Mapping the use of open data, software and infrastructure in transport research

Anja Fleten Nielsen, Institute of Transport Economics - Norwegian Centre for Transport Research
Key findings

- Open data will benefit science, sharing is still limited
- Sharing depends on field of science
- Share of open access articles increased
How to make TRC work?

- Clear definition of transport research data
- High quality metadata (DCAT)
- FAIR data (LERU)
- Infrastructure for Big Data (OpenAIRE)
- Governance
- Ethical issues
Open and FAIR data

Open data: data that can be freely used, modified, and shared by anyone for any purpose

FAIR data: Findable, Accessible, Interoperable and Reusable

EUROSTAT, OECD library, EU ODP
Research based on open and/or FAIR data (n=51)

Percentage of researchers based on open and/or FAIR data

<table>
<thead>
<tr>
<th>Percentage of research</th>
<th>Percentage of institutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-25%</td>
<td>58.8</td>
</tr>
<tr>
<td>26-50%</td>
<td>21.6</td>
</tr>
<tr>
<td>51-75%</td>
<td>17.6</td>
</tr>
<tr>
<td>76-100%</td>
<td>2</td>
</tr>
</tbody>
</table>
Advantages (n=51)

Main advantages for using open/FAIR data

- **Possibility of conducting time-series analysis**: 62.7%
- **Accessibility**: 29.4%
- **Low cost**: 23.5%
- **Time savings**: 19.6%
- **Non-users/unsure**: 5.9%
- **Other**: 7.8%
Challenges (n=51)

Main challenges for using open/FAIR data

- **RELIABILITY OF THE MATERIAL**: 62.7%
- **RELEVANCE OF THE MATERIAL**: 39.2%
- **ACCESSIBILITY**: 33.3%
- **NON-USERS**: 3.9%
Barriers for production (n=51)

Main barriers for producing open/FAIR data

- COST/TIME: 43.1%
- GDPR: 51%
- COMPETITIVENESS OF THE ORGANIZATION: 35.3%
- OTHER: 7.8%
Open software/source

Open source: information sources that can be used, modified and shared by everyone and are publicly available

Open source software: software with a source code that anyone can inspect, modify and enhance

Figshare, Github, GitLab
Open software use (n=45)

Percentage of time spent on open source computer software

<table>
<thead>
<tr>
<th>Percentage of institutes</th>
<th>Percentage of time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-25%</td>
<td>57.8</td>
</tr>
<tr>
<td>26-50%</td>
<td>35.6</td>
</tr>
<tr>
<td>51-75%</td>
<td>6.7</td>
</tr>
<tr>
<td>76-100%</td>
<td>0</td>
</tr>
</tbody>
</table>
Reasons for using open software (n=32)

Main reasons for encouraging use of open source software

- **Low Cost**: 75%
- **High Security**: 9.4%
- **Quality**: 21.9%
- **Transparency**: 40.6%
- **Stability**: 6.3%
- **Interoperability**: 31.3%
- **Flexibility**: 59.4%
- **Other**: 9.4%

0 10 20 30 40 50 60 70 80
Percentage

Reasons
Reasons for not using open software (n=13)

Main reasons for not encouraging use of open source software

- Low security: 53.8%
- Lack of knowledge: 38.5%
- Low quality: 23.1%
- Lack of documented efficiency: 23.1%
- No existing open source software significantly relevant for my institution: 46.2%
- Other: 7.7%
Programs used (n=45)

Open source software programs

- UNSURE: 6.7%
- OTHER: 6.7%
- CODING SOFTWARE: 44.4%
- DATA PROCESSING SOFTWARE: 53.3%
- SURVEY SOFTWARE: 40%
- COMMUNICATION SOFTWARE: 28.9%
- GIS SOFTWARE: 42.2%
- TRANSPORT MODELING SOFTWARE: 15.6%
- WORD PROCESSING SOFTWARE: 48.9%

Type of software

Percentage users
Infrastructure

Facilities, resources and related service that is used collaboratively

Laboratories, archives, databases
Types of open research infrastructure the organization have

- **Laboratories (hard infrastructure)**: 41.2%
- **Computing systems/models**: 33.3%
- **Databases**: 51%
- **Models**: 29.4%
- **Others**: 7.8%
- **None**: 13.7%
Pricing (n=26)

Type of pricing for shared infrastructure

- Other: 13.8%
- Two-part tariff: 6.9%
- Average cost pricing: 13.8%
- Marginal cost pricing: 27.6%
- No charge: 58.6%

Percentage
Barriers for sharing (n=22)

Barriers for sharing infrastructure

- Cost/Time: 36.4%
- GDPR/Privacy Restrictions: 45.5%
- Competitiveness of the Organization: 50%
- Other: 18.2%
Differences between research areas

<table>
<thead>
<tr>
<th>Key area</th>
<th>Open Data</th>
<th>Open Software</th>
<th>Open Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport planning</td>
<td>35 %</td>
<td>6 %</td>
<td>85 %</td>
</tr>
<tr>
<td>Technological</td>
<td>22 %</td>
<td>9 %</td>
<td>73 %</td>
</tr>
<tr>
<td>Other</td>
<td>9 %</td>
<td>6 %</td>
<td>82 %</td>
</tr>
</tbody>
</table>
Interface to the European Open Science Cloud

- Transport vocabulary in csv format for text mining
- List of organizations/institutes dedicated to transport research
- List of projects/calls/programmes related to transport (TRIMIS and cordis)
- List of repositories related to transport research such as RE3data.org
Governance and operational business models

Norway
Belgium
Greece