

REproducible Research In Transportation Engineering



rerite.org

@ The IEEE Intelligent Vehicles Symposium
Detroit, MI, United States

State of the Union: Open Science in Transportation 2026

Junyi Ji (Vanderbilt University)

Joint research efforts with Ruth Lu (MIT), Linda Belkessa (Université Gustave Eiffel), Liming Wang (Portland State), Silvia Varotto (ENTPE), Yongqi Dong (RWTH Aachen University), Nicolas Saunier (Polytechnique Montreal), Gregory S. Macfarlane (BYU), Bahman Madadi (ENTPE), Mostafa Ameli (Université Gustave Eiffel), Cathy Wu (MIT)

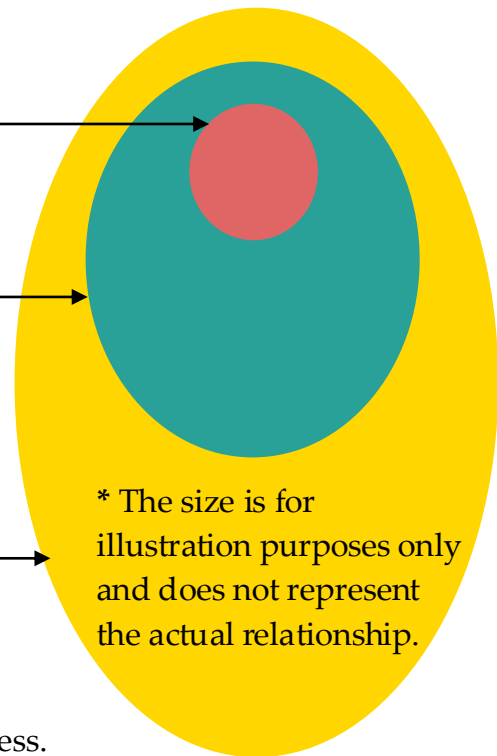
What is open science?

Open science: **Making science more accessible, inclusive and equitable for all** (United Nations Educational, Scientific and Cultural Organization, 2021).

Replicability - “obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data”

Reproducibility - “obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis. This definition is synonymous with *computational reproducibility*”

Availability - “Making artifacts used in the research process available to the community” **(focus of this talk)**



National Academies of Sciences, Engineering, and Medicine. 2019.

Reproducibility and Replicability in Science. Washington, DC: The National Academies Press.

Back in ITSC 2024 @ Edmonton : An informal study by Cathy

An informal study...

- How reproducible

- A low bar for (to the paper

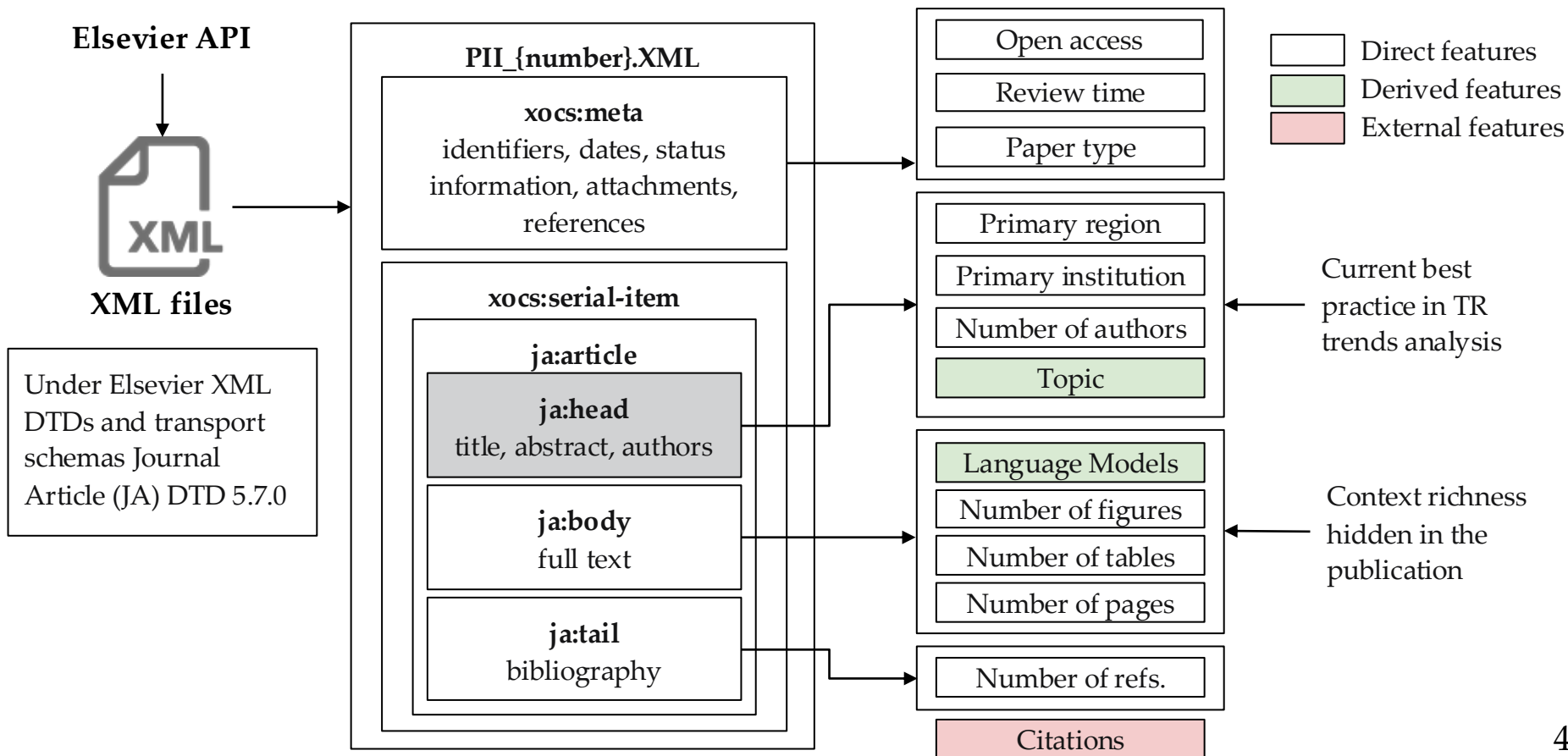
- Methodology:
 - “github”, “ava
 - Latest Neur
 - Latest issue
 - Latest issue

