TRA2020 – Rethinking transport
Towards clean and inclusive mobility • Helsinki 27–30 April 2020

Exploring the Establishment of a European Transport Research Cloud

#1170

J. Rod Franklin[a], Martin Böhm[b], Sarah Jones[c], Tatiana Kovacikova[d], Katarzyna Nowicka[e], Rafal Rowinski[f], Katerina Folla[g], George Yannis[g]

[a]Kühne Logistics University, Germany; [b]AustriaTech, Austria; [c]Digital Curation Centre, UK; [d]University of Žilina, Slovakia; [e]SGH Warsaw School of Economics, Poland; [f]European Commission, Belgium; [g]National Technical University of Athens, Greece

Hosted and organised by: Ministry of Transport and Communications, Traficom, Väylä, Business Finland, VTT
Co-organised by: European Commission
Together with: ADAPT, alice, CEDR, ECTP, ERRAC, ERTRAC, ETRA, WATERBORNE

Objectives and Methodology

Objective

- The exploration of the **state of the art, the barriers, the opportunities and the needs** for setting up a European Transport Research Cloud (TRC)

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

Based on the results of the research conducted for the development of the report on "Analysis of the state of the art, Barriers, Needs and Opportunities for Setting up a Transport Research Cloud" for the European Commission

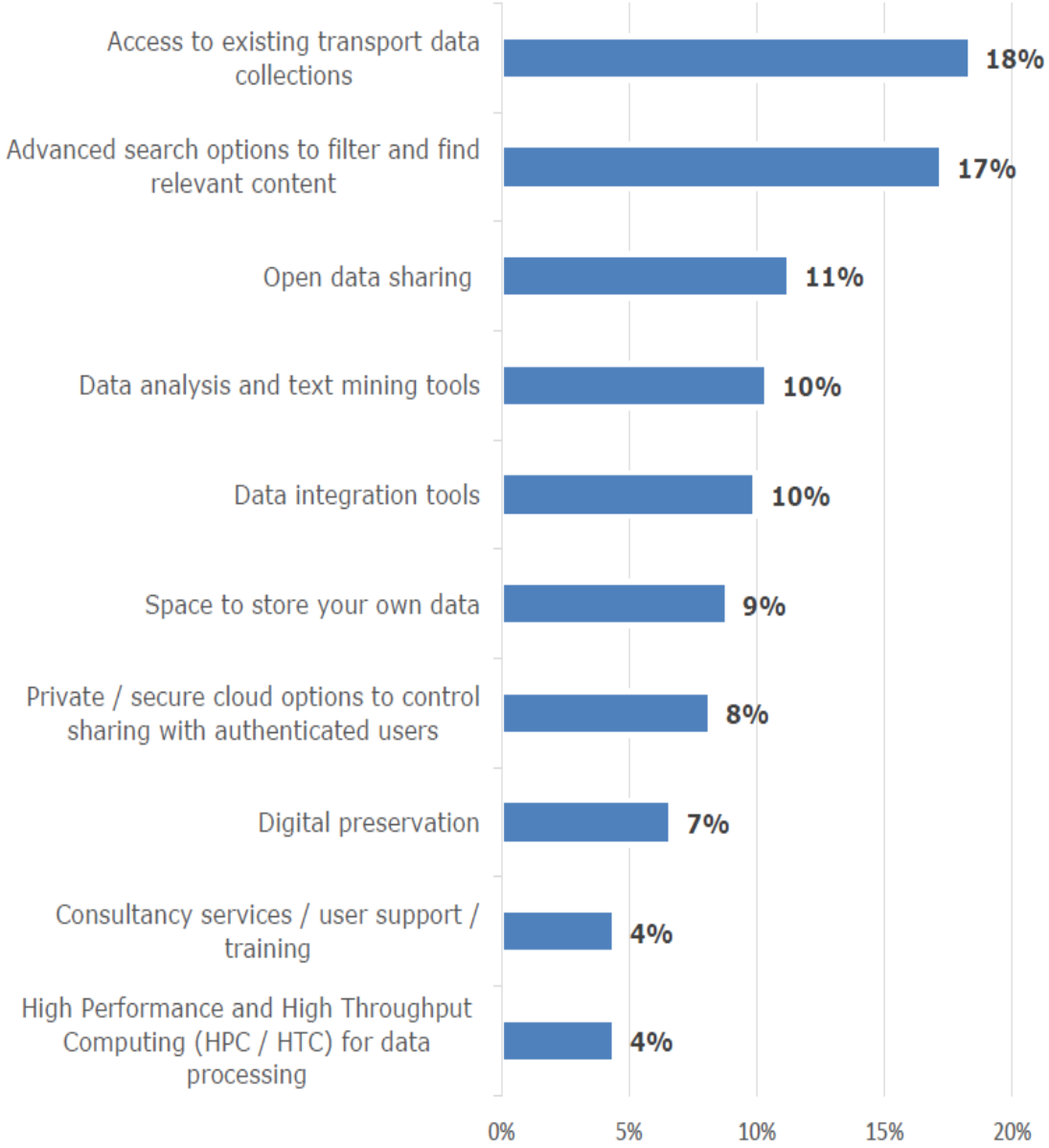
TRA 2020 HELSINKI

