Constructing tomorrow with data

NFDI4ING
Research data management in engineering sciences

Data Mesh for RDM in the Engineering Sciences

26. August 2025, 16:20 – 16:30

CoRDI 2025, Aachen
Track “RDM Infrastructure”

Mario Moser, Tobias Hamann, Anas Abdelrazeq, Robert H. Schmitt

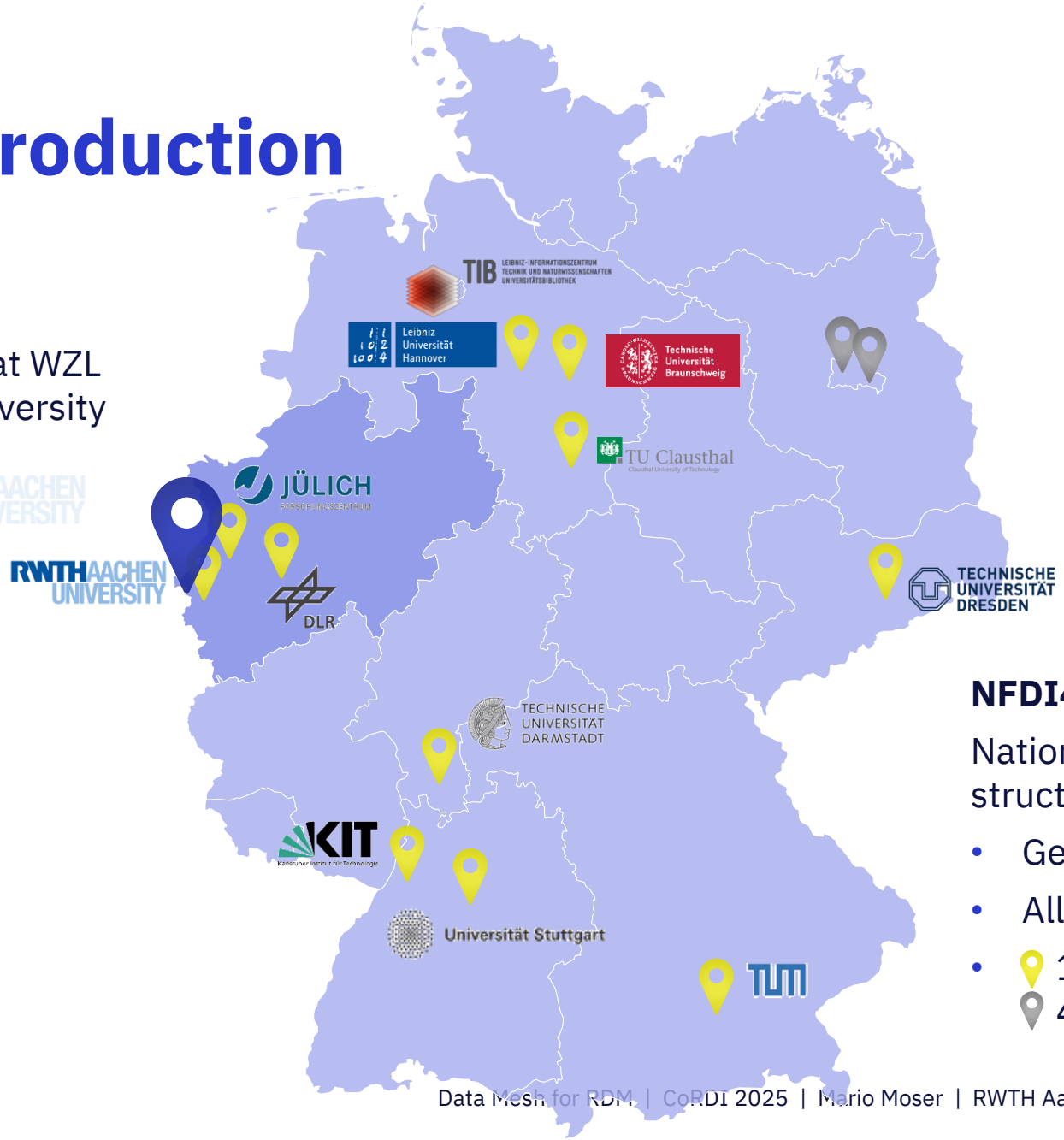
Werkzeugmaschinenlabor WZL of RWTH Aachen University

