

# Fachexpertise

## Strukturierten Daten und Graphen

# Agenda

- 1) Sammlungen, Objekte, Datenkompetenzen (SODa), (5 Min. Mark)
- 2) Speed Networking, (10 Min. alle)
- 3) Expertise und Spezialisierung, (30 Min. Canan)

\*\*\*Kurze Pause\*\*\*

- 4) Impulse Talks zur Kartierung von Objektdatenpools, (60 Min. Referent:innen)
- 5) FishBowl Diskussion zur Erhebung der Bedarfe, (20 Min. alle)
- 6) Zusammenfassung, offene Fragen und nächste Schritte, (5 Min. Canan)

# SODa Intro

Mark Fichtner



GERMANISCHES  
NATIONAL  
MUSEUM

# SODa – Datenkompetenzzentrum Sammlungen, Objekte, Datenkompetenzen

Centre for Collections, Objects, Data Skills

# Partner



I·IU Berlin, Koordinierungsstelle Sammlungen



Universität Erlangen-Nürnberg



IGSD e.V.



Germanisches Nationalmuseum



# Zielsetzung

Das “Zentrum Sammlungen, Objekte, Datenkompetenzen” ist ein Ort, an dem für die Arbeit in und mit wissenschaftliche Sammlungen notwendige und nützliche Datenkompetenzen vermittelt, kuratiert und weiterentwickelt werden.

Das Zentrum ermöglicht das Kennenlernen, die Diskussion und die Fortentwicklung von Standards, Good Practice und neuen Verfahren des sammlungsbezogenen Forschungsdatenmanagements und der datenbasierten Sammlungs- und Objektforschung.

Personal: 14 Positionen (2 für Programmentwicklung, 7 für Datenkompetenzen, wie z.B. 2D, 3D, Provenienz, ...),

