

The Personal Health Train

Johan van Soest

Clinical Data Science group
MAASTRO Clinic – Knowledge Engineering
Maastricht University Medical Centre+

Positions

- Reusable personal (health) data
- Machine interpretable at the core
- Distributed and federated data analysis approaches
- Data capture at the source
- Responsible use of personal (health) data
- An open ecosystem
- (Pay back) incentive for data providers

Barriers to sharing data

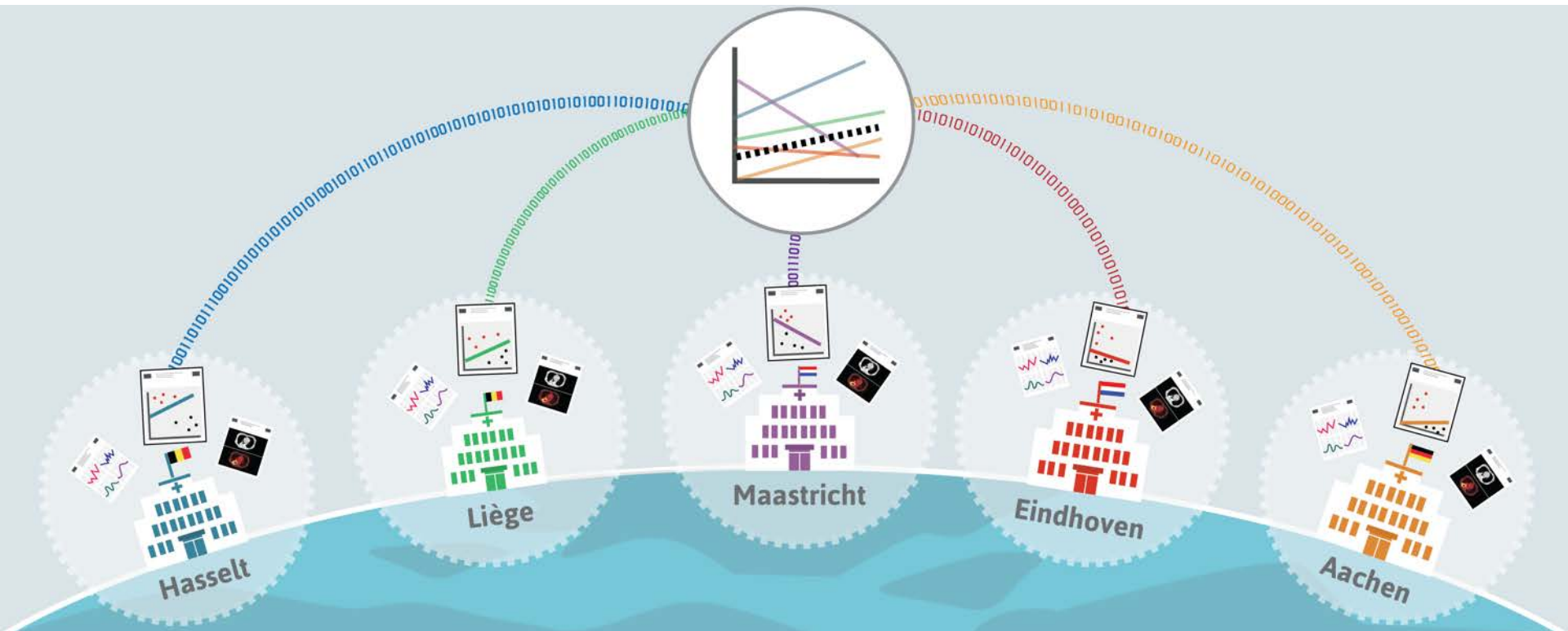
[..] the problem is not really technical [...].
Rather, the problems are **ethical, political, and administrative.**

Lancet Oncol 2011;12:933

1. Administrative (I don't have the resources)
2. Political (I don't want to)
3. Ethical (I am not allowed to)
4. Technical (I can't)



Sharing algorithms instead of data

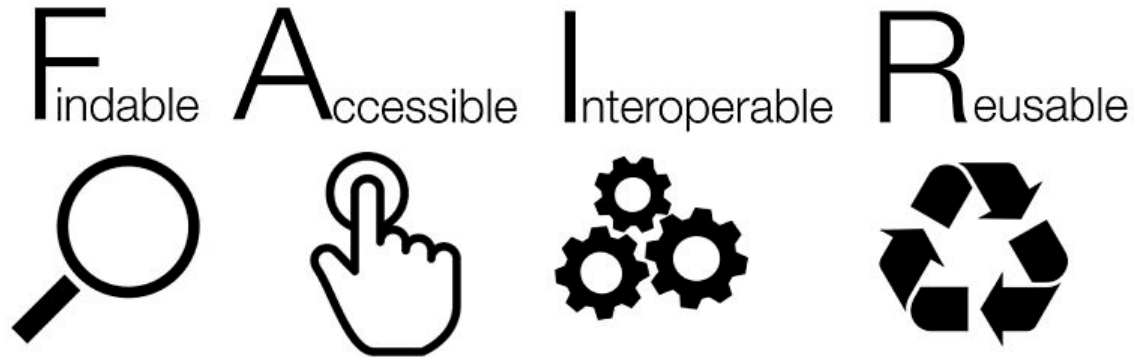


Application of Machine Learning for Multicenter Learning, [van Soest, Dekker, Roelofs, Nalbantov \(2015\)](#)
Infrastructure and distributed learning methodology, [Deist et al. \(2017\)](#)