Survey Characteristics

- A survey was undertaken in the summer of **2018**
- **Aim** to identify existing data documentation and sharing practices of transport researchers and potential expectations from a cloud service
- Four **thematic areas**:
  - transport research data,
  - cloud service requirements,
  - opportunities and barriers and
  - funding mechanism
- **87 responses** were collected from 29 countries
- Researchers from **academic institutions** (85%), representatives of **public authorities** (4%) and the **commercial sector** (6%) and others participated

Survey Results (1/3)

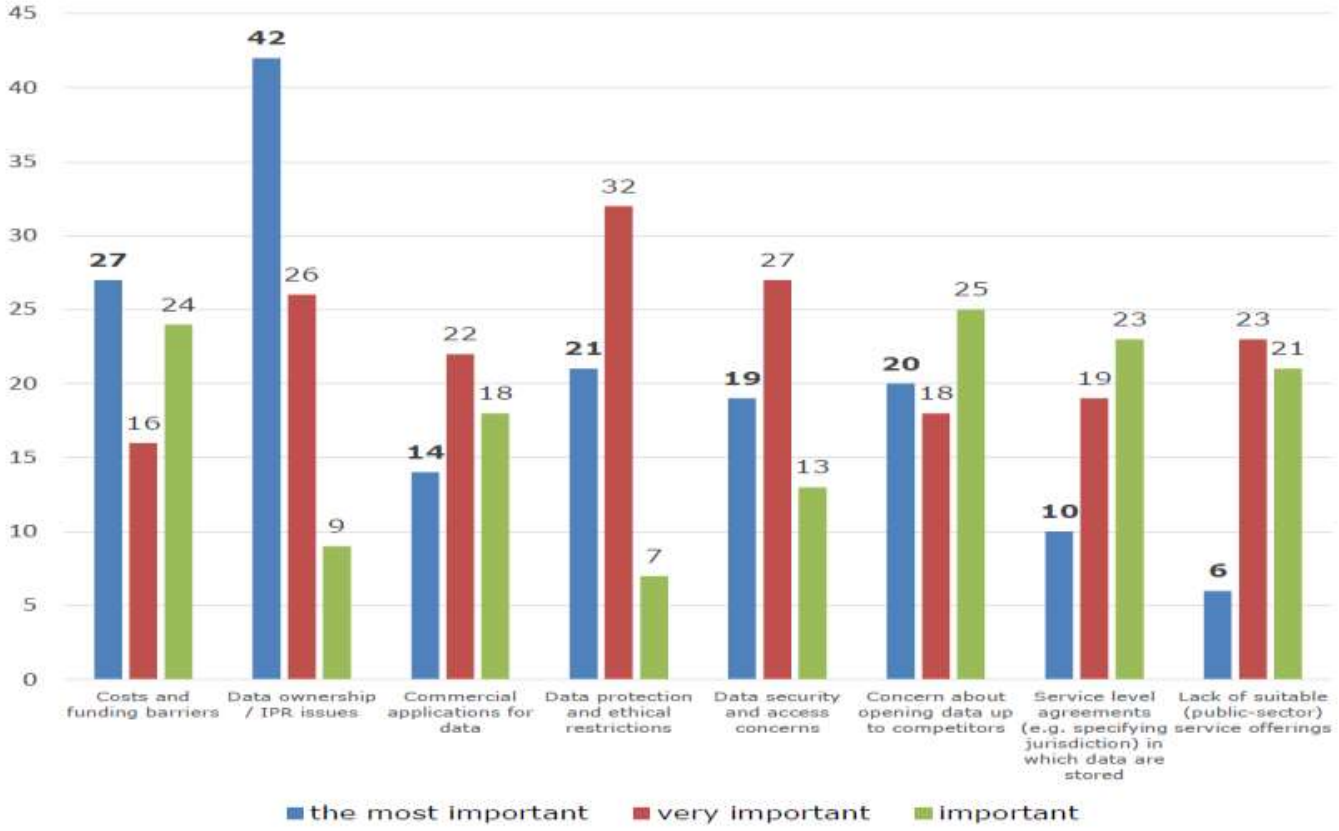
- **Relationship with transport data**: “analyse transport data” (28%), “use transport data” (25%) and “process data” (20%)
- A consensus was reached that **data should be available primarily for research purposes**, with some survey data being more sensitive due to privacy issues.
- 95% of respondents expect **data to be described** (documentation and metadata information).
- About 50% of the participants **would use the TRC depending on the quality of the service provided**
- Respondents would expect to have access to existing **transport data collections** (18%), advanced search options to **filter and find relevant content** (17%) and **open data sharing** (11%)



