CEIR | Center for Enterprise Information Research

Using Materialised Ontology-Based Data Access (MOBDA) for the Harmonisation of Trace Data from Enterprise Collaboration Systems

Lucas Schlömer, **Martin Just** and Petra Schubert

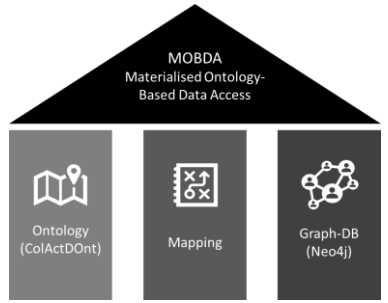
university of koblenz

AGENDA

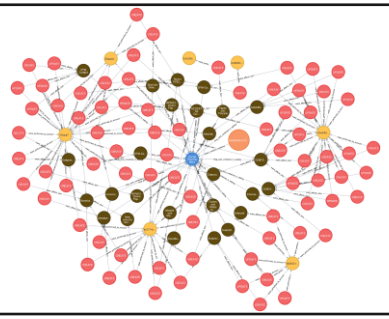
The path from chaos to structure

Introduction

MOBDA



Replication Study



Q&A

CEIR
Center for Enterprise Information Research

Susan Williams
Professor
CEIR Research Director

Enterprise Information Management

Petra Schubert
Professor
CEIR Research Director

Business Application Systems

Julian Mosen
CEIR Researcher
Head of IT

Social documents

Sebastian Bahles
CEIR Researcher
Manager UCT & IndustryConnect

Workspace Management

Simon Meier
CEIR Researcher

Ontology-based data access

Jens Alberts
CEIR Researcher

Work routines & coordinative practices

Jennifer Gerbl
CEIR Researcher

Hybrid Work and
Digital Ethnography

Martin Just
CEIR Researcher

Social Process Mining and
Cross-system analysis

Lucas Schlömer
CEIR Researcher

Content rich data spaces

Cornelia Mc Stay
Administrative Assistant

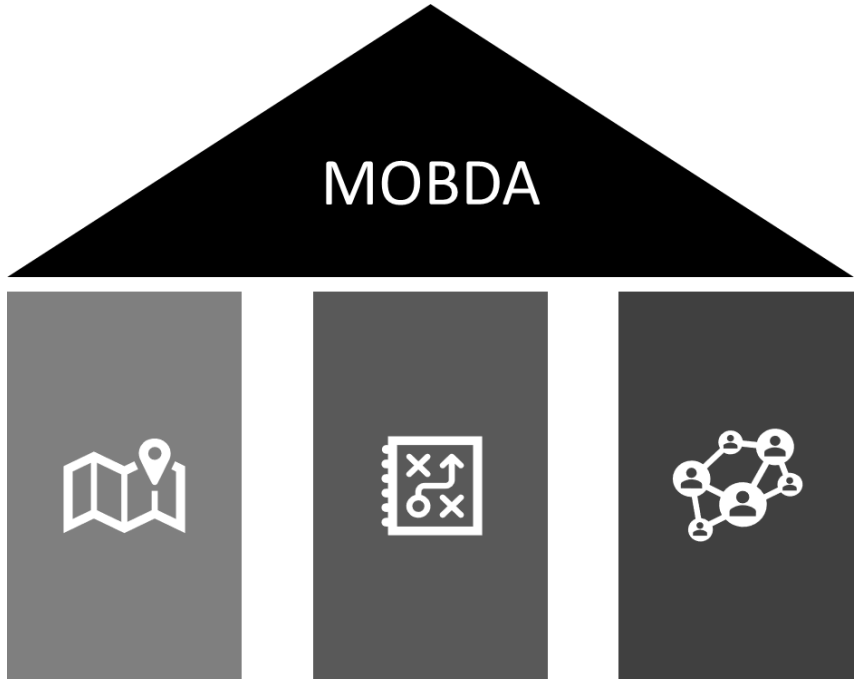
INTRODUCTION

Collaboration spans system boundaries

INTRODUCTION

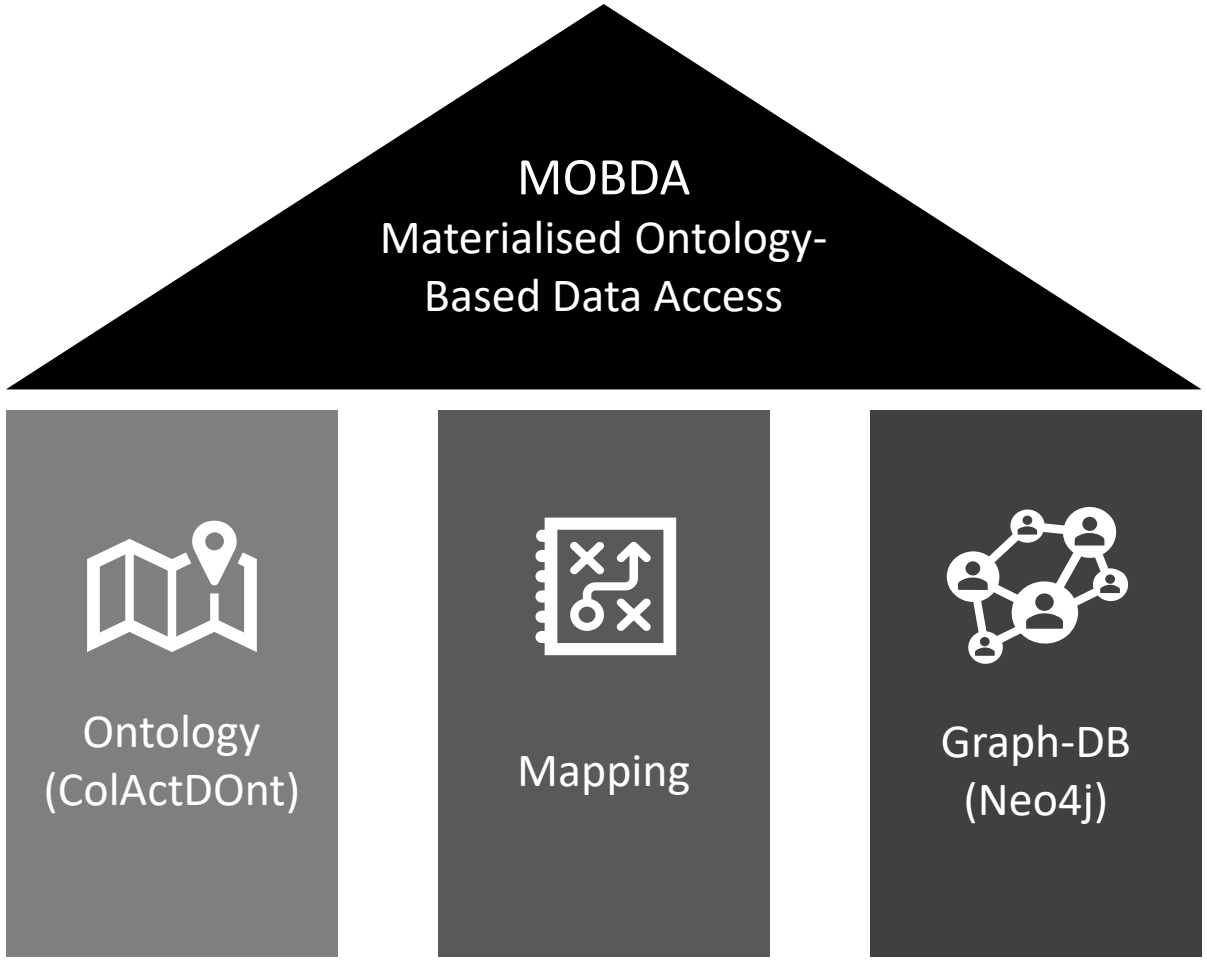
Objective and approach of MOBDA

- **Research question:** How can we collect, transform and **harmonise** the **digital traces** from **multiple heterogeneous collaboration systems** in a **joint datastorage** and create an **analysis method** for the continuous and longitudinal **analysis of collaborative work processes** that translates the queries of the analyst into the language (vocabulary) of the heterogeneous backend systems?
- **Research Methodology**:
  - Design Science Research (Vaishnavi et al., 2017)
  - Evaluation: Replication Studies (Dennis & Valacich, 2014)
- **Data Set**:
  - HCL Connections instance
    - ~6500 registered accounts
    - **~6 million event records**



MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

Building blocks



- OBDA approach for graph databases (Cysneiros & Salgado, 2016):
  - Derive data **structure** from ontology
  - Create **mapping** (source system → ontology)
  - Extract, prepare and **import** data
  - Data **access** via ontology terminology

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

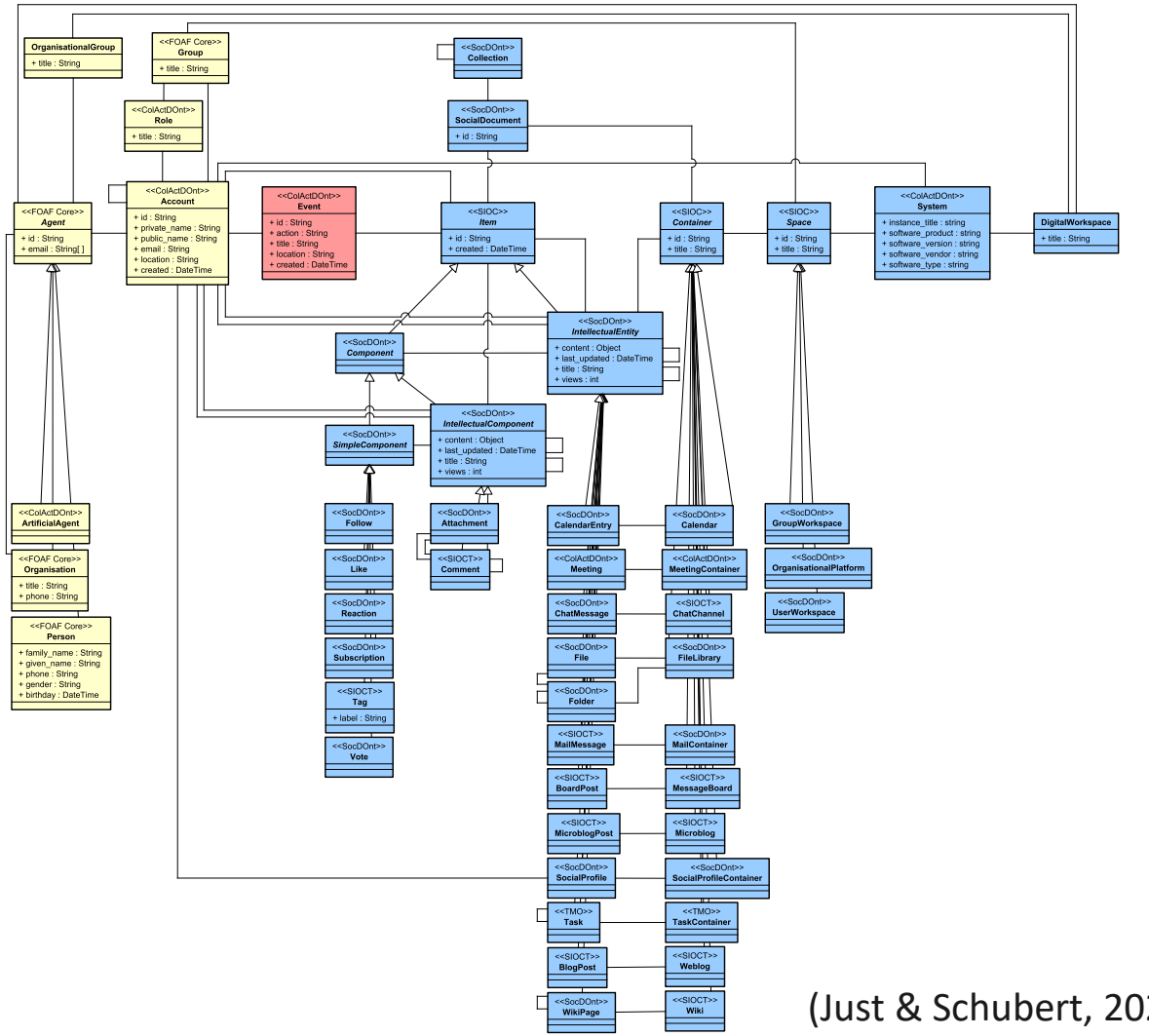
Ontology: The key to data harmony

- Formal and shared **specification of a knowledge model** for the standardised description of a knowledge domain (Gruber, 1995)

→ **Uniform terminology** and **data structure**

- Collaborative Actions on Documents Ontology (**ColActDOnt**) defines concepts for events in ECS (Just & Schubert, 2023)
- Describes content, transactional and organisational **Data** (Schwade & Schubert, 2018)

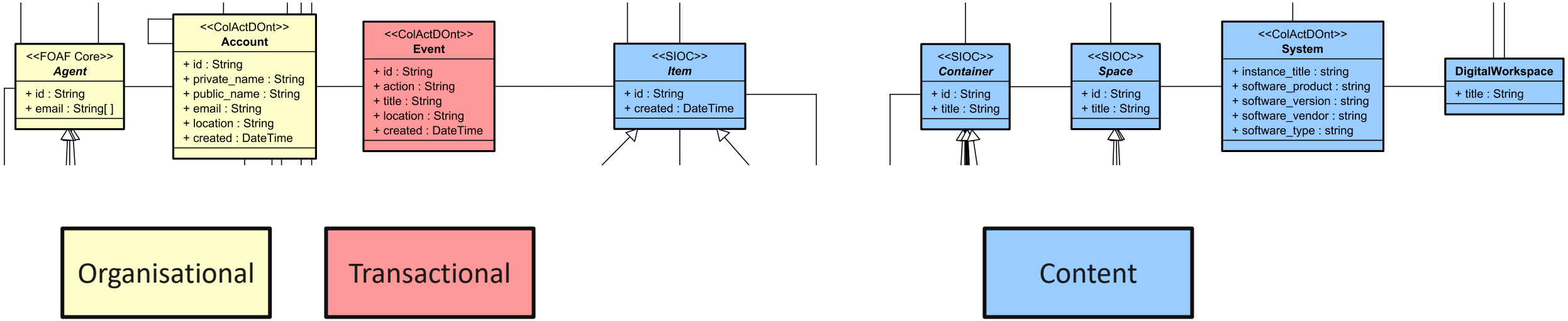
→ Enables **cross-system analyses**



(Just & Schubert, 2023)