Some informal results...

- NeurIPS 2023 proceedings (first 25 of 3600 papers)
 - 14.5 of 25 have Github repo (60%)
- T-ITS, Sept 2024 (first 50 of 177 papers)
 - 6 review articles
 - 3 of 44 have Github repo (7%)
- TR Part C, Oct 2024 (30 papers)
 - Hard to assess
 - 4 of 30 papers are available as open access
 - 2 of the 4 have Github repo (7%+)

How to scale it up without losing resolution?

Meta data from the XML files



Sharing is rare — and data deposit is rarest of all



5%

shared a
code repository

4%

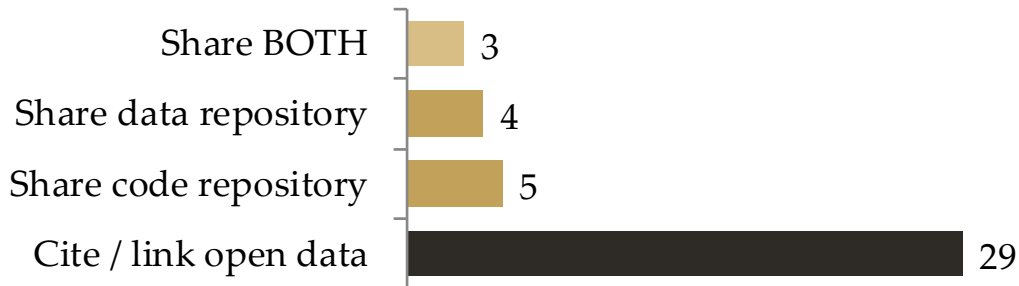
shared a
data repository

3%

shared
both

29%

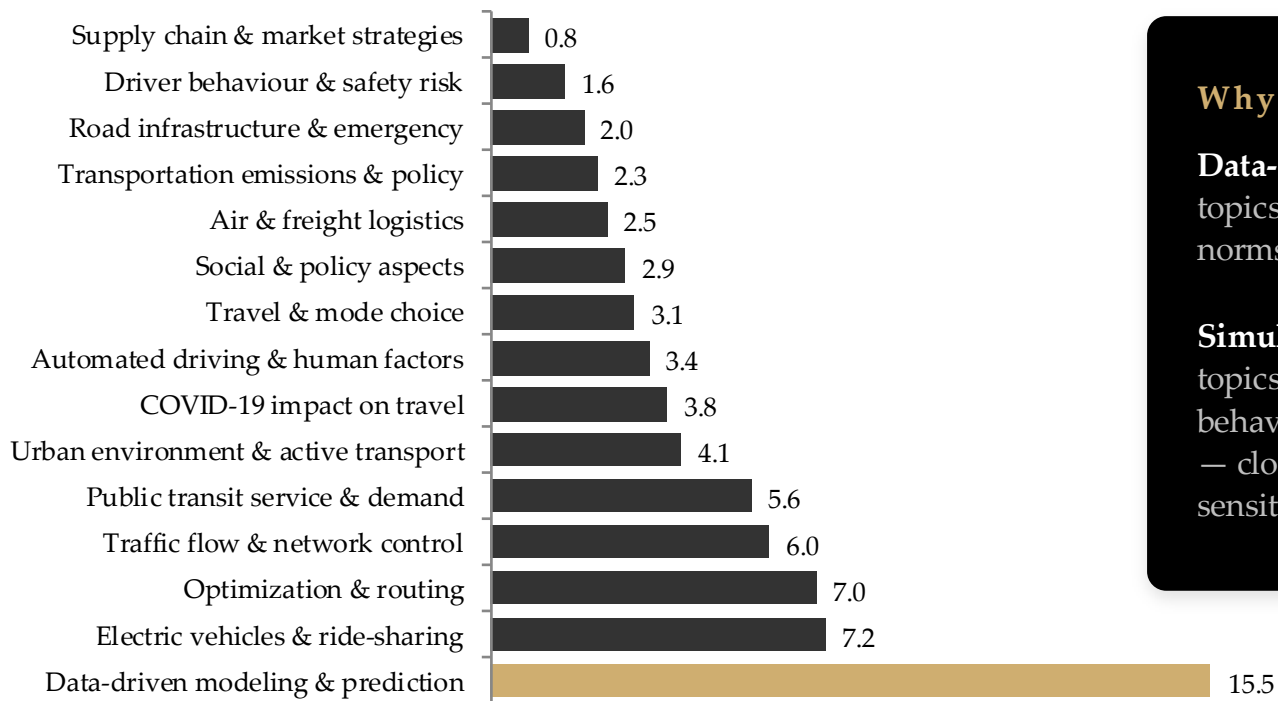
cite or link
existing open data



The takeaway

Many papers **benefit** from open data (29%), but few **contribute** new artifacts. Less than 4% release new data — so a handful of works underpin much of the field.

Data-driven research shares code 3× than average



Why the spread?

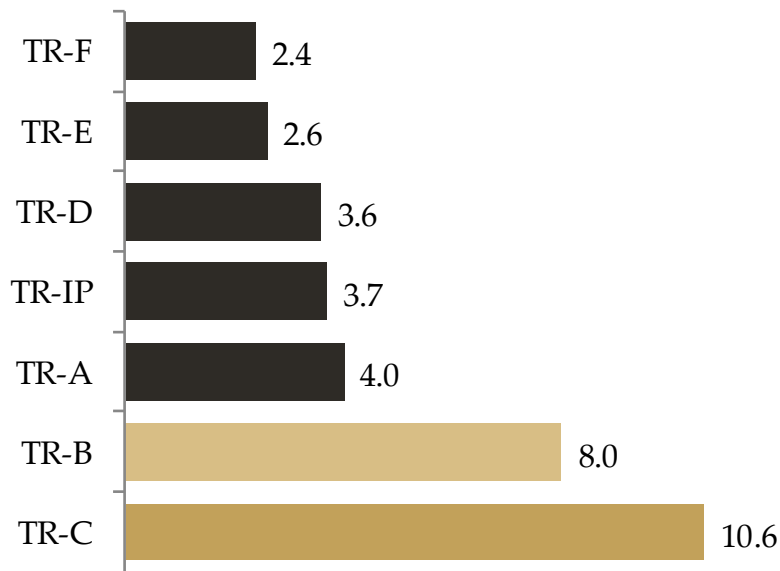
Data-driven & ML-adjacent topics inherit code-sharing norms from CS.

Simulation- & proprietary-data topics (supply chain, driver behaviour, logistics) share least — closed tools, privacy, sensitive data.