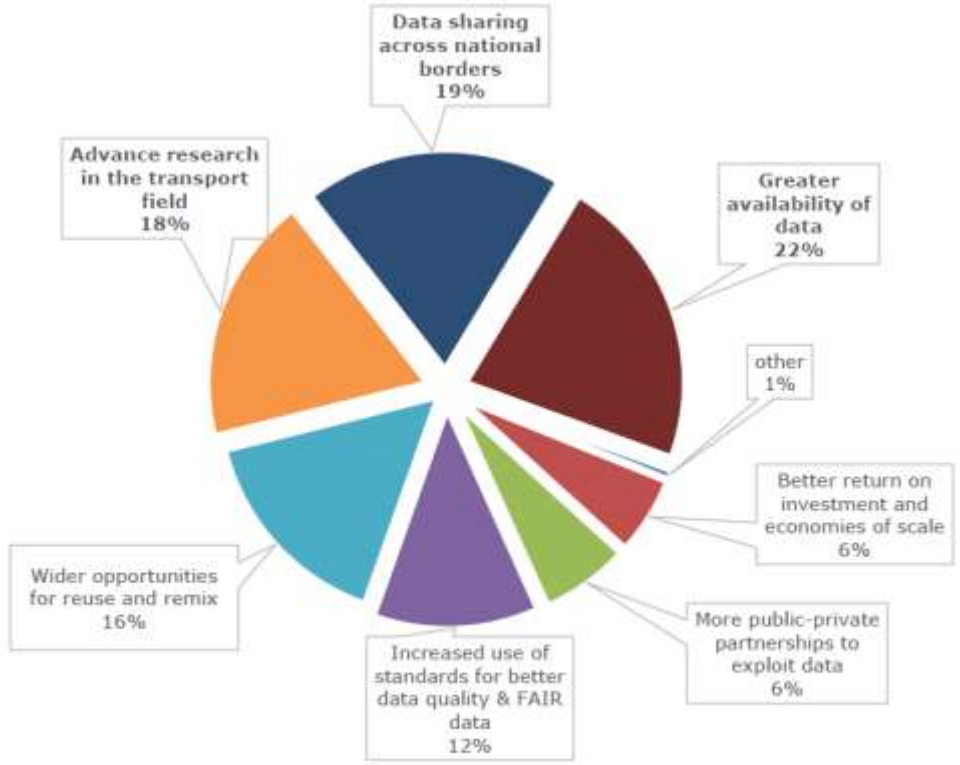
Q5: What functionality would you expect from a TRC?