DOI 10.5281/zenodo.16948822

Quick Introduction

Mario Moser

Research Associate at WZL
of RWTH Aachen University

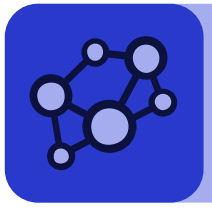


NFDI4ING

National Research Data Infra-structure for Engineering Sciences

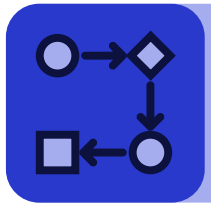
- Germany-wide
- All engineering disciplines
- 11 (Co-)Applicants,
  44 Participants

Characteristics and challenges of research data* and the RDM landscape, esp. in the engineering sciences

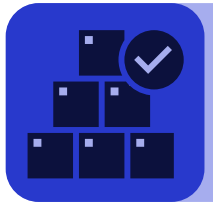


Decentralised w.r.t. to tools and organisation

Findability of data is a current issue



Reuse purpose of research data **unknown upfront**



'Brownfield' with **existing tools and services**

Research is **interdisciplinary**, even in engineering

Researcher's **time** at university << Retention **period** of data

Heterogeneous data and data sources

Research Data Management is **not a technical topic soley**

Under this characteristics, **how could a Research Data Infrastructure look like?**

* Scope specification: 'Data' in the sense of tabular and 'big' data, while software is excluded

The (industrial) Data Mesh concept

– Characteristics –

- **Industrial Data Management Approach**
- **Decentralised datasets**
- **Socio-technical approach**
- **Not a software installer, not a specific technology**

– Four Data Mesh Principles –

Domain Ownership

With their domain knowledge, domain teams are responsible for managing their data

Data (as a) Product

Data is ‘productised’ for future usage with metadata and infrastructure, accessible via API

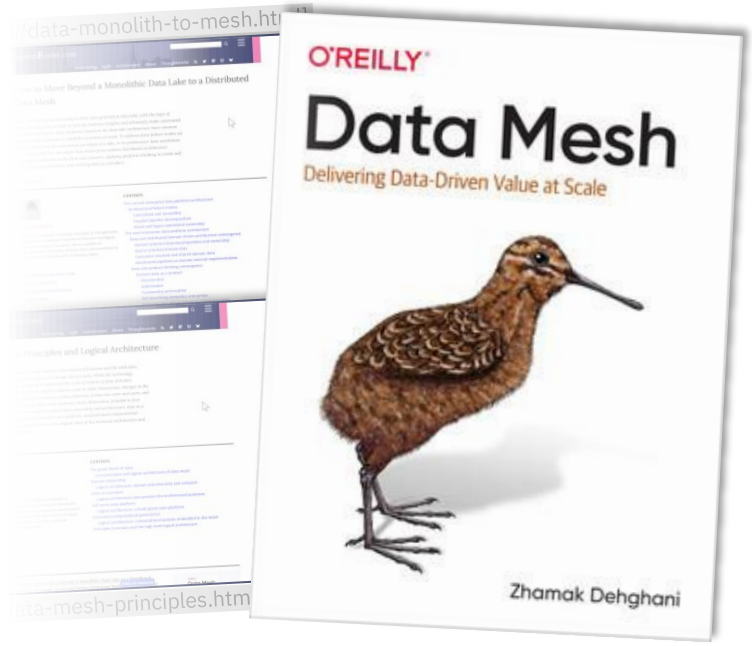
Self-Serve Platform

Datasets are findable and maintained on a central platform along their lifecycle, incl. KPIs