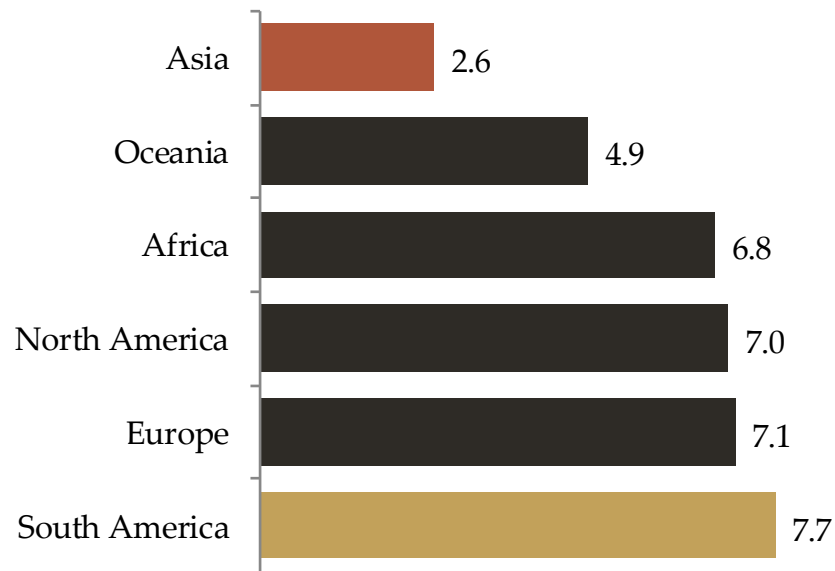
TR-C leads on journals and regional gaps exists



Code available by journal (%)



Code available by corresponding-author region (%)



TR-C explicitly promotes open science in its scope; regional gaps echo differences in funding policy, infrastructure and data-sharing experience.

Data & Code Availability Trends by journal

Share of articles that release a public dataset or publicly available code, traced year by year.

