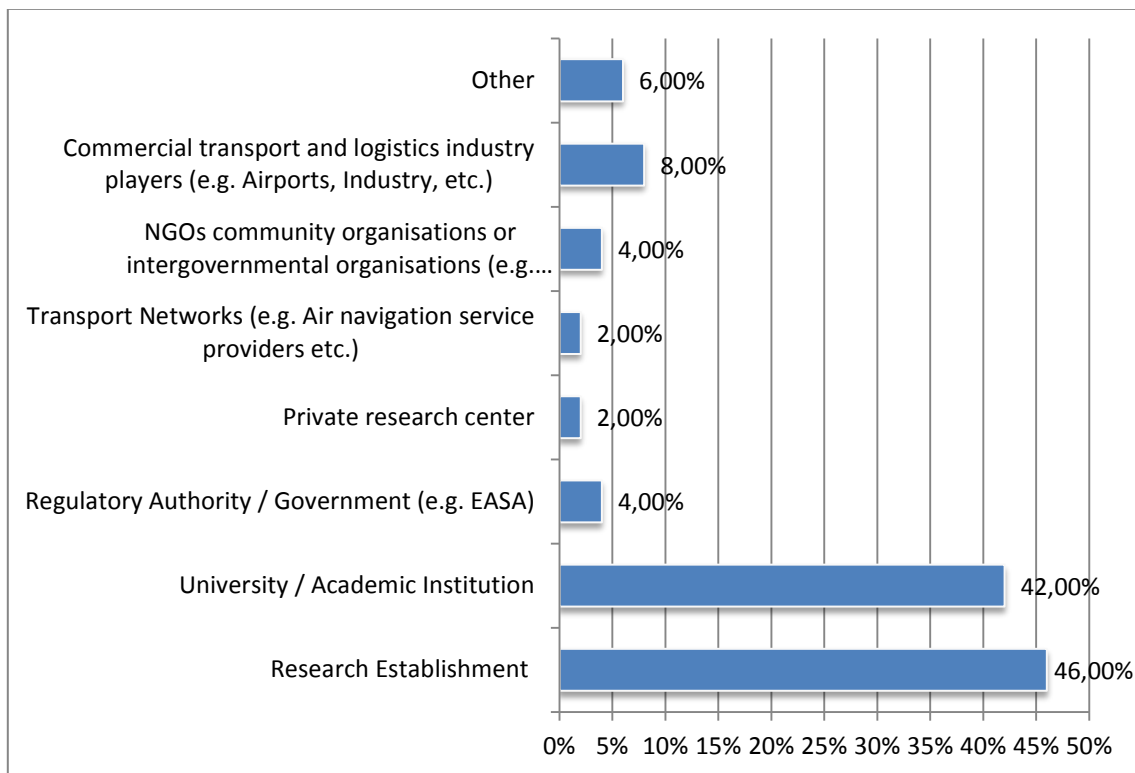


Figure 3: Type of organization

Furthermore the survey provides extra value as applicants from various competence areas are presented. A large amount of stakeholders, 35%, have a technological background followed by almost equal representation of socio-economic, transport planning and legal/regulatory expertise. The exact figures of the representation of each competence area can be found in Figure 4.

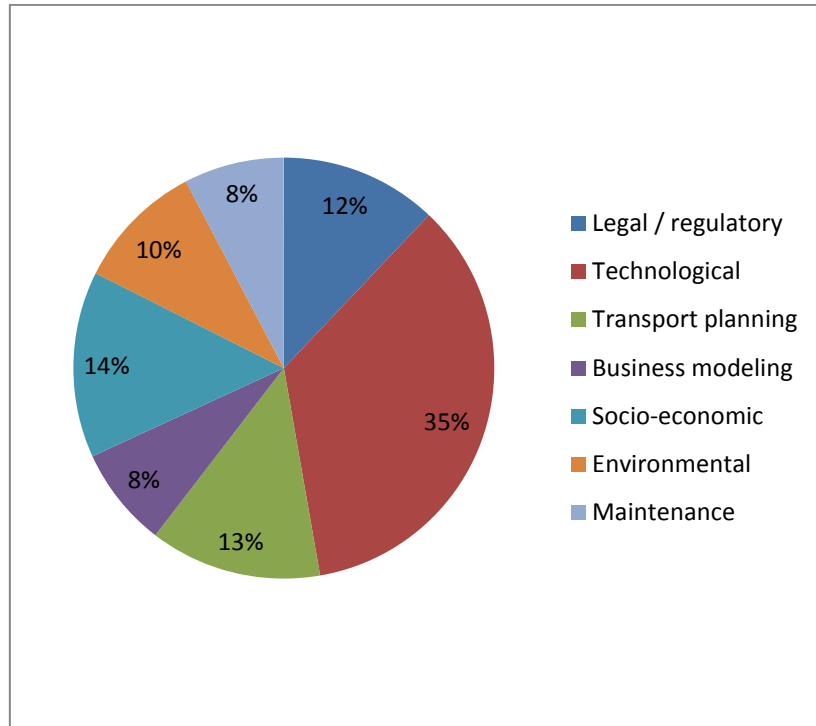


Figure 4: Competence areas

For the survey general statistics, 39 replies from male and 11 replies from female applicants were received. As seen in Figure 5 and Figure 6, the vast majority of the participants are very experienced professionals with more than 20 years of working experience (68%), aged between 31-50 (48%) and 51-65 (46%) years old (mostly men), coming from all modes of transport. A significant portion of participants (22, 9%) appear to be involved in more than one modes of transport (i.e. multimodal), as shown in Figure 7. The majority of participating organizations are of large (72%) and medium (22%) size, located in countries with established regulatory framework on open science (i.e. the Netherlands, Germany, France etc.), as shown in Figure 8.

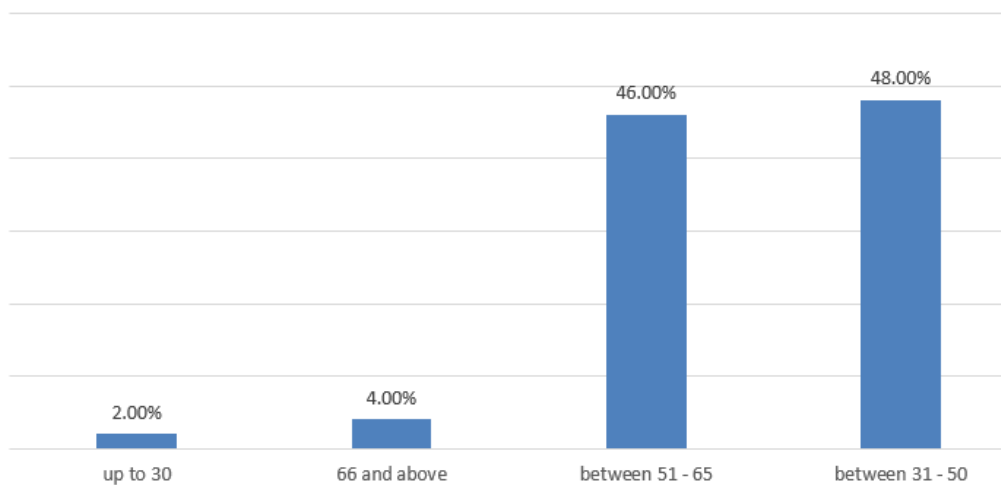


Figure 5: Age range of the participants

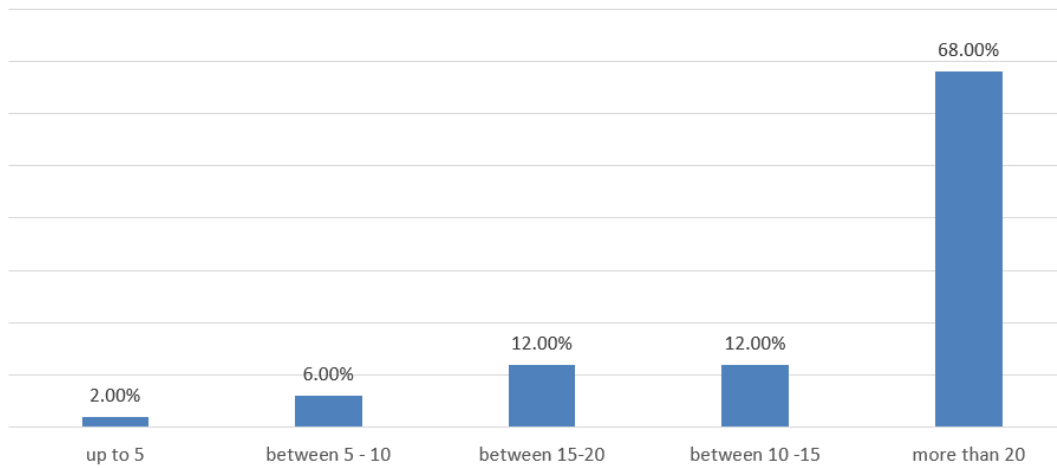


Figure 6: Experience level of the participants (years of experience)

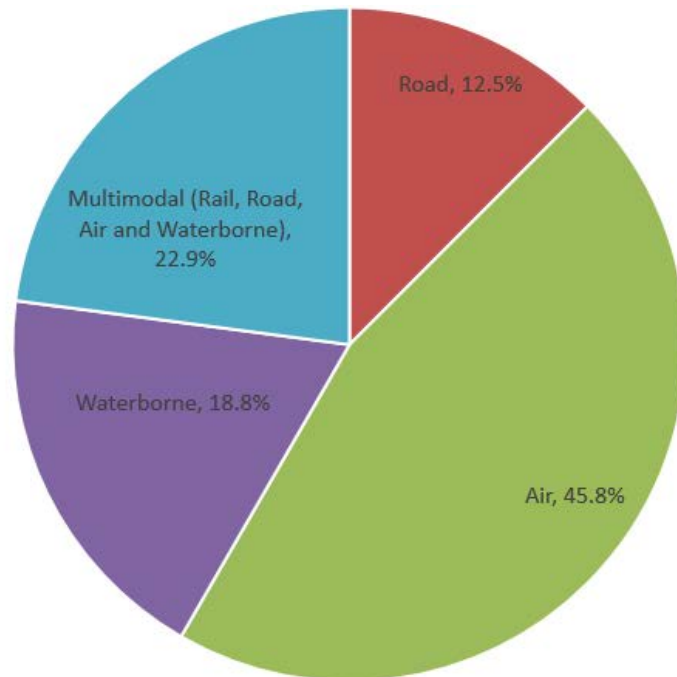


Figure 7: Modes of transport participating in the survey

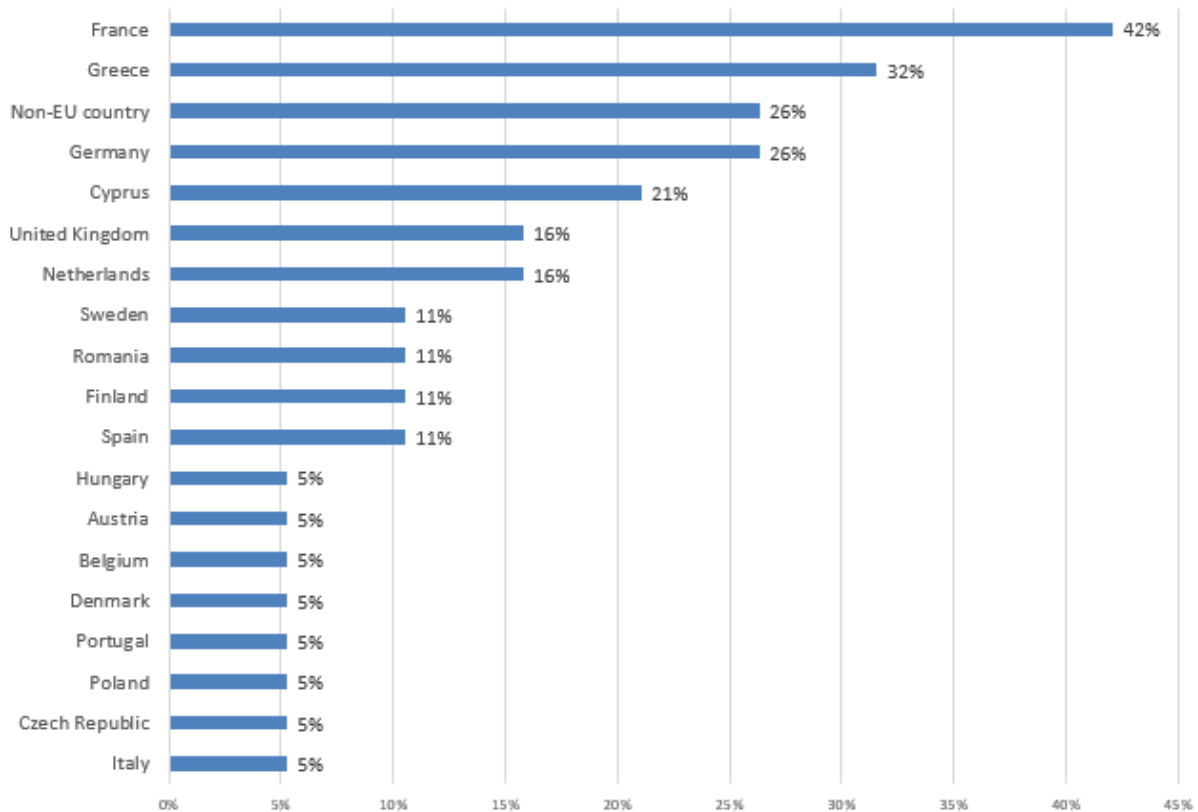


Figure 8 Countries of organizations participated in the survey

All the above give a tremendous value to the quality of the collected data (and consequently to the results) which come from reliable resources (i.e. from experienced professionals etc.) and cover a large market share dealing with all modes of transport (i.e. large organizations with a lot of accumulated knowledge etc.) and in a variety of competence areas.

3.2.1 Stakeholders experience with Open Science

It has been identified that stakeholder's commitment is important for making the best use of Open Science and to ensure the quantity of data made available. As a first step we wanted to map how well informed stakeholders feel and if it is their common practice to use:

- a. Open Access Publications.
- a. Open Software and Tools;
- b. Open Data;
- c. Open Science;

In general, stakeholders feel informed about all aforementioned topics. As it can be seen on Figure 9, stakeholder level of awareness is slightly higher for Open Software and tools and Open Access Publications, comparing to the Open Science and Open Data, where around 20% of applicants do not feel that informed about Open Science. This could be due to the fact that Open Science is a relatively modern general concept, embracing a variety of ideas and therefore stakeholders do not feel particularly confident about their knowledge of it. Because of the recent nature of Open Data, many of its platforms are either not fully operational yet or not known to a bigger audience. On the other hand Open Source Software has been used for longer time among developers hence it is widely

known and its use more extended. Also Open Access Publications are expected to be more widespread through conventional publishing platforms containing hybrid Open Access publications. There is a need to promote all Open Science topics, with a focus on Open data, in order to raise the level of awareness and to engage more stakeholders.

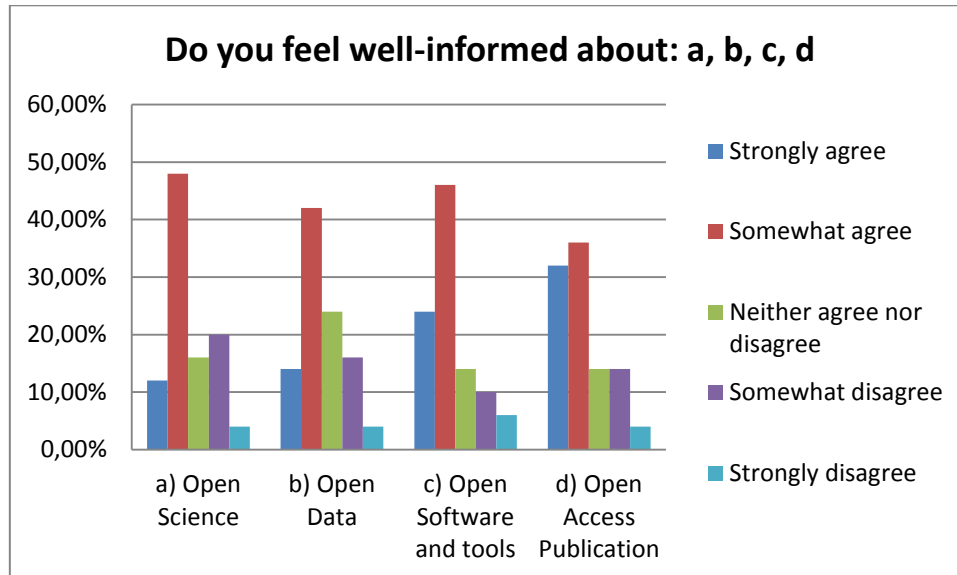


Figure 9: Stakeholders awareness for Open Science

The usage of Open Science, Open Data, Open Software and Tools and Open Access Publications is mainly dependent on how well informed stakeholders are, meaning the lower the level of awareness, the less it is being used. Stakeholders' responses reflect hesitation of the usage Open Data and Open Science as there was high percentage (30-40%) of neither agree nor disagree with the statement. The usage of Open Software and tools as well as Open Access publications is more common, as presented in Figure 10, where stakeholders have expressed stronger commitment.