Beginn: 15. November 23

# Fragen?

# Speed Networking



# Wir waren...



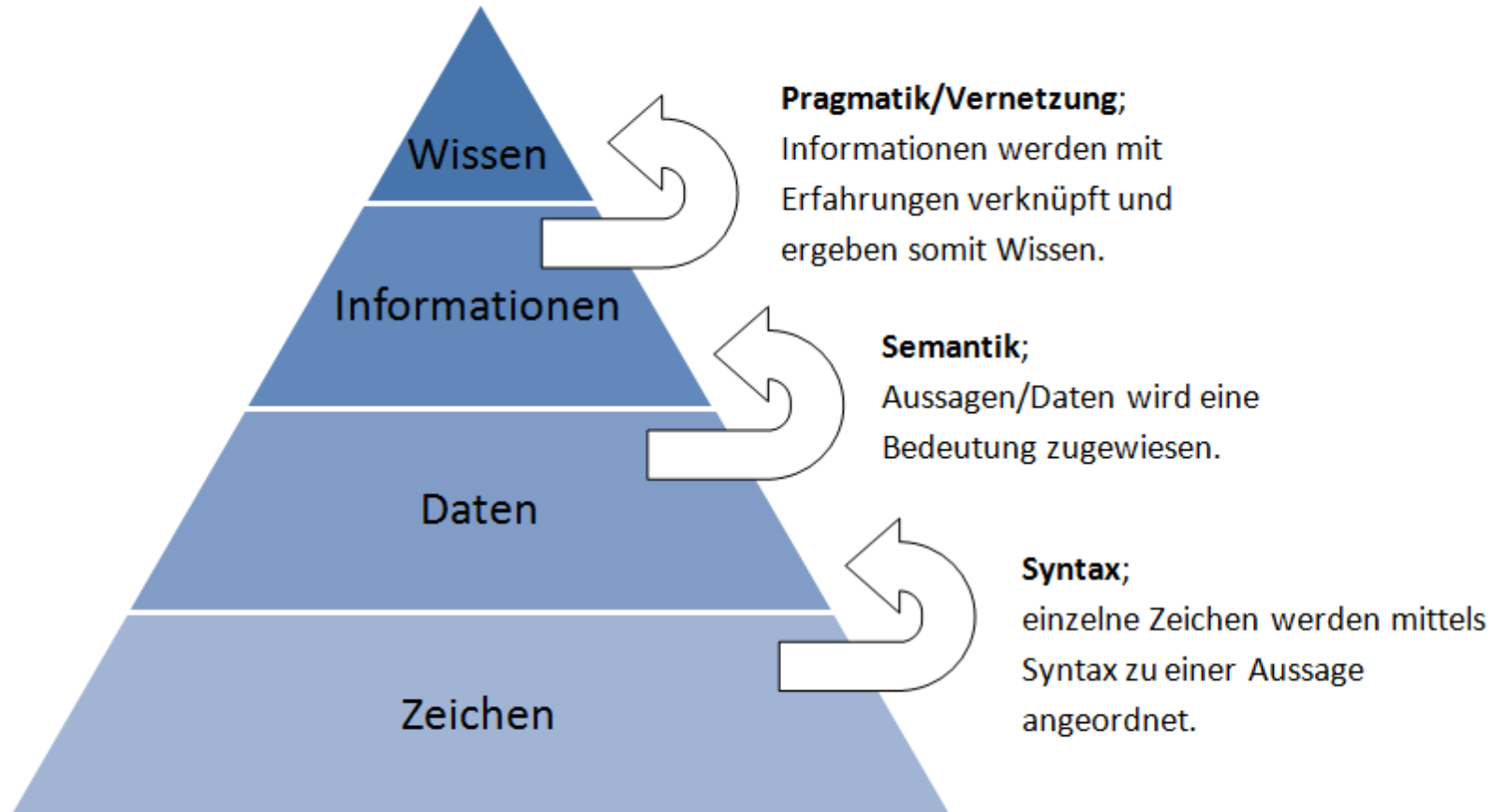
# Vorstellung der Fachexpertise

**Canan Hastik**

# Agenda

- 1) Daten, Information, Wissen
- 2) Strukturierte Daten
- 3) Graphdaten
- 4) Bedeutung & Relevanz
- 5) Anwendungsbereiche
- 6) WissKI-Tasting
- 7) Analyemethoden & Werkzeuge
- 8) Trends

# Daten, Information, Wissen



(c) 2012 Raffael Herrmann; [www.derwirtschaftsinformatiker.de](http://www.derwirtschaftsinformatiker.de)



# FAIR strukturierte Daten werden

- (1) ...mit Datenstandards dokumentiert.
- (2) ...in standardisierten Dateiformaten für unterschiedliche Dateitypen (Bild, Text, Video,...) organisiert.
- (3) ...in maschinenlesbaren Datenformaten gespeichert und geteilt.
- (4) ...in geeigneten Tools und mit Kontext korrekt interpretierbar.

# Lernziele „Arbeiten mit Daten“ (Auswahl)

## Lernende können...

- (1) Mit Datenstandards dokumentieren.
  - die Bedeutung von Datenstandards bei der Datenintegration benennen.
  - Techniken zur FAIRifizierung von Daten benennen und anwenden.
- (2) In standardisierten Dateiformaten für unterschiedliche Dateitypen (Bild, Text, Video,...) organisieren.
  - Dateiformate in verschiedenen Kontexten anwenden.
  - Werkzeuge zur Dateiformaterkennung und -validierung benennen und anwenden.
- (3) In maschinenlesbaren Datenformaten speichern und teilen.
  - Formate für Archivierung und Nachnutzung benennen und bewerten.
  - Daten in nachnutzbaren Formaten anlegen oder in entsprechende Formate überführen und konvertieren.
- (4) Mit geeigneten Tools und Kontext korrekt interpretieren.
  - Forschungsdatenmanagement-Werkzeuge, -Software und -Infrastrukturen benennen, anwenden und analysieren.
  - ...

# Graphbasierte Daten werden

- (1) ...mit Datenstandards dokumentiert.
- (2) ...in standardisierten Datenbankformaten organisiert, die die Struktur und Beziehungen der Datenpunkte widerspiegeln.
- (3) ...in maschinenlesbaren Datenformaten gespeichert und geteilt.
- (4) ...in geeigneten Tools und semantisch korrekt interpretierbar.

# Lernziele „Graphdaten“ (Auswahl)

## Lernende können...

- (1) Mit Datenstandards dokumentieren.
  - Semantische Datenmodelle konzipieren und modellieren.
  
- (2) In standardisierten Datenbankformaten organisieren, die die Struktur und Beziehungen der Datenpunkte widerspiegeln.
  - Generieren von flexiblen Datenmodellen
  
- (3) In maschinenlesbaren Datenformaten speichern und teilen.
  - anwenden von Graphen zur Integration verteilter und heterogener Datenpools.
  
- (4) In geeigneten Tools semantisch korrekt interpretieren.
  - beherrschen von schnellen und effizienten Abfragetechniken.



# Bedeutung & Relevanz

- Strukturierte Daten sind leicht verarbeitbar und analysierbar
- Graphen stellen Verbindungen/Beziehungen zwischen Objekten dar und
  - machen Beziehungen und Muster in Daten sichtbar
  - ermöglichen tiefergehende Einsicht
  - und lösen unterstützen bei komplexen Fragestellungen.

# Anwendungsbereiche

- Katalogisierung und Inventarisierung von Sammlungsobjekten
- Digitalisierung, Online-Zugänglichkeit und Publikation von Sammlungen
- Visualisierung und Analyse von Daten
- Automatisierung und Datenintegration
- Forschung und interdisziplinäre Studien
- Wissensgraphen
- Empfehlungssysteme
- Netzwerkanalysen

# Lernende können...

## Analysemethoden & Tools benennen, anwenden und bewerten.

- 🌐 Einsatz von Graphdatenbanken zur Untersuchung und Visualisierung komplexer Beziehungen
  - Modellieren und Visualisieren von komplexen Beziehungen
  - Erkundung von Daten
- 🌐 Analyse strukturierter Daten in Sammlungsdatenbanken
  - SQL Abfragen und Analyse relationaler Datenbanken
  - XML-basierte Strukturierung und Auswertung
- 🌐 Graphenanalyse mit spezialisierten Algorithmen
  - Verständnis für Algorithmen zur Clusterbildung
  - Effiziente Analyse von Netzwerkstrukturen in Sammlungen

# Trends



Pollin und Steiner (2024), <https://tinurl.com/dhcraft-vsd-krems> (CC BY)

# WissKi-Tasting



<https://objekte-im-netz.fau.de/projekt/>



<http://wisski-stak01.virt.uni-oldenburg.de/>



<https://tafelmalerei.gnm.de/>



<https://boehler.zikg.eu/>



<https://fuchsmagen.wisski.uibk.ac.at/>



<https://ais-dka.gnm.de/>



<https://konchylien.leopoldina.org/>



[https://editionhansposse.gnm.de/das\\_projekt](https://editionhansposse.gnm.de/das_projekt)



# Kartierung von Datensammlungen und Objektdatenpools

Praxisbeispiele von  
Hanna-Lena Meiners  
Olaf Simons  
Judith Blume  
Fritz Dross



Trägerinstitution/-initiative mit Ansprechpersonen, Verantwortlichen und Besitzverhältnissen

## Ziel ist es, ....

eine zentrale Infrastruktur an der GU aufzubauen, die die Sammlungen zur digitalen Erschließung nutzen können und so die Insellösungen loszuwerden

## Nutzung in der Forschung (Bitte ankreuzen X)

- Vergleichsforschung mit Bestimmungen
- Forschungsergebnis / Produkt einer Forschung
- Datengestützte sammlungsübergreifende Forschung (z.B. Mustererkennung)
- Kontextualisierung, Herstellen von Bezügen und Verbindungen
- Wissenschaftsgeschichte
- Typus- und Referenzsammlung
- Sonstiges\_Provenienzforschung\_\_\_\_\_
- Keine Forschungsnutzung

## Nutzung in der Lehre (Bitte ankreuzen X)

- Studiengang und/oder Studienmodul
- Lehrprojekt
- Haptische Erfahrung als Lehrmittel, Anschauungsmaterial
- Didaktisches Mittel
- Virtuelle Nutzung
- (Digitaler) Methodenerwerb, z.B. Objektbiographie, Objektdokumentation, Datenbasierte Erschließung, ...
- Erhaltung und Restaurierung
- Interdisziplinäres Lernen

## Weitere Nutzungsformen



## Diese Themen beschäftigen uns:



## Diese Personen möchten wir erreichen:



## Das erwarten wir von SODa:



## Welchen Handlungsbedarf gibt es?

Wie erreichen wir die relevante Zielgruppe mit passgenauen Angeboten?



Wie können Angebote leicht auffindbar und nutzbar gemacht werden?



Wie erreichen wir, dass existierende Angebote häufiger nachgenutzt werden?




Was fehlt, was ist wünschenswert, was läuft gut?

# Beispiel Steckbrief

# FishBowl Diskussion

## Bedarfe und Botschaften

Ein Archiv möchte eine umfassende Wissensbasis über ihre Sammlungen aufbauen und Forschern eine effiziente Recherche ermöglichen.

**Beispiel:** Forschende sollten in der Lage sein Graphdatenbanken und spezifischen Dateiformaten (wie RDF für Linked Data) zur Darstellung und Analyse von Beziehungen zwischen Sammlungsobjekten oder Forschungsergebnissen zu nutzen. Sie sollten in der Lage sein, einen Wissensgraphen, der die Verbindungen zwischen historischen Dokumenten, Autoren, Publikationsjahren, Themen und geografischen Standorten abbildet aufzubauen. Durch die semantische Modellierung und Visualisierung dieser Beziehungen können Forschende schnell relevante Informationen über bestimmte Themen oder historische Perioden finden und analysieren.  





# FishBowl Diskussion

## Bedarfe und Botschaften

Eine Sammlungsleitende möchte die Beziehungen zwischen Sammlungsobjekten und deren Einfluss auf die kulturelle Entwicklung einer Region verstehen.

**Beispiel:** Sammlungsleitende sollte in der Lage sein Verfahren zur automatischen Datenintegration mittels geeigneter Schnittstellen und Konvertierungstools, um Daten aus unterschiedlichen Dateiformaten (z.B. XML, CSV, JSON) zusammenzuführen und zu analysieren anzuwenden.



Ein Museum möchte seinen Besuchern personalisierte Empfehlungen für Ausstellungen und Sammlungsstücke geben, basierend auf deren Interessen und Besuchshistorie.

**Beispiel:** Nutzung eines Wissensgraphen, der die Beziehungen zwischen verschiedenen Sammlungskategorien, historischen Epochen, Künstlern und Kunstwerken modelliert. Durch die Analyse des Besucherverhaltens und der Vorlieben können Empfehlungen generiert werden, die auf ähnlichen Interessen oder bereits besuchten Ausstellungen basieren.



# FishBowl Diskussion

## Bedarfe und Botschaften

Ein Forschungsprojekt, ist auf eine umfassende Analyse verschiedener Datentypen angewiesen.

**Beispiel:** Forschende sollten geeignete Formate (z.B. CSV für tabellarische Daten, JSON-LD für strukturierte Daten) zur Nachnutzung und Analyse von Forschungsdaten nutzen und in der Lage sein die Auswahl von Verfahrensweisen für die Verarbeitung und Analyse sowohl quantitativer als auch qualitativer Daten vorzunehmen.



# FishBowl Diskussion

## Bedarfe und Botschaften

Ein Museum möchte seine Sammlungsobjekte digital erfassen und verwalten.

**Beispiel:** Strukturierter Datenformate (wie XML oder JSON) werden zur Speicherung von Objektinformationen eingesetzt, um eine systematische Katalogisierung zu ermöglichen. Sammlungsbetreuende sollten Werkzeuge zur Dateiformaterkennung und -validierung verwenden können, um die Integrität und Langzeitarchivierung dieser Daten sicherzustellen.



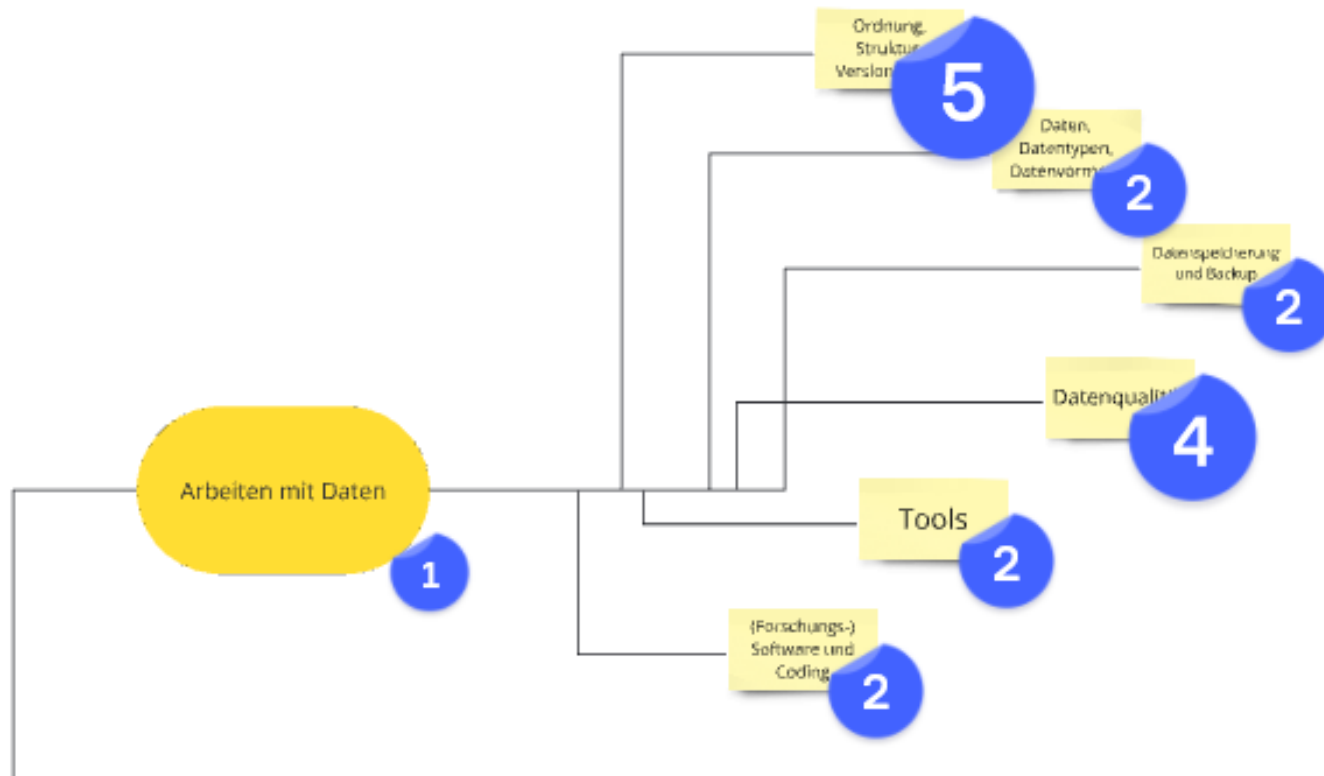
Ein Archiv strebt danach, historische Dokumente zu digitalisieren und online zugänglich zu machen.

**Beispiel:** Kuratorinnen, sollten geeigneter Dateiformate (wie PDF/A für Texte oder TIFF für Bilder) für die Archivierung und Langzeitnutzung kennen und erzeugen können. Sie sollten auch eigenständig Werkzeuge zur Dateiformaterkennung nutzen können, um die Qualität der digitalisierten Dokumente zu überprüfen und sicherzustellen.



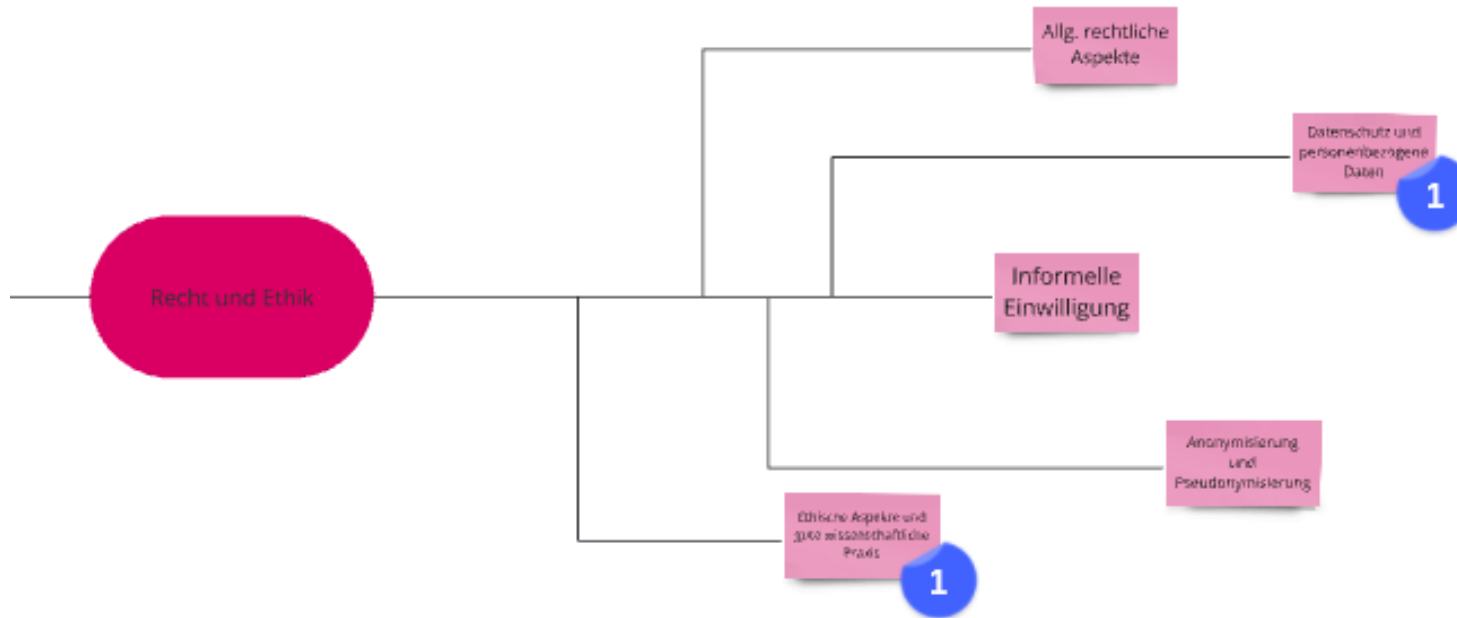
# FishBowl Diskussion

Welche FDM-Bereiche und -Themen sind von besonderem Interesse?



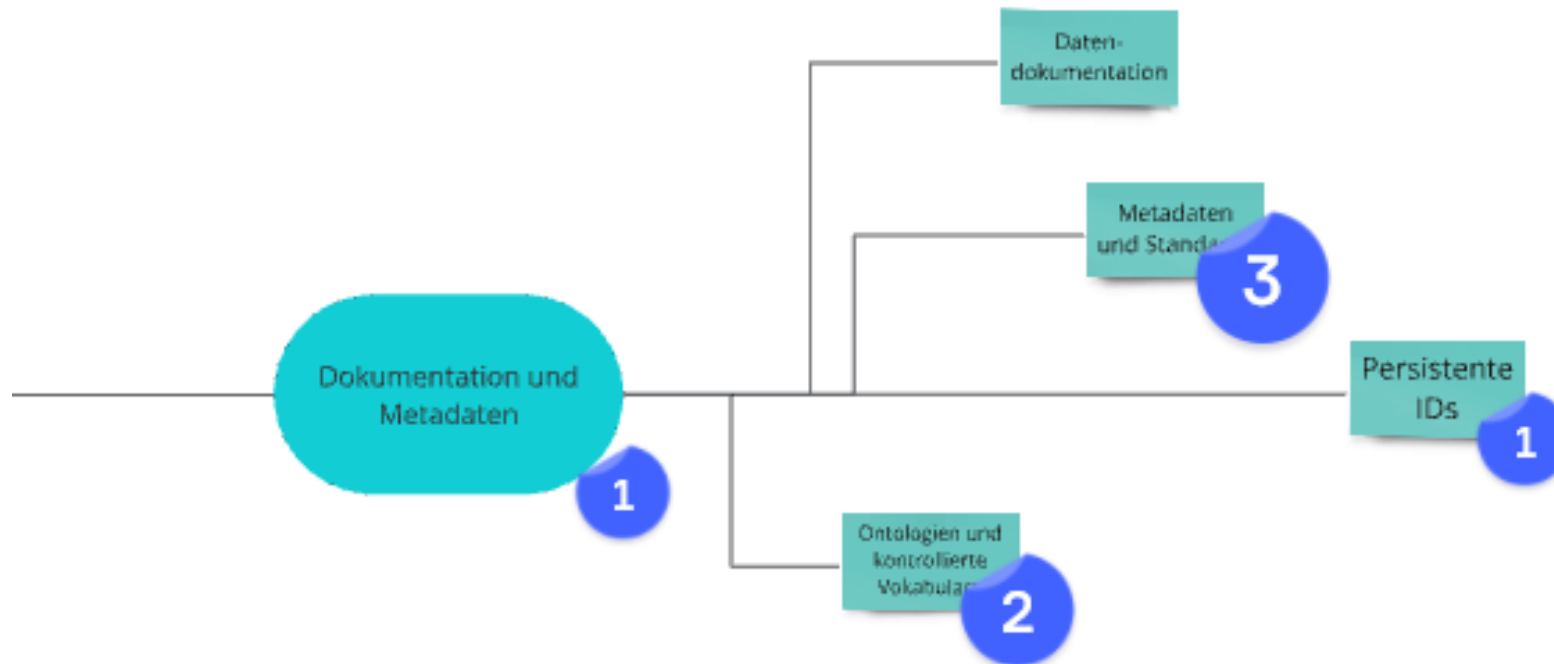
# FishBowl Diskussion

Welche FDM-Bereiche und -Themen sind von besonderem Interesse?



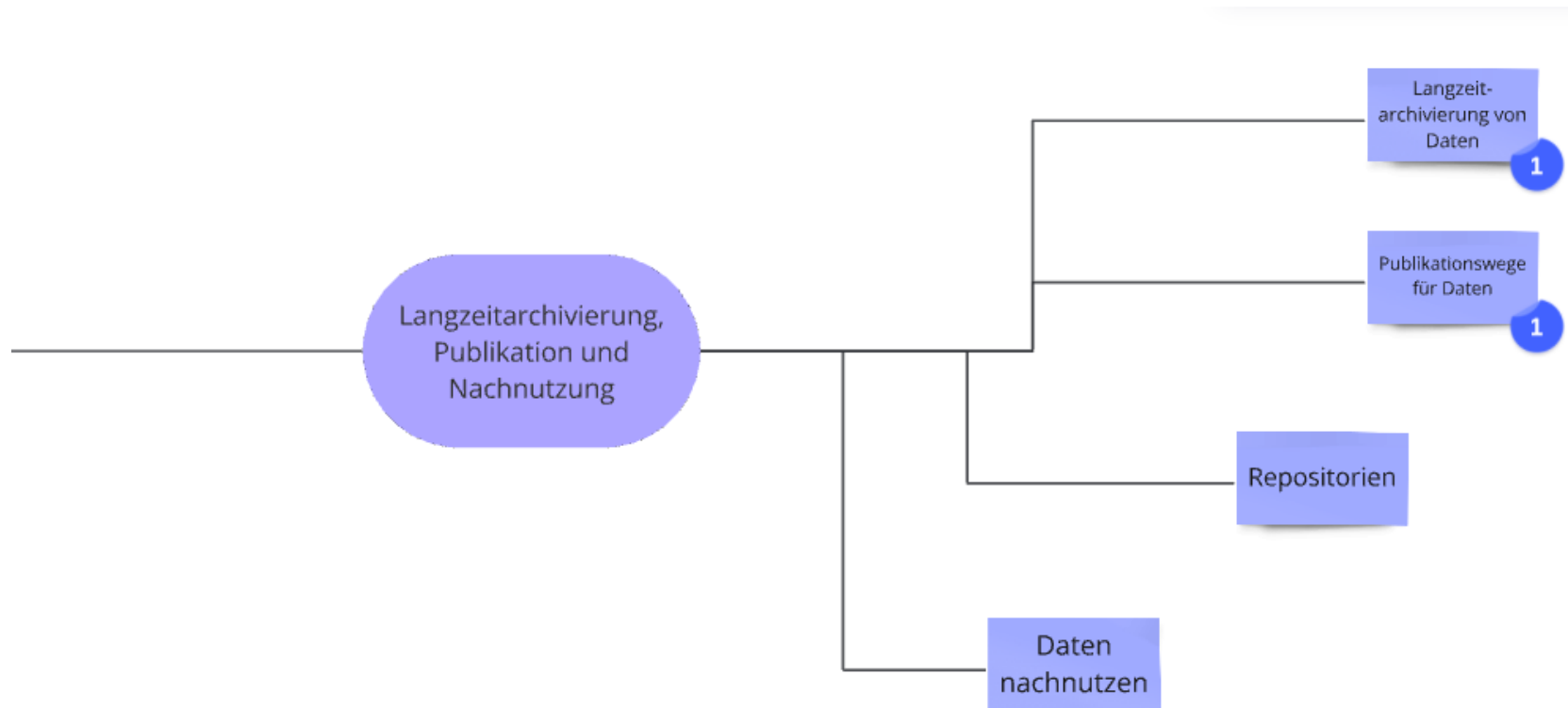
# FishBowl Diskussion

Welche FDM-Bereiche und -Themen sind von besonderem Interesse?



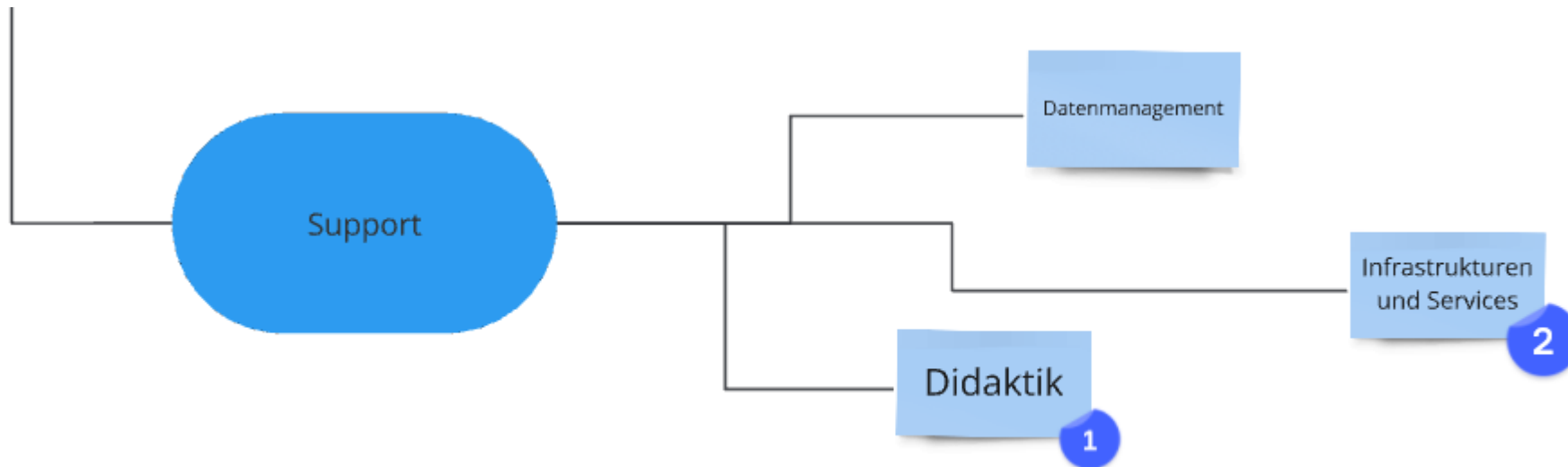
# FishBowl Diskussion

Welche FDM-Bereiche und -Themen sind von besonderem Interesse?

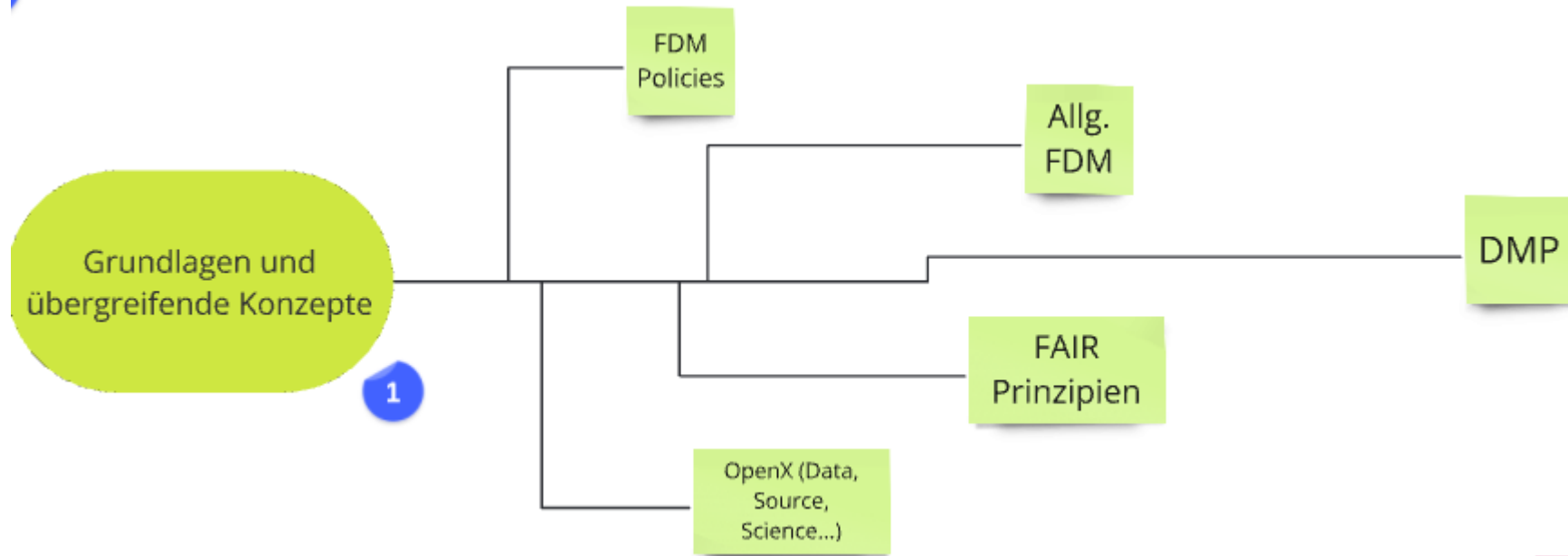


# FishBowl Diskussion

Welche FDM-Bereiche und -Themen sind von besonderem Interesse?







Canan Hastik | Sammlungen, Objekte, Datenkompetenzen (SODa)  
Fachexpertise Strukturierte Daten und Graphen  
Web: <https://sammlungen.io/de>

IGSD e.V.  
Interessengemeinschaft für Semantische Datenverarbeitung e.V.  
Irisstr. 23a  
80935 München  
E-Mail: [c.hastik@igsd-ev.de](mailto:c.hastik@igsd-ev.de)  
Web: <https://igsd-ev.de>

# Vielen Dank für die Aufmerksamkeit!