Reading each panel

—●— Code publicly available

—●— Open data released

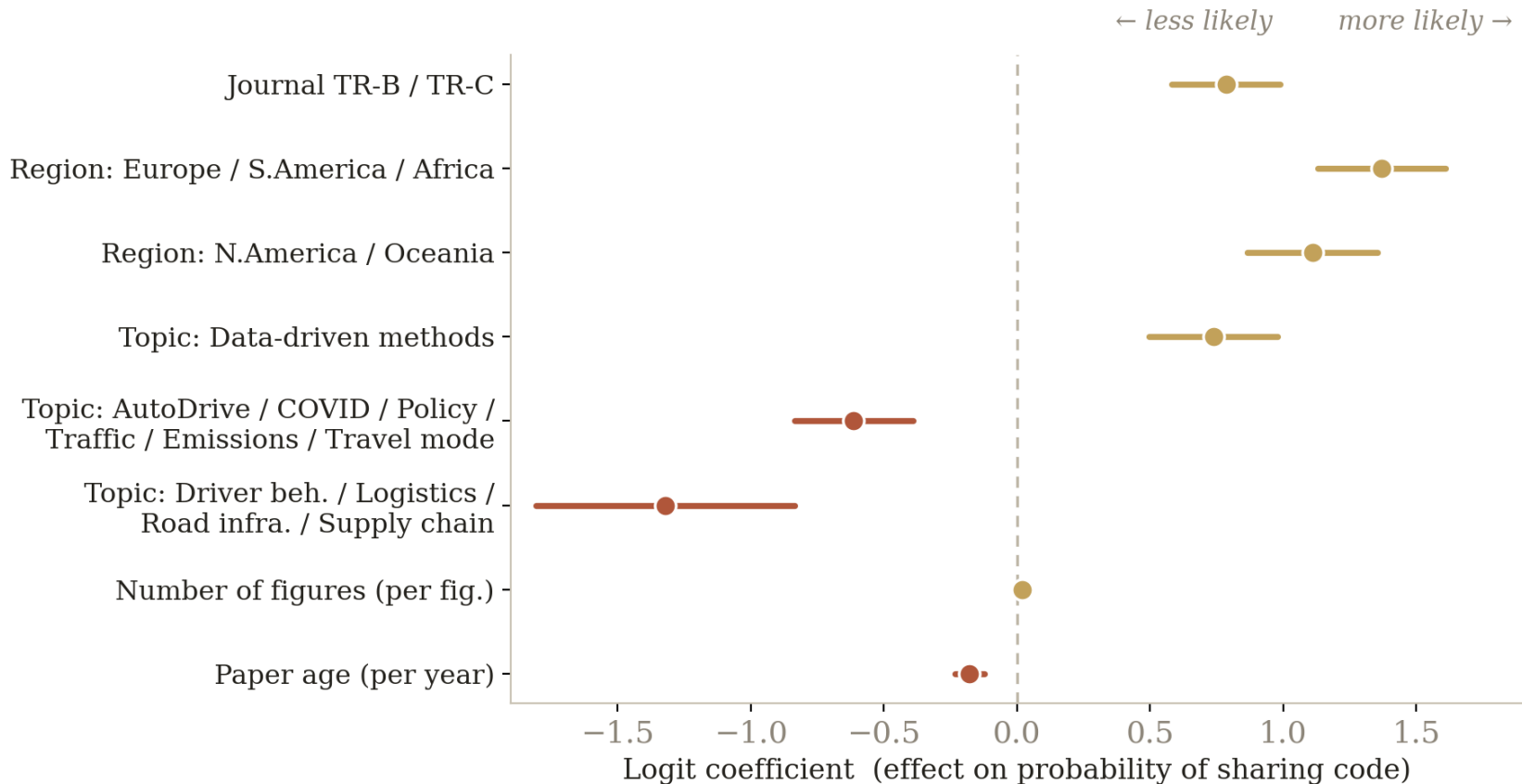


Open data = at least one dataset publicly available (repository / DOI). Public code = source code released and reachable.

N = 10,724 articles across seven Transportation Research journals, 2019–2024. Panels after the aggregate are ordered by overall openness.

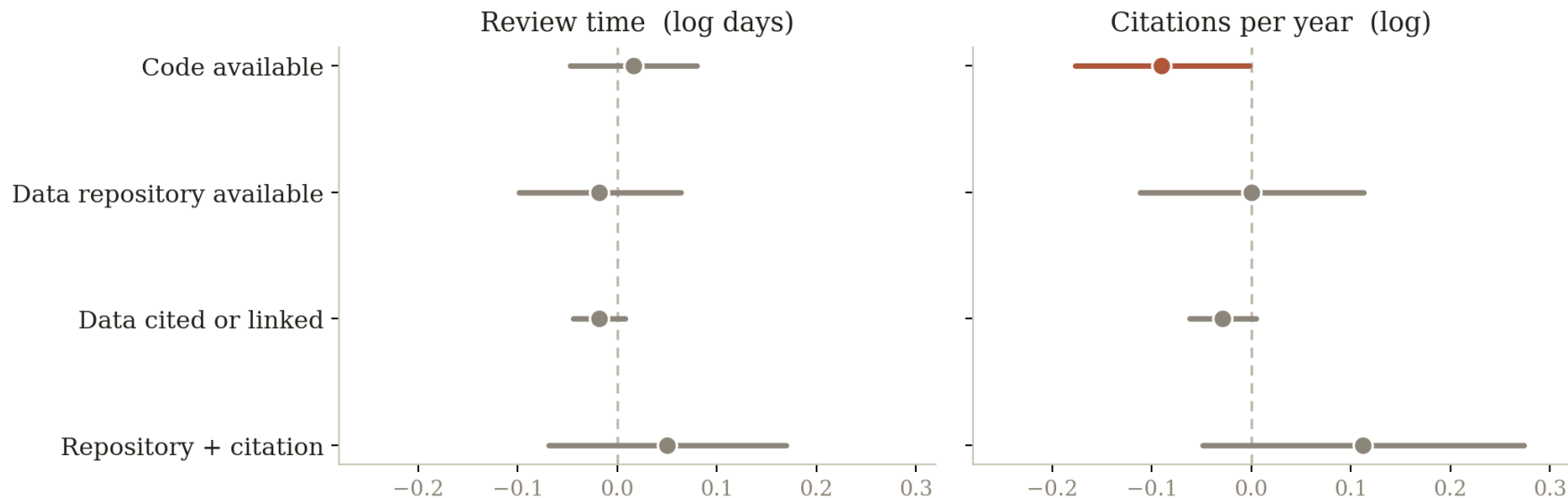
Source: author analysis of indexed article metadata

Choice modeling for code sharing



All coefficients $p < 0.05$. Reference: TR-D · Asia · Optimization. Adj. $\rho^2 = 0.112$, $N = 10,480$.

No statistically significant incentives for sharing



Coefficient with 95% confidence interval • grey = interval overlaps zero (no significant effect)

No faster review

Sharing code or data does not shorten review — if anything, code repos add ~2–27 days.