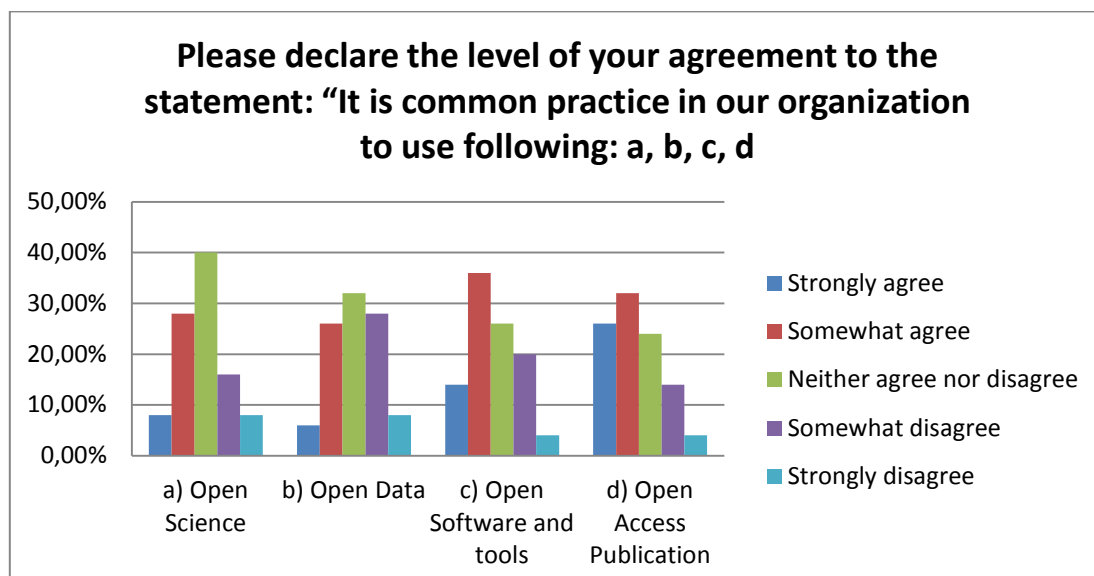


Figure 10 Level of usage of Open Science

The survey shows that the general stakeholders' awareness, especially for Open Science and Open Data needs to be raised in order to enhance the usage of Open Science in transport research data. To have a better understanding of the stakeholders' knowledge, and analysis based on mode of transport needs to be performed.

3.2.2 Mode specific study

For **Open Science**, in the road, air and multimodal transport, in general, around 50% of stakeholders feel informed. According to the results displayed in Figure 11, the level of awareness of Open Science is slightly lower in waterborne transport, where the percentage of stakeholders who somewhat disagree is relatively high, 21%. However, no solid conclusions are drawn from such observations as the total amount of interviewees from this sector is rather low, making the outcomes comparatively buoyant.

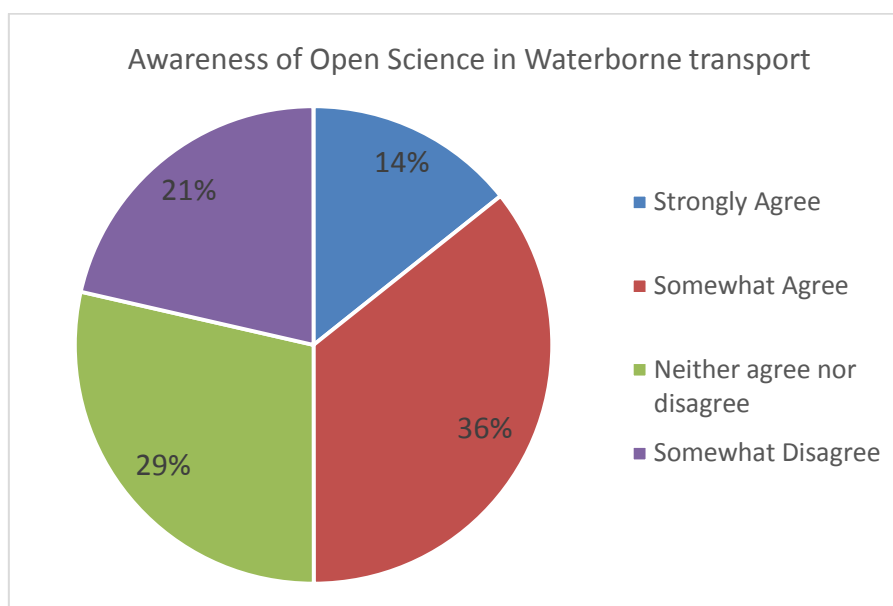


Figure 11: Awareness for Open Science in waterborne transport

It could be said that Open Science usage depends usually on the stakeholders' level of awareness. In multimodal transport, the usage of Open Science seems to be slightly more common comparing to other modes of transport, where 45% of stakeholders somewhat agree with using Open Science. On the other hand, in the air transport mode, the results from Figure 12 indicate that around 20% of the respondents somewhat disagree and around 40% of stakeholders neither agree nor disagree with the usage of Open Science in their field.

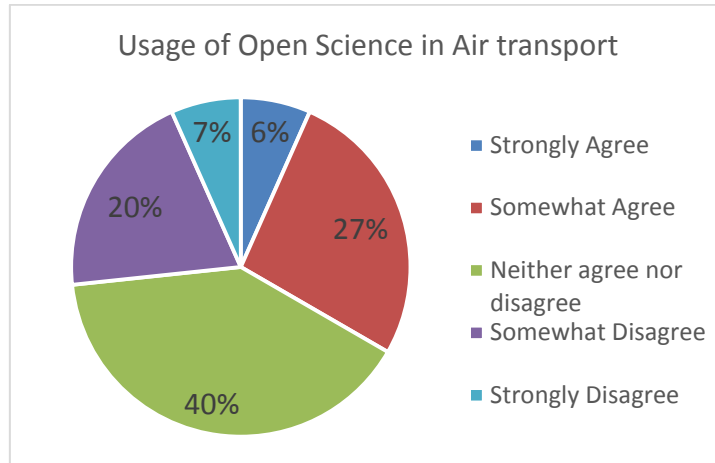


Figure 12: Usage of Open Science in air transport

Furthermore for waterborne transport (Figure 13), the majority of stakeholders, 57% neither agree nor disagree to use Open Science in their field of expertise. The aforementioned transport sectors are more restricted, with very strict safety rules in the field of aviation, and time consuming implementation of changes in waterborne transport. This could be why Open Science usage is not that common in these sectors and more research needs to be done to realize the opportunities of Open Science within them.

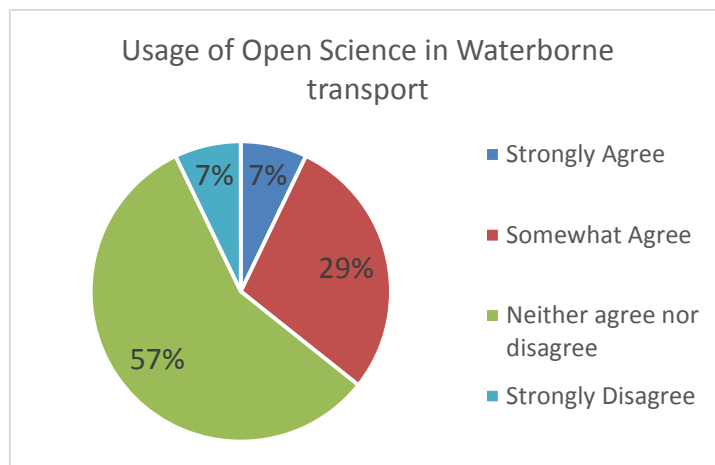


Figure 13: Usage of Open Science in waterborne transport mode

For **Open Data**, similarly to Open Science, 50% of the respondents of all modes of transport feel well informed. For multimodal transport (Figure 14), the stakeholders feel more informed about Open Data comparing to other transport modes. The amount of stakeholders, who feel somewhat informed about Open Data was slightly lower for air transport, 47%, and waterborne transport, 43%.

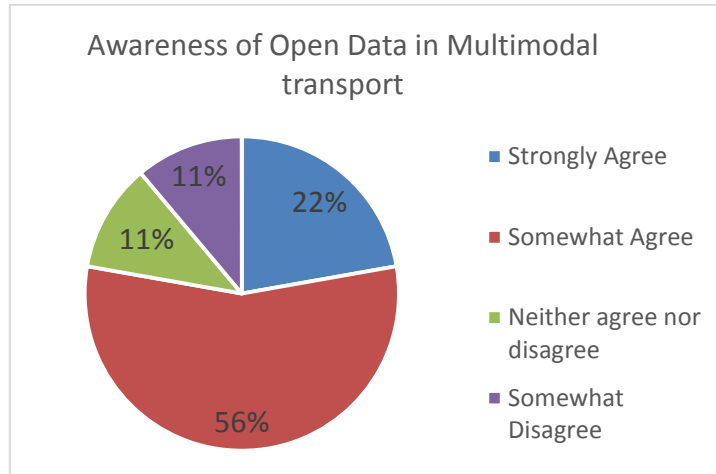


Figure 14: Awareness of Open Data in multimodal transport

In general **Open Data** is more used in multimodal transport. On the other side, for air transport (Figure 15), the amount of stakeholders (36%) who disagree of using Open Data is relatively high. It is worth to investigate further why for air transport the usage of Open Data is not that common while the amount of stakeholders who feel informed was higher. For waterborne transport (Figure 16), the amount of stakeholders who neither agree nor disagree with using Open Data is 50%. This could be because the data sharing between sectors is fragmented and it is unclear how all the stakeholders could benefit from making data openly available and re-usable.

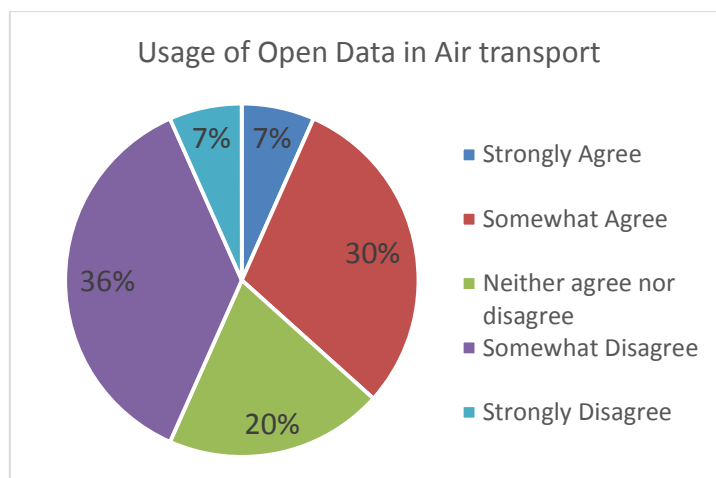


Figure 15: Usage of Open Data in air transport

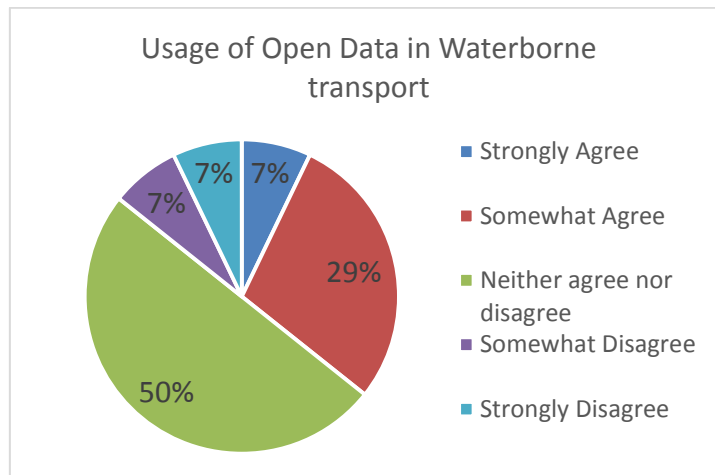


Figure 16: Usage of Open data in waterborne transport

Open Software and tools are in general more known among all transport modes. This could be because Open Software has been used for longer time and is more developed, comparing to Open Science and Open Data. It seems that the usage of Open Software and tools is less popular for the waterborne sector (Figure 17), where 29% of stakeholders somewhat disagree using Open Software in the sector. More research could be done to investigate how to make Open Software and tools more attractive to all stakeholders.

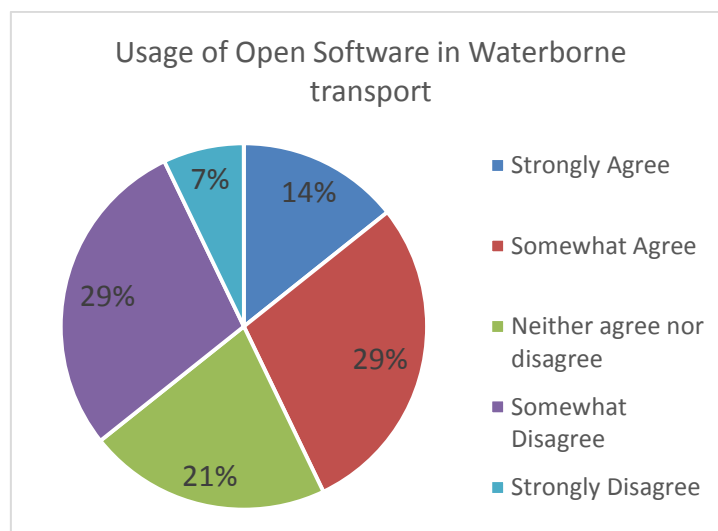


Figure 17: Usage of Open Software in waterborne transport mode

For **Open Access** publications the level of awareness is relatively high, where for example in the air transport (Figure 18), 20% of stakeholders strongly agree and 44% somewhat agree with being informed. For the road transport 44% of stakeholders feel strongly informed and 25% feel somewhat informed of Open Access publications.

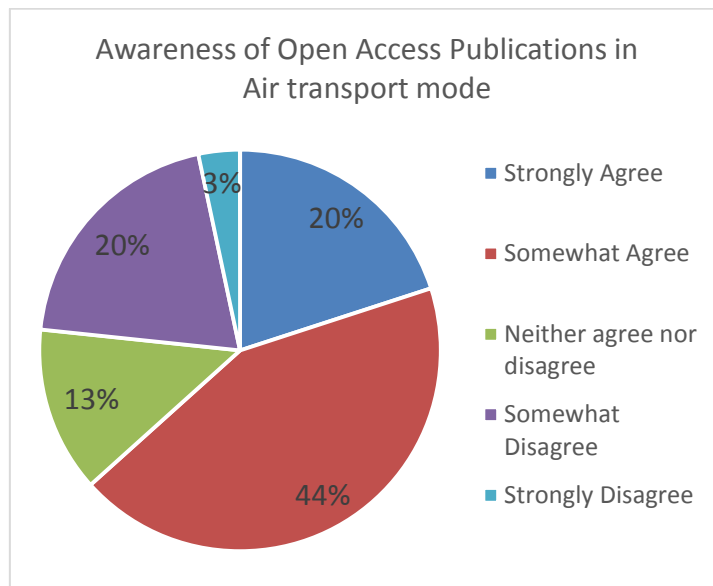


Figure 18: Awareness of Open Access publication in air transport mode

Open Access usage is also higher in all transport modes. It is interesting to note that for road transport (Figure 19) the amount of stakeholders who agreed to the usage of Open Access is coherent with the level awareness, where 25% somewhat agree and 44% strongly agree. This could mean that the stakeholders who are informed about Open Access are also using it.

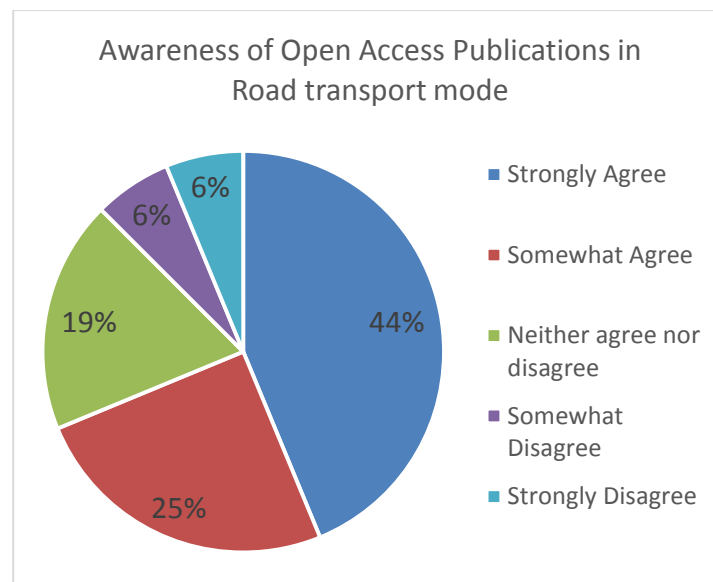


Figure 19: Usage of Open Access Publications in road transport mode

In general there is more commitment needed for all the transport modes, especially when it comes to Open Science and Open Data. Based on the knowledge and usage of Open data the goal of the survey was to specify how stakeholders work with **Open Data**. It is worth to note that 38% of stakeholders use the data, followed by 28% process and 22% create the data (Figure 20).