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

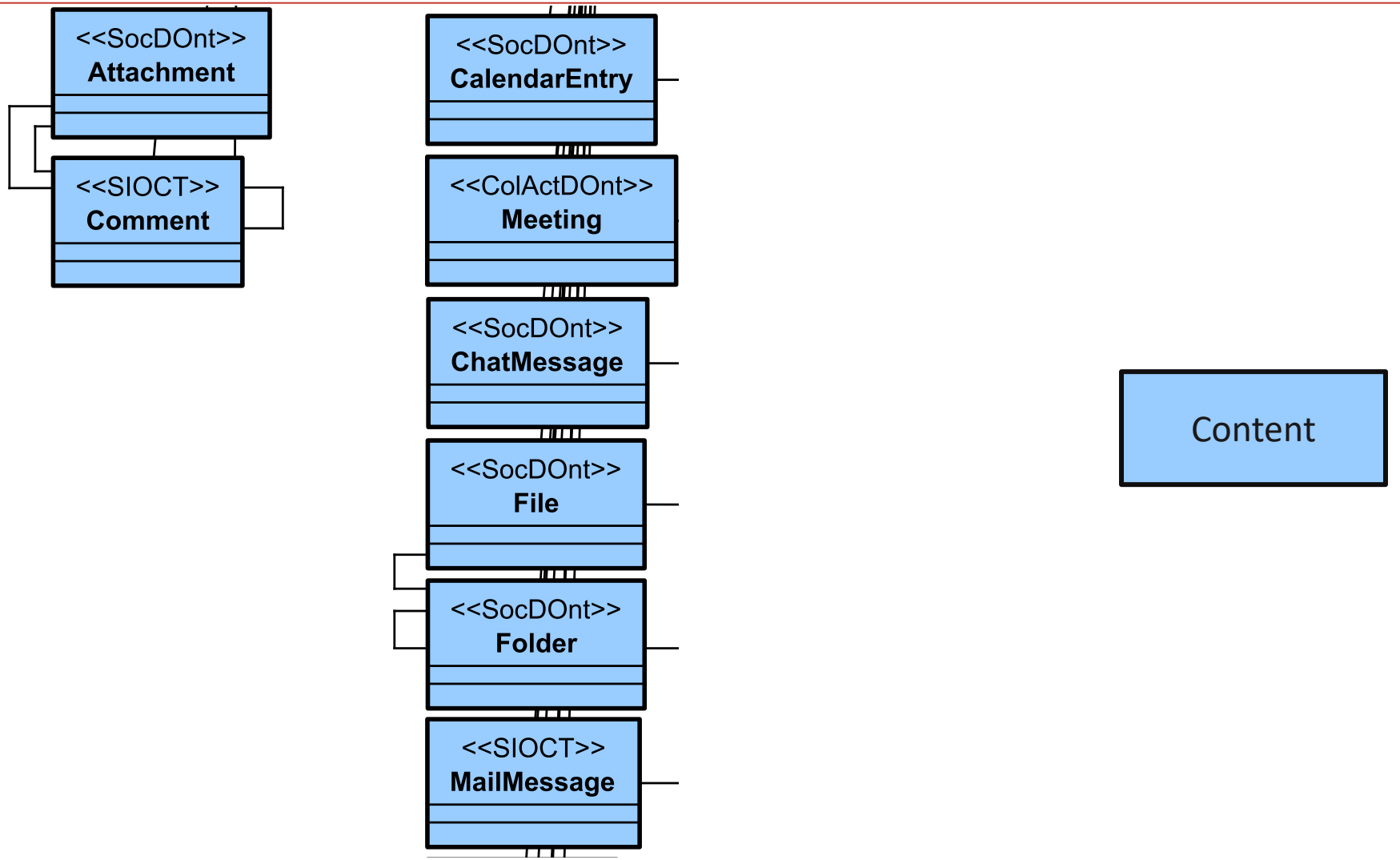
Ontology: Central concepts



(Just & Schubert, 2023)

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

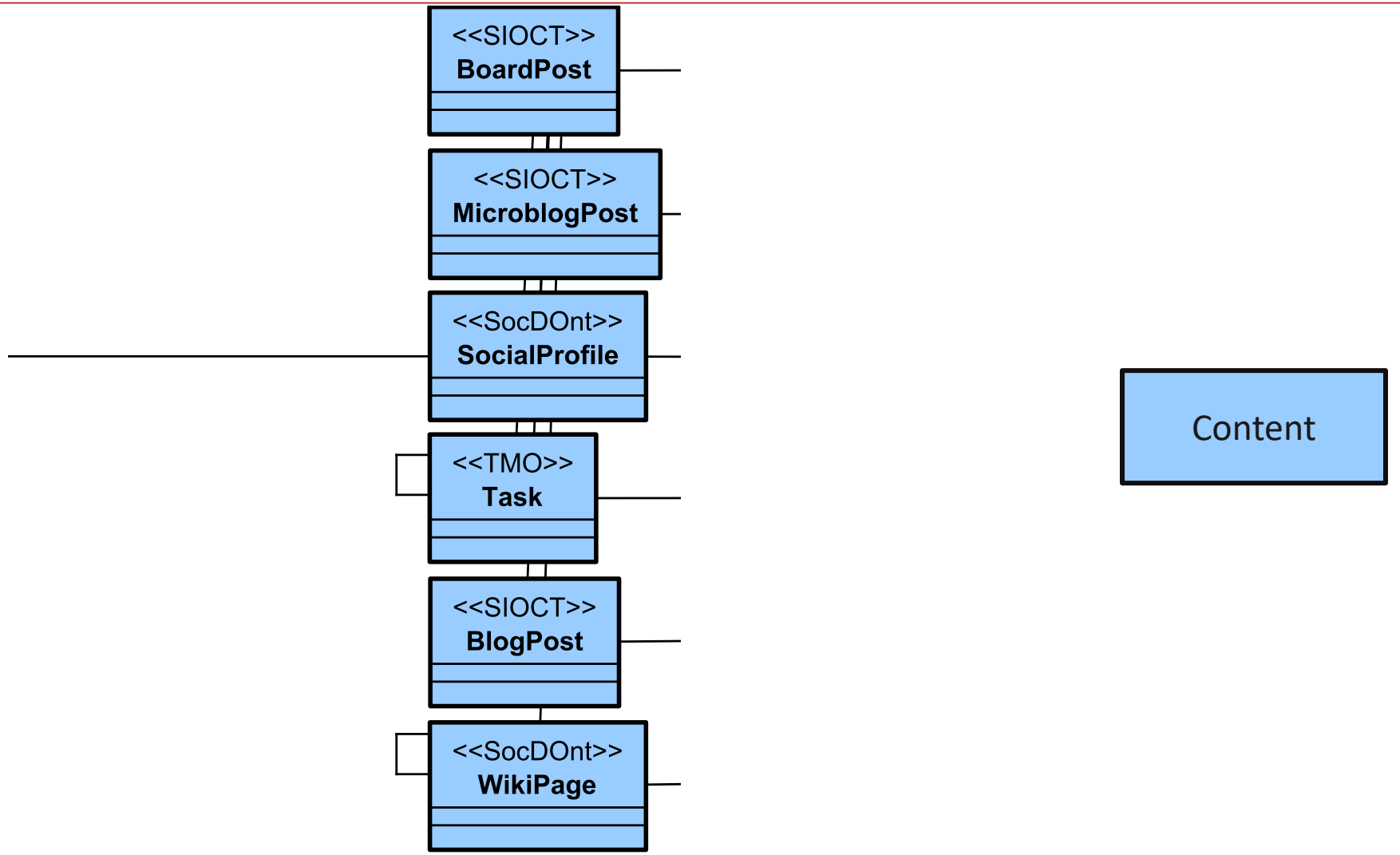
Ontology: Content concepts



(Just & Schubert, 2023)

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

Ontology: Content concepts



(Just & Schubert, 2023)