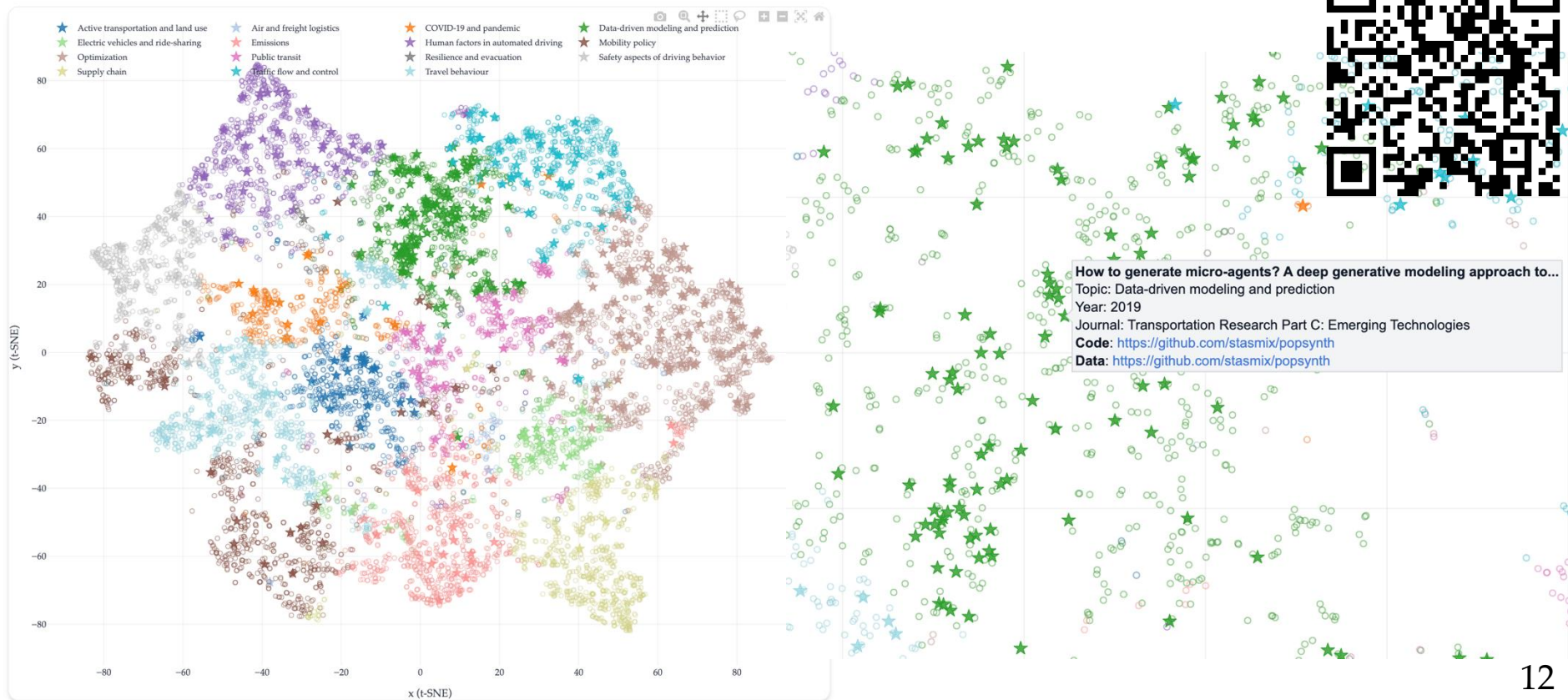
No citation premium

No significant lift in citations; the only interval off zero is slightly negative for code.

Interactive webpage to explore our analysis

<https://www.rerite.org/MOST>

RERITE



Interactive webpage to explore our analysis

<https://www.rerite.org/MOST>

RERITE



RERITE

Open Science Explorer for
Transportation Research

Click a dot to view details.