Survey Results (2/3)

- The **most important benefits** from the TRC: (a) “greater availability of data", (b) “data sharing across national borders” and (c) “advance research in the transport field”



Q7: What are the three biggest barriers that you see in using open data services?



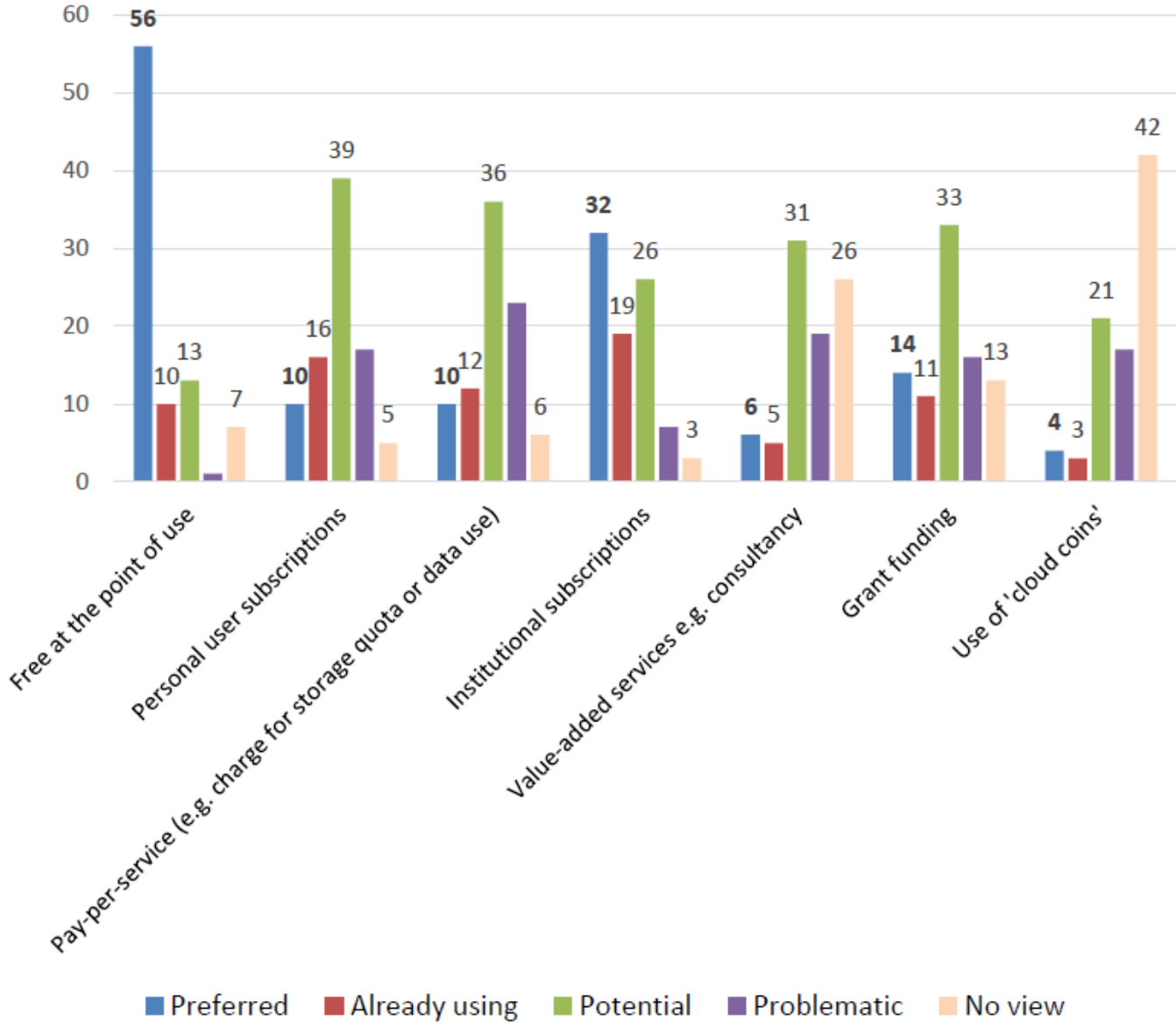
Q6: What benefits do you envisage from a TRC?

