BE OPEN

Kick off Meeting
14 February 2019
Brussels

Open Science in Transport

G. Yannis*, M. Böhm, J. R. Franklin, S. Jones,
T. Kovacikova, K. Nowicka

*Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece

Presentation Outline

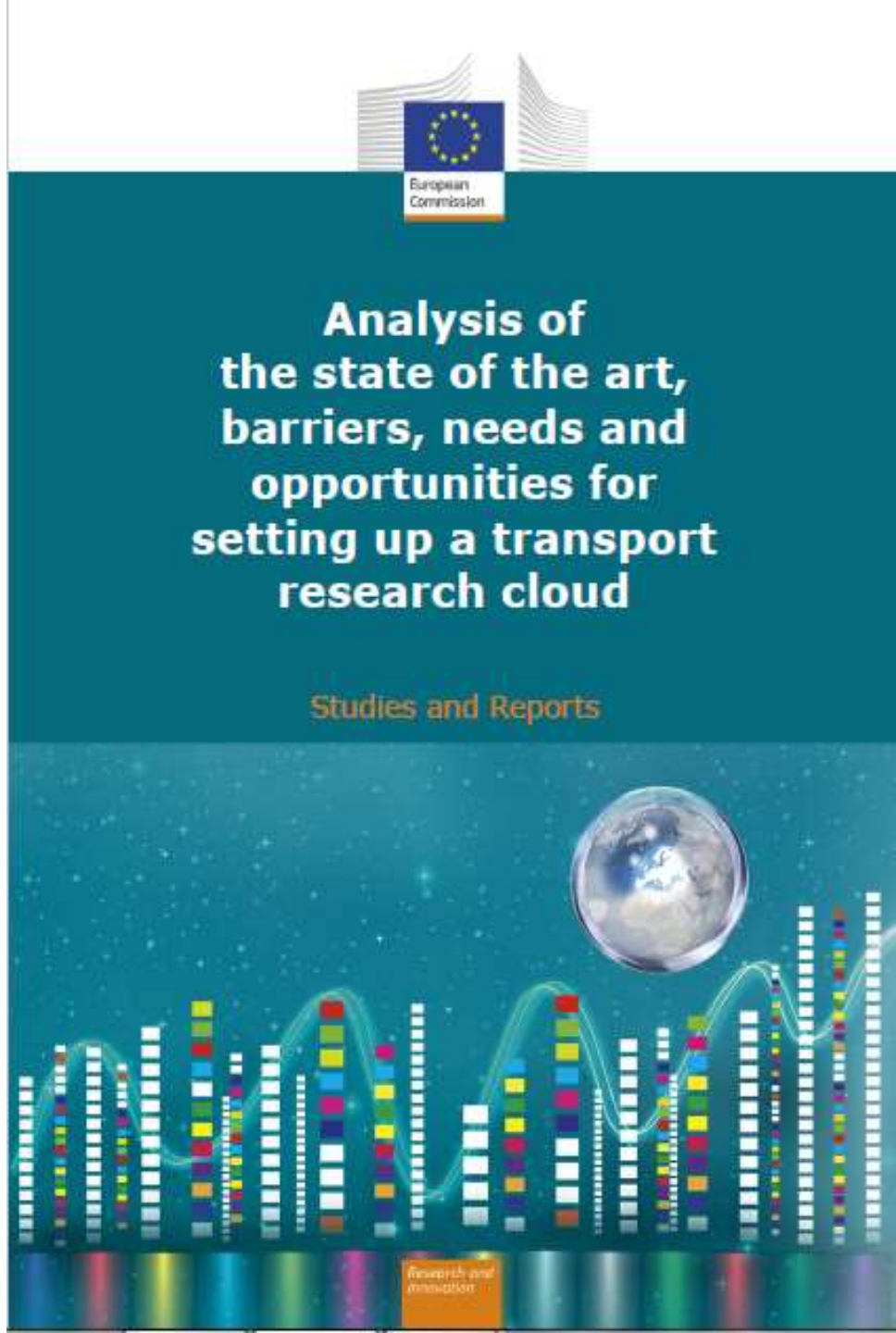
1. The Transport Research Cloud
2. Open Access in EU Countries
3. Global Approaches to Open Data
4. Setting up a Transport Research Cloud
5. Concluding remarks