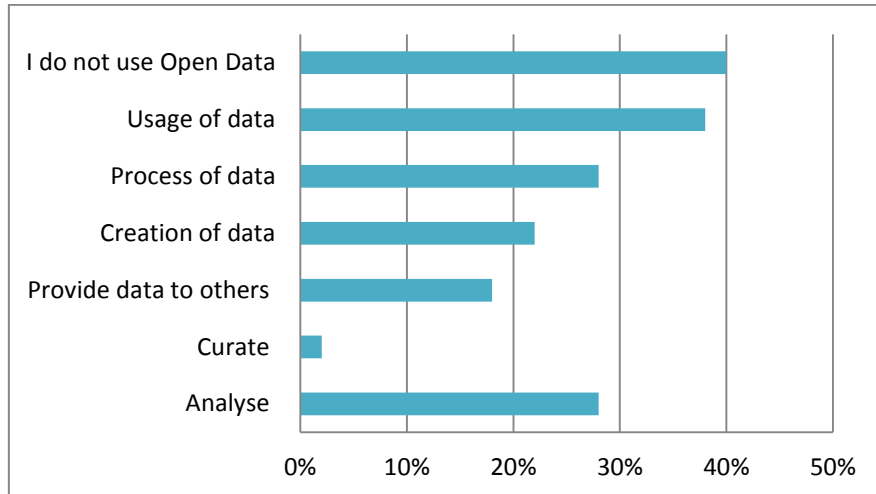


Figure 20: Usage of Open Data

Furthermore, a more specific analysis was performed based on transport modes. In order to provide an overview, a cumulative graphic was prepared in a way where modes of transports are shown on the Y-axis and the x-axis presents the absolute number – 1- which was calculated taking into consideration the amount of replies from stakeholders from the transport mode on the data analyses, data processing and data usage. It can be seen in Figure 21 and Figure 22 that the tendency of using and analyzing the data is more common for waterborne, road and multimodal transport. Data processing (Figure 23) seems to be more popular for waterborne and multimodal transport.

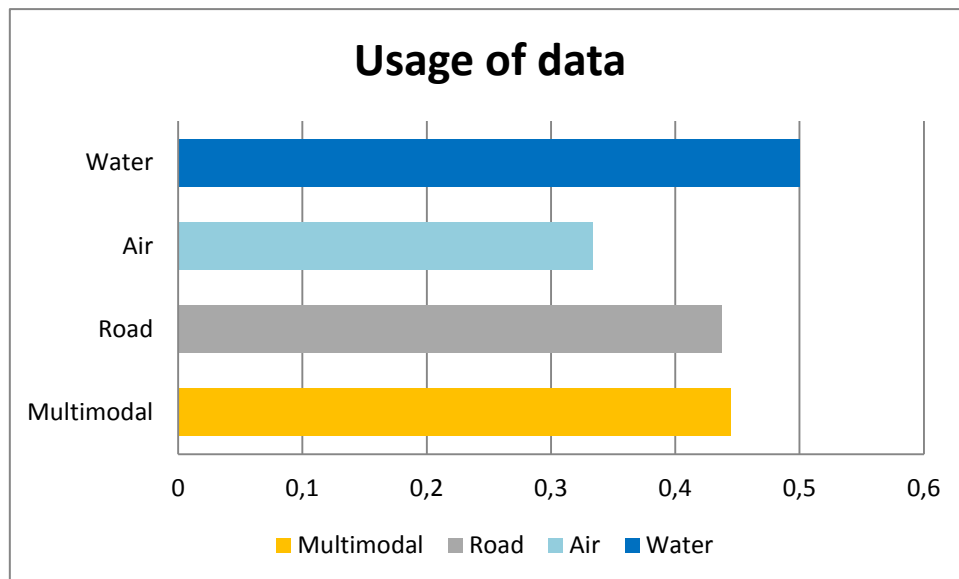


Figure 21: Usage of data based on mode of transport

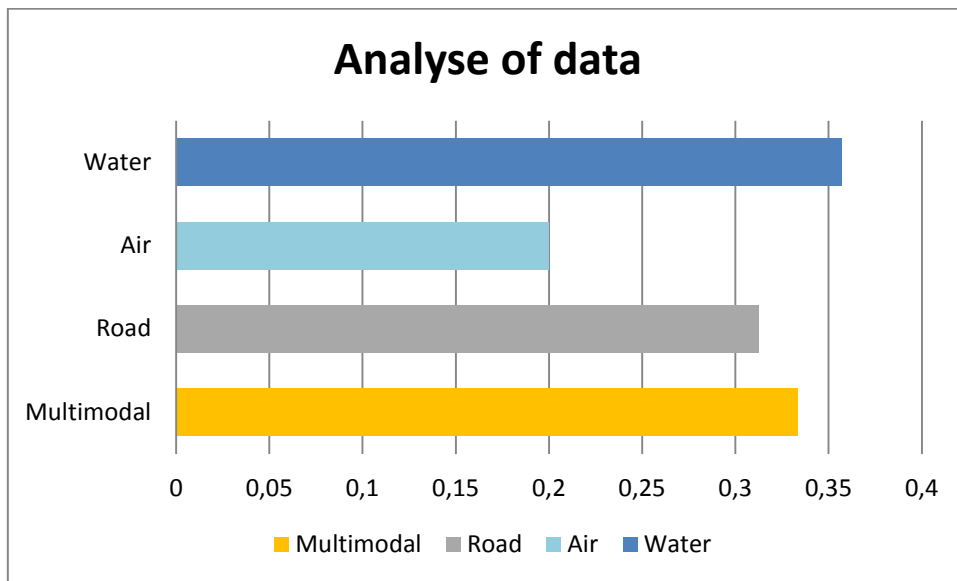


Figure 22: Result of data analyses in different sectors

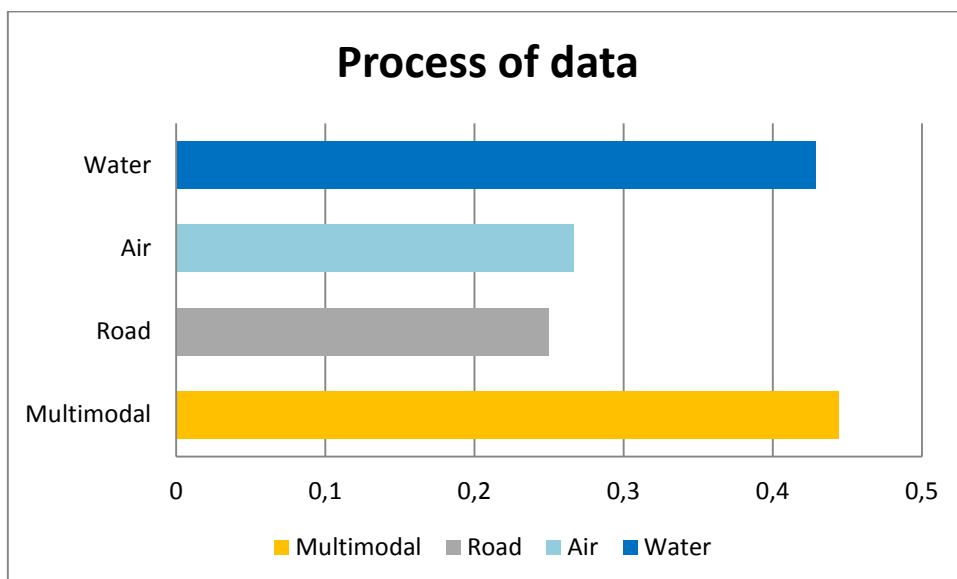


Figure 23: Result of process data in different sectors

Based on the stakeholders' feedback the data is being used for different purposes. For example:

- a) Measurement of reference datasets to benchmark and / or advance numerical simulation tools;
- b) Create and use open benchmark for algorithms, provide underlying data of publications to the public;
- c) Bases of new hypothesis, evidence and demonstration;
- d) Air traffic data and aircraft performance models;

- e) Maps creation for conducting research;
- f) Socio-demographic and time-use data;
- g) Performance models.

This means that stakeholders use Open Data for different purposes and there is a need to provide compatible Open Data, which would allow stakeholders to use the data faster and in a flexible way.

3.2.3 Open Innovation

In terms of Open Science, Open Innovation has an important role to foster knowledge. *“We should use our innovation talent the best way possible to be the lead in Open Science in the world.”* **Open Innovation** means transferring knowledge, expertise and even resources from one company to another. It is common knowledge that industries and researchers are still very protective when it comes to data sharing and a change of mind should be considered. It is important to work together towards dynamic, networked, multi-collaborative innovation ecosystems. Open Science alone does not ensure that research results and scientific knowledge are commercialized or transformed into socio-economic value. Using Open Innovation could help to connect and exploit the results of Open Science and facilitate the faster translation of discoveries into societal and economic value. (19)

One goal of the survey was to find out how organizations classify in relation to Open and Closed Innovation principles and also to see the preferred mechanism. Twelve statements from closed and open innovation, for example: *“The smart people in our field work for us”* against *“Not all the smart people work for us”* were proposed. In order to provide a better comparison and to evaluate stakeholders’ commitment to Open and Closed Innovation, an absolute number of 1 was calculated and a cumulative graph was prepared. The negative values on the x axis represent stakeholders commitment to Closed Innovation and positive represent commitment to Open Innovation principles. As an outcome, there was a tendency towards Open Innovation principles, as can also be seen on Figure 24. Among stakeholders in the transport sector more dissemination for Open Innovation needs to be done to foster knowledge.

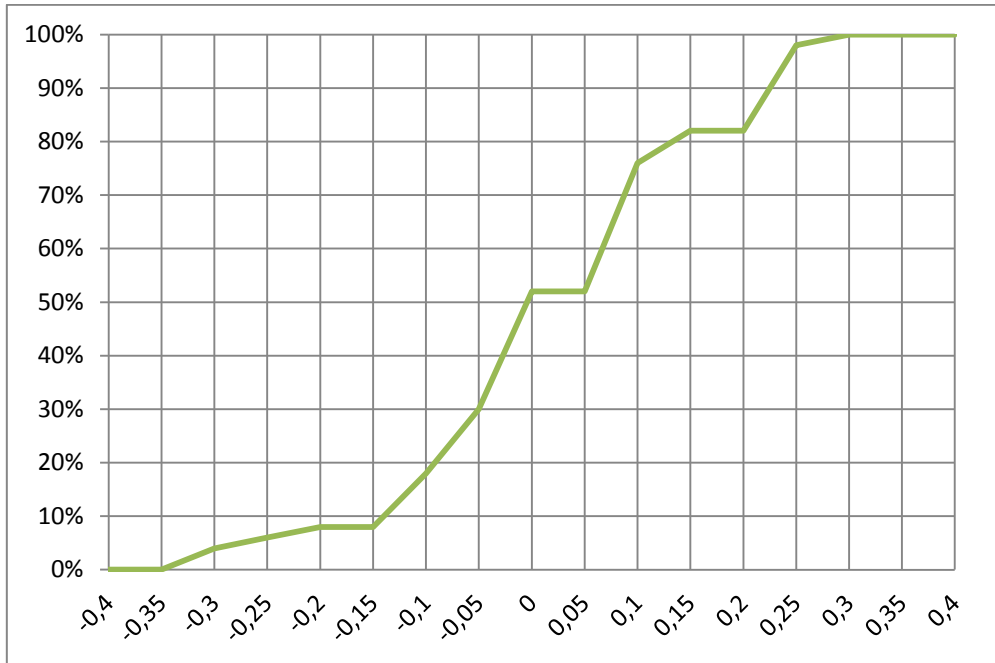


Figure 24: Open and Closed innovation

3.2.4 Derived Challenges concerning Open Science/ Open Data

The majority of stakeholders are representing the technological and socio-economic fields, therefore also 44% of the respondents are working with related data, as can be seen on Figure 25. Environmental and transport planning areas follow up, probably due to current exploration of new mobility concepts and increasing environmental and sustainability concerns.

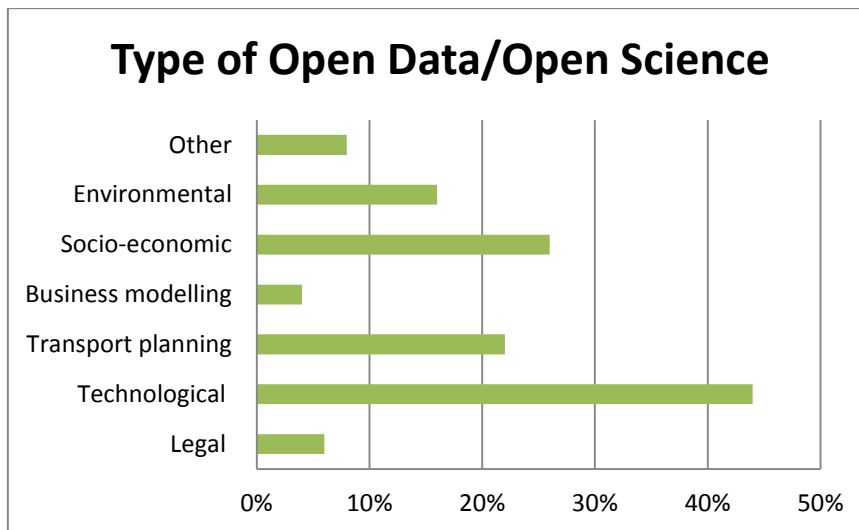


Figure 25: Main type of Open data used

The most commonly used data platforms are European (for example Open Science Cloud) and national platforms. What is interesting to note is that nevertheless European data platforms are very popular, a combination of data platforms is also being used, where 33% of stakeholders use national as well as European data platforms. It could be helpful for the users to have one central database to

work with the data. As can be seen on Figure 26, the usage of virtual data platforms such as GitHub, is very low among transport research users.

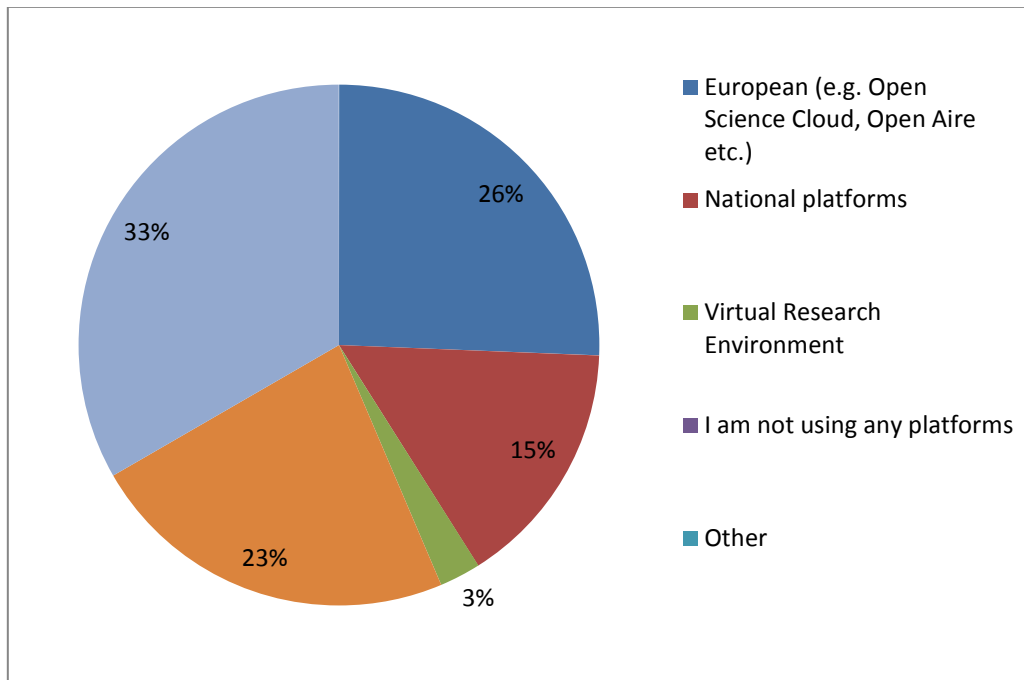


Figure 26: Usage of Data platforms

It is important to note that a large amount of stakeholders, 52%, see that Open data/ Open Science would bring benefit to their organization (Figure 27). Furthermore, 42% recognize that using Open Data would be an opportunity to precompetitive research. It could be due to the low level of awareness of Open science that the limitations are not clear.

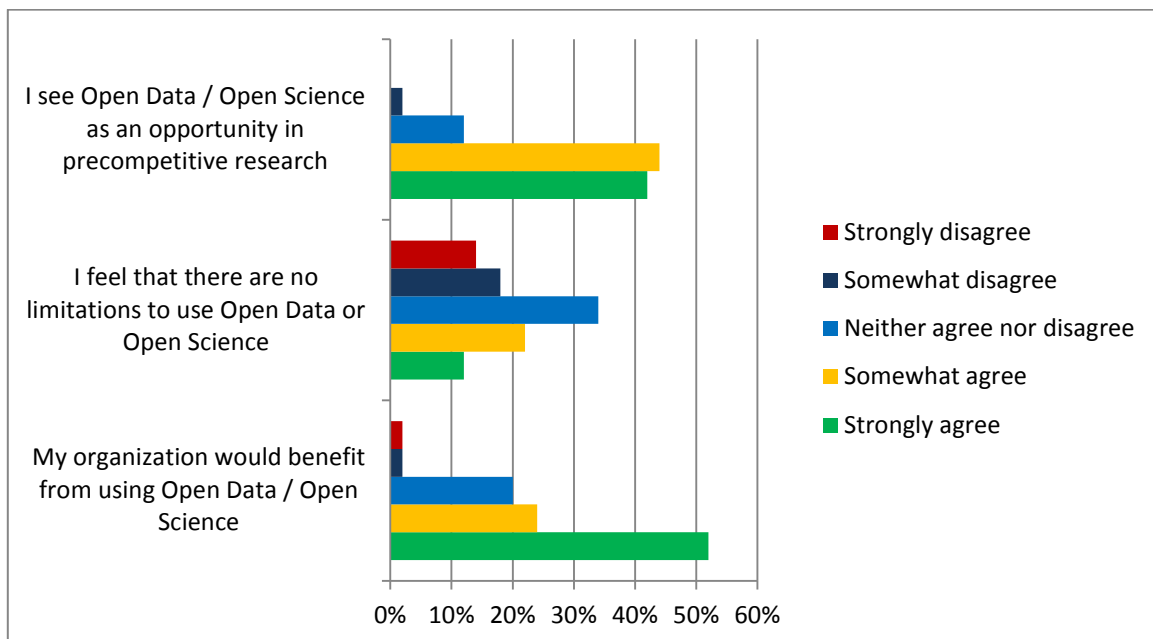


Figure 27: Possibilities for Open Science

On the other hand, 14% disagree, which means that some limitations are identified and need to be addressed. Based on the comments received from the stakeholders some of the limitations are identified as follows:

- a) Reciprocity is needed for opening data and science
- b) Concern how Open Data is being shared in the whole world e.g. USA and China
- c) Risk of losing IP
- d) High volume of workload and not enough funding available
- e) Lack of detailed description of metadata and datasets available in the platforms. That could lead to misusing the data in the future for other research

While listing the limitations, also stakeholders' needs should be mapped. Figure 28 brings out that stakeholders' main needs are Common policy and time factor (24%), followed by Monetary and immaterial benefit (18%). It is interesting to note that also the need for a Central repository (16%) is valued relatively high, which could be because of the data produced in different sectors is fragmented.