Federated Governance

Data governance is split on a global level for interoperability, and a local domain level for specificity

– Origins –



amak Dehghani (2022): Data Mesh
eilly Media, Inc.
N: 9781492092391
://learning.oreilly.com/library/cover/9781492092384/250w/]

Hypothesis: Data Mesh as Data Infrastructure for RDM in the engineering sciences

Domain Ownership — F A I R
- Socio-technical
- Decentralised
- Existing tools and services
- Researcher's not available

Data (as a) Product — F A I R
- Unknown reuse purpose
- Inter-disciplinarity
- Data (sources) heterogeneity
- Researcher's not available

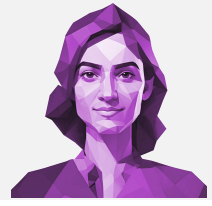
Self-Serve Platform — F A I R
- Findability
- Unknown reuse purpose

Federated Governance — F A I R
- Unknown reuse purpose
- Inter-disciplinarity

Mario Moser *et al.* (2025): **Envisioning and proposing Data Mesh for Research Data Management in the Engineering Sciences** https://preprints.inggrid.org/repository/view/53/ (Preprint)

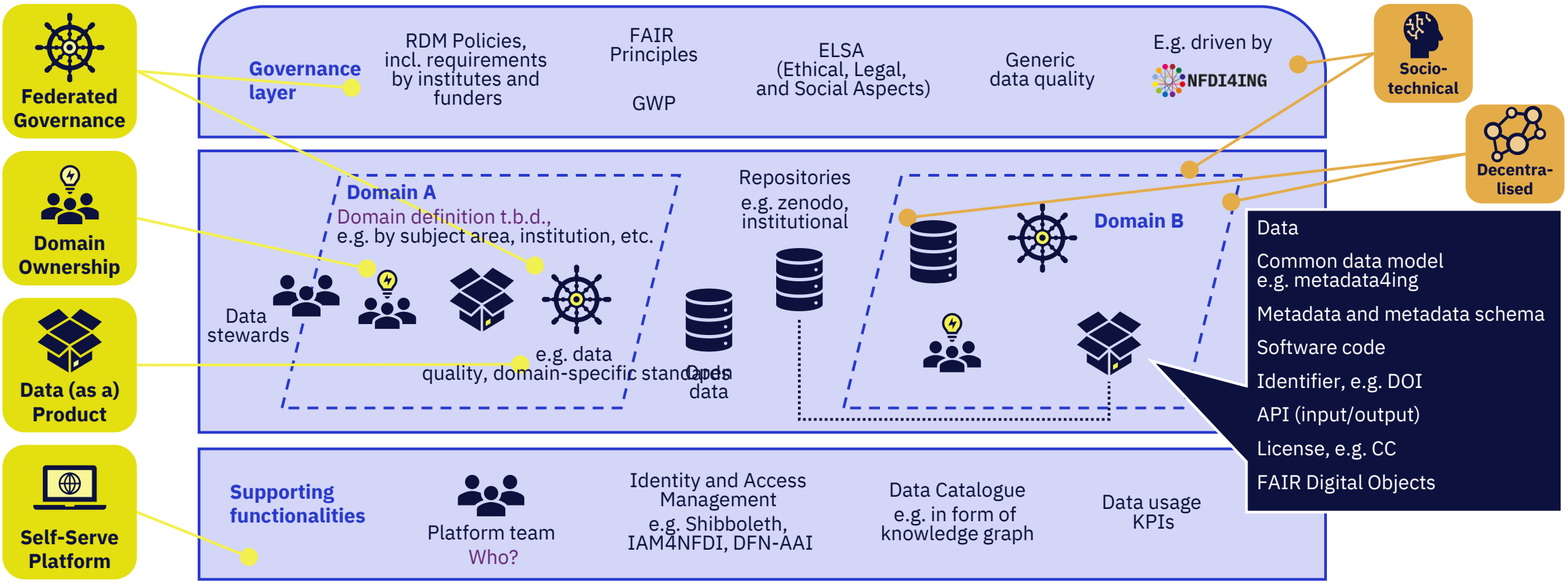
But: The (industrial) Data Mesh concept is not transferable 1:1 to RDM

