Connecting research and public sector data through FAIR principles
Round Table
March 19, 2021

FAIR Challenges in Transport Open Science

George Yannis, Katerina Folla
NTUA, Greece

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No 824323
Background

- **Transport research data is highly diverse** in content, structure, use, and degree of openness

- **All Transport Modes** (Road, Rail, Maritime, Air, Combined transport), **All Transport Types** (persons and goods, urban and interurban, national, European and international, operations and services)

- **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
BE OPEN project

- Capitalize upon **existing initiatives** enabling Open Science
  - Key actors will coordinate and support actions for promoting Open Science policies, services and infrastructures
  - Involve key actors in planning and implementation
  - Enable key actors to learn from direct experience, previous knowledge and other relevant stakeholders

- Facilitate a **common understanding** among actors
  - Promoting Open Science
  - Prioritizing existing initiatives and actions at regional, European and International level

- Monitor progress in order to facilitate continuous improvements in Open Science exploitation
  - Proper indicators will be developed for supporting Open Science purposes
  - A monitoring process will be used to address information management, internal coordination, external coordination, risk management and other relevant dimensions

**https://www.topos-observatory.eu/**

Promote, regulate and standardise Open Science (OS) in Transport

- Develop a framework of common understanding of OS in transport
- Map existing OS resources
- Facilitate an evidence-based dialogue to promote and establish OS in transport
- Provide policy framework and guidance for OS implementation in transport
- Engage a broad range of stakeholders in a participatory process for OS uptake

Duration: 30 months
Start Date: 01-01-2019

Call: H2020-MG-2018-SingleStage-INEA
Type of Action: Coordination and Support Action
GA Number: 824323
Transport Research Data

- **Original transport research data** (e.g., data from Field Operational Tests, Naturalistic Driving Studies, research results and research models)

- **Operational data** directly related to research (as accident data, transport volumes data, etc.)

- Data from **published research** (as presented in scientific journals, delivered at conferences, workshops, etc.)
### Transport Key Stakeholders

<table>
<thead>
<tr>
<th>Type of data</th>
<th>Legal/Regulatory</th>
<th>Technological</th>
<th>Transport planning</th>
<th>Business modelling</th>
<th>Socio-economic</th>
<th>Environmental</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Research Data</strong></td>
<td>• Policy makers</td>
<td>• Research centres and Universities</td>
<td>• Policy makers</td>
<td>• Research centres and Universities</td>
<td>• Research centres and Universities</td>
<td>• Research centres and Universities</td>
</tr>
<tr>
<td></td>
<td>• Public authorities</td>
<td>• Commercial transport and logistics industry players</td>
<td>• Public authorities</td>
<td>• Commercial transport and logistics industry players</td>
<td>• Public authorities</td>
<td>• Policy makers</td>
</tr>
<tr>
<td></td>
<td>• Transport networks</td>
<td></td>
<td>• Transport networks</td>
<td></td>
<td>• Transport networks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Commercial transport and logistics industry players</td>
<td></td>
<td></td>
<td></td>
<td>• Commercial transport and logistics industry players</td>
<td></td>
</tr>
<tr>
<td><strong>Operational data</strong></td>
<td>• Public authorities</td>
<td>• Public authorities</td>
<td>• Public authorities</td>
<td>• Public authorities</td>
<td>• Public authorities</td>
<td>• Public authorities</td>
</tr>
<tr>
<td>meats)</td>
<td>(international level)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data from published transport research</strong></td>
<td>• Policy makers</td>
<td>• Policy makers (international level)</td>
<td>• Policy makers</td>
<td>• Policy makers (regional level)</td>
<td>• Policy makers (regional and national level)</td>
<td>• Policy makers</td>
</tr>
<tr>
<td></td>
<td>• Public authorities</td>
<td></td>
<td>• Public authorities</td>
<td></td>
<td>• Public authorities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Transport networks</td>
<td>• Transport networks</td>
<td>• Transport networks</td>
<td>• Transport networks</td>
<td>• Transport networks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Commercial transport and logistics industry players</td>
<td>• Commercial transport and logistics industry players</td>
<td>• Commercial transport and logistics industry players</td>
<td>• Commercial transport and logistics industry players</td>
<td>• Commercial transport and logistics industry players</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Benefits of Data Sharing

- Open and easily accessible data will facilitate research across communities and countries.
- Promote more transport public-private partnerships, with commercial companies being encouraged to make their data available.
- Reduction of funding and effort requirements for development and operation of transport systems.
- Strengthen capabilities and capacities/gather intelligence.
- Foster transparency and innovation in the new digital transport era.
Challenges in Transport Data Sharing

- **Socio-cultural**
  - diverse approaches within the transport community
  - absence of incentives and rewarding systems