Figure 28 Stakeholders needs for sharing their experience

Based on the stakeholders collected comments, following issues have to be emphasized:

- Costs for data creation and curation would have to be covered
- Sustainable information and work procedure is needed
- Support from organization by clear guidelines and infrastructure
- Common Policy, at international as well as national level
- Detailed description of the source

To conclude, the stakeholders' main needs are to have a common policy and clear guidelines to ensure the quality of data. For better use of Open Science and Open Data, measures and actions for data protection and security have to be made clear. Funding schema to provide the necessary resources to support stakeholders in their contribution to Open Science needs to put in place.

Implementing technical solutions and focusing on automation processes would help to reduce stakeholders' cost and time. For all the aforementioned actors, support from European as well as national level is needed.

3.2.5 Conclusion

The organized survey and corresponding feedback received from different stakeholders gives an overview of Open Science and Open Data usage in the transport sector. It is interesting to point out that nevertheless the main respondents are from Universities and Research Institutions, the level of knowledge of Open Science, including Open Data, is relatively low. This could be because the usage of Open Data and many of its platforms are not fully operational or the options are not known to the stakeholders. In order to improve that, it could be interesting to offer workshops to promote Open Science. The Open Software and tools as well as Open Access Publications are more known to the stakeholders and the level of awareness and usage among stakeholders is higher, although more use could be expected.

The mode specific study brings out that there is a need to foster the knowledge, especially for Open Science and Open Data among all transport modes. Due to the lack of replies from rail transport and a low number of replies to the survey for waterborne sector, further investigation is needed to understand the level of awareness and usage of Open Science and Open Data as well as the measurements to foster the knowledge.

Stakeholders acknowledge that they could benefit from Open Science, where 52% could see that it would bring benefit to the organization and 42% strongly agree to the point that it would be an opportunity in precompetitive research. On the other hand, stakeholders also strongly pointed out that there are limitations for using Open Science and Open Data. The main listed challenges were lack of common policy, time and monetary as well as immaterial benefit. Furthermore, data security and quality was also brought up. There is a need to foster Open Science by finding solutions for the challenges raised by stakeholders - providing common understanding among all stakeholders, developing technical solutions for automatization and improving data quality as well as setting up the funding schema to support stakeholders' contribution. In addition, more work could be done at European as well as national level to promulgate the knowledge and to assure availability of quality data so there would be more usage of Open Science and Open Data.

Data for survey questions are listed in the Annex.

3.3 Main challenges

The main concept of using Open Science is to make research data openly available among all different stakeholders in order to use the research data to create new value. In the field of transport, one of the main goals is to provide a sustainable transport network. With the new innovative technologies and automation, a lot of data is being created by all different transport modes. This also sets challenges for Open Science, where differences in transport modes, fragmentation and legal as well as technological issues lead to a lack of data quality. The following challenges have been previously identified by the European Commission Studies and Report "Analysis of the state of the art, barriers, needs and opportunities for setting up a Transport Research Cloud" (20):

1. **Fragmented data and large variety of stakeholders:** there is a large variety of stakeholders from different modes of transport. They produce different types of data and it would be in their interest to re-use research data from all transport modes. Also, available research data is fragmented, being stored in different catalogues, repositories and platforms. All those factors are making data analyses and accessibility complex. Therefore interoperability between different databases and repositories is needed.
2. **Enhancing data security and privacy principles:** in the transport sector research data is being created by public as well as private stakeholders. The collected data could contain sensitive information. Storing sensitive information in a cloud creates challenges to protect and secure the data. Therefore there is a need to solve data privacy, legal liability and IP related issues. In addition, partnership agreements need to be put in place between original researchers before sharing the data so that it would be understandable and suitable for everyone—original researchers as well as the future researcher who would use the data.
3. **Funding:** making data publicly available needs participation of all the stakeholders. In addition, supporting infrastructures to allow easy access to research data and to make it reusable are required. There is a need to find resources – man power as well as funding - for the necessary infrastructures.
4. **Technological challenge:** technological solutions need to be optimized in a way to accommodate large volumes of datasets and to ensure interoperability of data from different sources. The importance of interoperability resides in the possibilities of reusable data. Reusable data allows for repeatability of studies, which generates more faithful, accurate and reliable outcomes, and for resource optimization, as different studies needing the same data could simply access it from the corresponding source. There is also a need for infrastructures and tools to store the data for longer time. Therefore there is a need for reliable archive systems, platforms and tools which could read the data.
5. In open science, **data quality** is extremely important to allow for data reuse and to engage stakeholders from different fields. Due to fragmentation of the data and lack of explicit guidelines to store the metadata and datasets in transport research, improvement through the creation of a common understanding is needed. A template could be created that would allow to standardize as much as possible the description of metadata as well as data sets for all different transport modes. One of the main concerns in sharing and using Open Data is the quality of available data. In order to build user trust and to promote Open Science in transport research, it is important to describe the available data as detailed as possible, including information on data collection and experimental procedures. The document should give an overview of the purpose for which the field tests or data collection were made, research questions, sample selection criteria and overall description of recruitment, overall description of used equipment, date and timing of different phases of study, test plan and execution and how contact was kept during the study. This gives an opportunity for different stakeholders to use the data and builds trust among the users. (13)

6. **Lack of skilled experts:** who would have the miscellaneous experience in the field of analytics together with the understanding of the best use of data sets to add value to the research data.
7. **Legal challenges:** collecting transport research data from different modes of transport, different countries and including sensitive data restrictions sets legal barriers. There is a need to standardize regional as well as international rules for data privacy including support from European legal regulation, covering security and privacy especially. Also, data made available online should be ethically used – it is important to set up a mechanism to observe that the regulations are being followed. The most important concept for all the stakeholders would be that the data should be made *“as open as possible and as closed as needed”* (4).

From the organized survey, there were several stakeholders supporting the previously defined challenges, where a strong need for common understanding in Open Science requires the creation of a supportive framework. It is important to harmonize legal requirements at European as well as national level to support usage of Open Science in the field of transport research data. In addition, it was acknowledged that more dissemination is required, especially because the level of knowledge among the transport network sector was not very high.

The lack of individual knowledge identified in the survey also leads to believe that if research data needs to be found based on individual knowledge of where it is stored, users are rather unlikely to find and access it, impairing the whole research process. Also noteworthy is that the quality of the data, in spite of its importance, is not necessarily taken care of in all platforms, which is why proper guidelines need to be defined. The Transport Research Cloud is expected to overcome these issues by granting access to a variety of curated Open Data sets. For optimal implementation Task 2.1 (2) recommends that its objectives should remain aligned with those of European Open Science and the EOSC.

Beyond the challenges that accessing research data presents in general, the field of transport entails bigger challenges of its own. The high volume of data collected in this domain, ever-growing due to the implementation of disruptive technologies such as automation and connected vehicles, makes Big Data a valuable resource for the sector. This type of data, however, requires special handling concerning storage space and computer processing capabilities, building up the technological challenge.

4 Open Science and transport research

In this section, different sources, platforms and models used in Open Science will be investigated to evaluate their potential for implementation within the proposed structure for a framework of common understanding. The major inputs to the structure, as shown in Figure 29, are therefore the challenges, which describe the main topics of interest to be addressed by the framework in order to satisfy the user needs and overcome anticipated issues, and Open Science sources, which will be proof of the potential for implementation of Open Science in transport research.

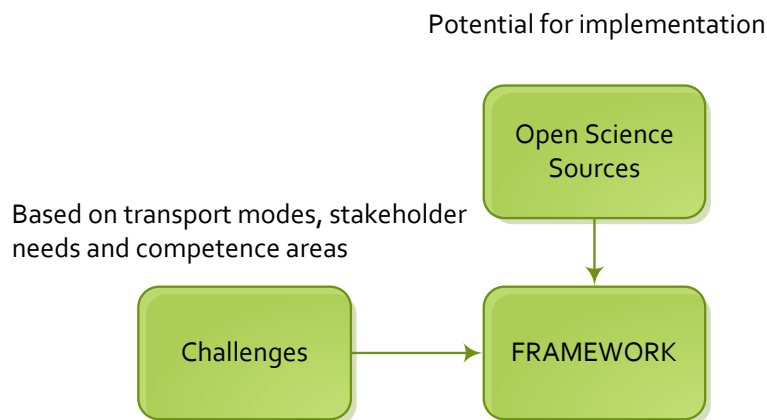


Figure 29: Schematic of framework inputs

Open Science refers to a global concept that goes beyond Open Data and Open Access. It implies a holistic approach to the different aspects and stages of research. In order to bring together the concepts of Open Science and transport research, the research process itself will be analysed, particularizing to the case of transport research through the different types of data used in it according to the literature (20). The different Open Science sources discussed will be sorted in terms of type of transport research data and step of the process, binding the Open Science platforms with transport research and mapping out which platforms provides which kind of data relevant at each step of the research process. Furthermore, data platforms will be categorised in terms of stakeholders and competence areas to reflect on all the main aspects related to research within the transport field.

As this study is focused on the implementation of Open Science in transport research, it is important to properly define research in order to evaluate the different tools used in the process and at which stage is it best to implement them. The whole research process can be understood as a self-developing cycle through which research outputs from certain studies act as well as inputs of subsequent activities, as depicted in Figure 30.

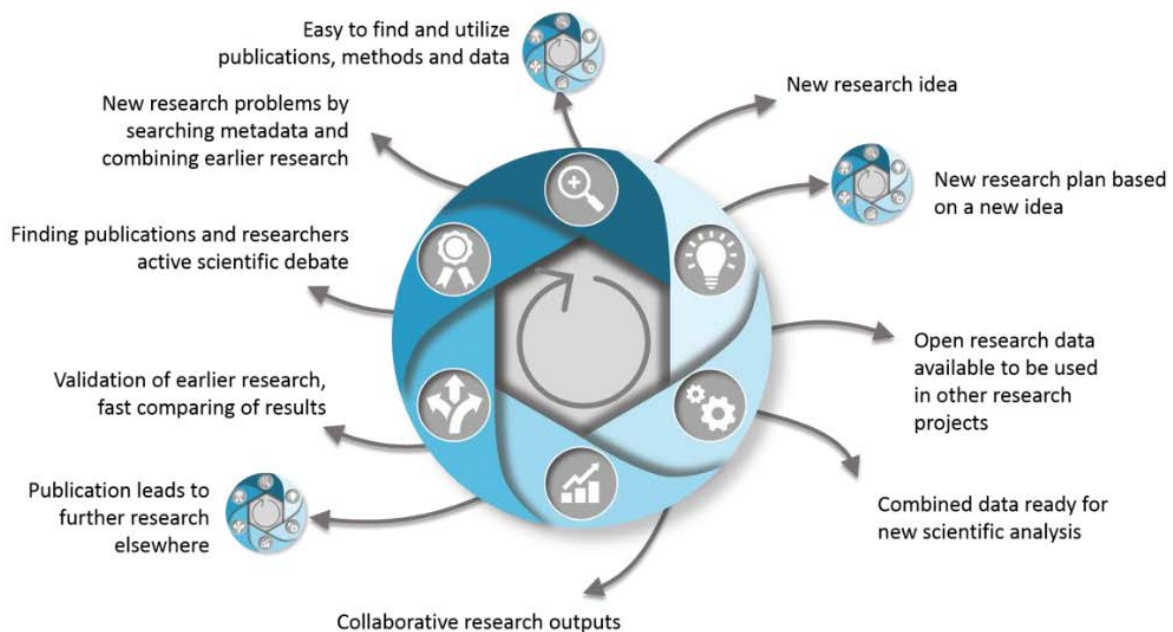


Figure 30. Research cycle (21)

Along the research process, six different main stages can be identified, namely (21):

- Planning research: this stage is normally focused on the development of an adequate methodology. Experiments, interviews and observations can be planned and a workflow management system arranged. The most important steps are:
 - Defining the research problem
 - Planning research
 - Specification of research environment
 - Funding and resourcing
 - Organisation of research
- Research: data collection, whether in the form of numbers, code, texts, images or sound builds the basis of research and Open Data. Documents are normally developed at this stage, mostly as drafts, so versioning control, storage and management is also important. Critical steps are as follows:
 - Choosing or developing research methods
 - Choosing and collating research data
 - Processing research data
 - Documenting research events
 - Discovering and identifying research findings
- Managing research outputs: to extrapolate conclusions from the data collected throughout the research, an analytical phase is to follow. Outputs are devised based on statistics, processes and analysis of both the data and documentation. Interactive computing is a useful tool at this stage to generate outputs and describe research findings. The storage and preservation of the methods and outputs must too be taken into account.

- Publication and dissemination: different types of research are made public in different forms, e.g. journal articles, dissertations, books, source code, etc. Whichever the nature of the output, it is important to:
 - Evaluate and publish research outputs
 - Distribute research outputs
 - Communicate research findings
 - Ensure long-term accessibility
- Assessing findings: for the sake of data quality, research outputs should be properly revised and validated. This can be done by means of peer reviews, which can also assess the preservation needs of different research outputs.
- Discover and reuse: at this stage is where new original ideas sprout, and innovation begins. Inspiration for such new ideas can be found on social media, reading a blog or looking through previous studies. The goal is to further utilise findings from previous activities to curate the outputs and foster new projects and development.

Focusing more on the field of transport, **transport research scientific data** are categorised in the following way (20):

1. **Original research data:** e.g., data from Field Operational Tests (FOTs), Naturalistic Driving Studies (NDS), research results and research models from published papers. This category of data is most likely to embody the basis of the TRC. A certain amount of information will be requested for each research project submitting data to the TRC. This information should include at least the following:
 - Experiment context (possibly included as metadata);
 - Data file description (possibly included as metadata);
 - Data access instructions (possibly included as metadata);
 - Raw dataset(s) used in the research;
 - Transformed and aggregated datasets: cleaned-up, derived, annotated data, generalized graphs and tables, data combined with other datasets, etc.;
 - Research models and research results
2. **Operational data directly related to research:** this classification is generally constituted of data from public authorities and could include accident data, vehicles registry, etc.
3. **Data from published transport research**

In the Deliverable 2.1 (2) and Table 2 the different resources used in the BE OPEN project are described in order to prepare the framework for building effective Open Science platforms and tools in the transport domain.

Type of Data	Resources used in BE OPEN
Original Research data	EC funded projects
	Industry and research projects and initiatives
	Research activities libraries (e.g. TRIMIS, ERTICO Observatory, etc.)
	University projects and initiatives
Operational data directly related to research	National public authorities
	European public authorities
	Other public authorities
Data from published transport research	Scopus
	Web of Science (WoS)
	Google Scholar
	Other citation database of peer-reviewed literature

Table 2: Scientific sources in Transport research

4.1 Open Science Platforms

In the following, a number of data platforms will be assessed and compared according to the different types of data they store, the main competence areas they cover, and the stakeholders involved. This will provide the necessary information to evaluate their potential for implementation within the proposed framework structure.

To perform this classification, the three main categories of data suggested to be included in the Transport Research Cloud (TRC) by previous documents will be used. The Transport Research Cloud is a project for a thematic pillar of the European Open Science Cloud addressing the specific needs of the transport research community.

Registry of Research Data Repository: is a worldwide registry of research data repositories covering various branches of knowledge. It provides repositories for the long-term storage of datasets to researchers, funding institutions, publishers, etc. The goal is to nurture a sharing culture, enhanced access and improved visibility of the data available (22). This platform is particularly useful regarding the assessment of research findings within the research process, as in order to evaluate the preservation needs the possibilities for long-term storage should be taken into account. Publication and dissemination are also enhanced through the use of this platform.