Scope for NFDI4ING's archetype in **TA FIONA**



Target picture of an engineering sciences Data Mesh within NFDI

Potential technologies and partners from NFDI(4ING) and beyond, without any alignment or commitment yet; without saking for completeness



Outlook: Next steps in NFDI4ING's archetype Fiona

Expert interviews on **needs and component**s for a RDM Data Mesh

Identification of **existing RDM components** for a RDM Data Mesh

Adaption of Data Mesh concept for research data

- Domain definition
- Specification of data products
- Characeristics of science
- Inter-org. data sharing

Designing a **prototypical RDM Data Mesh**

Six archetypes in NFDI4ING:

- Prototypical engineering researchers
- Can belong to more than one engineering subject area
- Have a certain research method or approach (bespoke experiments, research software, samples tracing, HPMC, simulations)
- Plus archetype **Fiona**: ***Data re-use, combination and enrichment*** (rf. https://nfdi4ing.de/fiona and https://doi.org/10.5281/zenodo.15647808)

Thank you.

Mario Moser

Research Associate

WZL of RWTH Aachen University

 mario.moser@wzl-iqs.rwth-aachen.de

 @mmoser (NFDI instance)

For all other NFDI4ING-related topics, check out or contact:

 https://nfdi4ing.de

 info@nfdi4ing.de

 https://questions.nfdi4ing.de/

 https://nfdi4ing.de/services

 https://nfdi4ing.de/fiona

 NFDI4ING

Acknowledgement

“ Mario Moser would like to thank the Federal Government and the Heads of Government of the Länder, as well as the Joint Science Conference (GWK), for their funding and support within the framework of the NFDI4ING consortium.

Funded by the German Research Foundation (DFG) - project number 442146713. ”

Availability of presentation DOI 10.5281/zenodo.16948822

Extended abstract: DOI 10.5281/zenodo.16736138

Let‘s exchange: Q&A and Feedback

Your direct Questions

Comments

Thoughts

Potential **inter-view** partner

Raise your hand
and let‘s talk directly

… or get into contact with me
(Contact details on previous slide)