View Mode

Code

Data

Topic

All

Code Availability

All

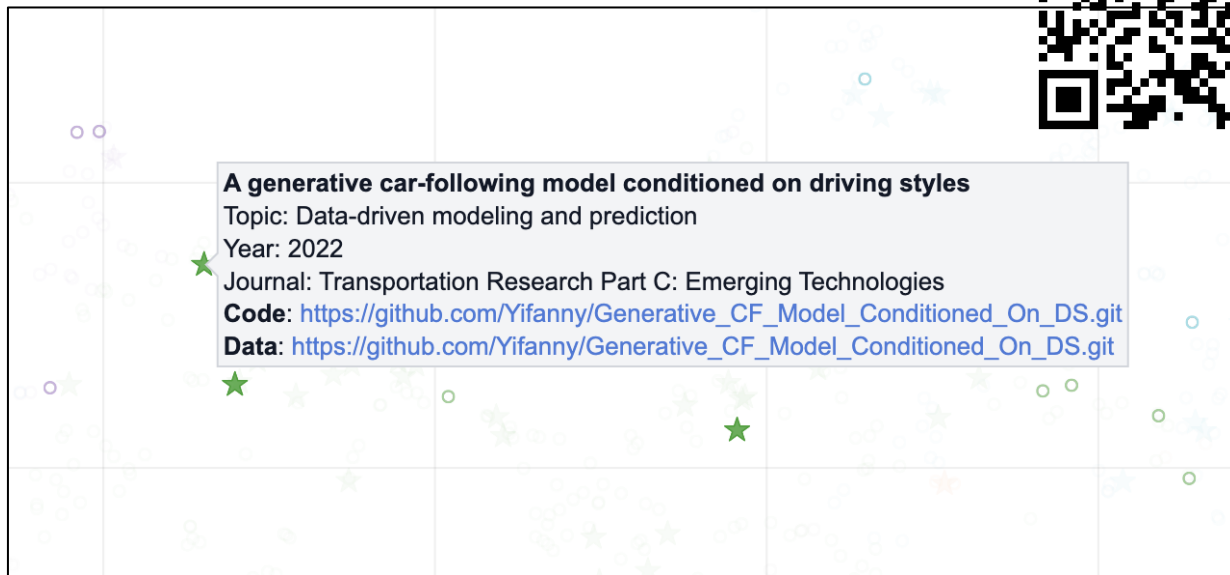
Search abstract (beta)

e.g., calibration

Random Walk

Discover a random paper with open data or code.

Search calibration



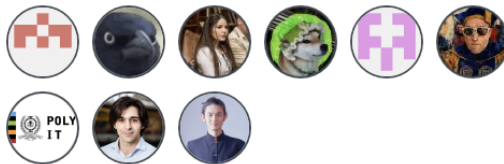
Challenges and lessons learnt from this project



Well... We're committed to code sharing.

It's **NOT** easy to maintain our GitHub repositories in a **fully reproducible** way with 9 active developers, multiple programming languages, 8 branches, 318 commits...

Contributors 9



Languages



Jupyter Notebook	77.9%
HTML	14.5%
Python	4.4%
R	1.6%
TeX	1.3%
Shell	0.2%
Other	0.1%

Awesome RERITE

<https://www.rerite.org/awesome-rerite>



Contents

- [Transportation Networks](#)
- [Discrete Choice Modeling](#)
- [Traffic Dynamics and Control](#)
 - [Vehicle trajectory data](#)
 - [Freeway traffic control](#)
 - [Urban traffic control](#)
- [Spatio-Temporal Data and Modeling](#)
- [Road Safety](#)
- [Optimization](#)
- [Transportation and AI](#)
- [Miscellaneous](#)

Traffic Dynamics and Control

Vehicle trajectory data

- [pNEUMA](#)
- Songdo dataset
 - [Songdo Traffic Dataset](#)
 - [Songdo Vision Dataset](#)
- [Waymo Open Dataset](#)
- [I-24 MOTION](#)
- [TGSIM I-294](#)
- [Zen Traffic Data](#)
- [OpenACC](#)
- [microSIM-ACC](#)

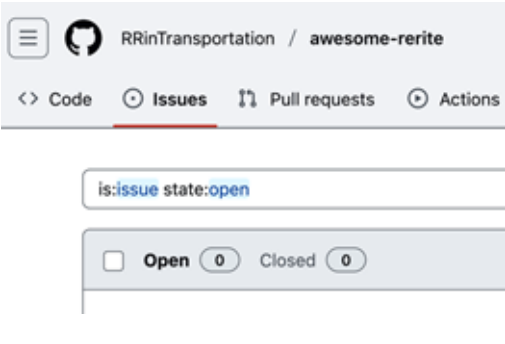
The Lord of the Rings

- [The Nagoya Traffic Jam Experiment](#)
- [The Dome Traffic Jam Experiment](#)
- [The Arizona Ring Experiments Dataset \(ARED\)](#)

Find more to share and contribute? It's easy!

We'd like to promote more and share more with the community!

Open an issue at the github page or email us!





rerite.org

REproducible Research In Transportation Engineering

- We are a working group in promoting Reproducible Research practices in Transportation;
- We are actively organizing tutorials, sharing resources within the community.

RR @ITSC 2024

[Home](#) [Requirements](#) [Future Activities](#)

Reproducibility in Transportation Research: A Hands-on Tutorial



Tutorial topics

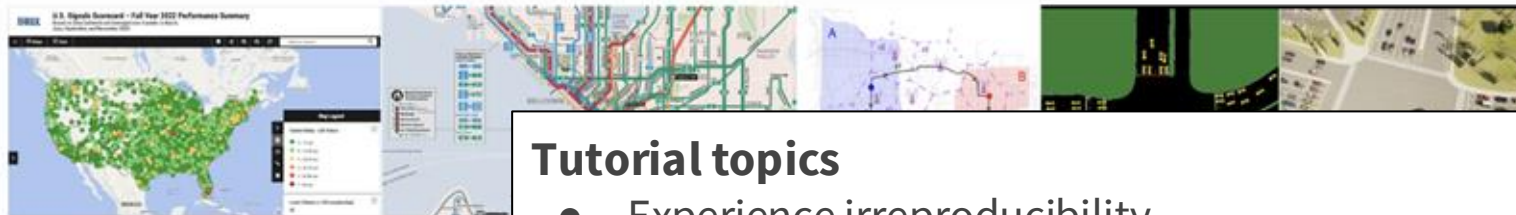
- Experience irreproducibility
- Step-by-step guide to making code reproducible on Github
- Making code pretty (understandable & reusable)

ITSC 2024, CPS-IoT Week 2025

ITSC 2025, TRB 2026



Reproducibility in Transportation Research: A Hands-on Tutorial



Tutorial topics

- Experience irreproducibility
- Step-by-step guide to making code reproducible on Github
- Making code pretty (understandable & reusable)

Year of 2026: Open science meets Agentic AI

The international journal of science / 8 February 2024

nature

Embrace open science – before it's too late

A UNESCO report laments the lack of progress towards making science more open. More awareness could aid efforts to achieve the UN's Sustainable Development Goals.

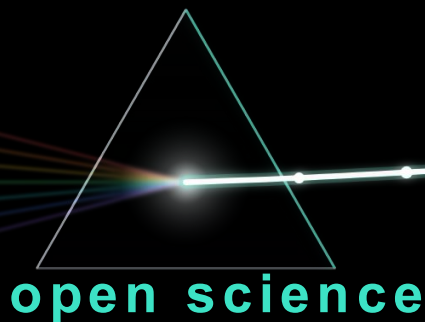
measures to monitor progress.

The European Commission, based in Brussels, was an early proponent of open science. Between 2002 and 2020, it increased its funding for 'societal engagement' projects from €88 million to €462 million – an amount that is now equivalent to US\$500 million. Moreover, a decade ago, all scientific publications arising from the European Union's €80 billion Horizon 2020 programme needed to be published open-access. Citizen science is another growing area in open science with much promise, UNESCO notes. By 2018, half of all records in the Global Biodiversity Information Facility – an international open-access data repository based in Copenhagen – were from citizen scientists, up from around 10% in around 2007.

- Agentic AI is now able to reproduce **87%** (N=92) end-to-end empirical reanalysis in economic publications based on **open data** (Xu and Yang, 2026).

// the field's open work
a whole spectrum of contributions

// concentrated —
the field moves forward →



// Run the prism in reverse: the field's scattered open work — **open datasets**, **transparent benchmarks**, **credible testbeds** — concentrates through shared, reproducible practice into one bright line of progress that carries the whole community forward.

**The need for research infrastructure
& Call for collaborations and community efforts**



RE3T

Thank you!
We love feedback.

Questions/comments

Email: junyi.ji@vanderbilt.edu

Joint research efforts with Ruth Lu (MIT), Linda Belkessa (Université Gustave Eiffel), Liming Wang (Portland State), Silvia Varotto (ENTPE), Yongqi Dong (RWTH Aachen University), Nicolas Saunier (Polytechnique Montreal), Gregory S. Macfarlane (BYU), Bahman Madadi (ENTPE), Mostafa Ameli (Université Gustave Eiffel), Cathy Wu (MIT)

RERITE



rerite.org



Project
homepage

REproducible Research In Transportation Engineering



rerite.org

@ The IEEE Intelligent Vehicles Symposium
Detroit, MI, United States

State of the Union: Open Science in Transportation 2026

Junyi Ji (Vanderbilt University)

Joint research efforts with Ruth Lu (MIT), Linda Belkessa (Université Gustave Eiffel), Liming Wang (Portland State), Silvia Varotto (ENTPE), Yongqi Dong (RWTH Aachen University), Nicolas Saunier (Polytechnique Montreal), Gregory S. Macfarlane (BYU), Bahman Madadi (ENTPE), Mostafa Ameli (Université Gustave Eiffel), Cathy Wu (MIT)