Open Science/Data Repository: Repositories are an important link of making data publicly available for different stakeholders in different countries, while providing access to a large amount of research outputs. They are acting as a hub in a bigger network where information stored in the repositories could be used for future research. The importance of E-research

infrastructure repositories is growing, with the aim of providing a more integrated and seamless suite of services to various communities. There is a need to standardize across the whole system vocabularies, metadata and indicators to enhance the interoperability. With the rapid growth of technologies and the need to continually monitor the ongoing developments, extensive updates of the repository platform technologies are needed. Furthermore, local implementation of guidelines and standards is required in the field of transport research, with future developments in the field of: autonomous transport systems, drones and low-altitude aerial mobility, Big and Open Data, data governance and data availability to support the next generation repositories are necessary. The difference between current and next generation repositories is depicted in Figure 31. The work has already started as the Next Generation Repository Working Group has been set. Together with COAR - Confederation of Open Access Repositories, whose mission is to enhance the visibility and application of research outputs through global networks of Open Access digital repositories, user needs and next generation repository goals have been defined. The COAR vision for next generation repositories is: *“to position repositories as the foundation for a distributed, globally networked infrastructure for scholarly communication, on top of which layers of value added services will be deployed, thereby transforming the system, making it more research-centric, open to and supportive of innovation, while also collectively managed by the scholarly community.”* (23)

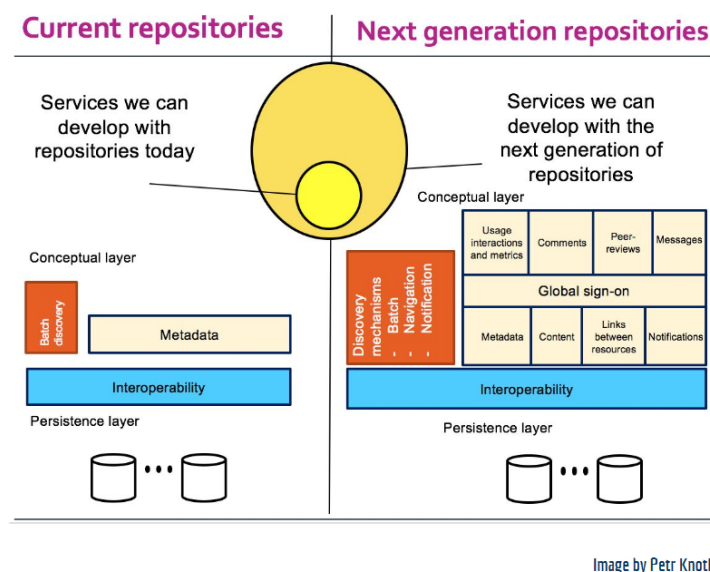


Image by Petr Knoth

Figure 31: Difference between current and next generation repository (23)

CORDIS: Community Research and Development Information Service, EC’s primary source of results from funded projects for research and innovation. It includes data from FP 1 to Horizon 2020. It aims at facilitating research outcomes to specialists in the different research areas to empower Open Science, foster innovation and nurture European growth. Its core is a structured public repository storing all project information in control of the European Commission, including reports and open-access publications, amongst others. Own CORDIS articles and publications are also included in the repository to facilitate the search, and web versions are available in a variety of European languages. The platform itself is funded within the framework of the Horizon 2020 programme (24). The

application of this platform is of paramount importance within the “Discover and reuse” phase, collecting all research outputs and findings together in a single platform accessible to other researchers that may use this information in their activities.

TRIMIS: Transport Research and Innovation Monitoring and Information System. The goal of the online platform is to oversee the implementation and effectiveness of the roadmaps developed by the Strategic Transport Research Agenda (STRIA), evaluating technology trends, research and innovation capacities and developments in the European transport sector. STRIA is the interface between transport and other sectors such as energy, information and communication technology. STRIA sets a roadmap on Connected and Automated Transport for road rail and waterborne and builds and integrates in seven thematic transport areas (12):

- Cooperative, connected and automated transport (CAT)
- Transport electrification
- Vehicle design and manufacturing
- Low-emission alternative energy for transport
- Network and traffic management systems
- Smart mobility and services
- Infrastructure

The monitoring nature of this initiative supports research at the stage of findings assessment, allowing evaluating the effectiveness of the research being undertaken and post-assessment corrections to ensure the achievement of its main objectives.

EU ODP – EU Open Data Portal: portal providing free access to a growing array of data from the EU. This data is made available for its re-use, both for commercial and non-commercial activities. Offering accessible, free-of-charge data could foster its further use in innovation, potentially yielding to additional economic advantages. Openness and accountability of the European institutions is also promoted through this platform (25). This portal is most valuable at the data collection stage of the process, making available a wide range of information that can be reused in new research activities.

EOSC: European Open Science Cloud. It provides a vision for a research data ecosystem in Europe empowering data-intensive science and interdisciplinary research. This ecosystem will federate previous research infrastructures and data practices across Europe. The European Cloud will enhance access to such infrastructures and ensure interoperability for a wider range of users. It will boost the use of FAIR data to materialize the benefits of Open Science. Its applicability covers the whole cycle of research. Several projects funded by the EC, e.g. the EOSC Pilot and e-Infrastructures such as eInfraCentral, EOSC-Hub and OpenAIRE, are defining the governance framework and establishing initial EOSC services for 2020. Figure 32 sketches the general structure.

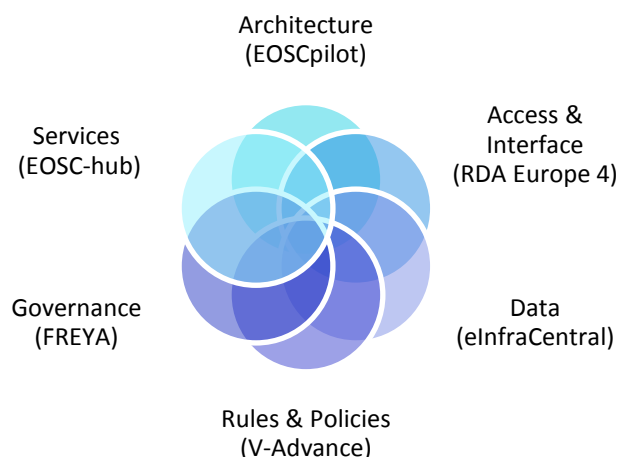


Figure 32: Structural scheme of EOSC

- EOSCPilot [ARCHITECTURE]: H2020 implementation project presenting a Governance framework for the EOSC. This includes areas where common policies are necessary, namely data management, service delivery and Open Science, and acknowledging the influence of the ethical and legal environment. The outcome will be a policy framework and its rules of participation for organisations to work within the EOSC. Its focus lays on the maturity of European Open Science policies and best practices through the cooperation of a wide range of stakeholders to determine actions that need to be undertaken and barriers to be overcome in the evolution of this vision. The barriers are of technical, social and scientific nature. The main driver of the EOSC should be research and therefore, this platform should implement instruments allowing solving scientific issues that cannot be effectively handled in the current situation. (26) EOSCPilot has been a success and using the same principle in the field of research data would be a good way to share the information, knowledge and datasets created. Due to the fact that transports research data is more complex, with big volumes of datasets being created and different stakeholders needing to use the data, the main challenge is the reliability and quality of the data, but also legal aspects. To understand and reuse data, it is important to create a common understanding in transport research based on Open Science that includes different stakeholders.
- eInfraCentral [DATA]: H2020 implementation project bridging research and e-infrastructures. The aim is to establish a single-entry point – the eInfraCentral portal- where end users can access the service catalogue. This will facilitate the supervision of Key Performance Indicators (KPIs) focusing on availability, quality of service and user satisfaction (26). E-Infrastructures foster Open Science by providing the digital services needed by researchers, e. g. networking, computing and data management. This project capitalises on previous investments to create a common catalogue following policy and sustainability guidelines to develop a European e-infrastructure “market place” including a wider range of tools and services. The mission is to ensure knowledge and access of more users to e-infrastructures by 2020 and to create a common attitude to the definition and monitoring of their services. (27)

- Research Data Alliance (RDA): global forum for the development of social and technical infrastructure to facilitate open sharing of data. The infrastructure should boost data-sharing and data-driven research to stimulate the growth of a cohesive data community across all boundaries. (28)
 - RDA Europe 4 [ACCESS & INTERFACE] – H2020 implementation project to build social and technical bridges. It is a European resolution to connect with the RDA to ensure active participation of European stakeholders within RDA activities at political, research, industrial and digital infrastructure level. This is achieved through the foundation of RDA National Nodes in Europe and the adoption of RDA results to ensure interoperability and global sharing of data. (26)
- V-Advance [RULES & POLICIES] – H2020 implementation project to unify metrics and open data. Its mission is to drift scholarly communication towards openness and transparency, finding innovative channels to communicate and oversee research. The range of services offered includes a unified catalogue, metrics dashboard and open data. A network of national nodes connects communities of different disciplines and regions to ensure that the whole society benefits from validated scientific knowledge. (26)
- FREYA [GOVERNANCE] – H2020 implementation project to implement persistent identifiers as core components of open research. The infrastructure for Persistent Identifiers (PIDs) is to be extended as the basis of open research to facilitate data discovery. A new service (PID Graph) will be created through the integration of PID systems into a network, which should also be implemented in the EOSC using disciplinary demonstrator systems. The RDA promotes a community of practice in this area. The project follows on from ORCID, DataCite Interoperability Network (ODIN) and Technical and Human Infrastructure for Open Research (THOR). (26)

The EOSC-hub and OpenAIRE-Advance projects started in January 2018 to play key roles in the development of EOSC. Their combined set of services and resources will supply almost 2 million researchers with data and multidisciplinary research. EOSC-hub provides access to a wide range of interoperable data, computing and networking resources while OpenAIRE-Advance creates a socio-technical network and infrastructure to support open scholar communication. Cooperation between both programmes is likely to generate a common catalogue of e-infrastructure services that cover the full life cycle of research, including data generation, accessibility, use and re-use of information. (29)

- EOSC-Hub [SERVICES]: H2020 implementation project to create an interoperative marketplace for service and data. Numerous service providers are summoned to create the Hub, avoiding the problem of service provisioning fragmentation and granting access to state-of-the-art services (26). The objectives of the project are (30):
 - Simplify access through an open and integrated service catalogue
 - Remove fragmentation of service provisioning through the technical integration and adoption of standards for interoperability of compute, storage, data and software platforms
 - Consolidate e-Infrastructures

- Widen the access and expand the user base
 - Provide a knowledge hub
 - Increase innovation capacity of research e-infrastructures.
- OpenAIRE-Advance: its mission is tightly connected to that of the European Commission, aiming to give unlimited, barrier free, open access to research outcomes funded with public capital in Europe. The tasks are (31):
 - Provide open science services
 - Link research
 - Monitor (open) science
 - Train for open science
 - Build global bridges
 - Facilitate open innovation

Transport Research Cloud: thematic pillar of the EOSC addressing the specific needs of the transport research community. The main goal is to supply researchers in the field of transport and logistics with access to sets of open data concerning their main topics of interest. If the availability and accessibility of data is improved, the possibilities for reuse and remix grow and the process for innovation is optimised. Creating such an environment can also foster the development of public-private relationships between industrial stakeholders and research teams who would otherwise need to find and contact the data source individually. The main challenges identified for such a platform would be the general sensitivity of transport data and its control and the observance of its ownership and use (20). As a thematic pillar of the EOSC, this platform would too cover the whole research process.

EUROSTAT: statistical office of the European Union in charge of handling and publishing reliable statistical data at European level. It strives to create a common understanding that includes technical standards, concepts, methods and structures. The collection of data is performed by the Member States through their respective statistical authorities, where data is validated and analysed before it is handed over to Eurostat. The office's task is to centralise the data and guarantee its interoperability in order to create faithful comparisons between countries and regions (32). Access to reliable statistical data facilitates the analytical aspect of research, optimising research outputs management and boosting the quality of the corresponding findings.

Directory of Open Access Journals (DOAJ): community-curated online catalogue listing open access, peer-reviewed journals. The target is to boost the visibility, accessibility, prestige and use of open access academic research journals irrespective of area of expertise, location or language, with a strong focus on the quality of such research. DOAJ is an independent organisation and is devoted to remain so to ensure keeping all the services available and metadata free to use (33). Publication and dissemination can be improved by means of such a directory, offering further possibilities in terms of findings assessment and availability for subsequent discovery and use by third parties.

Table 3 reviews the aforementioned Open Science sources in terms of competence areas and stakeholders. This is particularly important because it provides an overview of the possibilities that each of the sources presents for the different types of stakeholders with special focus on the

competence areas, so that stakeholders can assess which platforms are most suitable for a particular task regarding a particular competence area.

	Open Science Sources	Competence Areas	Stakeholders
Research Data	CORDIS	Technological, Transport planning Socio-economic, Environmental	Public authorities
	TRIMIS	Technological, Transport planning Business modelling, Socio-economic, Environmental	Public authorities
	EOSC	Technological, Socio-economic	Research
	Transport Research Cloud	Legal/regulatory, Technological Transport planning, Socio-economic	Public Authorities Research Industry
	EU Open Data Portal (EU ODP)	Legal/regulatory, Socio-economic Environmental,	Research
	Registry of Research Data Repositories	Technological, Socio-economic Environmental	Research Public Authorities
Operational Data	EUROSTAT	Legal/regulatory, Socio-economic	Public Authorities Research Industry
Published Research Data	Directory of Open Access Journals	Technological, Socio-economic Environmental	Research

Table 3: Review of diverse Open Science sources by competence areas and stakeholders

TOPOS: Transport fOrum/Observatory for Promoting Open Science. This will be a tool to build a strong knowledge base on the application of Open Science approaches in transport research. The outcomes of the present document and the Work Package it belongs to will build up to its development. The TOPOS fOrum for national and European stakeholders will provide an environment to exchange experiences and best practices for the implementation of Open Science principles in the field of transport. The TOPOS Observatory will be founded on existing efforts like the Open Science Monitor Framework developed in EOSCpilot, FAIR tools, TRIMIS, among others (34). TOPOS will be very useful in terms of publication and dissemination, although such an exchange environment will be most valuable as a platform to discuss and promote best practices when using Open Science in transport research, yielding to the optimisation of the overall system.

4.2 Open Software Sources

Besides Open Data, sources of Open Software should also be discussed, as the data collected for research can also be code in nature and the findings published from a research activity may include source code. Software has become an essential part of research and innovation, as well as of industrial activities. Due to that, certain platforms have become available to offer access to Open Software, the most relevant being:

Software Heritage: this foundation collects preserves and shares all publicly available software in source code form so that applications going from cultural heritage to industry and research can be developed. The origin and full evolution of the software is compiled for the prospective usage of its

metadata. The collected software will be made available as open source software, together with its technical architecture and processes to ensure the sustainability of the foundation. This would become the biggest software source code collection ever compiled, so everything will be indexed, organised and referenced to guarantee accessibility. Through unique identifiers fundamentally bound to the software, a resilient web of knowledge is to be devised over the platform to nurture the development of a variety of services. (35)

GitLab: software management tool consisting of a unique application concentrating the full DevOps cycle. It is a web-based Git-repository manager with an open source license that covers the management and security along the whole life cycle, from planning, through creation, to monitoring. (36)

GitHub: web-based project development platform. It offers particular features like wikis for the different projects and the creation of repositories to store and develop code. It involves a collaborative community and is widely used as host for open source software ventures. (37)

4.2.1 Open Internet

No less important is the role of the Internet within all these activities. To ensure its sustainable development and reliability, several ventures have been undertaken, highlighting:

IETF: the Internet Engineering Task Force is an open international community aiming to improve the functionality of the Internet through the production of high level technical documents. It is considered the premier Internet standards body, developing open standards through open processes. These documents focus on the design, use and management of the Internet. The driving principles are (38):

- Open process: every person can participate and influence the work and all results and procedures are made public online.
- Technical competence: the force issues documents on topics on which they hold the required competence and welcome technically competent input from all sources. The outputs are also expected to present engineering quality, following sound network engineering principles.
- Volunteer Core: both participants and leadership cooperate to further the mission of "making the Internet work better."
- Rough consensus and running code: standards are based on the combination of engineering judgement and real-world experience.
- Protocol ownership: the IETF takes accountability for all aspects of a protocol or function it takes ownership of and, conversely, avoids to exert control over those for which it is not responsible.

The IETF boosts new transport technology by working out new transport protocols for more effective and efficient movement of data. The central role transport protocols play in how the Internet works is reflected by the fact that Transport Area working groups often include participation by technologists involved in other Areas such as Security and Routing.