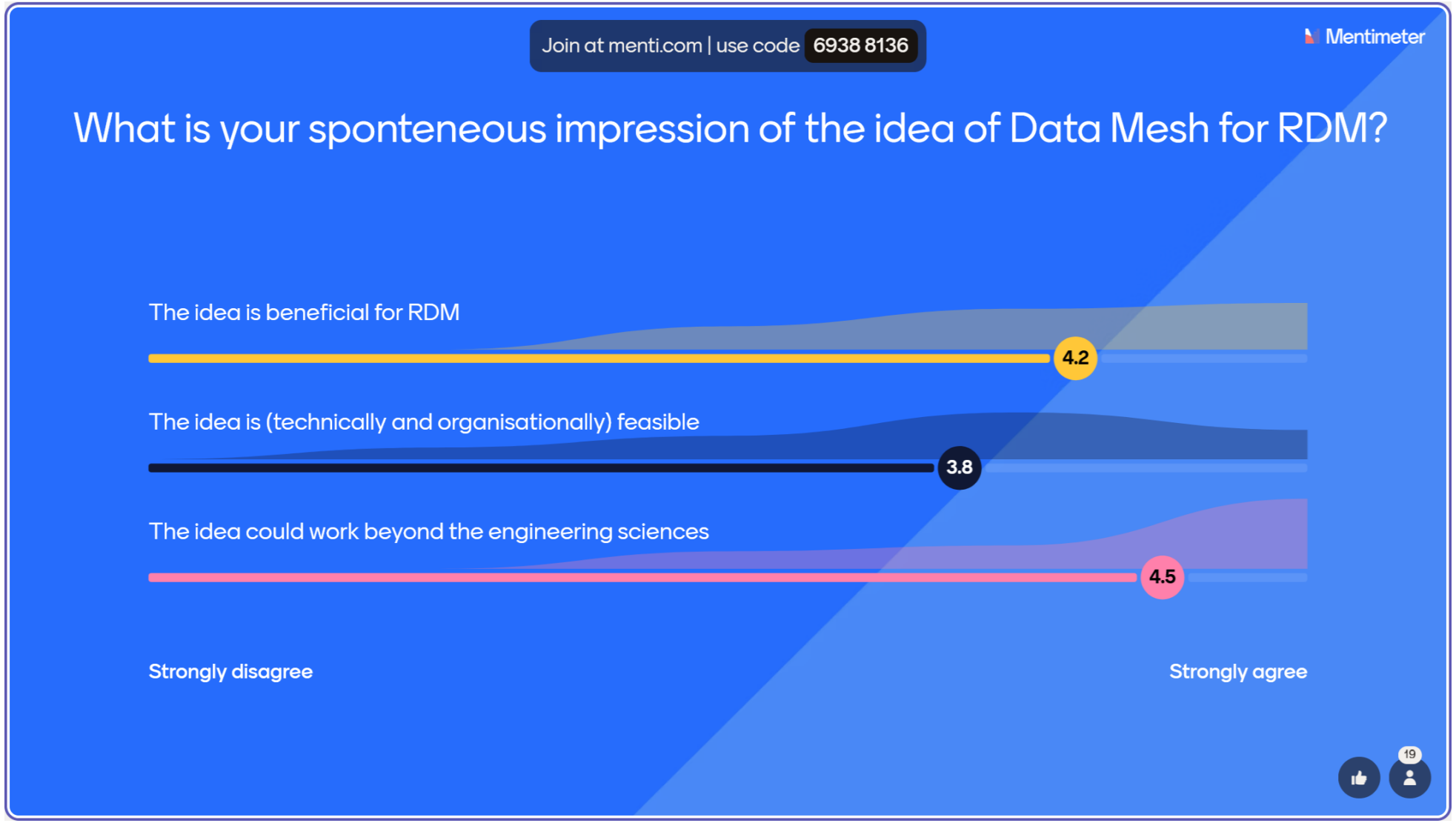
Scan the QR code and
give your input via Mentimeter

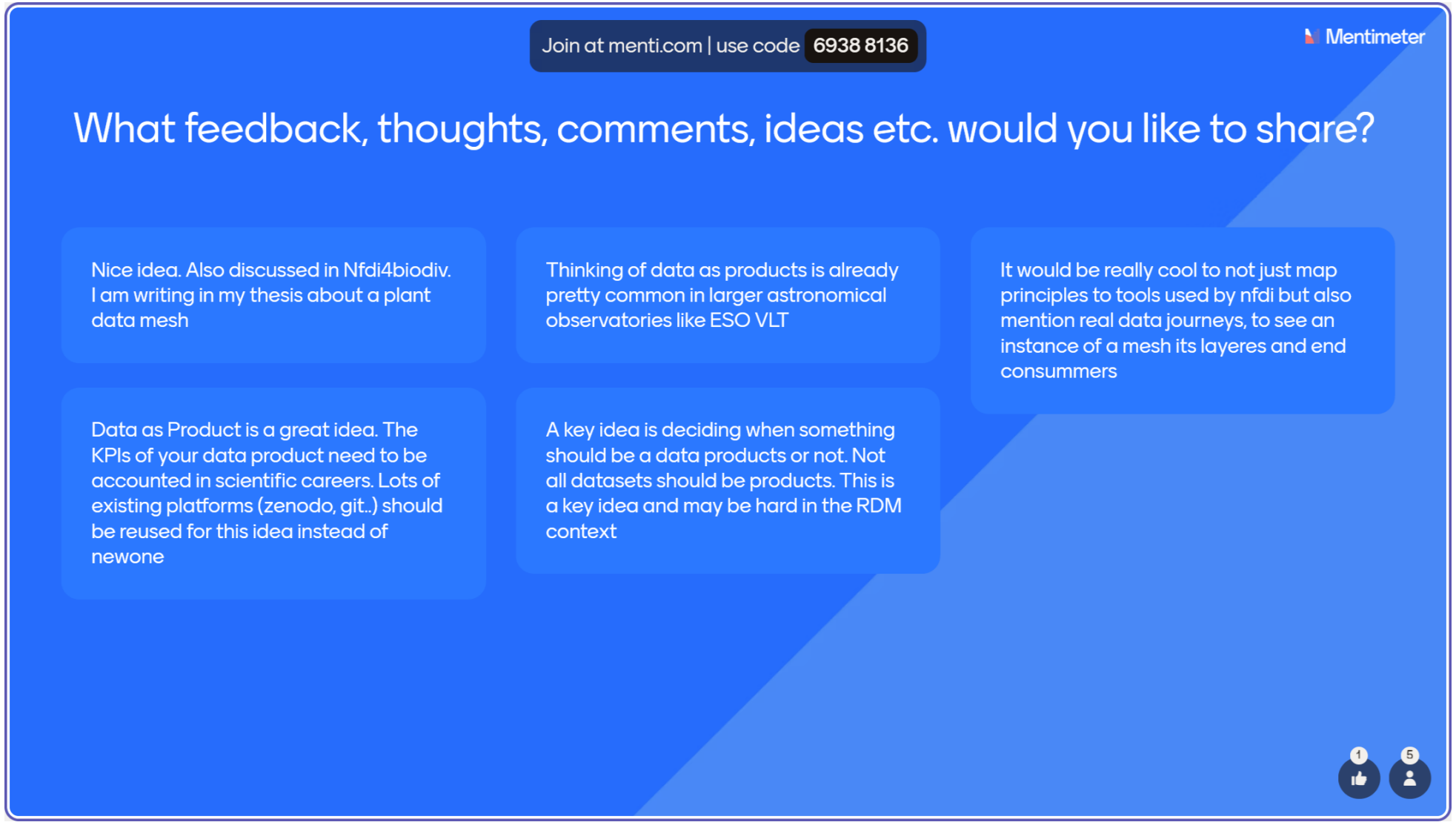
(Results will be shared publicly, link valid on 2025-08-26 only)

Ideas

For **Questions** as well

Feedback


Join at menti.com | use code 6938 8136
Mentimeter
What is your sponteneous impression of the idea of Data Mesh for RDM?
The idea is beneficial for RDM
4.2
The idea is (technically and organisationally) feasible
3.8
The idea could work beyond the engineering sciences
4.5
Strongly disagree
Strongly agree
19


Join at menti.com | use code 6938 8136
Mentimeter
What feedback, thoughts, comments, ideas etc. would you like to share?
Nice idea. Also discussed in Nfdi4biodiv. I am writing in my thesis about a plant data mesh
Thinking of data as products is already pretty common in larger astronomical observatories like ESO VLT
It would be really cool to not just map principles to tools used by nfdi but also mention real data journeys, to see an instance of a mesh its layeres and end consummers
Data as Product is a great idea. The KPIs of your data product need to be accounted in scientific careers. Lots of existing platforms (zenodo, git..) should be reused for this idea instead of newone
A key idea is deciding when something should be a data products or not. Not all datasets should be products. This is a key idea and may be hard in the RDM context
1
5