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

Ontology: Formal specification - w3id.org/ColActDOnt

CEIR Center for Enterprise Information Research

university of koblenz
continue discovering

ColActDOnt - Collaborative Actions on Documents Ontology

IRI / Latest Version:
https://w3id.org/ColActDOnt [JSON] [OWL/XML] [TTL]

Version IRI / This Version:
https://w3id.org/ColActDOnt/1.0 [JSON] [OWL/XML] [TTL]

Date:
06/2022

Current version:
1.0

Authors:
Petra Schubert, Martin Just

Publisher:
CEIR - Center for Enterprise Information Research

Visualisations:
PNG - SVG - WebVOWL (CEIR)

XES Extension:
Documentation - xesext

Publications:
Just, M., & Schubert, P. (2023). Collaborative Actions on Documents Ontology (ColActDOnt). Procedia Computer Science, 219, 294–302. https://doi.org/10.1016/j.procs.2023.01.293 [PDF]
Just, M., Schubert, P., Blatt, J., & Delfmann, P. (2024). Data Preprocessing for Cross-System Analysis: The DaProXSA Approach. Procedia Computer Science, 239, 1635–1644. https://doi.org/10.1016/j.procs.2024.06.340 [PDF]

License:
Attribution 4.0 International (CC BY 4.0)

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

Mapping: Introduction

- **Data models** of the systems are **mapped to** the **concepts** of the ontology
- **Harmonise heterogeneous data** sources **into** a **data model** described by the ontology

Mapping process for MOBDA

Ontology entity selection

Identification and **extraction** of system data

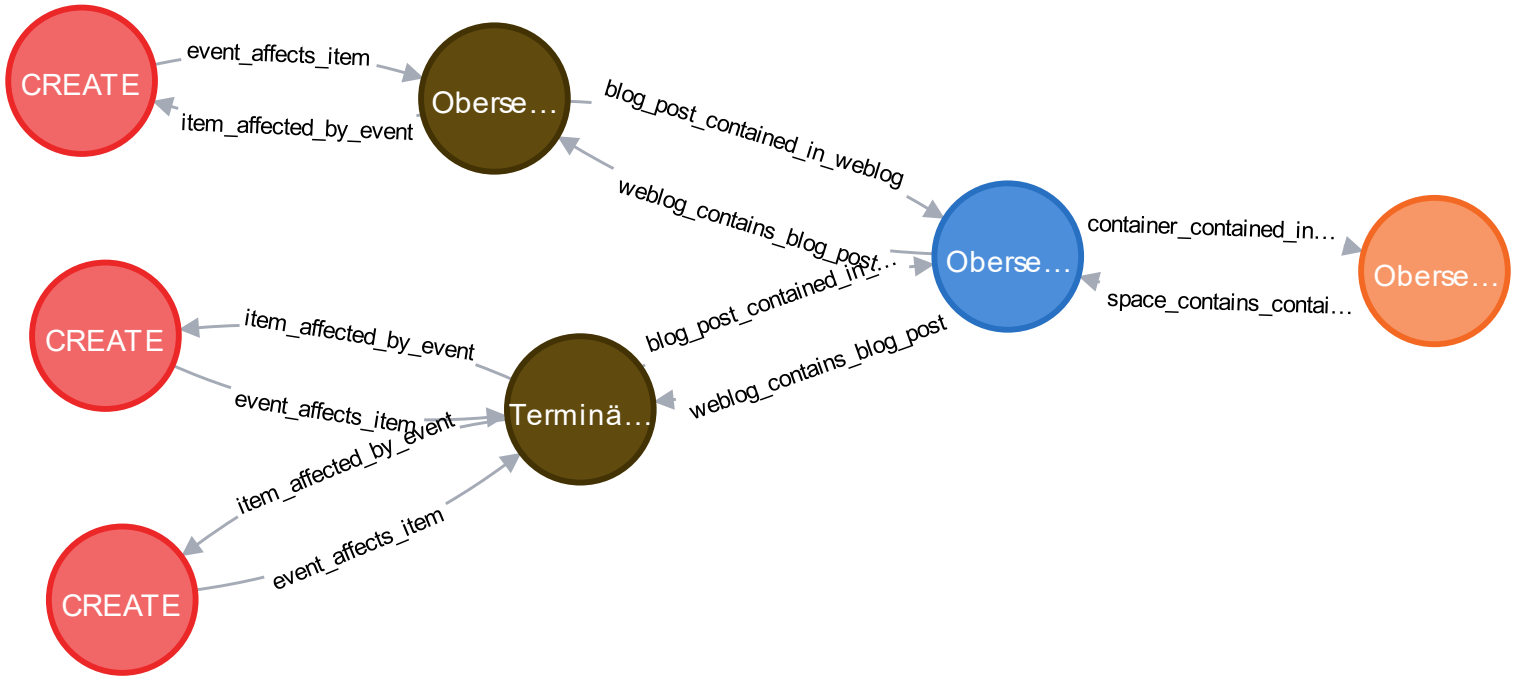
Transformation to ontology format

Import to graph database

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

Mapping: Graph databases

- **Advantages**:
  - **Visualisation** of data structures
  - **Materialisation** of ontology elements
  - High **performance**



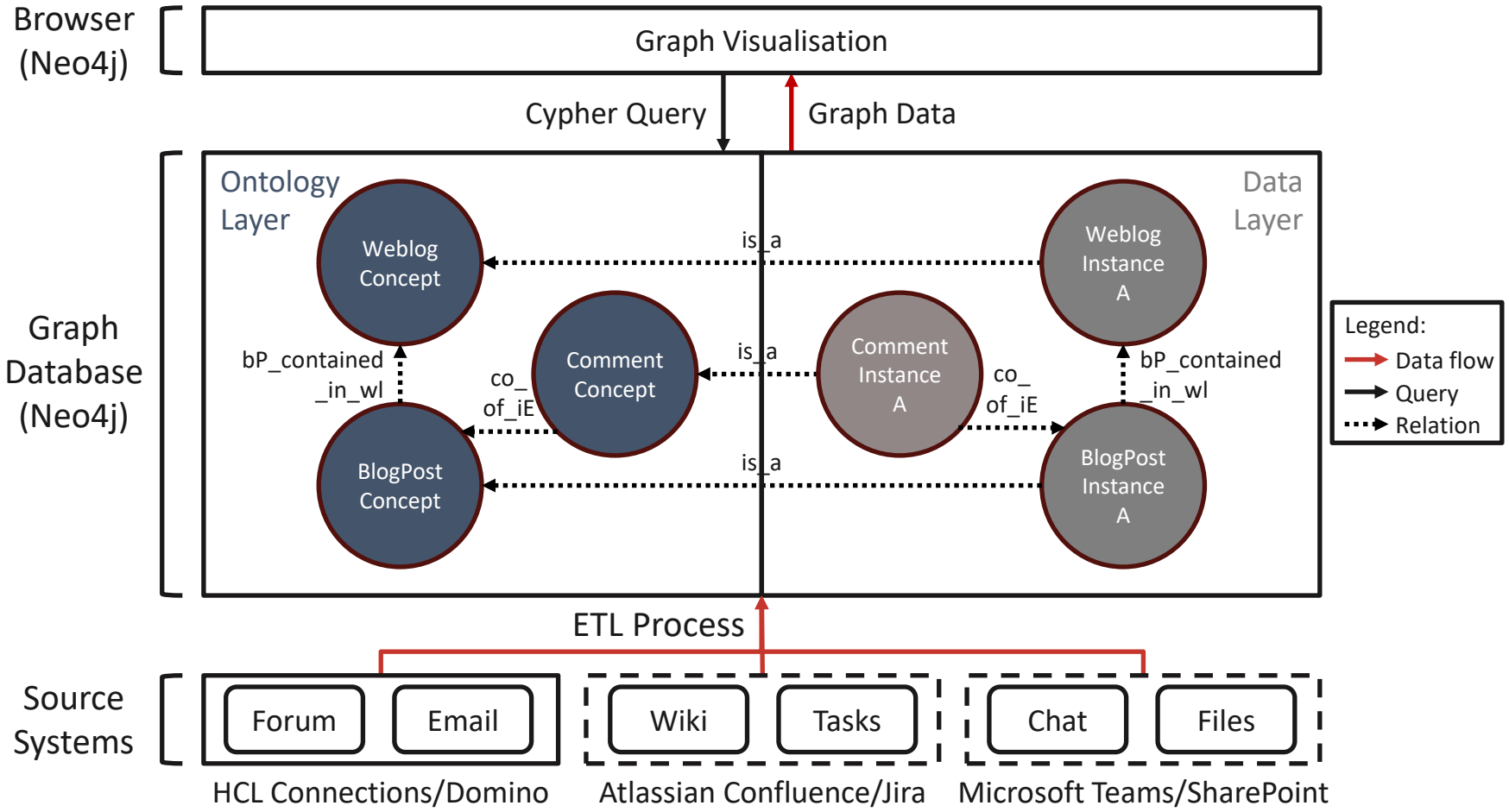
- **Neo4j Cypher** query

```
MATCH (e:Event)
-[rel1:event_affects_item]-(bP:BlogPost)
-[rel2:blog_post_contained_in_weblog]-(w:Weblog)
-[rel3:container_contained_in_space]-(gws:GroupWorkspace)
WHERE e.ACTION <> 'UNDEFINED'
RETURN e, bP, w, gws, rel1, rel2, rel3
LIMIT 3
```

MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

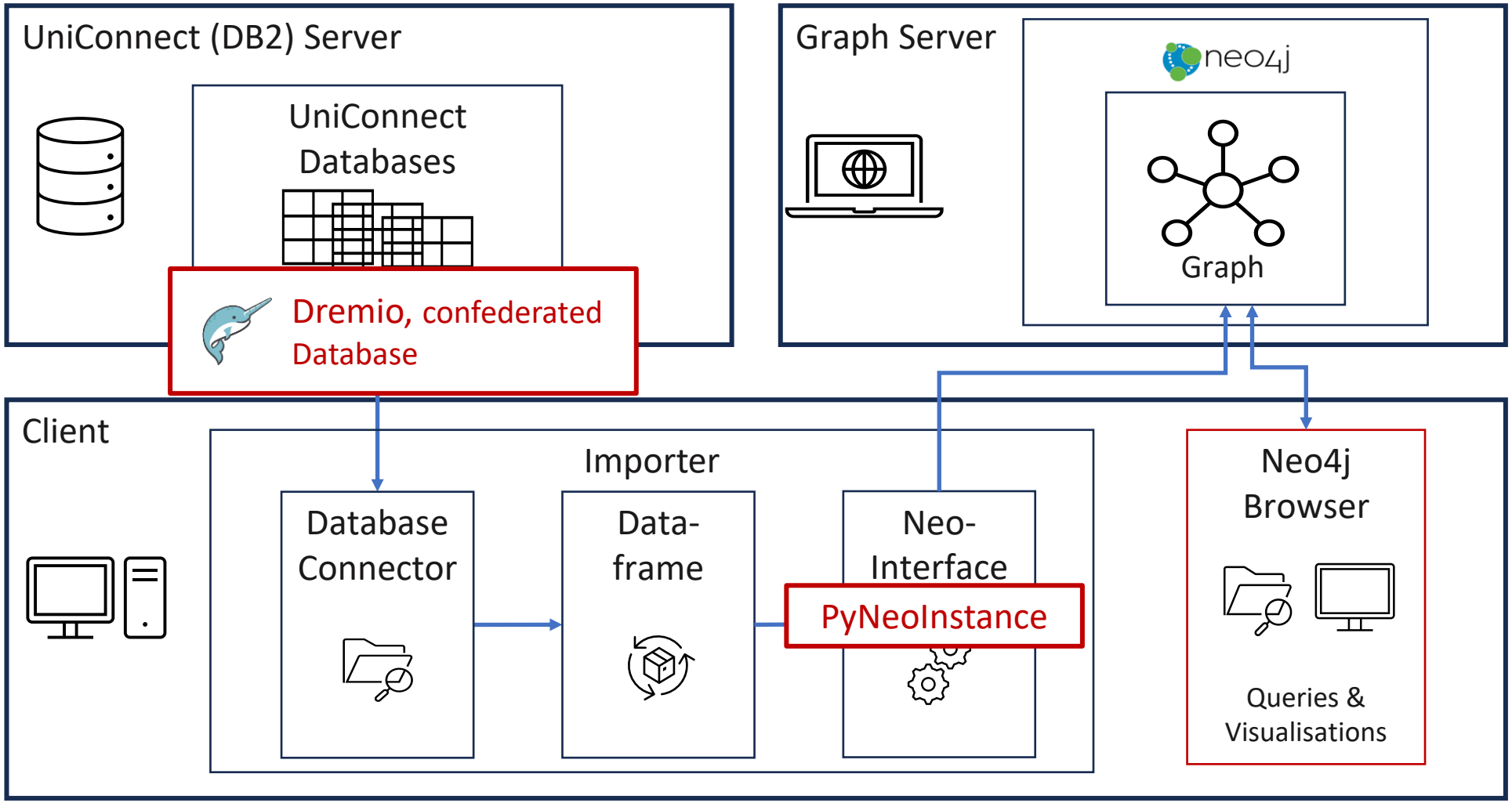
Mapping: Structure of the MOBDA datastore

- **Materialised linkage** of ontology and data
- Nodes: ~7 million
- Edges: ~32 million