- **Technological**
  - lack of awareness on communication technologies and European research e-infrastructures

- **Political**
  - different interests and needs of the various stakeholders
  - lower priority of Open Science in governmental agendas and resources allocation

- **Organizational**
  - lack of open research tools, workflows, units and services
  - lack of available human resources

- **Economic**
  - high initial investment costs on infrastructure and scientific personnel
  - long-term efficiency of research activities and resource allocation

- **Legal**
  - diversified legislation frameworks at global level
  - unclear legal environment on data privacy, ownership and security issues
Challenges, Opportunities and Barriers in Transport Research

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Opportunities</th>
<th>Barriers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Researchers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical challenges:</td>
<td>Openly sharing their data:</td>
<td>Openly sharing their data:</td>
</tr>
<tr>
<td>● Expertise in data security and privacy</td>
<td>● More co-operations/contacts</td>
<td>● Significant effort to produce dataset</td>
</tr>
<tr>
<td>● Expertise in data management</td>
<td>● Gain recognition</td>
<td>● Data protection and ethical restrictions</td>
</tr>
<tr>
<td>● Expertise in open licence practices</td>
<td>● Co-authorship to other researchers’ publications using their data</td>
<td>● Concern to opening up to competitors</td>
</tr>
<tr>
<td>● Expertise in database design and computer programming</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data management</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>● Data quality</td>
<td>● Accessibility to more data</td>
<td>● Insufficient documentation of the data</td>
</tr>
<tr>
<td>● Data protection and security</td>
<td>● More cross-disciplinary co-operations</td>
<td>● Not easy accessibility</td>
</tr>
<tr>
<td>● Complex nature of transport data and information</td>
<td>● New, original research results and products</td>
<td>● Poor data quality</td>
</tr>
<tr>
<td><strong>Public Transport Companies/Organisations</strong></td>
<td><strong>Research Institutions</strong></td>
<td><strong>Private Transport Companies/Organisations</strong></td>
</tr>
<tr>
<td>● Data ownership conflicts</td>
<td>● Advance of the science in the transport field</td>
<td>● Data ownership/confidential data</td>
</tr>
<tr>
<td>● Data protection, privacy and ethical issues</td>
<td>● Increased collaborations across institutional, national and disciplinary boundaries</td>
<td>● Conflicts regarding ownership/IPR</td>
</tr>
<tr>
<td>● Skilled personnel</td>
<td>● Increased collaboration between companies and research infrastructures</td>
<td>● Protection of personal data</td>
</tr>
<tr>
<td><strong>Use of open data:</strong></td>
<td>● Improve transport operations and performance</td>
<td>● Protection of commercial/confidential data</td>
</tr>
<tr>
<td>● Accessibility to more data</td>
<td>● Foster data-based decisions</td>
<td>● Conflicts regarding ownership/IPR</td>
</tr>
<tr>
<td>● More cross-disciplinary co-operations</td>
<td>● Transparency</td>
<td>● Protection of personal data</td>
</tr>
<tr>
<td>● New, original research results and products</td>
<td></td>
<td>● Limited financial resources</td>
</tr>
</tbody>
</table>
| | | | Delphi Survey on Challenges, Opportunities and Barriers in Transport Research

- **Legal and ethical issues** (GDPR, privacy issues, IPR, etc.) were assessed as the main challenges for both research institutions and transport companies.

- **Lack of skilled personnel** was also highlighted as a significant challenge for research institutions and public transport companies.

- **Commercial competition** was also identified for the private transport companies.

Requirements for Open Transport Data

- **FAIR data**: Findable, Accessible, Interoperable and Reusable
- Ensure **data quality**, i.e. relevance, accuracy, credibility, timeliness, accessibility, interpretability, coherence
- **Standards** are needed for the data collection and data formats
- High quality **metadata** describing properly the data
- Appropriate **formats** of the metadata, so that search engines easily find and characterize data
- Appropriate **infrastructure services** so that both data providers and data users easily use open data platforms
- **Data Management Plans** should be developed in all projects, based on online tools conforming to common methodologies
Conclusions

- Open Science could increase the current **great potential** of Transport Systems (new infrastructure, services, governance) with:
  - more data and knowledge
  - broader geographical coverage

- **Data sharing** will allow the verification of the scientific results, foster collaborations among researchers and promote more public-private partnership

- Further work needs to be done in setting standards and understanding the needs of related stakeholders

- Key policy issues are needed to be tackled, concerning the conditions in which data are provided, curated, maintained and accessed with **new and innovative business models**
Connecting research and public sector data through FAIR principles
Round Table
March 19, 2021

FAIR Challenges
in Transport Open Science

George Yannis, Katerina Folla
NTUA, Greece

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No 824323