GIPO: Global Internet Policy Observatory. The European Commission supports an open internet. GIPO is an online platform to facilitate comprehension of the influence and development of the Internet. It shows the support of the European Commission for an open internet. The GIPO tool will guide users through the labyrinthine Internet policies being discussed on Internet governance international events. The concept is based on the crowd-sourcing of knowledge and expertise to create a tool to assist users in understanding and monitoring changes in technology and Internet-policy regulations worldwide. Its impact is expected to be biggest on stakeholders from developing countries, to build their capacity and governance expertise on Internet governance issues. (39)

IGF: the Internet Governance Forum gathers equitably stakeholders from various groups to discuss public policy matters regarding the Internet. The outcome is undefined, as the aim is to inform those with policy-making capacity. It enables a common understanding on how to optimise the possibilities of the Internet and mitigate the associated risks. The mandate of the Forum was established by the UN General Assembly on 2015 covering the following (40):

- Discuss public policy issues related to key elements of Internet governance
- Facilitate discourse between bodies dealing with different cross-cutting policies
- Interface with appropriate inter-governmental organizations and other institutions
- Facilitate the exchange of information and best practices
- Advise all stakeholders in proposing ways and means to accelerate the availability and affordability of the Internet in the developing world
- Strengthen and enhance the engagement of stakeholders
- Identify emerging issues and, where appropriate, make recommendations
- Contribute to capacity building for Internet governance in developing countries
- Promote and assess the embodiment of WSIS (World Summit on the Information Society) principles in Internet governance processes;
- Discuss, inter alia, issues relating to critical Internet resources;
- Help to find solutions to the issues arising from the use and misuse of the Internet
- Publish its proceedings

IoT: Internet of Things is the concept of connecting any device with an on and off switch to the Internet. IoT can be applied to things like transportation networks: "smart cities" which can help us reduce waste and improve efficiency for things such as energy use; this helping us understand and improve how we work and live. (41)

4.3 Open Science Models

For scholar articles publications publishing data under Open Science, there is Open Access. **Open Access** refers to online, free of cost access to peer reviewed publications with limited copyright and licensing restrictions. When using Open Access, the economy and society benefit from it, while helping to improve quality, faster progress to market and greater transparency. International discussions on the main guidelines needed to manage access to research data have been taking place for over 15 years. In 2007 the OECD Principles and Guidelines for Access to Research Data from Public Funding was established and adopted, bringing out the relevance of accessibility to research data and reuse (42). Protecting the available data in the process is vital too. Open Access plays a key

role within the framework ensuring controlled access to the data, which is of paramount importance regarding data sensitivity, proper reuse, intellectual property rights and other legal concerns.

The following models, with different restrictions, are used for publishing in Open Access: (7)

- *gold* – primary publication, article is immediately available. Gold Open Access journals must be listed in the DOAJ, providing open access, peer-reviewed journals;
- *green* – self archiving, mostly not a final version
- *hybrid* - subscription based journals in which individual articles have free Open Access

The next step towards Open Access to peer-reviewed scholarly publications from research funded by public and private grants is to implement Plan S. With defined strict requirements, the aim is to accelerate full Open Access publication. It applies to members who have agreed to **implement Plan S** – an alliance of Funders and Stakeholders called cOAlition S. They have agreed to issue Plan S principles and to provide guidelines for implementation by the end of 2021. It is strongly encouraged to make all research data and other research outputs as open as possible and as closed as necessary. **cOAlition S** supports the DORA Declaration on Research Assessment principles which is focused on improving the way in which outputs of scholarly research are evaluated. (43) (44)

There is also a need to collect and publish the data. There are a large number of databases created for various fields of science and technology. Due to the fact that many of the freely available databases are not separately classified there is poor awareness and accessibility. The objective of the **Listing of Open Access DataBases (LOADB)** is to create a web-enabled, linked, classified and categorized collection of Open Access Databases which one can access from a single portal. Although initial focus is on science and technology subjects, the ultimate aim is to include all subject areas. (45)

5 Common understanding of Open Science in Transport sector

It is important to realise the opportunities that Open Science has to offer. Open Science requires the involvement of all the stakeholders, from all modes of transports, in order to maximise its benefits. Combining identified challenges and available tools and models we can build a framework for the implementation of Open Science in transport research to create a common understanding among stakeholders from different transport modes.

At this stage all the necessary inputs to build the proposed structure for a framework have been analysed. Based on the challenges extrapolated from the mapping of the current situation, the main topics to be addressed by the framework will be identified. Furthermore, the structure will comply with the intrinsic features of Open Science being collaborative and transparent, while providing controlled access, univocal referencing and analytics to capture impact. Once the structure is properly defined, a correlation between the different Open Science sources and their potential for implementation along the different stages of the transport research process will be presented, bringing out the challenges they can overcome. Figure 33 shows a schematic of the inputs and requirements of the framework.

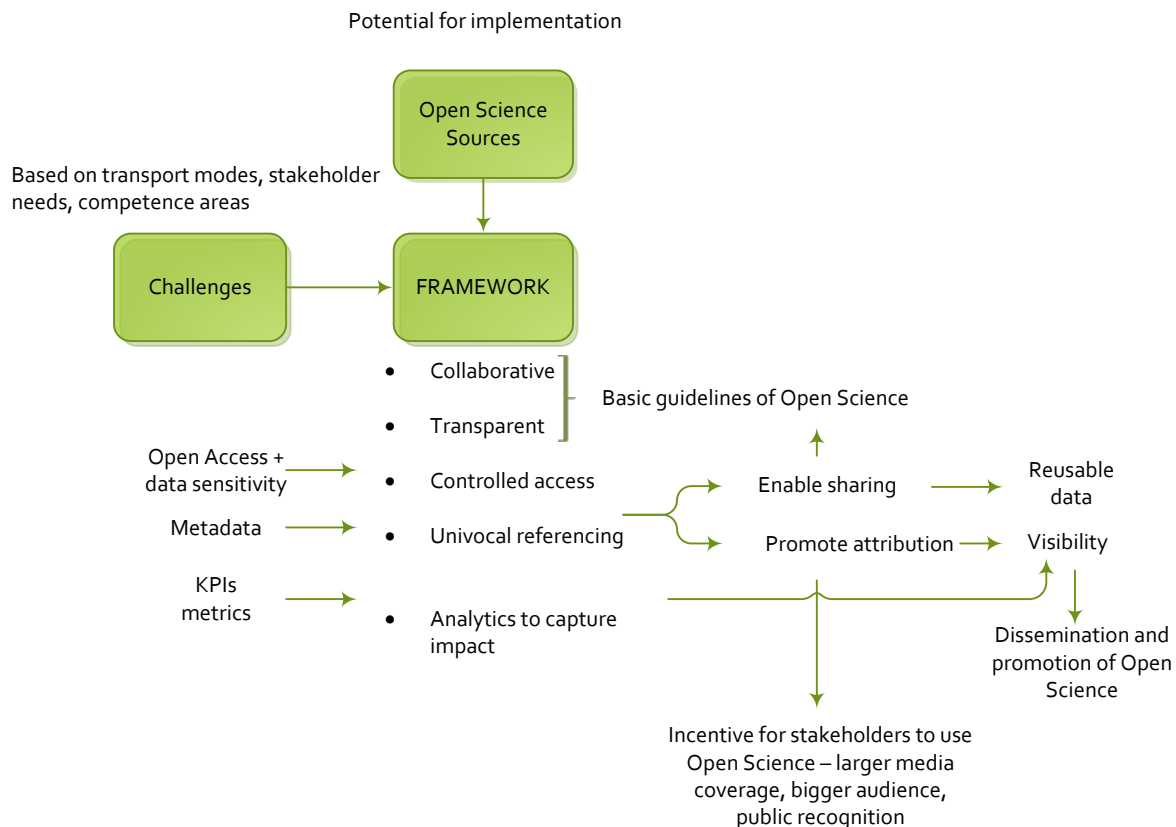


Figure 33. Framework inputs and requirements

5.1 Framework

The main goal of Deliverable 1.2 is to create common understanding and to propose a feasible framework which fosters it, addressing identified challenges by means of available opportunities to use Open Science in the field of transport research, including the interests of all the mode of transport. The framework proposed in the Deliverable 1.2 includes/builds upon the European Commission recommendations to foster and to remove barriers from Open Science, develop research infrastructure and to embed open science in the society (46):

1. The EC should provide clear guidelines for the responsible use of metrics, in the frame of next framework programme (FP9);
2. The EC should encourage the development of new indicators, and assess the suitability of existing ones, to measure and support the development of open science;
3. Assess the likely benefits and consequences as part of a programme of 'meta-research' before introducing new metrics into evaluation criteria;
4. The adoption and implementation of open science principles and practices should be recognised and rewarded through the European research system;
5. The EC should highlight how the inappropriate use of indicators (whether conventional or altmetrics or next generation metrics) can impede progress towards open science;
6. In EU research policymaking, funding and evaluation, metrics derived from private platforms should always be accompanied by open metrics to enable proper validation;
7. Realising the vision for the European Open Science Cloud (EOSC) will rely on linked meta-data that can become the basis for open, publicly available data infrastructure;
8. The European research system and Open Science Cloud should adopt ORCID as its preferred system of unique identifiers, and an ORCID iD should be mandatory for all applicants and participants in FP9;
9. The EC should encourage scholarly publishers across Europe to reduce emphasis on journal impact factors as a promotional tool, and only use them in the context of a variety of metrics that provide a richer view of performance;
10. The EC should identify mechanisms for promoting best practices, frameworks and standards for responsible use of metrics in support of open science;
11. The agenda of this Expert Group should be taken forward by a European Forum for Next Generation Metrics;

12. Over the next 18-24 months, the European Forum for Next Generation Metrics should focus on FP9 and the design of a next generation research data infrastructure addressing the main bottlenecks and listing available tools.

Based on the identified challenges, the proposed framework of common understanding of Open Science in the transport research addresses seven topics (13):

- (1) Policy;
- (2) Guidelines (data and metadata description);
- (3) Data protection and security;
- (4) Training requirements;
- (5) Support and research services;
- (6) Financial models and
- (7) Dissemination of Open Science in transport research

Figure 34 shows the structure of the proposed framework.

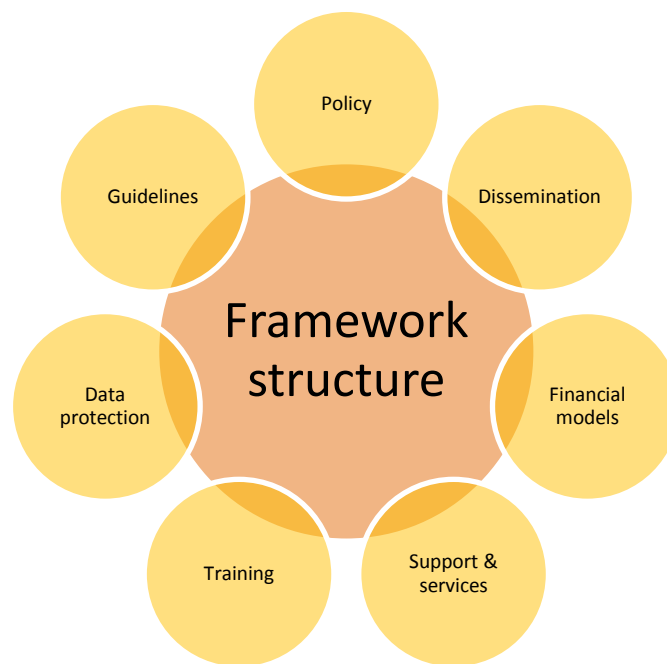


Figure 34: Proposed framework structure

1. Policy

With the geographical barriers between the various stakeholders using Open Science in the transport research, legal challenges have to be addressed. The most important matter is to create a common understanding on using Open Science in the field of transport research among different stakeholders, taking into consideration cultural differences. Secondly, agreements on different levels should establish a legal framework for data sharing, including intellectual property rights. Also, rules for sensitive data need to be considered, where the concept: *“making data as open as possible and close as needed”* (4) applies. Requirements for data should be addressed in the project agreements. The main actors to consider are: the grant agreement, where ownership and access to data should be

defined and distribution of the data established; access methods; research and commercial areas where data usage will be allowed; post project re-(use) of data; post-project financing description of the work; the participant agreement and agreements with data providers. For sharing the data and making it available during and after the project the following criteria have to be described (13) :

- What data is collected
- Where the data will be stored
- Who is responsible for the data
- Who will have access to what data and on what conditions
- The access procedures

This topic concerns primarily the legal/regulatory competence area, requiring the active involvement of policy makers and public authorities while subsequently affecting transport networks, commercial and logistic industry players and ultimately researchers. It also affects the whole research environment, impacting all the steps of the research process, but, most importantly, it ensures transparency, which is a key feature of Open Science.

From the European Commission there are several working groups, initiatives, tools and platforms available to support implementation of Open Science. For example: European Open Science Cloud Strategic Implementation Plan, which paves the road for setting up a global structure for standardization and data repositories with relevant data. (47)

High-Level Advisory Group 'Open Science Policy Platform" (OSPP), whose report was published in 2016 with the main goal to further develop and practically implement open science policies (48)

Other tools available to support policy are: Open Science Policy Platform, V-Advance, EOSCPilot and GO CHANGE (within GO FAIR), which is focusing on priorities, policies and incentives for implementing FAIR

2. Guidelines

Common guidelines enhance common understanding. Explicit guidelines and harmonization for data, metadata and datasets are required for the re-use of data. Furthermore, interoperability is the key word when using Open Science, meaning that data made openly available is easy to find and re-use.

Datasets need to be handled with care and they have to be well defined. There are different types of metadata available: a) Descriptive, b) Structural and c) Administrative – all those aspects need to define the dataset and include a description of measures, performance indicators, time and location segments and their associated values. Also, a detailed description of external resources and the way how data was generated and processed need to be included. This enables univocal referencing, which in turn boosts sharing and promotes attribution. Proper attribution leads to visibility and citations, which enhances the dissemination and promotion of Open Science as a practice, with all the benefits it entails. Attribution can also be considered as an incentive for stakeholders to use Open Science, as they receive larger media coverage, reach a bigger audience and receive public recognition.

Reusable data enhances research advancement, as it allows for validation of previous research activities, minimizes the work to collect data which is already available and enhances collaboration. For re-use of data, the following selection criteria for which data to store could be used:

1. potential for re-use from scientific and also maybe business perspectives
2. efforts needed to store the dataset
3. quality and amount of data
4. funding requirements to plan
5. long-term data preservation and management
6. availability of funding

Well-defined guidelines will affect all stakeholders through all competence areas, facilitating collaboration and transparency. As a result, the whole research process could be optimized. The following tools are supporting the implementation of these guidelines: EU ODP; EUROSTAT; FAIR; European Commission Open Research Publishing Platform; TRIMIS.

3. Data protection and security

To create trust among data providers and researchers the key features are data protection and security. Furthermore, clear data protection rules encourage stakeholders to use Open Science. Data is roughly divided into: sensitive and non- sensitive. Sensitive data is divided into:

- a. **Personal Data** – sensitive personal data, for example research accident databases
- b. **Proprietary data** – if revealed could potentially harm the company

Hereby it is important to carefully analyze the data and restrict as little as possible while taking into consideration the sensitivity of personal and proprietary data. For example, the vehicle and traffic research accident database often includes sensitive personal data. (13)

In addition to data protection rules, data security needs to be addressed. Storing sensitive data in the Cloud creates challenges for security and cyber security rules, which have to be developed to protect the data.

For these reasons, it is important that the proposed framework structure offers controlled access, reflecting on the legal/regulatory competence area.

At European level EOSCPilot uses a stakeholder driven mechanism to support policy implementation and actions required to provide advice and recommendations on any cross-cutting issue affecting Open Science within BE OPEN.

4. Training

Training is important to be aware of the level of data protection needed to handle data securely. It is important when it comes to sensitive data and uploading/handling complex datasets with Intellectual Property Rights. Proper training for all professionals involved is beneficial for all stakeholders in all competence areas, as it enhances the good functioning of the overall system, making all the instruments available throughout the research process. This builds up to a collaborative and transparent framework that ensures adequate control and usage.

At European level there is a working group set in place for skills development, one of the main tasks being to Introduce Open Science education and training. On their latest report they proposed the European Skills and Qualifications Matrix for Open Science and defined the following skills for researchers in order to practice open science (49):

- Can publish under Open Access
- Manage (open) data
- Conduct professional research
- Engage with citizen science

Within Horizon 2020 there are projects fostering the practical implementation of Open Science. The FOSTER webpage is providing e-learning possibilities on Open science. FOSTER Plus contributes to existing materials and co-produces new training content while addressing training activities to relevant stakeholders in the European research area (50). In addition, for FAIR data the initiative GO TRAIN has been started with the aim to coordinate data FAIR awareness and skills (51).

5. Support and research services:

Users of Open Science in transport research may need assistance to familiarise themselves with the system and to implement it to their activities, both as data supplier and user. This includes industry and research stakeholders.

In order to support the services in the aforementioned categories the following tools are available:

- Implementation Roadmap for the European Science Cloud – Communication European Cloud: basis for further consultation with the transport stakeholders for coordinating, supporting and promoting services and tools of Open Science in transport research, an opportunity to overcome the challenge of accessibility and enhance cooperation.
- Interoperability between different databases and repositories. With the technology development and new innovative solutions in the field of transport, next generation repositories need to be developed to overcome constraints to allow for data re-use.
- One of the most important projects to ease reusable data is FREYA, which is funded by the European Commission under Horizon 2020. The main goal of the project is to improve discovery, navigation, retrieval and access to research resources. FREYA together with European Open Science Cloud will contribute to Open Science by open and seamless services for storage, management, analyses and re-use of research data. In addition FREYA collaborates with the EOSC-projects OpenAIRE Advance and EOSC-hub. (52)
- Other tools available: Transport Research Cloud (TRC); eInfraCentral; Next generation repository; EOSC-Hub, OpenAIRE-Advanced; GO-BUILD – coordinating FAIR technology. Designing and building the technical standards, best practices and infrastructure components needed to implement the FAIR data principles; Transport Research Cloud (TRC); Registry of Research Data repositories; EOSC-Hub – services (H2020); RDA access and interface (H2020)

6. Financial support

The collection of large datasets involves a huge amount of effort and resources. To further benefit from these datasets and make better use of the invested resources, the datasets could be made available for data re-use. Funding is extremely important for large datasets. For example, the following actors are part of the data management costs:

- data selection, enhancement of documentation
- management & coordination personnel costs
- IT operations
- (metadata), creation of entries in relevant data catalogues
- analysis or data handling facilities
- analysis support services
- promotion and advertisement
- standardization and collaboration regarding dataset formats (optional)

Research funding agencies have high impact on this topic, as it is them implementing the mechanisms and requirements that need to be complied with so that one may receive funding for research. This is of particular importance in the area of business modelling, involving too policy makers and public authorities as potential sources of funding, industry stakeholders who may both provide and require financial support and research institutions who may be granted funds to expand their capabilities or support individual researchers expand theirs. Within the research process, most of the impact will concentrate in the stage of research planning, assisting on funding and resourcing research.