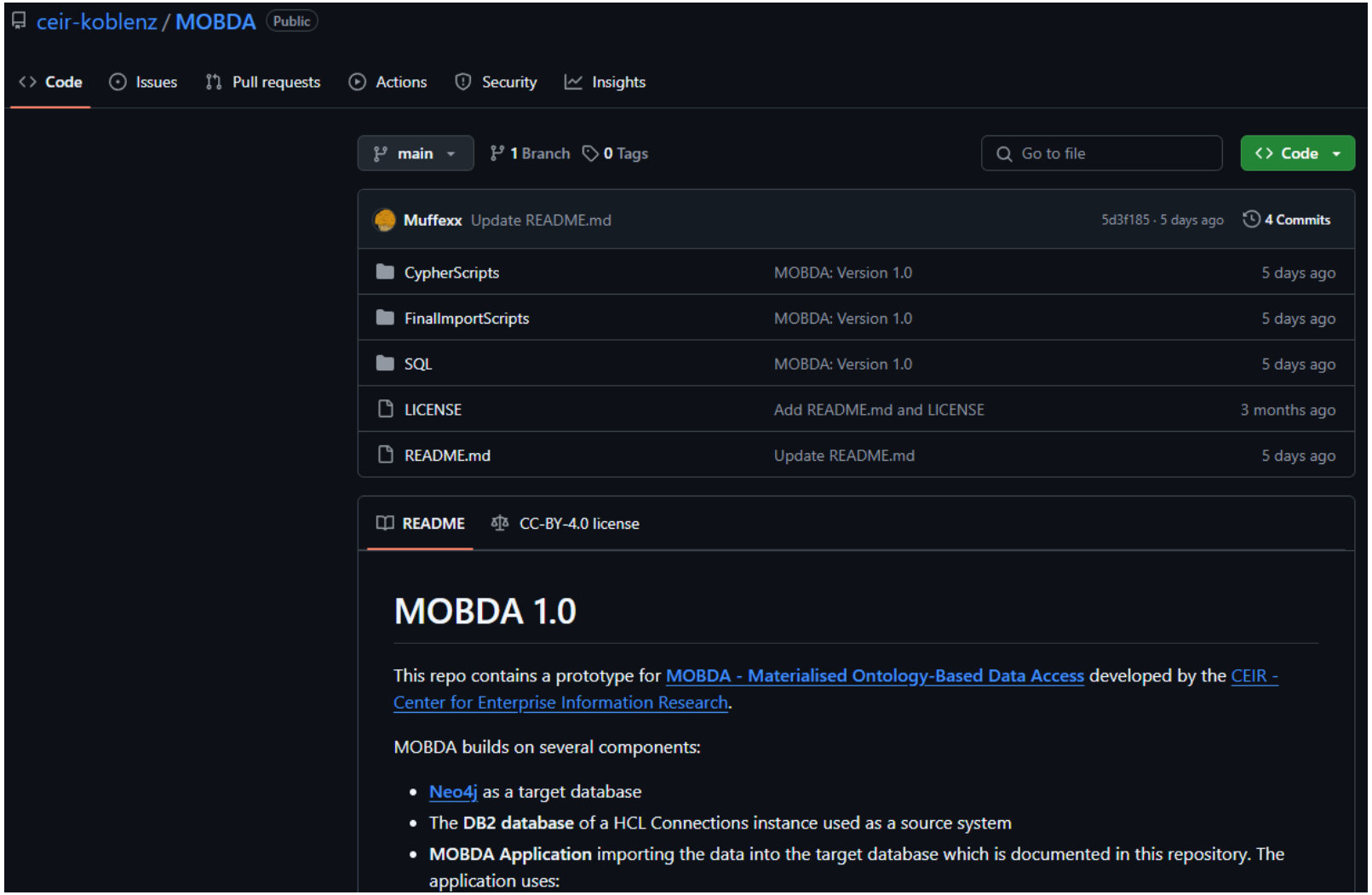
MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

System architecture: **Foreshadowing**



MOBDA – MATERIALISED ONTOLOGY-BASED DATA ACCESS

MOBDA: Open Access - w3id.org/CEIR/MOBDA



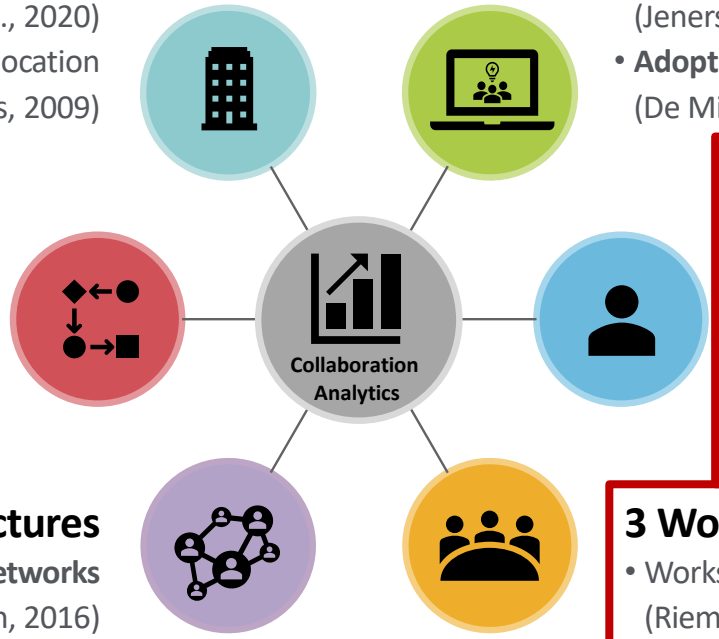
REPLICATION STUDY

Collaboration Analytics (CA): The six areas of CA

6 Organisational impact
- **Geographic** effects
(Warshaw et al. 2016; van Osch et al., 2017)
- **Ideas** and **social capital**
(Riemer et al., 2015; Bertolotti et al., 2020)
- **Expertise** location
(Nasirifard & Peristeras, 2009)

5 Activity sequences
- **Sequence** analysis
(Bøving & Simonsen, 2004)
- **Search** patterns
(Millen et al., 2007; Shami et al., 2011)
- Social **Process Mining**
(Naderipour, 2011; Blatt et al. 2023)

4 Network structures
- **Networks**
(Smith et al., 2009; Wiesneth, 2016)
- **Connectedness** of users
(Wu et al., 2010, Kügler et al., 2015)
- Influence of **hierarchies**
(Meske et al., 2020; Creamer et al., 2022)



1 System usage
- **General** usage metrics
(Steinhueser et al., 2015; Stocker & Müller, 2016)
- **Action-specific** metrics
(Jeners et al., 2013; Hacker et al., 2016)
- **Adoption** issues
(De Michele et al., 2018)

2 User participation
- **Intensity** of use
(Muller et al., 2012, Schwade and Schubert, 2019)
- **Nature** of use
(Hacker et al., 2017; Cetto et al., 2016)
- **Socialness**
(Guy et al., 2016; Meske et al. 2020)

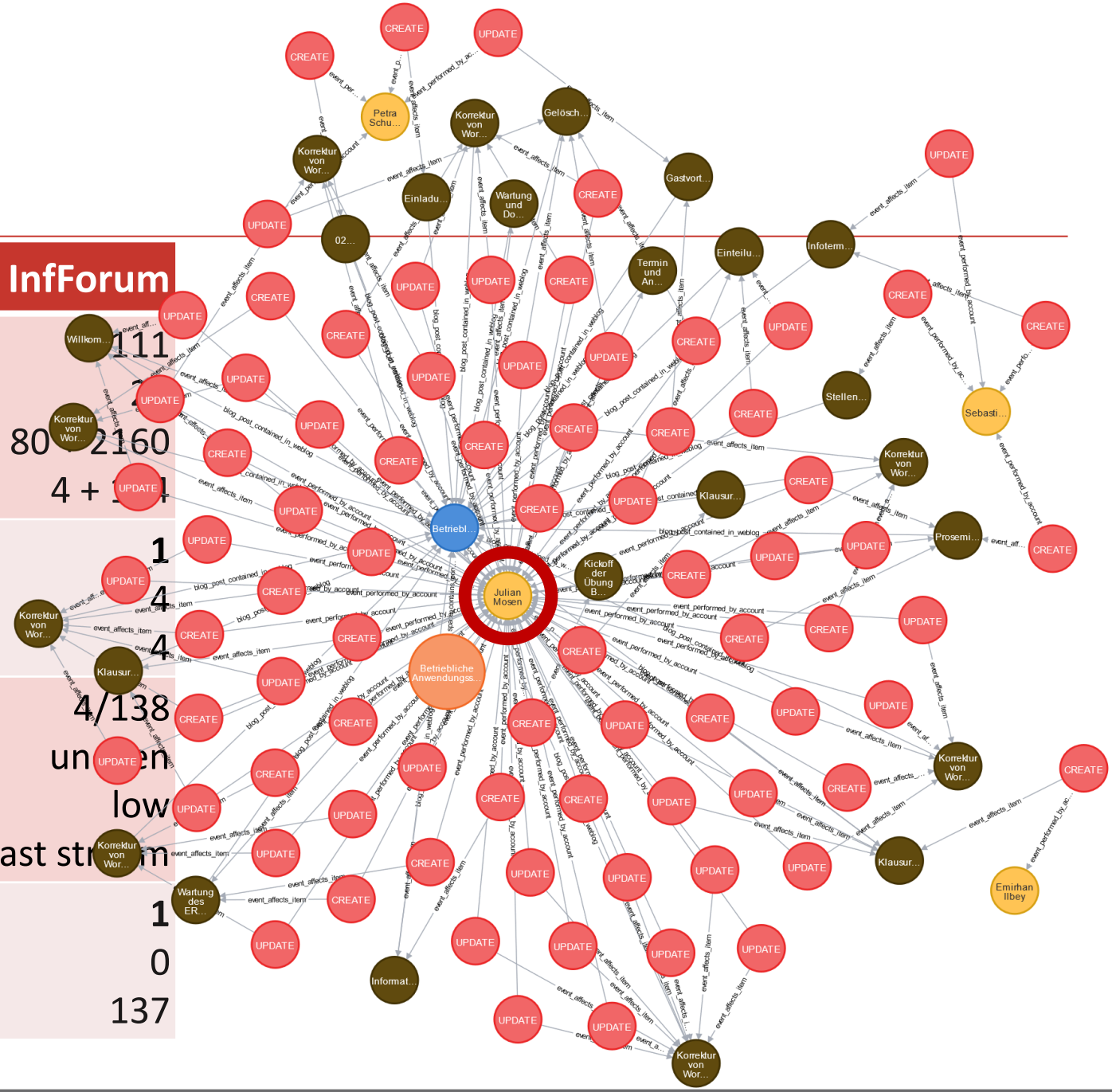
3 Workspaces
- Workspace **type**
(Riemer et al., 2020; Bahles et al., 2022)
- Community **health and success**
(Matthews et al., 2013; Aumayr and Hayes, 2017)
- **Content** analysis
(Williams et al. 2020, Schubert et al. 2020)