Overview

"Analysis of the State of the Art, Barriers, Needs and Opportunities for Setting up a Transport Research Cloud" – available at publications.europa.eu

A Report prepared for the European Commission in October 2018 by a Group of Experts:

- **J. Rod Franklin**, Kühne Logistics University – Chairperson
- **Martin Böhm**, AustriaTech
- **Sarah Jones**, Digital Curation Centre
- **Tatiana Kovacikova**, University of Žilina
- **Katarzyna Nowicka**, SGH Warsaw School of Economics – Rapporteur
- **George Yannis**, National Technical University of Athens



Objectives

"Analysis of the State of the Art, Barriers, Needs and Opportunities for Setting up a Transport Research Cloud"

The main objectives of the Report are:

- to identify requirements for data sharing within the transport research community and
- to examine the potential of a Transport Research Cloud (TRC) as subset of the European Union's European Open Science Cloud (EOSC) initiative.

Methodological Approach

- Leverage knowledge and networks of expert group to form a baseline of understanding on the topic.
- Detailed literature review to understand the general background and specific issues related to the transport domain
- Conduct a survey of transport researchers to obtain a first hand understanding of current understanding, issues, and requirements

Main Goals of TRC Group

- Identify the main needs, obstacles and opportunities for data sharing and the open science cloud model with regard to data protection, security, FAIR data principles, governance, funding models, competitiveness and IPR issues.
- Identify the scope and the characteristics of data to be included in the TRC with attention to the distinction between research and operational data and IPR issues.
- Identify data mining and analysis requirements and tools to maximize the usability of big data.
- Identify a common set of cloud-based services with application to a possible TRC.
- Assess the relevance of international dimension of a possible TRC.
- Provide recommendations for different options of funding of a possible TRC.

Open Access in EU Countries

Open Access in EU Countries (1/6)

Austria

- Open Knowledge Österreich (www.okfn.at) is a community, where everybody can as a person participate.

- Open Access Network Austria (www.oana.at) is operated by the Austrian Science Fund together with the Platform of Austrian Universities.

- E-Infrastructures Austria projects (www.e-infrastructures.at), started in 2014 and aim to implement infrastructure for „eScience“, following the FAIR principle.

Open Access in EU Countries (2/6)

Germany

- In 2003 the Max Planck Gesellschaft sponsored the Berlin Declaration on Open Science in the Sciences and Humanities.
- In 2008 the Alliance of Science Organizations in Germany launched its Priority Initiative "Digital Information".
- 300 open access journals are currently being published in Germany.
- All public universities are required to have an open access policy by the Federal Ministry of Education and Research.
- German standards organization, the DINI, has developed a set of standards for open access repositories and any services that will use thee repositories.

Open Access in EU Countries (3/6)

Greece

- A web portal (www.openaccess.gr/el) has been developed by the National Documentation Centre.
- The Organization For Open Technologies (https://ellak.gr/) supports a campaign for the promotion of Open Technologies and Science.
- 36 Greek Open Access Repositories are listed in the Directory of Open Access Repositories (OpenDOAR).
- NTUA examples of Open Science:
  - SafeFITS (https://unecetrans.shinyapps.io/safefits/)
  - SafetyCube (https://www.roadsafety-dss.eu/)
  - Africa RSO (http://www.africanroadsafetyobservatory.org/)
  - PRACT (http://www.pract-repository.eu/)
  - DaCoTA (http://safetyknowsys.swov.nl/Home/about.html)
  - ERSO (https://ec.europa.eu/transport/road_safety/specialist_en)
  - NRSO (www.nrso.ntua.gr)

Open Access in EU Countries (4/6)

Poland

- ~1000 scientific journals in the Open Access, including the Reading Room of the PAS.
- Centre of Open Science - CeON (otwartanauka.pl).
- Open Science Platform (pon.edu.pl).
- Federation of Digital Libraries (fbc.pionier.net.pl).
- Different actors and projects for open science: Open Education Coalition, Open Science Library, Portal Unleash Science, Citizens of Science, Young Science Foundation.
- Scientific PlatonTV (tv.pionier.net.pl) – implements and publishes professional video materials in public in the field of results of scientific research, conferences etc.
- Open Science Day – organized every autumn.

