Open Data in Transport Modeling

Lessons Learned from Santiago de Chile
Some Thoughts on Open Data
Open Data: What?

• "Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control." (Auer et al., 2007)

• The idea comes along with other “open movements” such as open source for computer code or open access for scientific material.

• Goals include transparency, the generation of social and commercial value, as well as participation and engagement. (Open Data Handbook, 2016)
Open Data: Where?

A COMPREHENSIVE LIST OF 2500+ OPEN DATA PORTALS AROUND THE WORLD
Transport Data and Models in Germany

Public authorities

The crowd

Consultants

Google, Apple, TomTom, Telekom, DB, BVG, …

Citizens

Transport data

- Network data
- Travel diaries, GPS tracks
- Flows
- Schedules
- Travel times, reliability

Recommendation

Aggregated results

Detailed results

Transport model
Transport Data and Models in Chile (and other countries)

- Public authorities
- The crowd
- Google, Apple, TomTom, Telekom, DB, BVG, …

Transport data

- Transport model
  - Scientists
  - Consultants

Results

Recommendation

Citizens
Showcase: MATSim Santiago de Chile

Credits to:  
Daniel Hosse  
Kai Turner  
Alejandro Tirachini
MATSim Santiago: Open Data Requirements

- **Car supply:**
  Openstreetmap ("The crowd")

- **PT supply:**
  Google Transit Feed ("Google")

- **Travel demand:**
  Travel diaries etc. ("Public authority")

- Large-scale (8m+ individuals)
- Activity-based demand
- Dynamic traffic flow model
- Utility-based learning cycle

![Diagram showing data flow from network data and population data to daily plans, execution, re-planning, scoring, and analysis.]

DLR.de • Folie 9
MATSim Santiago: Overview

Raw data:
- 60'054 individuals (~1%)
- 113'591 trips

MATSim:
- 42'459 individuals
- 103'055 trips
MATSim Santiago: Movies

Activity distribution:
[show movie]

Car traffic:
[show movie]

Public transport:
[show movie]
MATSim Santiago: Further Steps

- You can run the scenario, improve it, play around (see https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/)

- Ideal for BSc or MSc theses; the idea is to integrate every improvement into the current state (feeding back)

- Studies on policy interventions, business cases
Summary
Open Data in Transport Modeling: Opportunities

- Increase in transparency

- Better citizen engagement/empowerment?

- **Social value**: better transport planning through competition
  - identification of weak spots in the system
  - policy studies (e.g. provision of public infrastructure)
  - environmental/social analysis
  - …

- **Commercial value**: creating a platform for innovative mobility-based services
  - car sharing systems
  - supply chain/location planning
  - delivery/logistic planning, navigation
  - …
Open Data in Transport Modeling: Challenges

- **Privacy**: Anonymization of data at the source?

- **Quality**: Development of standards for data, data provision and simulation results; Who?

- **Feedback**: Unclear how to integrate improvements (performed by different stakeholders) back into the open data

- Need for a shift from confrontation (holding back information) to cooperation (sharing information) between authorities, citizens and private companies

- Increasing vulnerability to attacks?
Conclusions

• It is possible to create a freely available transport planning scenario within a few months if open data is available.

• For stimulating competition and innovation in transport modeling, the provision of open data seems a necessary mission for the public sector (e.g. through a clearinghouse).

• Better standards for data and data provision would safe a lot of time (see Ben Wellington @TEDxNY). Standards are also necessary for the provision of simulation results.

• Public authorities could/should additionally maintain a (reference) simulator including technical staff (common but not everywhere, see BMVI).

• There is potential to generate social and commercial value: better transport planning, citizen engagement and the development of new businesses.
Thank you.
Sources: Open MATSim Scenario of Santiago de Chile

• Documentation: https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/

• Runnable MATSim scenario, no MATSim installation necessary: https://svn.vsp.tu-berlin.de/repos/public-svn/matsim/scenarios/countries/cl/santiago/

• Code: https://github.com/matsim-org/matsim/tree/master/playgrounds/santiago/src/main/java/playground/santiago

• Further data sets available:
  • Operational data for colectivos, real PT (GPS), …
  • Smartcard tap-on data
  • …
Sources: Definitions and Open Data Portals

- [http://link.springer.com/chapter/10.1007%2F978-3-540-76298-0_52](http://link.springer.com/chapter/10.1007%2F978-3-540-76298-0_52)
- [http://opendatahandbook.org/](http://opendatahandbook.org/)
- [https://www.data.gov/open-gov/](https://www.data.gov/open-gov/)
- [https://data.gov.uk/](https://data.gov.uk/)
- [https://www.govdata.de/](https://www.govdata.de/)
- [http://daten.berlin.de/](http://daten.berlin.de/)
- [http://data.deutschebahn.com/](http://data.deutschebahn.com/)
- [http://transitfeeds.com/](http://transitfeeds.com/)
- [https://www.opendatasoft.com/a-comprehensive-list-of-all-open-data-portals-around-the-world/](https://www.opendatasoft.com/a-comprehensive-list-of-all-open-data-portals-around-the-world/)
- [https://youtu.be/6BTg8OXhEZk](https://youtu.be/6BTg8OXhEZk)