(Schubert, 2024, p. 6)

REPLICATION STUDY

Analysis – **Blog - User Example**

	Descriptions	InfForum
Graph elements (+ read events)	Nodes	111
	Blog entries	[illegible]
	Events	80[illegible]2160
	Accounts	4 + [illegible]1
User roles (Hacker & Riemer, 2021)	***Power users***	**1**
	Conversation starters	4
	Team members	4
Workspace types (Riemer et al., 2020)	Active users	4/138
	Activity distribution	un[illegible]en
	Reciprocity	low
	Workspace archetype	broadcast str[illegible]m
User types (Schwade, 2021)	***Creators***	**1**
	Contributors	0
	Consumers	137



REPLICATION STUDY

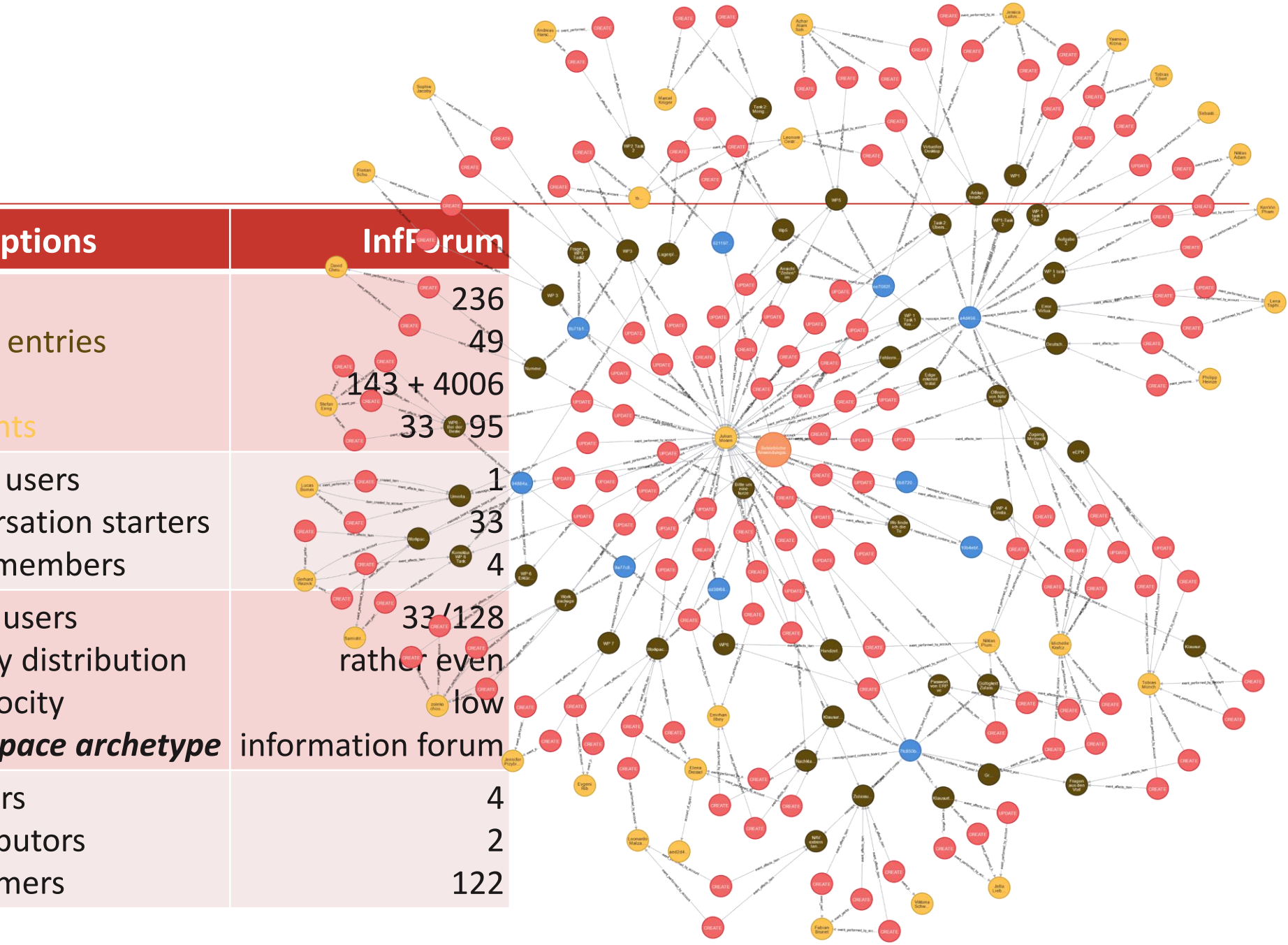
Analysis overview – **Forum**

	Descriptions	InfForum	PracticeCom	ProjTeam
Graph elements (+ read events)	Nodes	236	137	223
	Forum entries	49	19	45
	Events	143 + 4006	95 + 780	168 + 943
	Accounts	33 + 95	17 + 48	6 + 8
User roles (Hacker & Riemer, 2021)	Power users	1	0	2
	Conversation starters	33	17	5
	Team members	4	17	1
Workspace types (Riemer et al., 2020)	Active users	33/128	17/65	6/14
	Activity distribution	rather even	rather even	even
	Reciprocity	low	low	low
	Workspace archetype	information forum	information forum	Information forum
User types (Schwade, 2021)	Creators	4	2	4
	Contributors	2	6	1
	Consumers	122	57	9

REPLICATION STUDY

Analysis overview – **Forum**

	Descriptions	InfForum
Graph elements (+ read events)	Nodes	236
	Forum entries	49
	Events	143 + 4006
	Accounts	33 + 95
User roles (Hacker & Riemer, 2021)	Power users	1
	Conversation starters	33
	Team members	4
Workspace types (Riemer et al., 2020)	Active users	33 / 128
	Activity distribution	rather even
	Reciprocity	low
	Workspace archetype	information forum
User types (Schwade, 2021)	Creators	4
	Contributors	2
	Consumers	122

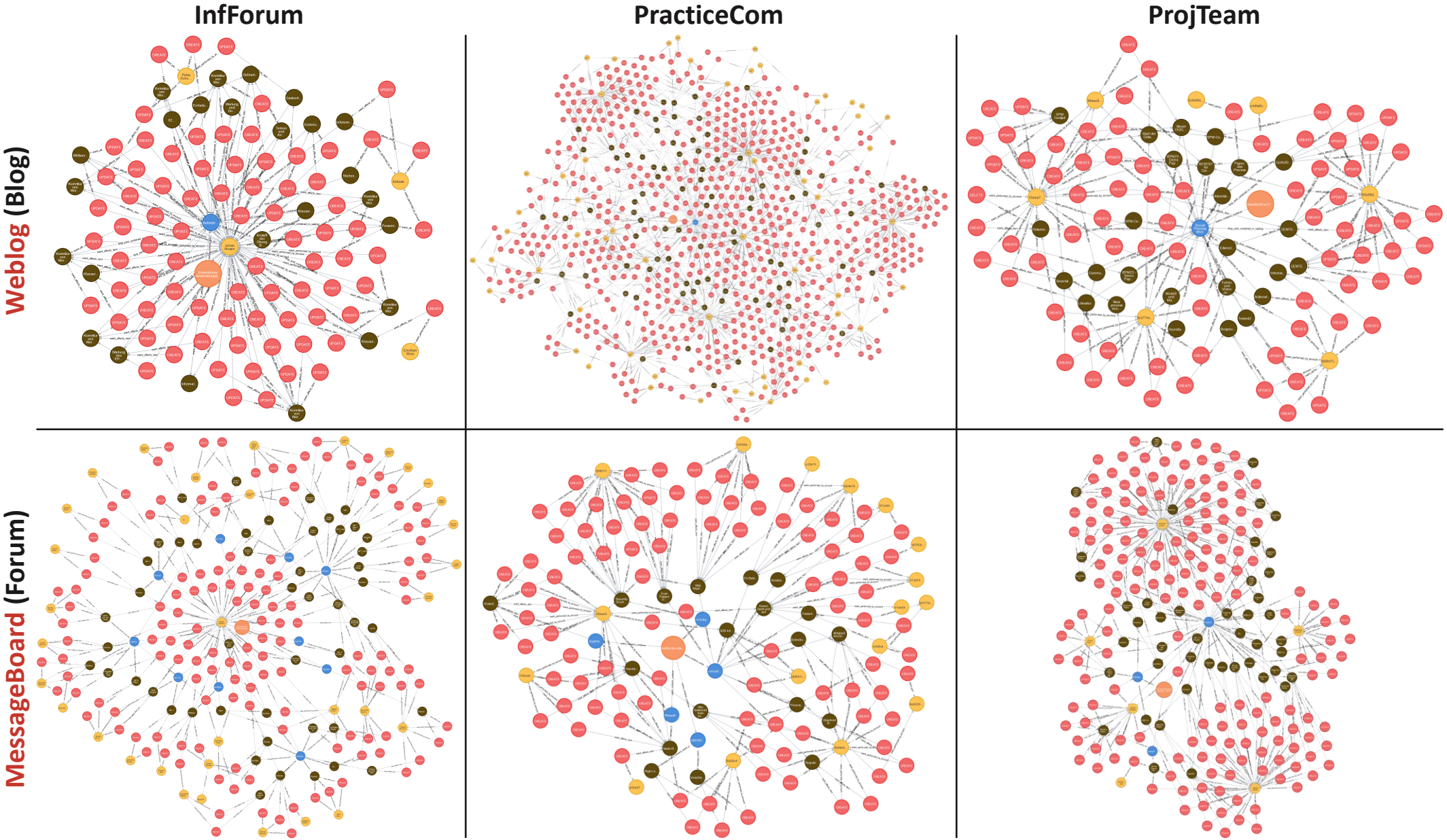


REPLICATION STUDY

Analysis overview – **Blog**

	Descriptions	InfForum	PracticeCom	ProjTeam
Graph elements (+ read events)	**Nodes**	111	821	113
	Blog entries	25	91	26
	Events	80 + 2160	669 + 3176	78 + 353
	Accounts	4 + 134	59 + 51	7 + 5
User roles (Hacker & Riemer, 2021)	**Power users**	1	5	4
	Conversation starters	4	18	6
	Team members	4	33	5
Workspace types (Riemer et al., 2020)	**Active users**	4/138	59/110	7/12
	Activity distribution	uneven	uneven	even
	Reciprocity	low	high	high
	Workspace archetype	broadcast stream	community of practice	project team
User types (Schwade, 2021)	**Creators**	1	10	4
	Contributors	0	23	1
	Consumers	137	77	7

REPLICATION STUDY



CONCLUSION

MOBDA as a concept for materialised data harmonisation

The path from chaos to structure

Further research

Q&A

REFERENCES

- Cysneiros, N. C., & Salgado, A. C. (2016). Including hierarchical navigation in a Graph Database query language with an OBDA approach. *2016 IEEE 32nd International Conference on Data Engineering Workshops, ICDEW 2016*, 109–114. https://doi.org/10.1109/ICDEW.2016.7495627
- Dennis, A., & Valacich, J. (2014). A Replication Manifesto. *AIS Transactions on Replication Research*, *1*, 1–4.
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, *43*(5–6), 907–928.
- Hacker, J., & Riemer, K. (2021). Identification of User Roles in Enterprise Social Networks: Method Development and Application. *Business & Information Systems Engineering*, *63*(4), 367–387. https://doi.org/10.1007/s12599-020-00648-x
- Just, M., & Schubert, P. (2023). Collaborative Actions on Documents Ontology (ColActDOnt). *Procedia Computer Science*, *219*, 294–302. https://doi.org/10.1016/j.procs.2023.01.293
- Just, M., & Schubert, P. (2024). A Dashboard for the Visualisation of Areas of Collaboration Analytics. *International Conference on ENTERprise Information Systems (CENTERIS)*, 1–9.

REFERENCES

- Riemer, K., Lee, L. L., Kjaer, C., & Haeffner, A. (2020). Identification of Enterprise Social Network (ESN) Group Archetypes in ESN Analytics. *Australasian Journal of Information Systems*, *24*, 1–20.
- Schubert, P. (2024). Areas of Collaboration Analytics. *International Conference on ENTERprise Information Systems (CENTERIS)*, 1–14.
- Schwade, F. (2021). Social Collaboration Analytics Framework: A framework for providing business intelligence on collaboration in the digital workplace. *Decision Support Systems*, *148*, 113587.
- Schwade, F., & Schubert, P. (2018). Social Collaboration Analytics for Enterprise Social Software: A Literature Review. *Multikonferenz Wirtschaftsinformatik 2018*, 1–12.
- Vaishnavi, V., Kuechler, B., & Petter, S. (2017). Design Science Research in Information Systems. *Association for Information Systems*, 1–66.

Thank you for your attention!

Lucas Schlömer

Researcher

Business Application Research Group
University of Koblenz

lucschloemer@uni-koblenz.de

Martin Just

Researcher

Business Application Research Group
University of Koblenz

martinjust@uni-koblenz.de

Petra Schubert

Professor

Business Application Research Group
University of Koblenz

schubert@uni-koblenz.de