



Building up the National Research Data Infrastructure (NFDI)

The story so far (Part I)

York Sure-Vetter

September 22nd, 2025 @ Edinburgh

Vision

Data as a shared resource

"Data as a shared resource for excellent research, organised by the scientific community in Germany"



Investment into NFDI

Up to 90 Million EUR per year from 2018 to 2028

90% funding from German federal state, 10% funding from German states

NFDI in figures

As of September 2025

Stakeholders



- > 1,500 contributors
- 322 association members
- 26 consortia
- 1 basic service consortium
- 5 sections
- > 390 participating organisations in the



Services

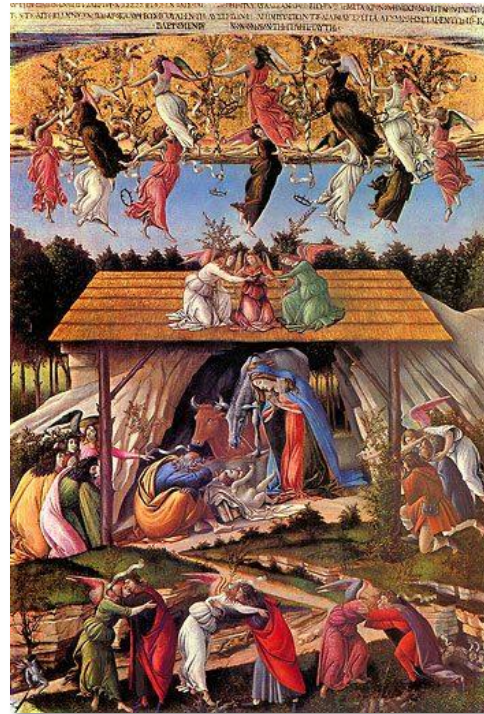
- > 650 active services
- > 180 planned services
- 8 basic services developed

Networking



- > 780 organised events
- 10 regular community meetings
- 31 working groups
- 681 participants in CoRDI 2023

View on societies



Source: Wikimedia Commons

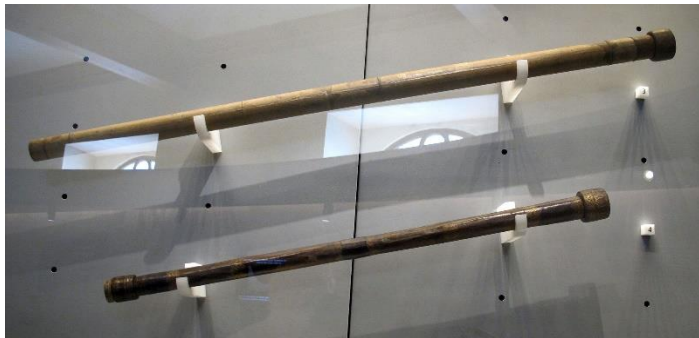
„Birth of Jesus Christ“
Sandro Botticelli



Source: destatis

Census 2011

View to far horizons



Source: Sailko CC BY-SA 3.0

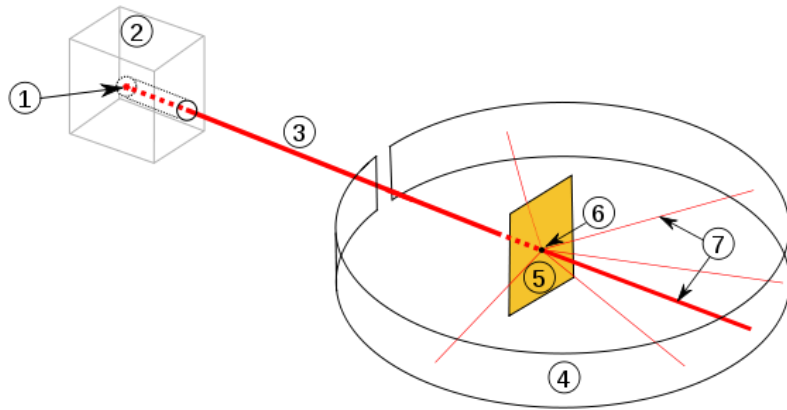
Telescope
(invented 1608)



Source: NASA Wikimedia Commons

Hubble Space Telescope
(launched April 25th, 1990)

Glimpse into insight



Source: Wikimedia Commons

Rutherford Scattering (1909)



Source: CERN

Large Hadron Collider (LHC, since 2008)

More and more research results are achieved
by using **already existing** research data.

see FAIR principles: <https://doi.org/10.1038/sdata.2016.18>

Basic idea of NFDI

Research data sets are FAIR*.
For all**. Forever***.