It is also worth noting that the use of Open Science in transport research can take ideas faster to the market, which in turn lowers development costs.

The EC has been funding several projects related to Open Science, but a funding strategy for its implementation could hold further potential. On the other hand, the European Commission report about “Cost-benefit analysis of FAIR research data” brings out that the EU could save €10.2 billion per year by using FAIR data, therefore commitment to support Open Science could be considered financially beneficial for all the stakeholders. (53)

7. Dissemination of Open Science in transport research

It is important to bring different stakeholders from different modes of transport together and to encourage the use of Open Science. Dissemination and promotion of Open Science in transport research increases its potential impact, speeding up innovation and discovery, which may lead to solutions to societal challenges in the different competence areas. This yields to taxpayers’ return on investment, as public funds are used to find such solutions that benefit society as a whole.

The following tools are available to support the dissemination: Research Data Alliance (RDA) - RDA Europe 4; Implementation Roadmap for the European Science Cloud – Communication European Cloud; FREYA; V-Advance; Transport fOrum/ Observatory for Promoting Open Science - TOPOS which is in the scope of the BE OPEN project.

5.2 Metrics

Metrics are used to capture the impact of the processes implemented. The evaluation and measurement of processes and activities is important in order to make changes in the system. To optimise the processes, metrics will help to identify areas where changes should be implemented.

In the field of transport, transportation metrics are used for evaluating the efficiency of transportation processes. On Table 4 the list of basic transport metrics is provided in relation to competence areas. (54)

Competence area	Metrics	Description
Business modelling	GHG emission factors:	Calculated by the division of emitted amount of GHG (CO ₂ , methane and nitro-oxide products) in grams with the kg of fuel consumed per passenger-km.
	Time effect	The comparison of the different means from the time effect point of view must be based on the door to door speed, or total traveling time.
	Cost evaluation	Direct and indirect costs determined for total time of usage.
Environmental	Weather factors:	Showing all factors affecting environmental conditions like temperature, wind speed, humidity, UV exposure rate etc.
	GHG emission factors:	Calculated by the division of emitted amount of GHG (CO ₂ , methane and nitro-oxide products) in grams with the kg of fuel consumed per passenger-km.
	Noise Exposure	Number of people exposed to significant noise (regardless of whether their houses or apartments have been sound-insulated). Significant aircraft noise levels are currently defined as values greater than or equal to Day-Night Average Sound Level (DNL) 65 decibels
Legal/ Regulatory	Successful quality inspection/evaluation rate:	Shows the successful maintenance or quality control inspections made divided by the total number of quality inspections.

Socio-economic	Economic values of investment:	Define the outcome of existing investments focusing on translating social, environmental and technological impacts in monetary values. These indicators are return on investment (ROI), vehicle life cycle etc.
	Traffic accidents:	Adjusted according to transport mode and is measured with type of accident (light, heavy, fatal) per a defined time period (month, year, decade etc.).
	Healthcare factors	Measure and collect all relevant data on human body health behaviour (blood pressure, heart rate etc.)
Technological	Capacity utilization:	Identifies the usage level of existing physical and digital infrastructure and assets.
Transport planning	Transport performance measurement factors	Used to measure: Time utilization for vehicles and for Transport Trips, Transport way utilization, Weight of transported freight, Fuel usage relating to vehicles and transport trios
	Cost evaluation	Direct and indirect costs determined for total time of usage.

Table 4: Transport metrics in relation to competence areas (54)

Measuring efficiency in the transportation system is a difficult task as in the different transport modes, different characteristics have influence. Efficient could be meant as performing or functioning in the best possible manner. In general, efficiency could be divided into two different levels: 1) level of interest and 2) level of dimension. Within the first category, evaluation of the transport system can be done from: energy, cost individual/personal requirements, society and community point of views. Figure 35 brings out different levels of efficiency in the transportation system:

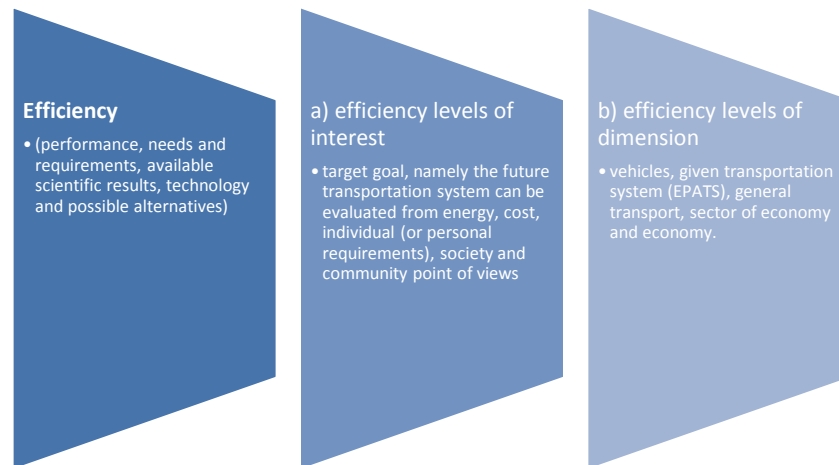


Figure 35. Efficiency measurements in transportation systems (55)

Efficiency is important for sustainable economy and implementation of new, integrated transport systems. The other important factor is cost, which could be calculated by taking into account different fields: energy, individual interest, society and community. For example users choose a transport system as a combination of different transportation means, minimizing the traveling cost. Figure 36 illustrates the cost evaluation hierarchy, where sustainable economy accounts for all costs, of transportation of both people and goods.

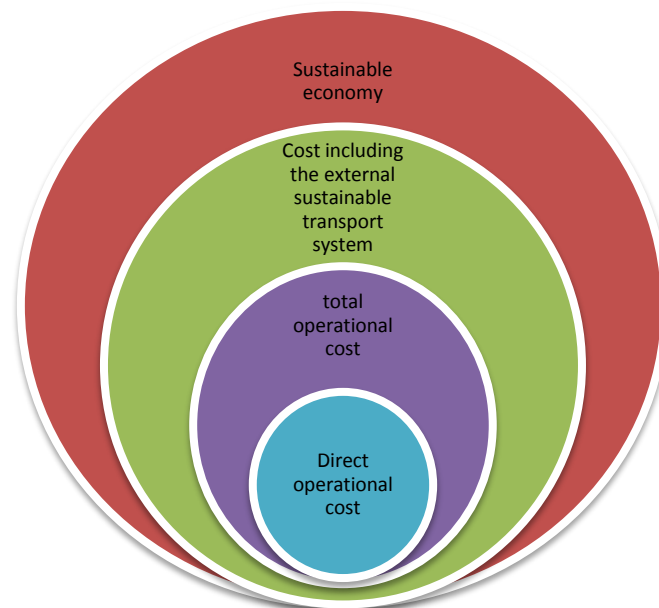


Figure 36: Cost evaluation hierarchy (55)

Beyond classical efficiency calculations, there are new technologies and platforms to help us understand modern transport networks. One of them is the IoT – Internet of Things - environment, with multiple sensors and devices monitoring different metrics and producing huge amounts of data (56). The Figure 37 presents a typical topology of an IoT system. According to this topology, any IoT system is equipped with private devices such as smart phones and cars as well as shared devices such as Raspberry Pis deployed in the field. The IoT systems are also equipped with sensors. These sensors

can be dedicated to particular devices, such as a speed calculator and a GPS sensor dedicated to a smart car, and blood pressure and heart rate sensors linked to a patient’s smart phone. Some other sensors can be shared by multiple devices such as traffic jam detectors and surveillance cameras. Dedicated sensors are typically connected directly to their devices (without the need for a network), while shared sensors can be connected to any device through a network. For example, a smart phone can connect to a surveillance camera through a network, while it has direct access to the blood pressure sensor (through a local or private network, which is not represented in Fig. 11). Likewise, a smart car has direct access to its speed calculator and GPS sensors while it connects to traffic jam detectors through a network. Consequently, dedicated sensors do not use any of the network bandwidth capacities of their dedicated devices. All devices are provided with resources to collect, process, and disseminate IoT data. For example, devices are equipped with resources for computing (CPUs), for storage (such as disks), for communication (e.g., network interfaces), and so on.

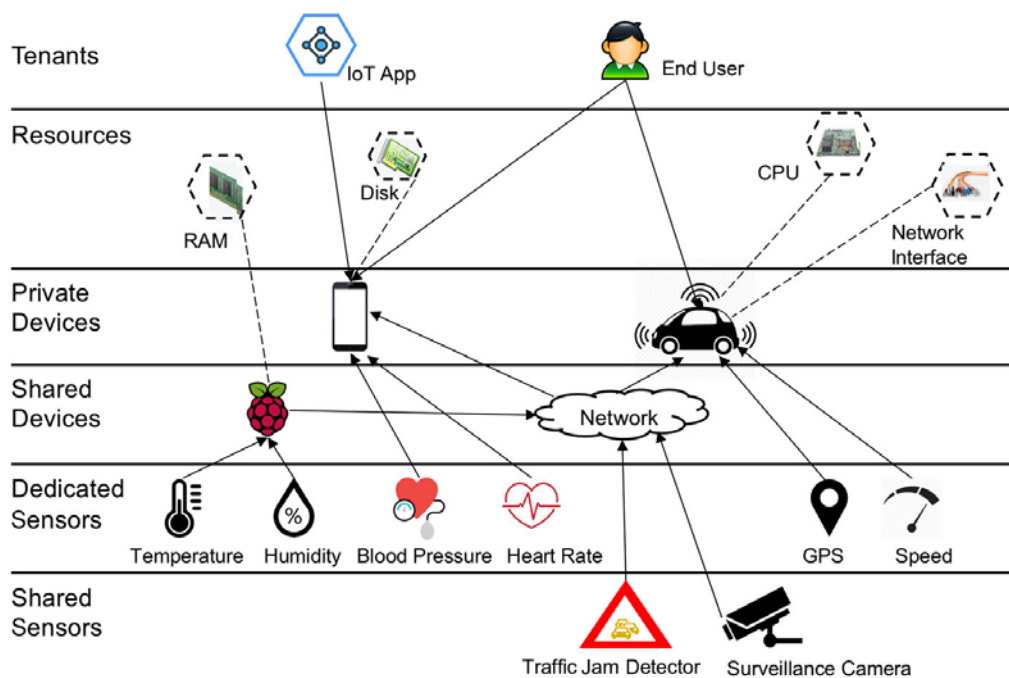


Figure 37: Typical topology of IoT systems (57)

Metrics usage for Open Science is also highly relevant, especially for the usage of publications, but also for tracking non-traditional publications (posts, blogs) and for the re-use of open data or open software. Several publishers, (e.g. Springer Nature, IEEE, ACM) display the number of downloads of the specific article from their platform. Elsevier’s Science Direct in cooperation with Mendeley provides information to researchers on the number of downloads of their publications from the Science Direct platform. These indicators serve as valuable function: when they are used responsibly, they are the best quantitative measures available to assess scholarly influence, mainly of publications in journals currently available. The role of metrics in supporting open science is (46):

- a) monitoring the development of the scientific system towards openness at all levels;
- b) measuring performance in order to reward improved ways of working at group and individual level;

These goals require the development of new indicators, as well as prompting the use of existing metrics in a more responsible fashion.

In Open Science the most used metrics are Alternative metrics, also known as **Altmetrics**, research indicators based on social media. The other metrics used are: **Bibliometrics** - Citation and content analysis used in Open Science, **Semantometrics** – with the goal to exploit primarily the full text of manuscripts to assess the value of a publication and **Webometrics** - The science at which there is an attempt to compute the extent of the World Wide Web. The interrelation of these metrics is depicted in Figure 38.

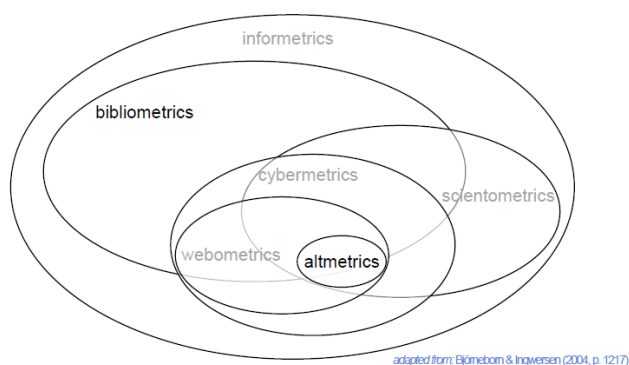


Figure 38: Basket of metrics for evaluation of science (46)

One of the central issues of Altmetrics is the identification of the engagement of the community with scholar content in social media and therefore the understanding of users and usage of social media in the context of scholarly communication and research data. (58) Altmetrics benefits can be grouped into three categories:

- a) formats of relevance: altmetrics can identify new formats of scholarly products to measure, which have not been considered in research assessments before, e.g., research data and software;
- b) forms of impact: these refer to the new audiences captured, who interact with or react to scholarly products and scenarios related with that, e.g., policy makers and policy documents;
- c) targets and uses: these reflect the purposes for which altmetrics can be used, e.g., budget allocation or self-assessment and career development.

Altmetrics are not sufficient for the use in Open Science and they need to be complemented with metrics and framework. The challenges of altmetrics are: lack of robustness, legal constraints, lack of free access, unclear basis and structural problems. Therefore there is a need to improve the available metrics in Open Science. The new metric “open metrics” would complement altmetrics with following the criteria (46):

- a) research products and data sources for metric development need to be logically selected, open documented, and chosen in line with the disciplinary norms;
- b) data that underlies metrics, indicators, and measurements need to be open and accessible (preferably via automatic processes, e.g. API);
- c) provision of software that was used for calculations;

- d) logical, scientific, and documented explanation of how data were derived and metrics calculated.

In other words, there remains scope for improvement and a need for what we call next generation metrics that measure, reward and create incentives for open science. Based on the review on Next generation metrics for Open Science it is concluded that metrics for open science is a complex system, where one size fits all solution is not applicable. There is need to perform more research for making better use of available metrics. In addition many altmetrics used are not reflecting transition towards open science and need to be supported by other metrics. Development of next generation metrics should be based on user needs and defined in terms of: robustness, humility, transparency, diversity and reflexivity. In addition, we should focus on dynamic changes in the capabilities and popularities of different social media and other web platforms.

EC has developed the European Open Science Agenda to implement the following recommendations (10):

1. foster open science
2. remove barriers to open science
3. develop research infrastructures for open science
4. embed open science in society

From some national journals and national or institutional repositories, altmetrics are not widely used in the Member States. Nevertheless in different countries there is work ongoing to develop new approaches for scientific evaluation. On Table 5 some examples from different countries' experience is provided.

Country	Initiative	Description
France	SJS	<i>Self-journal of Science</i> is an open repository as well as a new paradigm of scientific publication. With very low cost SJS implements an optimal peer review, which itself becomes a measurable process, and builds an objective and unfalsifiable evaluation system. (59)
	HAL	French National repository, which is used for archiving and dissemination of scientific literature, published or un-published, from universities or research institutions in all disciplines of human and social sciences. (60)
	Open Edition Books	Web platform for books in the humanities and social sciences. More than half of them are available in Open Access. Additional services are available through subscribing libraries and institutions. (61) Open Edition Books is part of HIRMEOS project (62)
Croatia	CRIS model	Current Croatian research information system, which contains structured descriptions of its basic components, and at the same

		time enables open access to archived publications
	CROSBIB	Croatian Scientific Bibliography, which collects the data on scientific output of the current research projects financed by the Ministry of Science, Education and Sports (MSES) and to make them publicly available in open access. (63)
	FULIR	Full-text Institutional Repository of the Ruđer Bošković Institute, which contains full-text publications written by Institute staff. (64)
Belgium	University of Liège's Green open access repository	Which the evaluation is only based on what has been submitted to the repository by the researcher him- or herself.
Portugal	Liège model	Good example of adopting open science. Within this, the deposit and availability of research outputs in the institutions' repositories are inextricably linked (conditionally) to research assessment procedures for career progression purposes.
Slovenia	COBISS/SciMet	Through this service, researchers can monitor the performance of their publications by using different altmetrics and more traditional metrics. This was considered as positive example. (65)

Table 5: Open Science usage in Member States

6 Main Outcomes

As stated in the beginning of section 4, Open Science comprises all of the research process, from discovery to publication and dissemination. In such an ecosystem, stakeholders from all levels should be included to foster a framework of common understanding that accounts for the needs of all of them, open to exchange and discussion to further exploit its potential. In terms of transport research, this exchange between stakeholders from different transport modes and competence areas can speed up the development and implementation of new transport concepts based on interoperability that can satisfy the user's needs at every step of the journey, building up on sustainable development and assisting in overcoming societal challenges. For this reason, Deliverable 1.2 connects transport modes, stakeholders and competence areas to extrapolate the main challenges to be addressed by the framework structure. Deliverable 1.2 is therefore complementary to Task 1.1 (1), where key actors, namely industry, research, public authorities and society in general together with competence areas were analysed.