- As the three **most significant barriers**: (a) “data ownership / IPR issues”, (b) “cost and funding barriers” and (c) “concern about opening data out to competitors”.

Survey Results (3/3)

- 76% of respondents **do not pay for storing data** in an open data service

- **Charging models** that would allow to use open data storage services most effectively:
  (a) “free at the point of use”,
  (b) “institutional subscriptions”
  (c) “grant funding”



Q8: What charging models would allow you to use open data storage services most effectively?

TRA 2020 HELSINKI

Characteristics and scope of transport research data

- **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.
- Transport **research data is diverse** in content, structure, use, and degree of openness.
- **3 main categories of research data** are suggested to be included:
  - Original research data
  - Operational data directly related to research
  - Data from published research in scientific journals, conferences, workshops, etc.
- Detailed **description of the available data** based on machine readable metadata
- **Data management systems and infrastructure** to ensure a simple access and reuse of data

TRA 2020 HELSINKI

Current approaches to support data sharing

- There are practically as many **operational models** for open data platforms as there are platforms
- **Key to success**: well defined business models, a clear understanding of their stakeholder value propositions, sufficient start-up funding
- Platform **management** varies by how and for what purpose the platforms were established, e.g. library function, domain specific platforms, governmental platforms etc.
- **Platform sustainability** requires a business model that generates value for stakeholders of the platform and financial sources providing long -term funding

Opportunities and barriers to transport data sharing

- Open and easily accessible data will **facilitate research across communities and countries**
- **Promote more public-private partnership**, with commercial companies being encouraged to make their data available
- Reduction of **funding and effort requirements**

- **Fragmentation** of data ownership and a **lack of interoperability** between datasets and platforms
- Transport data is often **ethically or commercially sensitive** requiring tight controls concerning access to the data
- The diversity of data sources affecting **data quality**
- Variations in **hardware and software** used for collecting the data
- **Lack of expertise** in machine learning, data mining, and data management

Needs for transport data and the TRC

- **Policy issues** are needed to be tackled, concerning the conditions in which the data are provided, curated, maintained and accessed and funding of the service.

- Significant **infrastructure** is required to ensure the proper management of data.

- Management, support, operations, storage, marketing, education, engineering, integration, and other ongoing **costs will need to be foreseen**, as well as sources of revenue for the TRC.

Recommendations (1/2)

- **Reusable research data**
  - Define what research data is
  - Researchers should be motivated to reuse existing datasets
- **Data as a public good**
  - Data collected within projects and processes paid by public money needs to be made available to the public
  - Certain data, due to privacy or secrecy requirements would need to be excepted from this requirement
- **Standards**
  - Research data should be available in a standardised format
  - Standards and formats of metadata should be defined

Recommendations (2/2)