<https://www.youtube.com/watch?v=mktAtHmy-FM>

FAIR services



A set of 15 principles that apply to *any* digital resource and their metadata.

data, data repositories, software, web services, scholarly publications

SCIENTIFIC DATA

The FAIR Guiding Principles for scientific data management and stewardship

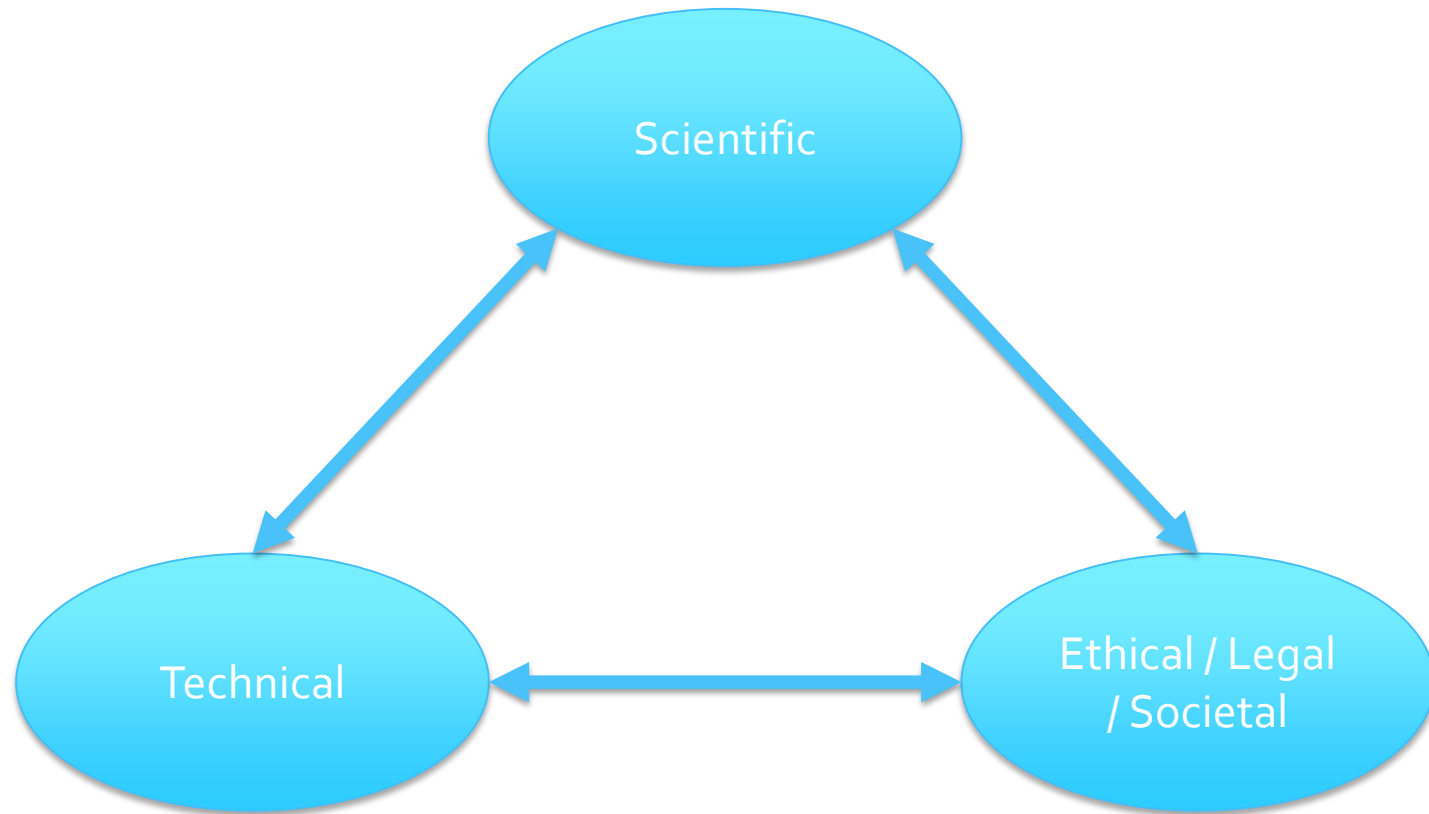
Mark D. Wilkinson, Michel Dumontier [...] Barend Mons

[Affiliations](#) | [Contributions](#) | [Corresponding author](#)

Scientific Data 3, Article number: 160018 (2016) | doi:10.1038/sdata.2016.18
Received 10 December 2015 | Accepted 12 February 2016 | Published online 15 March 2016

About Scientific Data
Scientific Data is an open-access, peer-reviewed journal for descriptions of scientifically valuable datasets. Our primary article-type, the **Data Descriptor**, is designed to make your data more discoverable, interpretable and reusable.

Triangle of constraints



Current projects



**Center for
Big Data Statistics**

Towards Smart Statistics



Maastricht University



Maastricht UMC+