The aim of the structure of a framework proposed is to provide the resources necessary to overcome the main challenges. To validate the potential of Open Science to achieve that, different sources of Open Science were analysed and categorised in terms of the type of transport research data they provide, the key actors and competence areas involved and at which step of the research process they can be of assistance. Issues like governance models, legal concerns, data security, infrastructure and training requirements and standards are addressed in this manner. Finally, the different sources are classified in terms of the challenges they can help overcome, proving the potential benefits of Open Science implementation in transport research.

The outcomes of Deliverable 1.2 can help build up prospective parts of the BE OPEN project. Capitalizing upon the initiatives of Open Science in transport research at European level as identified in D1.2, Task 3.2 "Set up of the European Open Science in Transport Forum" and Task 3.3 "Set up of Open Science in Transport Observatory" will provide a transparent forum and observatory, TOPOS, to act as an evidence-based, community driven sharing of knowledge and experiences in order to enable strategic planning and networking between different countries and research programmes covering the entire transport research network. Furthermore, Task 5.1 "Identification of main challenges, opportunities, constraints and bottlenecks of Open Science in transport research" can use the main challenges identified here as a starting point for their work and Task 5.2 "KPIs for Open Science in transport evaluation" will further develop the necessary tools to analytically capture the impact of Open Science in transport research. Deliverable 1.2 introduces relevant metrics in terms of transport research and Open Science as a means to that end, but it is important to develop specific KPIs to properly evaluate performance and assess critical success factors such as transport research policies, transport research publications, transport research infrastructures and software, etc. This will subsequently be connected to Task 5.3 "Impact assessment of Open Science in transport". Moreover, the outcomes of the survey can be used as a starting point for Task 1.3 "Transport stakeholder needs and objectives".

This Task proposes a structure for a framework of common understanding that addresses the main challenges encountered in the implementation of Open Science in transport research, analysing

initiatives with the potential to overcome them. However, further work is required concerning the actual implementation of these ideas and its operational aspects.

6.1 Potential for implementation in transport research

The proposed framework defines relevant topics to overcome the identified challenges and points out opportunities that can unravel the potential of current Open Science initiatives within the topics. Such opportunities are summarised in Table 6.

Identified Challenge	Framework Topics	Opportunities
Lack of skilled experts	Training requirements	GO-Train, European Skills and Qualifications Matrix for Open Science. FOSTER Plus
Legal challenges	Policy	Policy development to create common understanding EOSCPilot Open Science Policy Platform V-Advance GO CHANGE EU ODP EUROSTAT
Data quality	Explicit guidelines	FAIR European Commission Open Research Publishing Platform TRIMIS Transport Research Cloud (TRC) eInfraCentral
Technological challenge	Support and research services	Next generation repository FREYA EOSC-Hub OpenAIRE-Advanced GO-BUILD – coordinating FAIR technology
Funding	Financial schemas	EC initiative to support Open Science
Enhancing data security and privacy principles.	Data protection and security	Cyber security framework EOSCPilot



Fragmented data and large variety of stakeholders:	Policy and Guidelines, Dissemination of Open Science in transport research data	Research Data Alliance (RDA)/ RDA Europe 4 Transport fOrum/ Observatory for Promoting Open Science - TOPOS Implementation Roadmap for the European Science Cloud – Communication European Cloud FREYA V-Advance
---	--	---

Table 6: Classification of Open Science sources according to their potential impact

7 References

1. **Michela Fioretto (FIT), Afroditi Anagnostopoulou (CERTH).** *D1.1: Taxonomy of actors, terminology and experimental tools.* 2019.
2. **George Yannis (NTUA), Katerina Folla (NTUA).** *D2.1: Open access publications and the performance of the European transport research.* 2019.
3. **Natalia Manola (ATHENA), Nikolaos Angelopoulos (Konnekt-able).** *D2.2 Open/FAIR data, software and infrastructure in European transport research.* 2019.
4. **EUROPEAN COMMISSION, Directorate -General for Research & Innovation.** *Guidelines on FAIR Data Management in Horizon 2020.* 2016.
5. **Center for Open Science.** Open Science Framework (OSF) - Resource Review. *Journal of the Medical Library Association.* 2017.
6. **Merja Hoppe, Thomas Trachsel.** *Future transport research agenda: Guidelines for a forward looking transport.* s.l. : ResearchGate, 2018.
7. **European Commission.** Foster Open Science. [Online] 2014. <https://www.fosteropenscience.eu/>.
8. **European Union.** Mobility4EU. [Online] <https://www.mobility4eu.eu/project/>.
9. **European Commission.** *Horizon 2020: Work Programme 2018-2020.* 2019.
10. —. *EOSC Strategic Implementation Roadmap 2019-2020.* 2018.
11. **EUROSTAT; European Commission.** *Illustrated Glossary for Transport Statistics 4th edition.* 2009.
12. **European Commission.** *STRIA Roadmap on Connected and Automated Transport - Road, Rail and Waterborne.* Brussels : s.n., 2019.
13. **Helena Gellerman, Erik Svanberg, Yvonne Barnard.** *Data sharing of transport research data.* s.l. : Science Direct, 2016.
14. **European Union.** EuroFOT. [Online] <http://wiki.fot-net.eu/index.php/EuroFOT>.
15. **European Commission, DG Research and Innovation.** UDRIVE. [Online] <http://www.udrive.eu/index.php/about-udrive/overview>.
16. **European Union.** <https://www.fosteropenscience.eu/>. [Online] 2017.
17. **OECD.** *Measuring the Digital Transformation: A Roadmap for the Future.* Paris : s.n., 2019.
18. —. *Technology and Industry Scoreboard 2015: Innovation for growth and society.* Paris : s.n., 2015.
19. **European Commission.** *Open Innovation, Open science, Open to the World - a vision for Europe.* Brussels : s.n., 2016.

20. **European Commission.** *Analysis of the state of the art, barriers, needs and opportunities for setting up a transport research cloud.* Brussels : s.n., 2018.
21. **The Open Science and Research Initiative (ATT).** *Framework for Open Science and Research.* 2016.
22. Registry of research data repositories. [Online] <https://www.re3data.org/>.
23. **COAR - Confederation of Open Access Repositories.** *COAR Roadmap - Future Directions for Repository Interoperability.* 2015.
24. **European Commission.** CORDIS. [Online] <https://cordis.europa.eu/>.
25. **European Union.** EU Open Data Portal. [Online] <https://data.europa.eu>.
26. **The European Open Science Cloud for Research Pilot Project.** *Principal Findings from the EOSCpilot Project.* 2019.
27. **eInfraCentral - European Commission.** eInfraCentral. [Online] 2017. <http://project.einfracentral.eu>.
28. **Research Data Alliance.** Research Data Alliance. [Online] <https://rd-alliance.org>.
29. **EOSC-hub - European Commission.** EOSC-hub - collaborations. [Online] <https://eosc-hub.eu/collaborations/openaire-advance>.
30. —. EOSC-hub. [Online] <https://eosc-hub.eu>.
31. **OpenAIRE.** OpenAIRE - mission and vision. [Online] <https://www.openaire.eu/mission-and-vision>.
32. **EUROSTAT - European Commission.** eurostat - your key to European statistics. [Online] <https://ec.europa.eu/eurostat>.
33. **Directory of Open Access Journals (DOAJ).** <https://ec.europa.eu/eurostat>. [Online] 2019. <https://doaj.org>.
34. **BE OPEN - European Commission.** BE OPEN project - TOPOS. [Online] 2019. <https://beopen-project.eu/topos>.
35. **Software Heritage.** Software Heritage - mission. [Online] <https://www.softwareheritage.org/mission/>.
36. **GitLab.** GitLab. [Online] <https://about.gitlab.com/>.
37. **GitHub.** GitHub. [Online] <https://github.com/>.
38. **IETF.** Internet Engineering Task Force. [Online] <https://www.ietf.org/>.
39. **European Commission.** Global Internet Policy Observatory - GIPO. [Online] <https://ec.europa.eu/digital-single-market/en/global-internet-policy-observatory-gipo>.

40. **IGF - United Nations.** Internet Governance Forum. [Online] <https://www.intgovforum.org>.
41. *A simple explanation of "the Internet of Things".* **Morgan, Jacob.** s.l. : Forbes, 2014.
42. **OECD.** *Principles and Guidelines for Access to Research Data from Public Funding.* 2007.
43. DORA. [Online] <https://sfdora.org/>.
44. **Science Europe AISBL.** Coalition-S. [Online] <https://www.coalition-s.org/principles-and-implementation/>.
45. **Unit for Research and Development of Information Products (URDIP)** . Lisitng of open Access Databases. [Online] http://www.loadb.org/Control.do?_idx.
46. **European Commission.** *Next-generation metrics.* 2017.
47. **EOSC - Executive Board.** *European Open Science Cloud (EOSC) - Strategic Implementation Plan.* 2019.
48. **European Commission.** Open Science Policy Platform. [Online] <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>.
49. —. Working Group on Skills. [Online] https://ec.europa.eu/research/openscience/index.cfm?pg=skills_wg.
50. **FOSTER.** FOSTER Open Science. [Online] <https://www.fosteropenscience.eu>.
51. **GO FAIR.** GO FAIR initiative. [Online] <https://www.go-fair.org/go-fair-initiative/>.
52. **FREYA - European Commission.** Project FREYA. [Online] <https://www.project-freya.eu>.
53. **European Commission.** Open Science. [Online] <https://ec.europa.eu/research/openscience/index.cfm>.
54. **Studies, GY Kovács - Polish Journal of Management.** *DEVELOPMENT OF PERFORMANCE EVALUATION SOFTWARE.* 2017.
55. **Rohacs, Prof. Dr. Jozsef.** *EVALUATION OF THE AIR TRANSPORT EFFICIENCY.*
56. **Tata, Samir, Mohamed, Mohamed and Megahed, Aly.** *An Optimization Approach for Adaptive Monitoring in IoT Environments.* 2017.
57. **Aly Megahed, Jennifer A. Pazour, Ahmed Nazeem, Samir Tata, Mohamed Mohamed.** Monitoring services in the Internet of Things: an optimization approach. *Computing, Spiegel Nature.* 2019.
58. **Sugimoto, Cassidy, et al., et al.** Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology.* 2016.
59. **Bon, Michael.** *Principles of the Self-Journal of Science: bringing ethics and freedom to scientific publishing.* 2015.

60. **HAL-SHS Sciences de l'Homme et de la Société.** HAL-SHS Sciences de l'Homme et de la Société. [Online] <https://halshs.archives-ouvertes.fr/>.
61. **Open Edition Books.** Open Edition Books. [Online] <https://books.openedition.org/?lang=en>.
62. **European Commission.** High Integration of Research Monographs in the European Open Science infrastructure. [Online] <https://cordis.europa.eu/project/rcn/206340/factsheet/en>.
63. **Croatian Scientific Library. BIB.** [Online] <http://bib.irb.hr/?lang=ENand>.
64. **FULIR.** Full-text Institutional Repository of the Ruđer Bošković Institute. [Online] <http://fulir.irb.hr/information.html>.
65. **SciMet.** SciMet. [Online] <http://scimet.izum.si/>.
66. **European Commission.** *Cost of not having FAIR research data.* 2018.
67. **Cassidy R. Sugimoto, Sam Work, Vincent Larivière, tefanie Haustein.** *Scholarly use of social media and altmetrics: a review of the literature.* 2016.
68. **SCM wiki.** [Online] <https://scmwiki2012.wordpress.com/r/transportation-metrics/>.
69. **Organisation for Economic Co-operation and Development.** Organisation for Economic Co-operation and Development. [Online] 2019. <http://www.oecd.org/science/inno/open-science.htm>.
70. **OECD.** *Measuring the Digital Transformation: A Roadmap for the Future.* Paris : s.n., 2019.



8 Annex – BE OPEN Survey

List of survey questions



Dear Colleagues, We would like to invite you to take part in the BE OPEN project by completing this survey on challenges and opportunities of Open Science in transport research.

The BE OPEN project is a coordination and support action (CSA) sponsored by the European Commission in the Horizon 2020 framework. BE OPEN aims to create a common understanding of the practical implementation of Open Science in transport research and to develop a code of conduct for the entire research family. The Objective of this survey is to determine the state of play and to identify challenges and opportunities related to Open Data, Open science and Open access in the field of transport where:

Open Data - Research data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike; Open Science - The practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods and Open Access - Published research outputs that are made available online, free of charge, and free of most copyright and licensing restrictions. We would like to learn more about your experience and expectations concerning Open Data, Open Science and Open Access. Participation in this survey is of course voluntary and can withdraw at any point during the study. Your contact data will be retained for the duration of the project BE OPEN but it will neither be disclosed nor passed on to any third parties. The survey takes about 10 minutes to complete. If you have any questions about this survey or project, please e-mail to kristel.palts@dlr.de. We look forward receiving your feedback by the 24-th of June. Thank you very much for participating!

Please provide some personal information

A1. Gender

male

female

A2. Age range

up to 30

between 31 - 50

between 51 - 65

66 and above



A3. Years of professional experience

- up to 5
- between 5 - 10
- between 10 -15
- between 15-20
- more than 20

Please provide information for your organisation

B1. Please choose the mode of transport

- Rail
- Road
- Air
- Water
- Other

Other

B2. Your organisation

B3. Please choose type of your organization

- Research Establishment
- University / Academic Institution
- Regulatory Authority / Government (e.g. EASA)
- Private research center
- Transport Networks (e.g. Air navigation service providers etc.)
- NGOs community organisations or intergovernmental organisations (e.g. EUROCONTROL, etc.)
- Commercial transport and logistics industry players (e.g. Airports, Industry, etc.)



Other



Other

B4. Please choose the country of your organization

Please choose only ONE. If your organization operates in more than one country please select the one which represents the largest proportion of staff employment.

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxembourg
- Malta
- Netherlands
- Poland
- Portugal
- Romania





- Slovakia
- Slovenia
- Spain
- Sweden
- United Kingdom
- Non-EU country

B5. Please choose the size of the entity

- Small (10 to 50) people employed
- SME (50 to 249) people employed
- Large (250 and above) people employed

B6. Please specify your area of expertise

- Legal / regulatory
- Technological
- Transport planning
- Business modeling
- Socio-economic
- Environmental
- Maintenance

Please provide some information about your experience with Open Data, Open Science and Open Access

C1. Please indicate how strong you and your organisation relate to the following statements

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
The smart people in our field work for us.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not all the smart people work for us. We need to work with smart people inside and outside our company.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
To profit from R&D, we must discover it, develop it, and ship it ourselves.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
External R&D can create significant value. Internal R&D is needed to claim some portion of that value.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If we discover ourselves, we will get to market first.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
We don't have to originate the research to profit from it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Strongly agree Somewhat agree Neither agree nor disagree Somewhat disagree Strongly disagree

The company that gets an innovation to market first will win.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Building a better business model is better than getting to market first.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If we create the most and the best ideas in the industry, we will gain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If we make the best use of internal and external ideas, we will win.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
We should control our IP, so that our competitors do not profit from our ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
We should profit from others use of our IP, and we should buy others IP whenever it advances our own business model.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C2. What type of mechanism is used in your organisation?

Outside-In (Licensing in, acquisition, short term fellowships, spinning-in, venturing, user driven innovations, innovation driven by supplier competitions)

Inside-out (Licensing out, divestments, R&D for other companies, spinning-out)

Coupled Processes (Alliances, cooperation, joint venture, joint R&D , publicly funded R&D, private R&D, Co-Creation)

C3. Do you feel well-informed about:

Strongly agree Somewhat agree Neither agree nor disagree Somewhat disagree Strongly disagree

Open Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Open Data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Open Software and tools	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Open Access Publication	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C4. Please declare the level of your agreement to the statement: "It is common practice in our organization to use following: a, b, c, d"

Strongly agree Somewhat agree Neither agree nor disagree Somewhat disagree Strongly disagree

a) Open Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Open Data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Open Software and tools	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Open Access Publication	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C5. Do you work with Open Data? If yes, please specify

Analyse

Curate

Provide data to others



Creation of data

Process of data

Usage of data

I do not use Open Data

C6. How do you use Open Data within your organization / within your field of expertise? Please give some examples.

C7. What type of Open Data / Open Science are you working with?

Legal

Technological

Transport planning

Business modelling

Socio-economic

Environmental

Other

Other

C8. Which type of Open Data platforms would you prefer to use?

European (e.g. Open Science Cloud, Open Aire etc.)

National platforms

Virtual Research Environment

I am not using any platforms



Other



Other

C9. Please declare the level of your agreement to the following statements:

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
My organization would benefit from Open Data / Open Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I see Open Data / Open Science as an opportunity in precompetitive research	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I feel that there are no limitations to use Open Data or Open Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
My organization would benefit from using Open Data / Open Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C10. Would you have any additional comments on the benefits, opportunities and limitations of Open Data / Open Science?

C11. Are there any examples of Open Data / Open Science in Your field?

C12. What would be your requirements for making Your experience publicly available?

Monetary benefit



Comment



Central repository



Comment

Time



Comment

Common Policy



Comment

Immaterial benefit



Comment

Other



Other

D1. Thank you for participating in BE OPEN! Your input is very valuable for successful continuation of Open Data/ Open Science in the field of Transport. If you would like to receive the results of our project, please leave your e-mail address. If you would have any comments on Your experience with our survey, please do not hesitate to leave a comment below.



Thank you very much for participating in our survey!