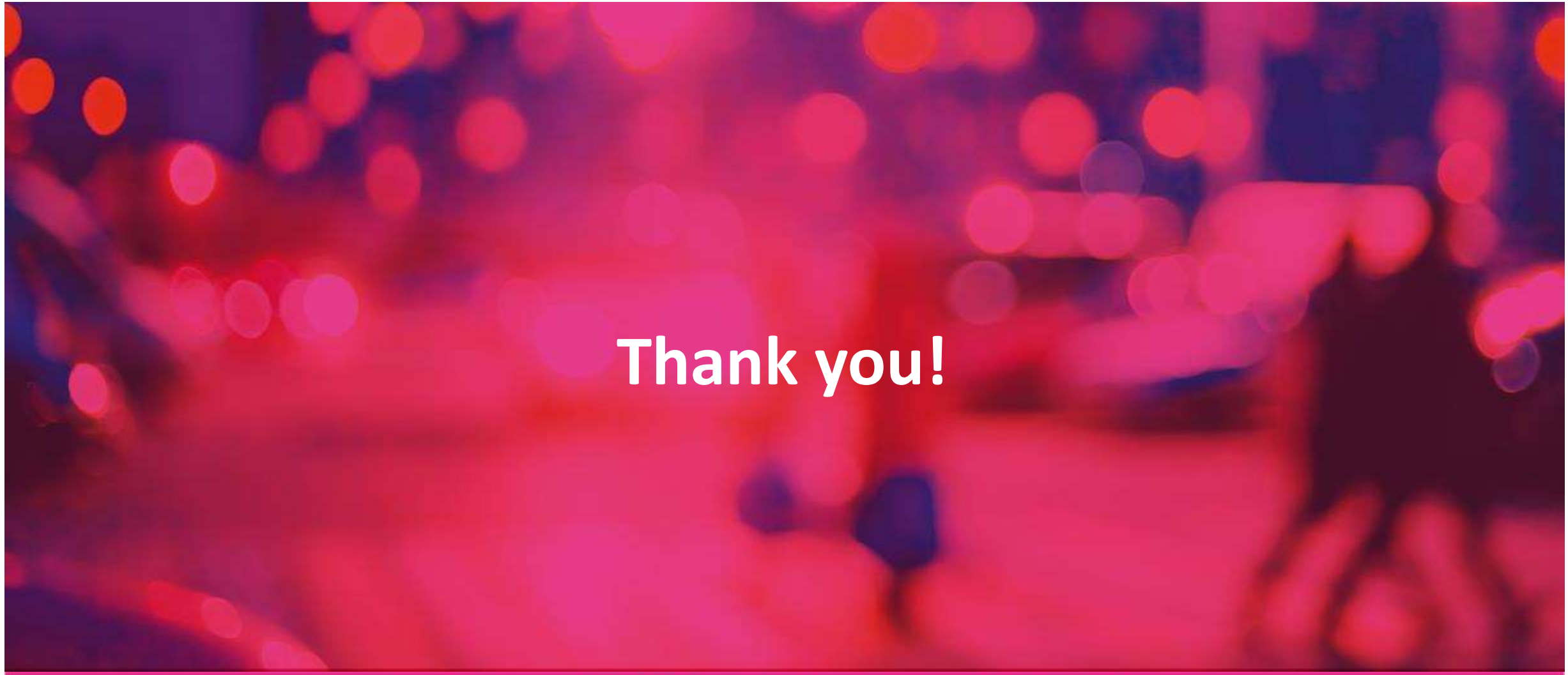
- **Infrastructure**
  - The infrastructure and operating requirements for a TRC should be defined
  - A sustainable business model should be built and the requirements for making the TRC a "go to" place for researchers should be developed
- **Incentives, education, and training**
  - EU policies for academic promotion, training, publication, and knowledge generation at public universities should be examined
  - Universities are recommended to provide proper incentives to their faculty and researchers to make their research data openly available

Conclusions

- Further work needs to be done in setting standards, understanding the proposals and needs of related stakeholders, identifying the infrastructure that could be used and its requirements for the **proper operation of a TRC**

- Important information can be drawn from the development of the **European Open Science Cloud (EOSC)**

- TRC should be considered as a **sustainable long-term project**, linked with open science and transport research projects.



traconference.eu #TRA2020 #rethinkingtransport @TRA_Conference

Hosted and organised by:

Co-organised by:

Together with: