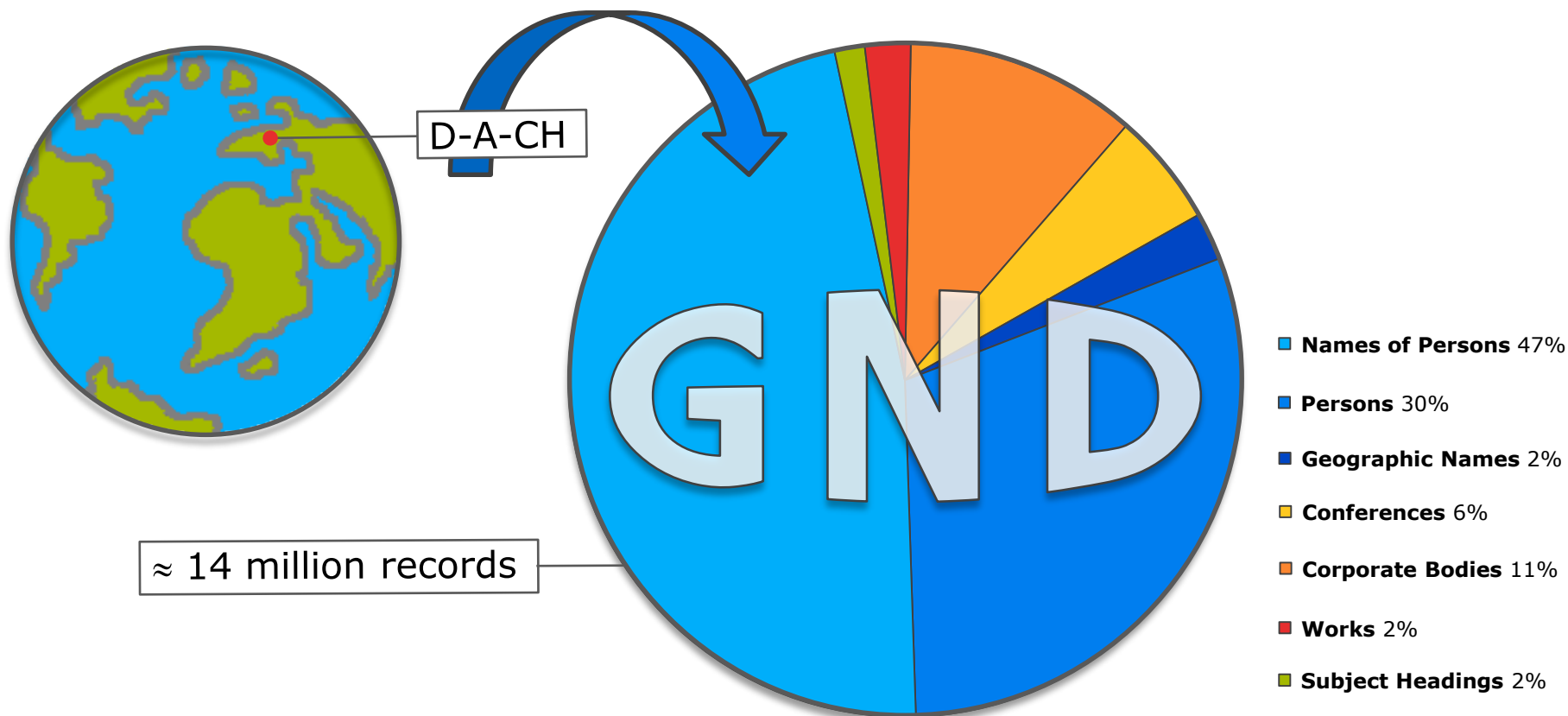


Sarah Hartmann, Jürgen Kett



# The GND initiative 2017-2021:

## Developing a backbone for the web of cultural and scientific data

# The Integrated Authority File = Gemeinsame Normdatei (GND)



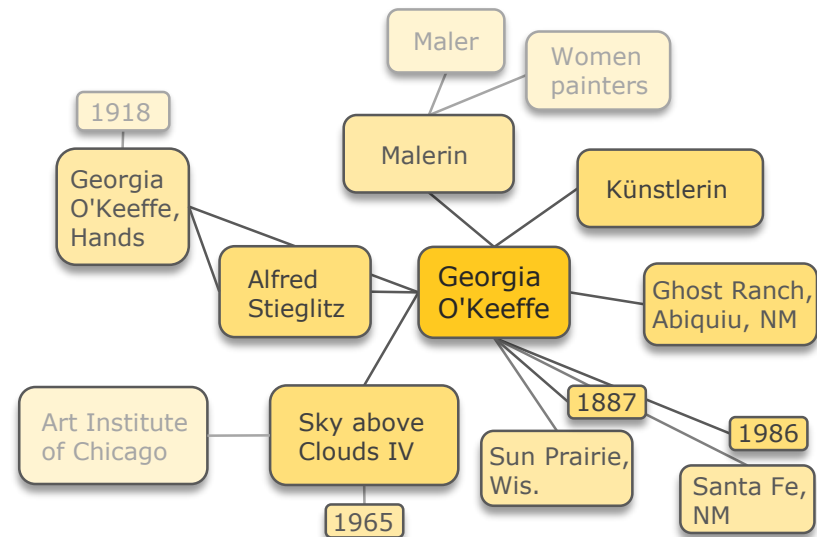
# Interfaces and formats

GND  	PICA	MARC 21 XML	RDF XML	Turtle	JSON	
<b>DNB OPAC</b>		•	•			passive
<b>Data Service FTP</b>		•	•	•		
<b>OAI-PMH</b>		•	•			
<b>SRU</b>		•	•			
<b>Linked Data Service</b>		•	•	•	•	
<b>Entity Facts</b>					•	active
<b>Cataloguing Client</b>	•					
<b>API SRU Record Update</b>		•				
<b>API Webcat (Persons)</b>	•					

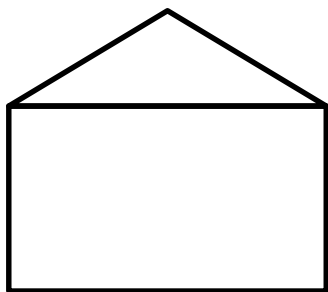
# GND characteristics and sustainability

- each record describes one entity (exception: names)
- Unique, persistent Identifier ( $\Rightarrow$  basis for URI)
- Entities have attributes and relationships to other entities
- Relations are designated by codes
- Modular data structure