- * Findable, Accessible, Interoperable, Reusable
- ** According to data access categories
- *** Alice: „How long is forever?“
White rabbit: „Sometimes, just one second.“

Early AI visions

Machine intelligence increases

Herbert Simon predicts in 1957 that a computer will be **world chess champion** in less than 10 years.

Formal knowledge increases

John McCarthy proposes in 1958 to transform the **complete knowledge of humankind** in first-order logic.

Vision: machine intelligence increases

Herbert Simon predicts in 1957 that a computer will be **world chess champion** in less than 10 years.

+40 years (1997):
IBM Deep Blue becomes
Chess champion



The story: <https://fivethirtyeight.com/features/the-man-vs-the-machine-fivethirtyeight-films-signals/>

Vision: machine intelligence increases

Herbert Simon predicts in 1957 that a computer will be **world chess champion** in less than 10 years.

+40 years (1997):
IBM Deep Blue becomes
Chess champion

+14 years (2011):
IBM Watson becomes
Jeopardy champion



The story: <https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html>

Vision: machine intelligence increases

Herbert Simon predicts in 1957 that a computer will be **world chess champion** in less than 10 years.

+40 years (1997):

IBM Deep Blue becomes
Chess champion

+14 years (2011):

IBM Watson becomes
Jeopardy champion

+5 years (2016):

Google DeepMind AlphaGo
becomes **Go** champion



The story: <https://www.youtube.com/watch?v=WXuK6gekU1Y>

Vision: formal knowledge increases

John McCarthy proposes in 1958 to transform the **complete knowledge of humankind** in first-order logic.

+32 years (1990):
World Wide Web,
the **global information space**



Source: https://de.wikipedia.org/wiki/Tim_Berners-Lee

Vision: formal knowledge increases

John McCarthy proposes in 1958 to transform the **complete knowledge of humankind** in first-order logic.

+32 years (1990):
World Wide Web,
the **global information space**

+11 years (2001):
Wikipedia,
the free Internet **encyclopedia**



*60+ million pages in
300+ different languages
that anyone can edit*

Source: <https://de.statista.com/statistik/daten/studie/195081/umfrage/anzahl-der-artikel-auf-wikipedia-weltweit/>

Vision: formal knowledge increases

John McCarthy proposes in 1958 to transform the **complete knowledge of humankind** in first-order logic.

+32 years (1990):

World Wide Web,
the **global information space**

+11 years (2001):

Wikipedia,
the free Internet **encyclopedia**

+11 years (2012):

WikiData,
the free **knowledge base**



*118+ million multilingual data items
that anyone can edit*

Source: <https://www.wikidata.org>

Knowledge Graphs are
meaningful to humans
and to machines.

Data as a common good for
excellent research,
organized by science in Germany.

Vision of NFDI

Big picture



Better research!

Big picture



Better research!



Scientist



Infrastructure
Staff



Data Steward

People / Roles

Big picture



Better research!



Scientist

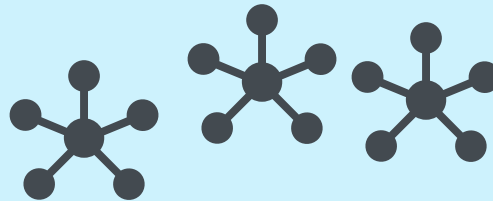


Infrastructure
Staff



Data Steward

People / Roles



Data & Metadata

Big picture



Better research!



Scientist

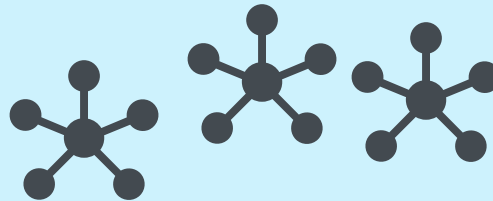


Infrastructure
Staff



Data Steward

People / Roles



Data & Metadata



Infrastructure



Better research!

Processes, Standards, Tools, Ethical & Legal Frameworks ...



Scientist

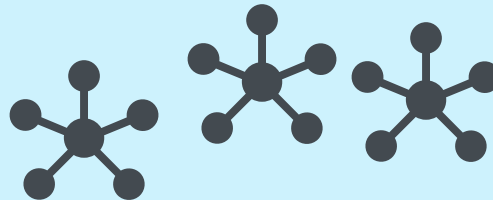


Infrastructure
Staff



Data Steward

People / Roles



Data & Metadata



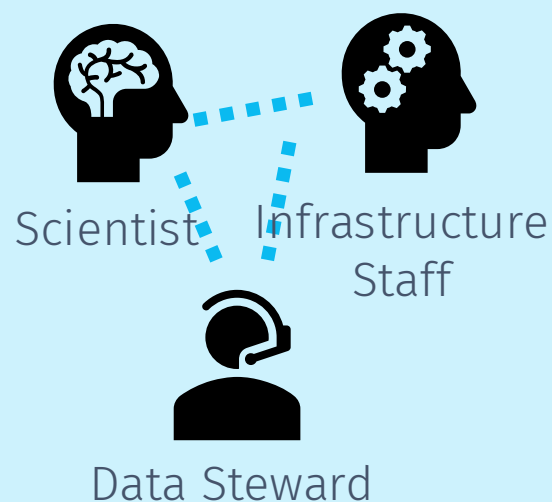
Infrastructure

Big picture



Better research!

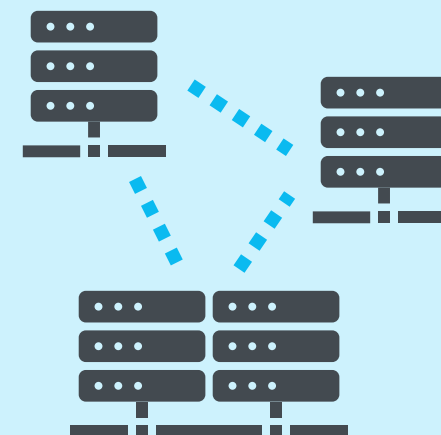
Processes, Standards, Tools, Ethical & Legal Frameworks ...



People / Roles

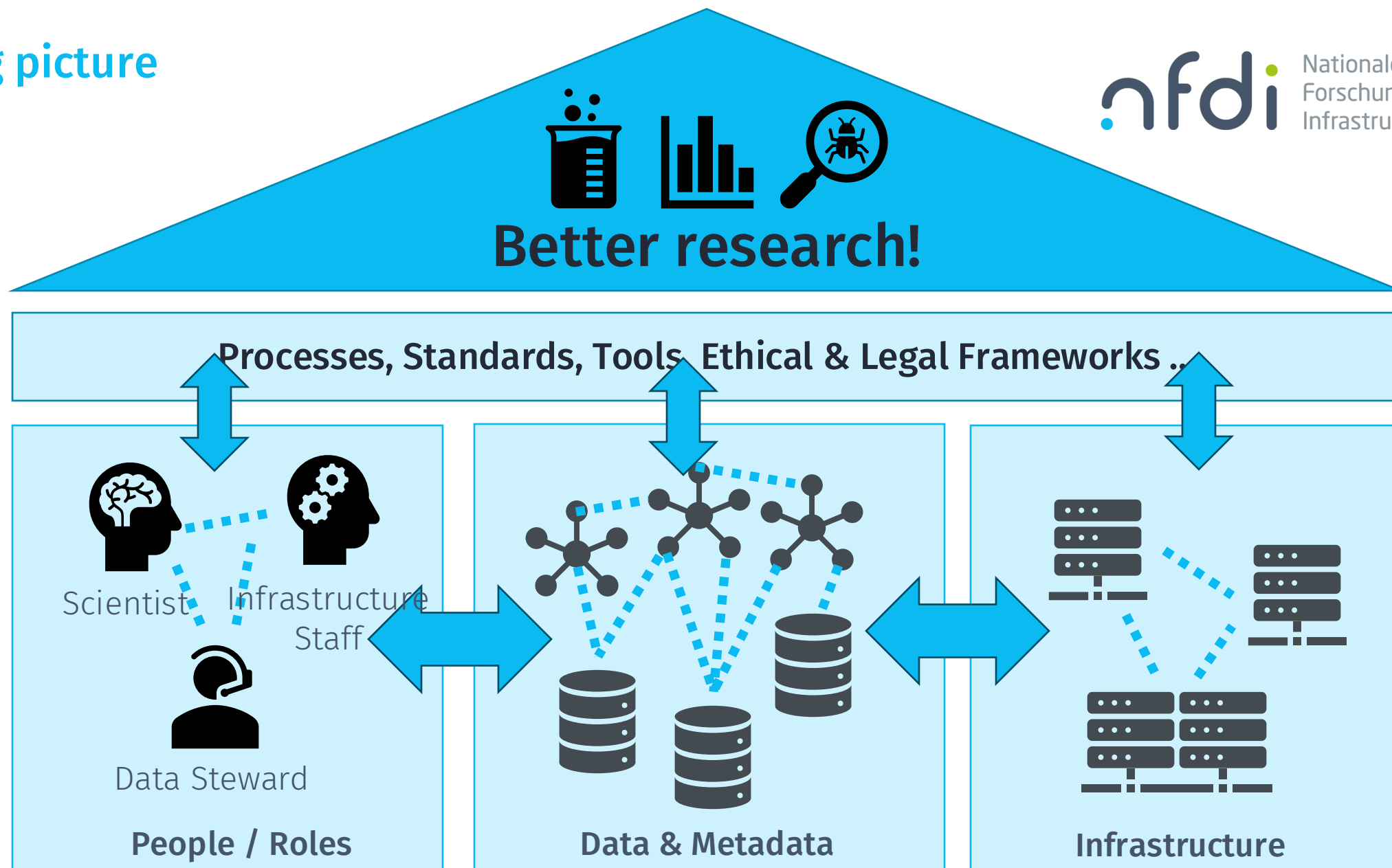


Data & Metadata

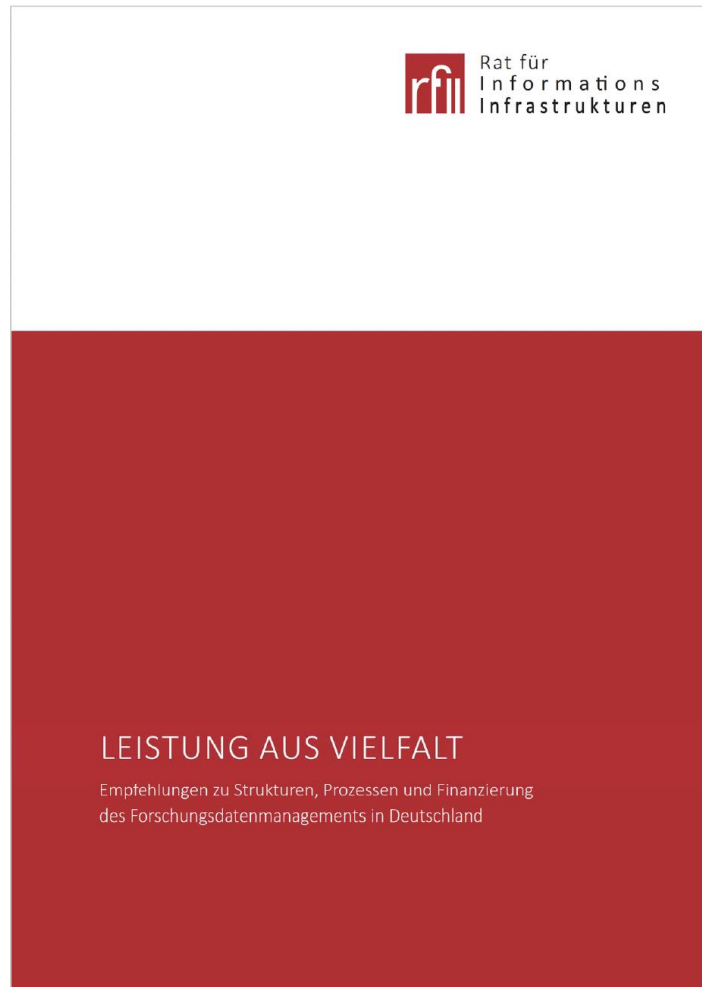


Infrastructure

Big picture



Recommendation to found a „NFDI“ of the Council on Information Infrastructures (2016)



Rat für Informationsinfrastrukturen:
Leistung aus Vielfalt. Empfehlungen zu
Strukturen, Prozessen und Finanzierung
des Forschungsdatenmanagements in
Deutschland, Göttingen 2016, 160 S.

Source: Rat für Informationsinfrastrukturen (RfII)

Goals of funding NFDI

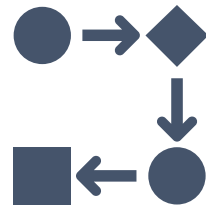
according to the agreement between German Federal Government and German States (2018)

Building up



a networked
information
infrastructure &
research data
management

Establishing



standardized
processes for
research data
management

Developing



a sustainable
portfolio of
services and
meta data
standards

Connecting



with European
platforms (in
particular EOSC)

Strengthening



the subsequent
use of existing
data &
the integrity,
quality and
protection of data

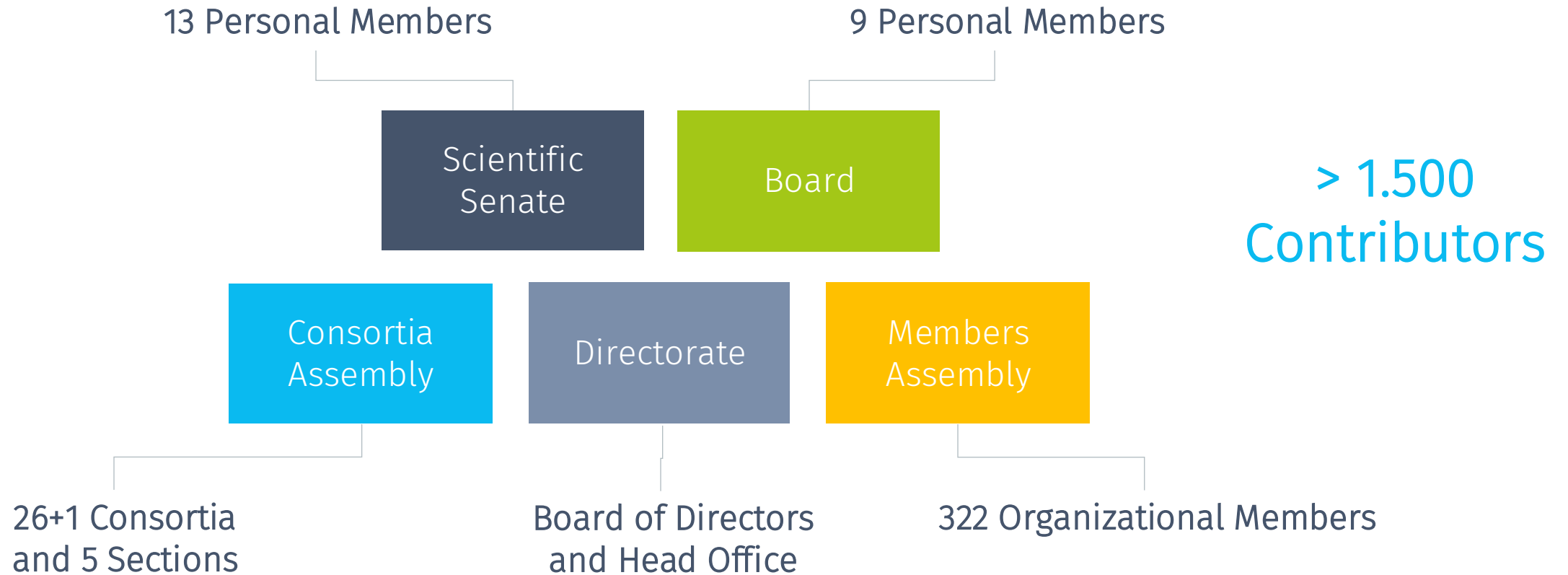
Foundation of NFDI Association in Hannover (2020)



Founding assembly on October 12th, 2020, in Hannover (Source: NFDI Assoc.)

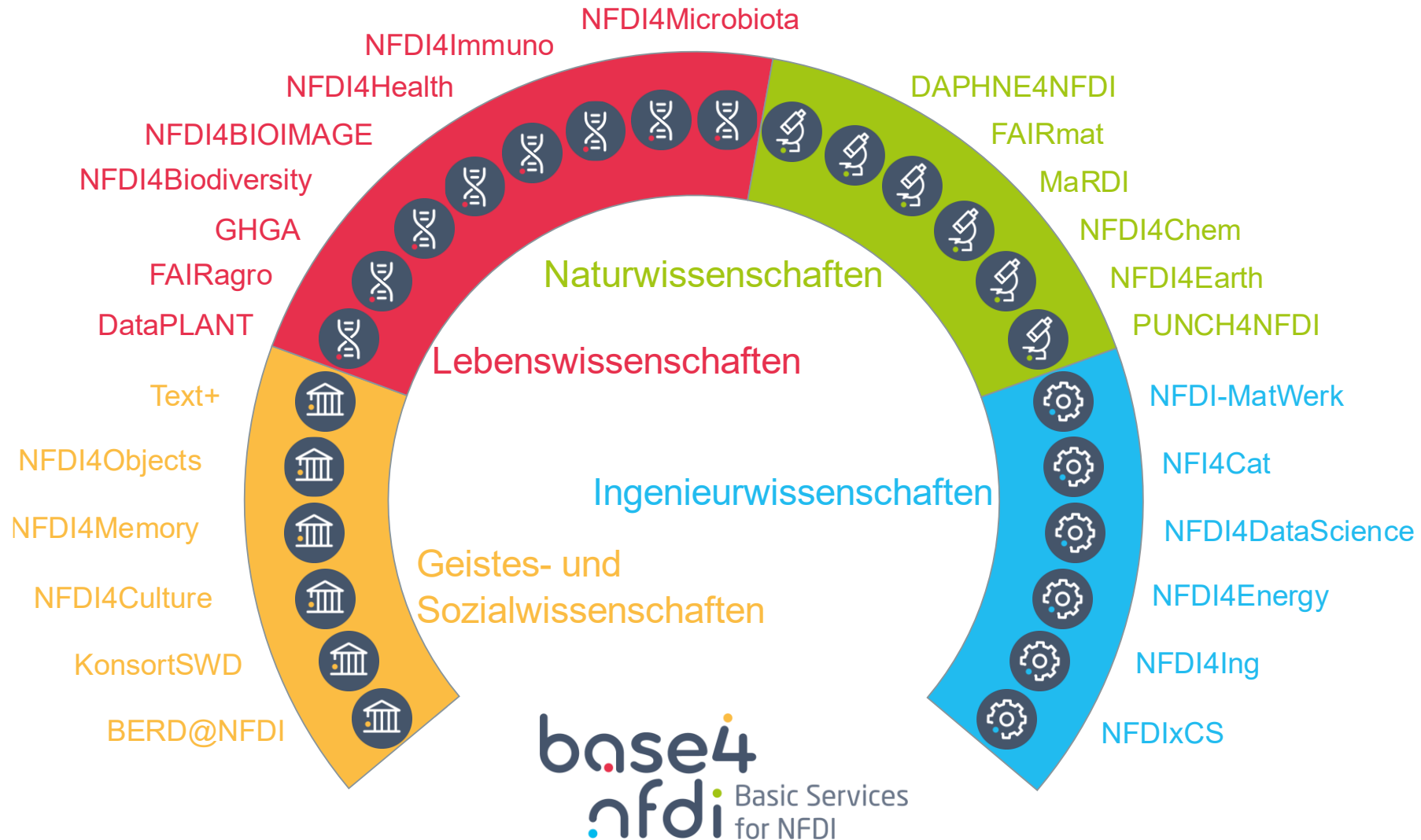
Organisational Structure (1/3)

consists of five organs, ...



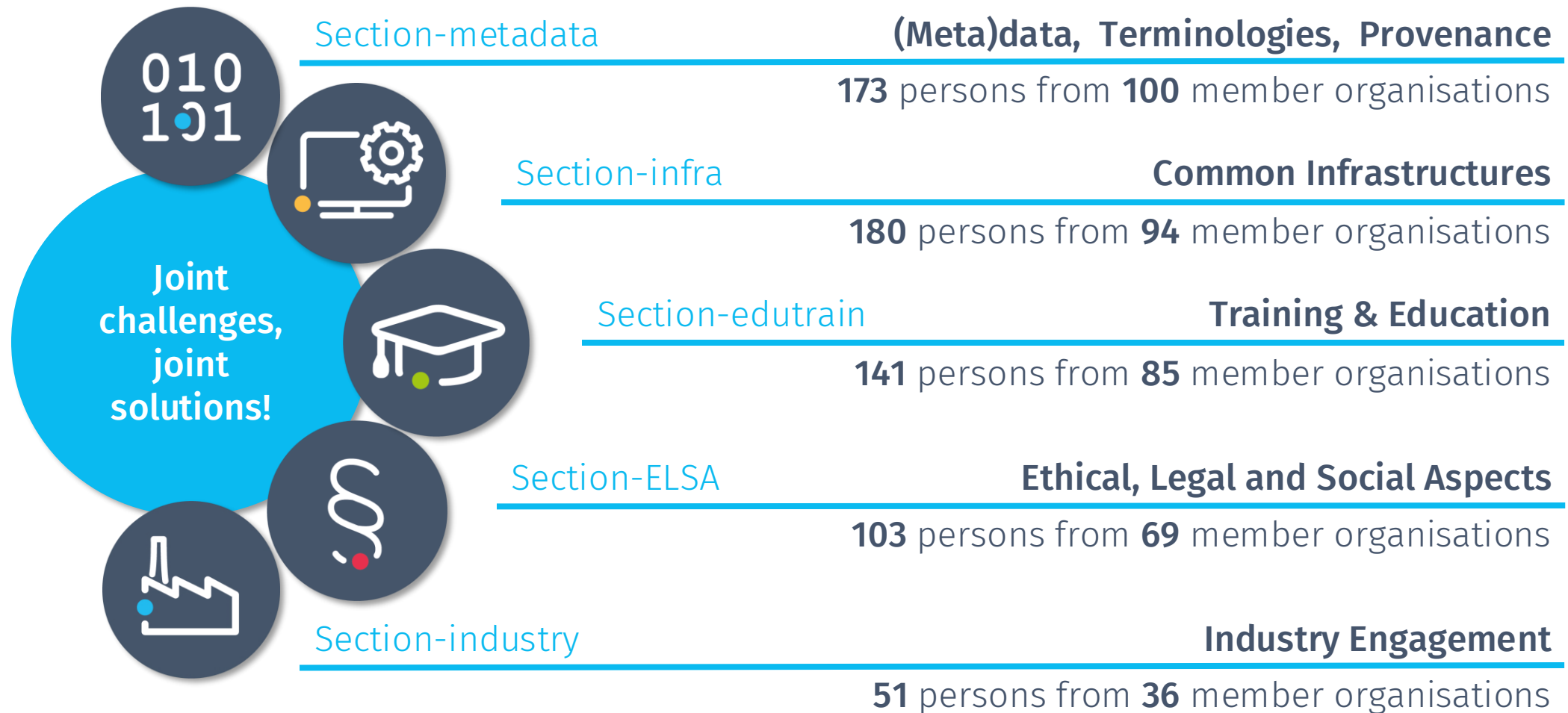
Organisational Structure (2/3)

... 26+1 consortia (incl. one for basic services) ...



Organisational Structure (3/3)

... and five sections



Base4NFDI delivers basic infrastructure services

... more services advancing

IAM4NFDI

- Identity- und Access-Management (IAM)
- Goal: unified access to data, software and computing resources for collaborative work

PID4NFDI

- Persistent Identifier Services (PID) for the German National Research Data Infrastructure
- Goal: adapting existing PID systems to specific needs of NFDI

TSA4NFDI

- Terminology Services 4 NFDI
- Goal: harmonising terminologies from various research communities

Cooperation agreement between DIN and NFDI

NFDI (22 October 2024)

Objective of the cooperation agreement:

*"Standardisation and standardisation of research data and data infrastructure for science, standardisation and normalisation for science, industry and society and thereby increasing the **efficiency of the German scientific system** and **the innovative strength of Germany as a business location.**"*



- joint information events
- Joint projects for the standardisation of research data and data infrastructures
- Joint prioritisation of topics with stakeholder groups
- Mutual membership in relevant committees
(e.g. in DIN: FOCUS.digital; in NFDI: sections, e.g. metadata)

Standardisation: Success Story

The German Human Genome Phenome Archive (GHGA)



Highly sensitive data from different sources is brought together:

- Health data from clinical studies
- Locally collected genome data sets

Standardisation in data processing:

- data protection compliant through a data protection framework and standardised usage agreements
- Interoperable through standardised metadata

Enables:

- Efficient linking of data to gain new insights
- Use of modern methods: machine learning / artificial intelligence
- Investigation of rare genetic diseases

→ <https://www.ghga.de>



Connecting Internationally with the European Open Science Cloud (EOSC)

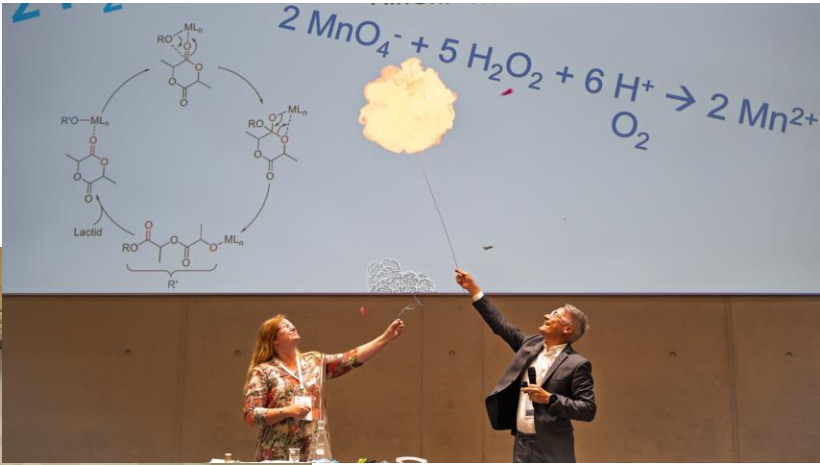
- European Open Science Cloud
 - The goal of EOSC is the development of a "*Web of FAIR Data and Services*" for science in Europe.
 - 256 members (28 mandated organisations)
- NFDI is mandated member of EOSC
 - Coordinating the activities of German EOSC members
 - Active role in EOSC events:
EOSC Symposium 2024 in Berlin – slogan „NFDI meets EOSC“



Quelle: www.flickr.com/photos/195720756@N05/52924773718/in/album-72177720308560519
© EOSC Association

6th NFDI Members Assembly 2nd CoRDI Conference

August 25th-28th, 2025
RWTH Aachen, Germany



Nationale Forschungsdateninfrastruktur (NFDI) e.V.
6th Members Assembly (Source: NFDI Assoc.)

The road ahead (Part II)

The German Council on Science and Humanities Council (WR)

(the oldest science policy advisory body in Europe)

conducted a structural evaluation of NFDI in 2024/2025.

Their recommendations:

- Institutionalize NFDI, beginning in 2029
- Strengthen and centralize the governance structure
- Transform NFDI into an AI enabler

**Nationale Forschungsdaten-
infrastruktur (NFDI) e.V.**

Albert-Nestler-Straße 13
76131 Karlsruhe

info@nfdi.de
<https://nfdi.de>

Ansprechpartner

Prof. Dr. York Sure-Vetter
york.sure-vetter@nfdi.de