Open Access in EU Countries (5/6)

Slovakia

- In 2016, the Open Access Working Group was established.
- "The Action Plan for Open Government in Slovak Republic, 2017 - 2019" was approved by Government in 2017, including a special section on Open Access to results of Science and Research.
- There are no Open Access repositories in Slovakia.
- The Central Registry of Theses and Dissertations provides full text access to some theses and dissertations.
- There are 42 open access journals published in Slovakia.

Open Access in EU Countries (6/6)

United Kingdom

- In 2005, a joint statement was released by the research councils in the UK on open access to publications.
- UK works towards open government data to ensure public sector information is made widely available for broad reuse.
- OpenDOAR, the directory of open access repositories, lists 258 repositories in the UK.
- The Directory of Open Access Journals list 1330 titles.
- Universities launched a group to bring together funders, institutions, publishers and other stakeholders on Open Access.
- Open Knowledge is a global non-profit network that focuses on realising open data's value to society.
- The Open Data Institute was co-founded in 2012 to advocate for the innovative use of open data to affect positive change across the globe.

Global Approaches to Open Data

Global Approaches to Open Data (1/4)

North America

- The implementation of open data portals has become commonplace.
  - ad hoc, without adherence to any formal standards
  - no single portal that one can access exists
- Collaborative scientific platforms provide domain partners with access to shared data, but cross domain sharing is difficult due to differing standards of practice.
- The US Department of Transportation, maintains a catalogue of open data sets on the US data.gov website (https://catalog.data.gov/dataset)
- The Canadian Department of Transportation maintains a catalogue of its open data sets on the open.canada.ca website (https://open.canada.ca/data/en/dataset).

Global Approaches to Open Data (2/4)

Latin America

- Several countries (e.g. Brazil, Mexico, Argentina) have begun to embrace a more open approach to government generated data.
- These centralized governmental data portals provide researchers with the ability to access statistics similar to those collected in most Western countries.
- Transport data is included in all of these portals.
- The Latin America Open Data Initiative has begun to work on various projects to more broadly encourage governmental sharing of data and the opening of data for use.

Global Approaches to Open Data (3/4)

Europe

- Europe has been a leader in the open data discussion establishing a number of open data portals.
- Open data in the transport area within Europe is maintained at the EU level, state level, region within state level and city level.

Russia and the CIS

- Russia and CIS have been very hesitant to adopt the open science concept.
- In 2014, the Russian government established the national open data portal (http://data.gov.ru).
- This portal provides access to transport researchers interested in macro-data concerning transport statistics.
- However, limited data at city and region level is available.

Global Approaches to Open Data (4/4)

Africa

- Many countries do not make country statistics or data available to the general public.
- Initiatives, like the African Development Bank and the openAFRICA project, are attempting to provide a single access point for open data sets.

Far East

- Countries such as India, Singapore, Japan and South Korea have government open data sites very similar to those of Western countries.
- Countries, such as China, while providing some open data, are just in the beginning of the process.

Oceania

- Australia and New Zealand have been at the forefront of the open data initiative.
- Australia's open research data portal is an example for the EOSC.
- Transport research is not one of the focus areas.

Setting up a Transport Research Cloud

Main Findings (1/3)

- Transport research data is diverse in content, structure, use, and degree of openness.
- Transport domains differ significantly in the data they collect, how they refer to the data, the analyses they perform, and in their views on open data.
- All transport researchers surveyed understood the value of open data, but only 5% currently openly share data.
- Issues with sharing involved privacy, trust, lack of mechanisms or incentives to share and a lack of understanding.
- Open data platforms should make it easier for the researcher to find and use data in research and this should be done at minimal cost.