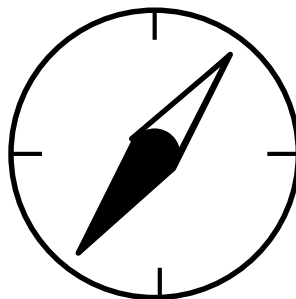
**$\rightarrow$  semantic network**



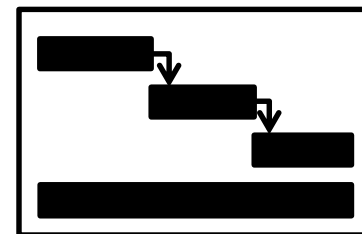
# The GND initiative 2017-2021



Organisation

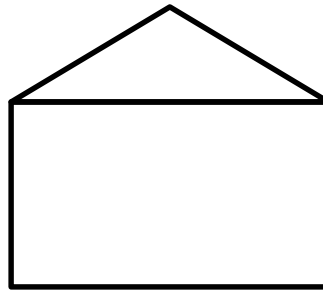


Guidelines

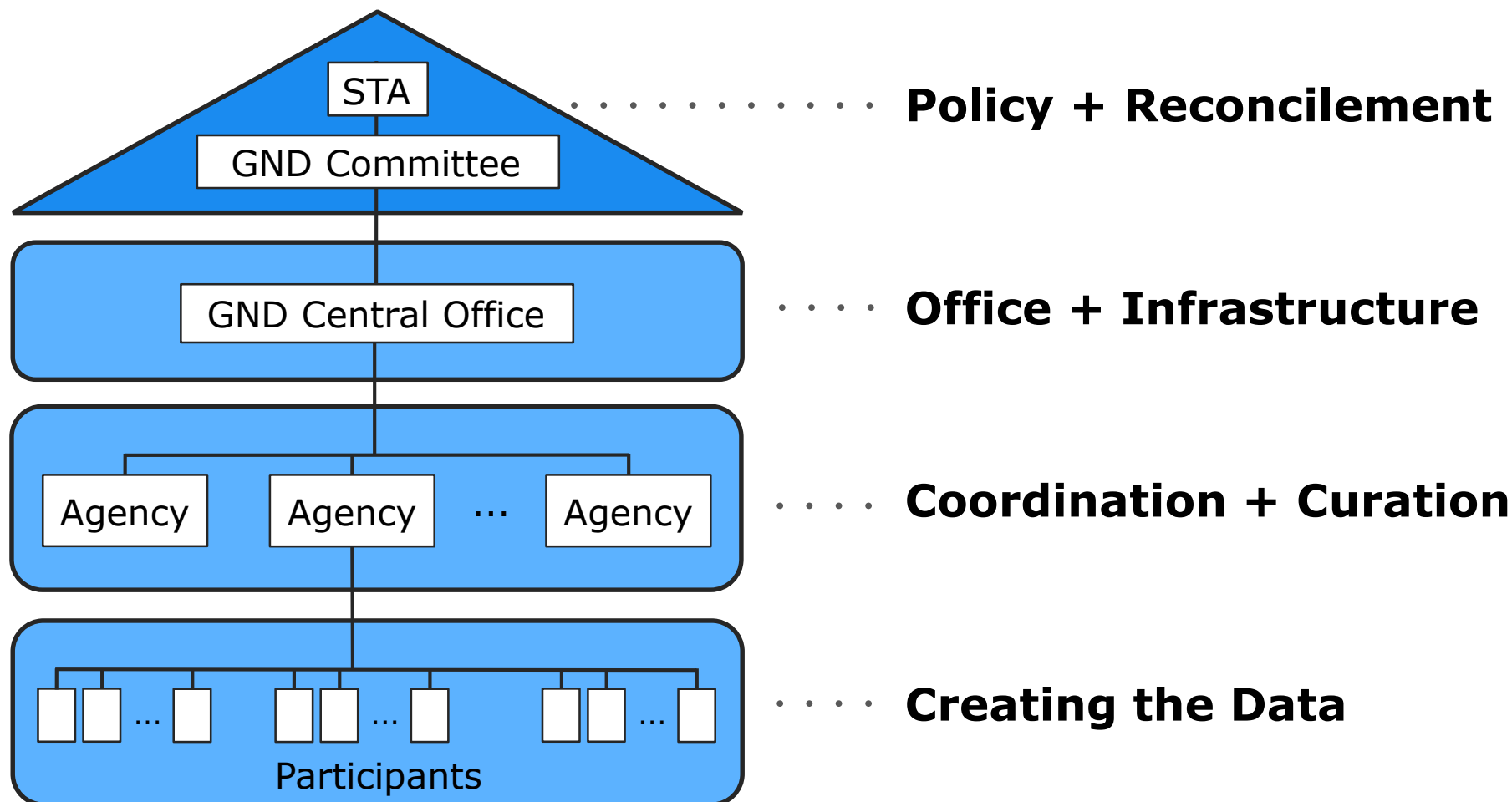


Work program

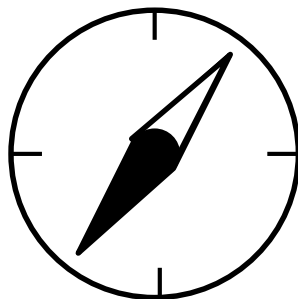
# Organisation



## GND cooperative – organizational structure



# Guidelines





# Guidelines

trusted  
quality

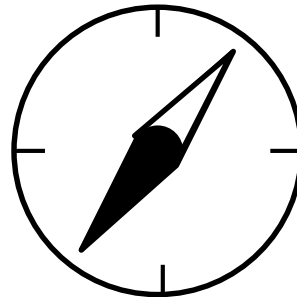
neutral

domain  
transcending

transparent

permanently  
accessible

stable



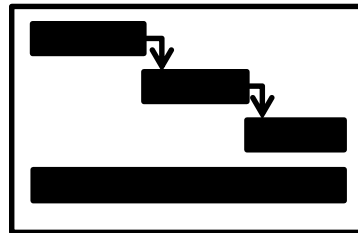
run  
cooperatively

obliging rules

open and free

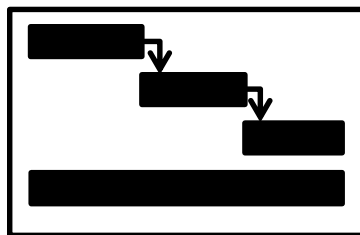
unambiguous

# Work program



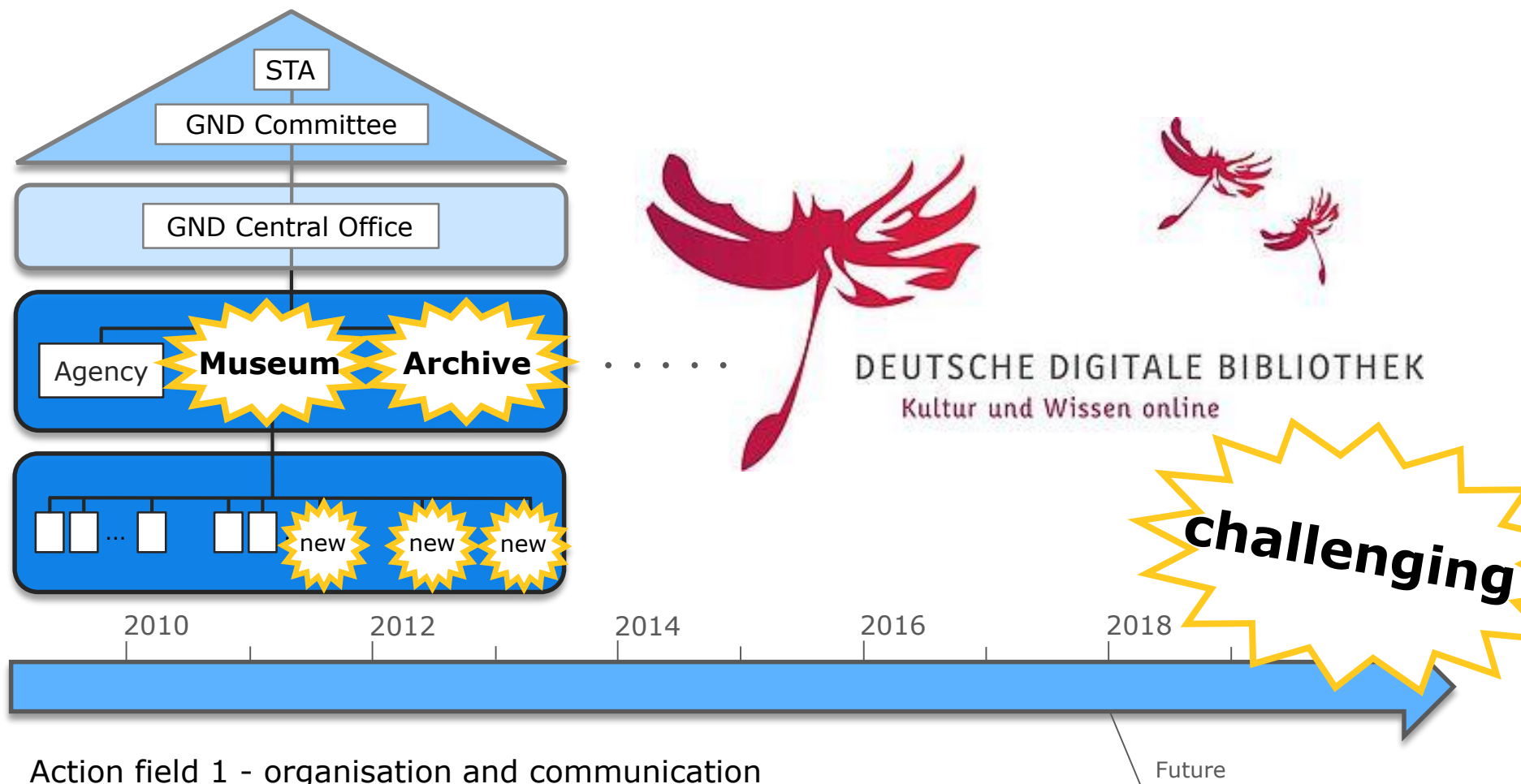
## Work program

- Action field 1 – organisation and communication
- Action field 2 – data management, maintenance, standardisation
- Action field 3 – import and datamining

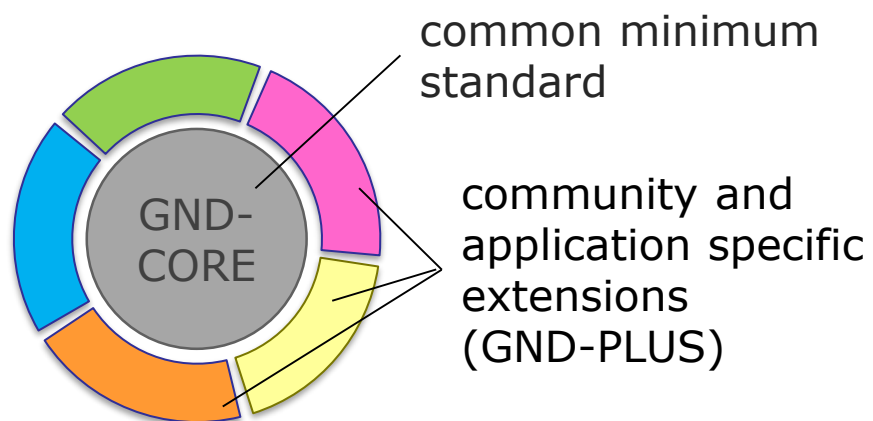


- Action field 4 – visualisation and end user applications
- Action field 5 – data supply and cataloging processes
- Action field 6 – collaboration with other communities

# Opening up to museums and archives

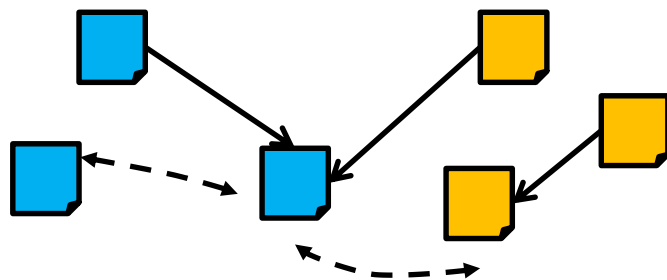


# Develop a domain transcending authority data system



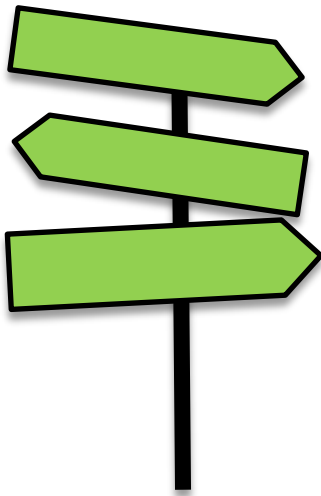
- modularize the data structure
- easy to use APIs and interfaces
- optimize data management
  - tracking of changes and provenance
  - suggestions for modification, add comments

# Analyze, linking and integrating data



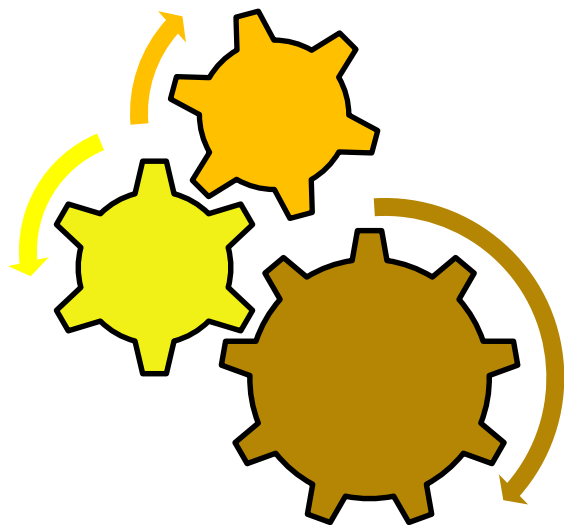
- „GND assistants“
- extend semantic links (internal and external)
- monitoring data quality
  - clearing up inconsistency, errors, doublets
- clustering (e. g. works)

# Improve access to GND network



- function as a „signpost“
- modules for visualization and navigation
- (test) implementation of semantic search system

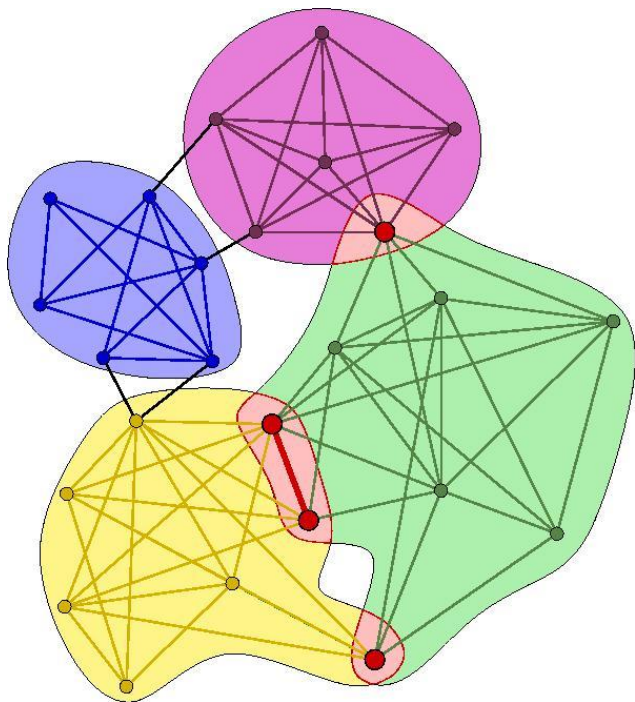
# Cooperative data supply



- modernize the infrastructure of data supply
- integration of GND in current systems
- interlocking of the cataloging and indexing workflows



# Expand user groups and applications



- deepen cooperations  
(Wikipedia/Wikidata, research, publishers ...)
- expand linking to other data sources or identifier systems
- provide applications to participate easily

## Next steps

Project ORCID DE: GND↔ORCID

Start project DDUP:  
quality control

Establishment of GND  
cooperative

2017

Project ARACHNE:  
infrastructure for linking data

Project GND4C:  
GND for cultural data

*In the pipeline: project  
GND for publishers*

# Thanks!

Questions?



[s.hartmann@dnb.de](mailto:s.hartmann@dnb.de)

[j.kett@dnb.de](mailto:j.kett@dnb.de)

<http://www.dnb.de/EN/gnd>

<http://d-nb.info/standards/elementset/gnd#>

# Back-up

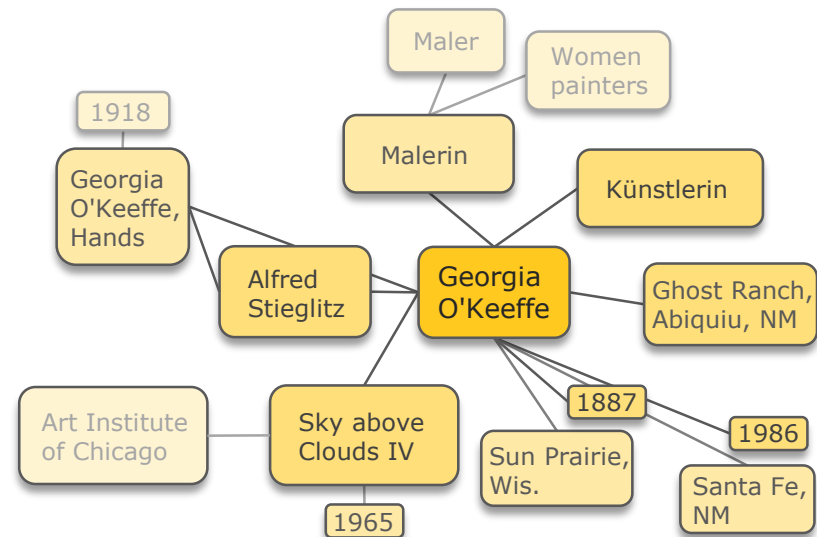
## **Challenges Develop a backbone for the web of cultural and scientific data -**

- Integration of new partners
  - other requirements of cultural and scientific data providers
- discussions about
  - rules, different kinds of „views“, quality levels
- responsibilities
  - who is allowed to create and update/correct which records or values
  - need for provenance data

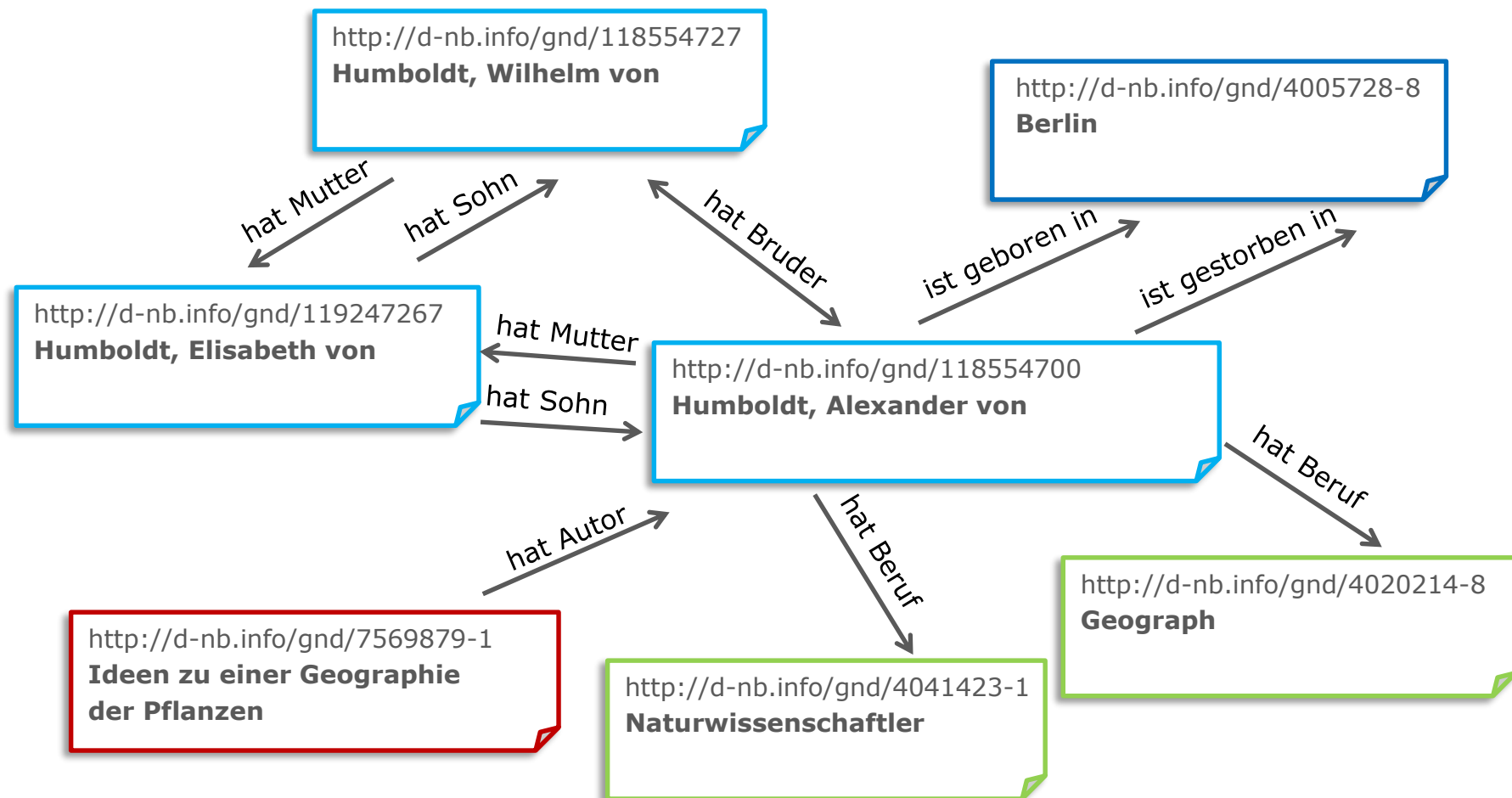
# GND characteristics and sustainability

- each record describes one entity (exception: names)
- Unique, persistent Identifier ( $\Rightarrow$  basis for URI)
- Entities have attributes and relationships to other entities
- Relations are designated by codes
- Modular data structure

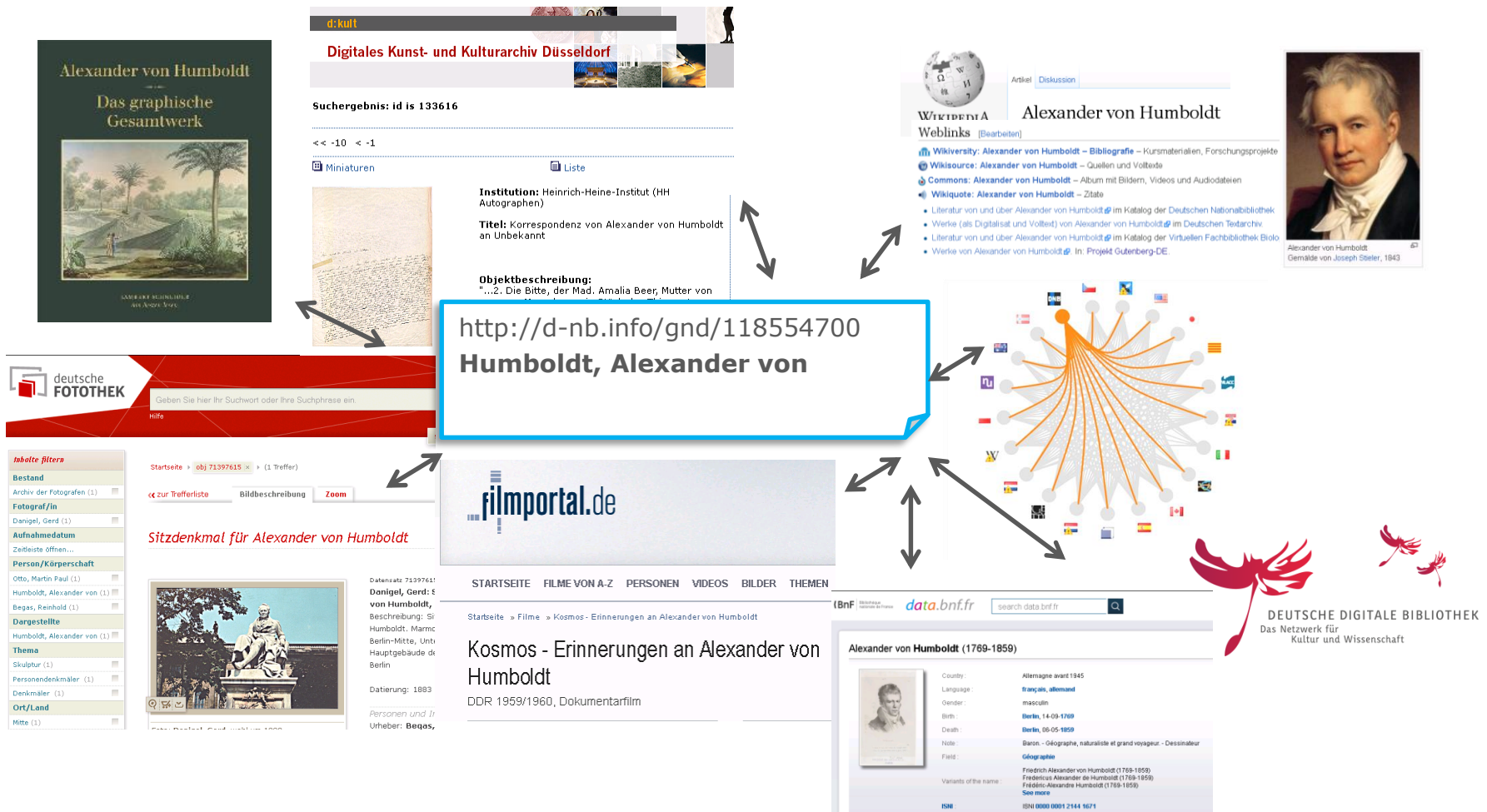
**$\rightarrow$  semantic network**



## Vernetzung innerhalb der GND



# Vernetzung zu externen Quellen

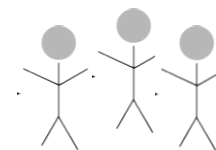




## Personen in der GND

- zwei Satzarten
  - Personennamen (Tn) und Personen (Tp)
- Personennamen
  - **nicht-individualisierte Datensätze**
  - verwendet für beliebig viele Personen mit diesem (bevorzugten) Namen
- Personen
  - **individualisierte Datensätze**
  - verwendet für genau eine Person
  - monatlicher Zuwachs ca. 20.000 Datensätze

<http://d-nb.info/gnd/100799892>  
**Müller, Heinrich**



<http://d-nb.info/gnd/Müller, Heinrich 1886->



## Voraussetzungen zur aktiven Mitarbeit

- Unterzeichnung der Kooperationsvereinbarung
  - (inkl. Leitlinien)
- [ISIL \(International Standard Identifier for Libraries and Related Organizations \(ISIL\)\)](#) oder [MARC Organization Code](#) zur eindeutigen Identifikation

# Actions

- Create new records
- Add additional information in existing records
  - z. B variant names, other attributes used for identifying the entity
- Correction of existing values
- Merging records / heading replaced by another (Umlenkung)
- in case of insufficient permission
  - „adding“ hints or requests to correct a record / values

# User groups in GND

- Different level mark provenance of the records
  - Level see MARC 079 \$c
- Level
  - 1 curation / quality team of library network
  - 2 local curation / quality team
  - 3 trained users
  - 4 untrained users
  - 5 other, non-librarian users
  - 6 legacy data, not edited by curation team
  - 7 automatically generated records
- Special responsibilities
  - For musical works (level 1)
  - For transcription or names in non-latin script

# Actions according to level

- Create new record
  - any user (with assigned level)
- add additional information
  - any user
- corrections of existing records (elements, values)
  - same or lower level (certain elements)
- Replace / delete / split
  - any user
  - but if lower level than 3: just a request for replace / delete / split























































