Main Findings (2/3)

- Research data must be reusable.
- Not all transport data should be classified as research data.
- Researchers need to be motivated to use existing or future datasets.
- Data is a public good.
- Data from publicly funded projects should be treated as a public good.
- Clear guidelines are needed to define where data as a public good ends and the intellectual property developed by researchers using these data begins.
- Standards are necessary.
- Search and knowledge development costs are currently too high for researchers to reuse data as these data are not collected in a standard manner.

Main Findings (3/3)

- **Cloud platforms** must be fit for purpose
  - Transport data is complex and "big" requiring sophisticated curation and documentation.
  - Passive data warehouses provide little value, and thus incentive, for researchers to reuse their stored data.
  - The TRC, while focused on the transport domain, is a subsidiary entity to the EOSC and thus must conform to the standards established for and by the EOSC.
- **Incentives and education** are required
  - Current education and promotion processes fail to encourage researchers to provide their data to other researchers in an open manner.
  - An extensive outreach program is needed to convince existing transport researchers of the value of open data, and the university accreditation organizations of the need to incent researchers, via appropriate promotion policies.

Recommendations (1/5)

Reusable research data

- The Commission should bring together researchers, research data users and data generators to define what constitutes transport research data.

- The Commission should conduct a detailed study among transport researchers in order to identify the objections behind limited use of data collected by others and develop recommendations on how to overcome these objections.

Recommendations (2/5)

Data as a public good

- Any and all data collected under contracts paid for by tax payer funds by default should be classified as public data.

- All data collected under a publically funded project should include a clear demarcation between the Intellectual Property created by the individual researchers in analyzing the data and the data itself.

Recommendations (3/5)

Standards

- The Commission should bring together members of the transport research community, governmental entities generating transport data used in research, infrastructure operators and commercial consumers of transport research and transport research data to define:
  - the standards that will be necessary for the collection of transport data by public institutions,
  - the data formats these data should adhere to,
  - the metadata that must be used to describe the data,
  - and formats of this metadata so that automated search engines can easily find and characterize the data.

Recommendations (4/5)

Infrastructure

- The Commission should conduct a detailed study on the infrastructure and operating requirements for a TRC to ensure that an appropriate level of service can be provided at a cost that is understood by all stakeholders.

- The Commission should conduct a detailed study of what the potential user and stakeholder communities would require from a TRC in order to make it the "go to" place for doing cutting edge transport research.

- The TRC should be modelled after the EOSC and governed in a manner similar to the EOSC taking into account the particularities of the transport domain.

Recommendations (5/5)

Incentives, training, and education:

- EU policies for academic promotion, training, publication, and knowledge generation at public universities to ensure that:
  - researchers are uniformly trained in the process of placing their research data into the EOSC
  - universities provide incentives to their faculty and researchers, so that their research data is placed into the EOSC, and to ensure that proper credit for the generation of the reused data is given to the individuals who originally collected the data.

- An analysis of training requirements should be conducted by the Commission and formal training/education programs should be developed for researchers, libraries, data curators, and other individuals who will be needed to carry out the development of a mind-set of open data by default.

Concluding Remarks

- Transport research data is currently collected for multiple purposes based on domain and mode interests without regard to standards or reuse.
- There is significant value, perceived by the transport research community, in the reuse of data.
- However, certain issues inhibit data reuse, e.g. incentives/ rewards, search costs, data validity and data standards.
- For a transport research cloud to be successful it must be easy to use, provide value to both data suppliers and data users, be understood and be a part of the day-to-day research routine of researchers.

BE OPEN

Kick off Meeting
14 February 2019
Brussels

Open Science in Transport

G. Yannis*, M. Böhm, J. R. Franklin, S. Jones,
T. Kovacikova, K. Nowicka

*Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece