

*39<sup>th</sup> International Conference on Conceptual Modeling*

*ER 2020*

*November 3-6, 2020*

# **PAST TRENDS AND FUTURE PROSPECTS IN CONCEPTUAL MODELING – A BIBLIOMETRIC ANALYSIS**

Felix Härer and Hans-Georg Fill

*University of Fribourg, Switzerland*



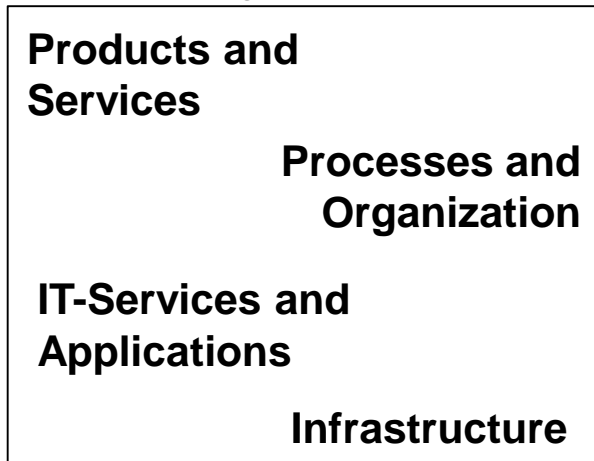
**UNIVERSITÉ DE FRIBOURG  
UNIVERSITÄT FREIBURG**

# AGENDA

1. Motivation
2. Data Collection
3. Results
  - Descriptive Analysis
  - Content-Based Analysis
4. Past Trends and Future Prospects

# MOTIVATION – DIGITAL TRANSFORMATION AND CONCEPTUAL MODELING

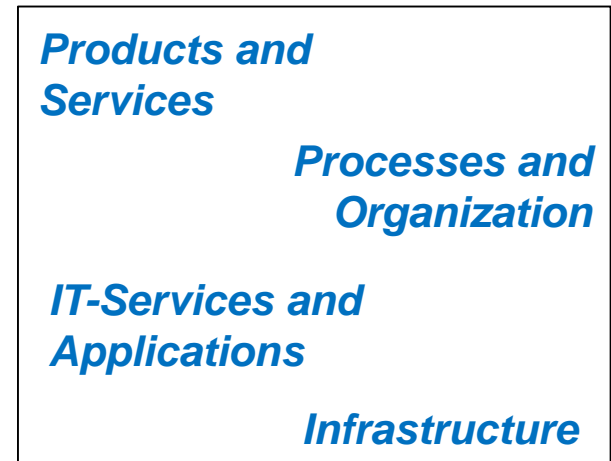
Traditional Organization



*Digital Transformation*



Digital Organization



Conceptual Modeling in the Past



*Prospects for Conceptual Modeling in the Future*

# MOTIVATION – ER TOPICS

*Proceedings ER 2005*

## Table of Contents

Specific Approaches  
Process Modeling and Views  
Conceptual Modeling in eLearning  
Managing Models and Modeling  
Requirements and Software Engineering  
Ontologies  
Web Services and Navigational Models  
Aspects of Workflow Modeling  
Queries and OLAP Summaries  
Temporal and Spatial Modeling

*Evolution  
of  
Topics*



*Proceedings ER 2019*

## Table of Contents

Invited Talks  
Conceptual Modeling  
Big Data Technology I  
Process Modeling and Analysis  
Query Approaches  
Big Data Technology II  
Domain Specific Models I  
Domain Specific Models II  
Decision Making  
Complex Systems Modeling  
Model Unification  
Big Data Technology III  
Requirements Modeling

# MOTIVATION – ANALYSIS REASONS

- *Chen, Song and Zhu* in 2007 published an analysis of ER conference papers between 1979 and 2005 [13]
  - No analysis has been performed since then to our knowledge
- *Jaakkola and Thalheim* suggest the possibility of a correlation between organizational cultures and the use of certain diagrams [7]
- Topical evolution of conceptual modeling and its direction

***“Those who remember the past ... can predict the future” ?***

[13] *Jaakkola, H., Thalheim, B.: Culture-adaptable web information systems. In: Information Modelling and Knowledge Bases XXVII. Frontiers in Artificial Intelligence and Applications, vol. 280 (2016)*

[7] *Chen, C., Song, I.Y., Zhu, W.: Trends in conceptual modeling: citation analysis of the ER conference papers (1979–2005). In: Proceedings of the 11th International Conference on the International Society for Scientometrics and Informatics, pp. 189–200. CSIC (2007)*

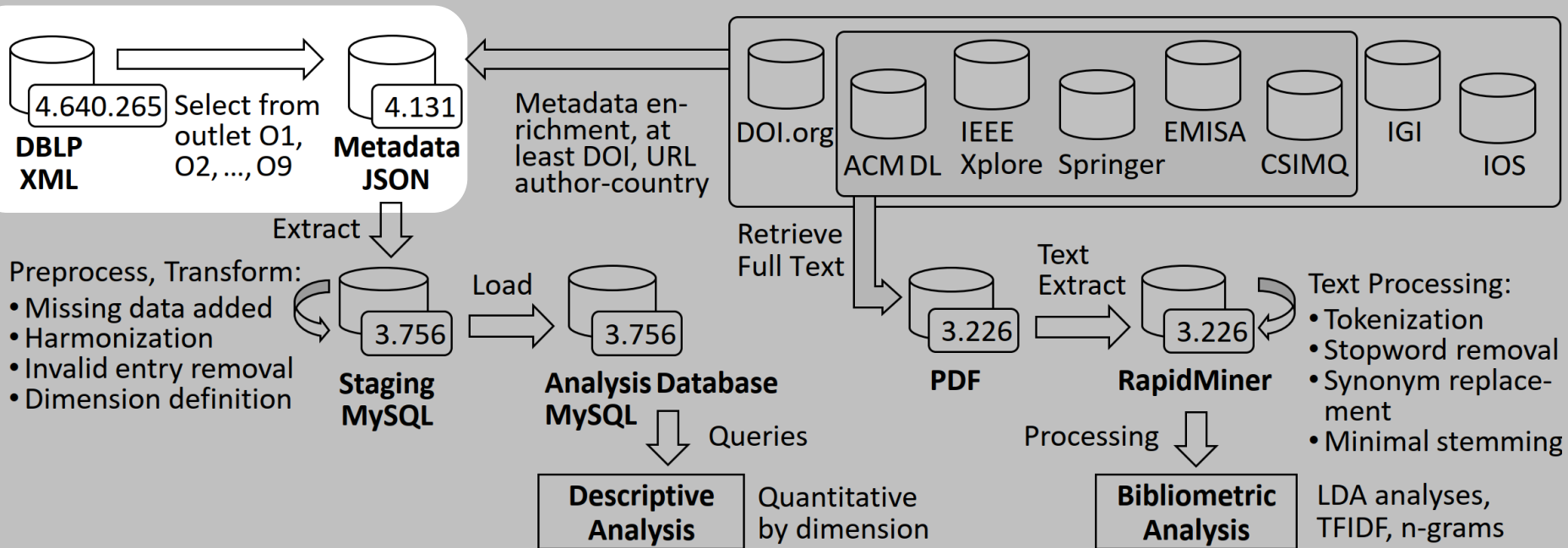
# RESEARCH QUESTIONS

## Resulting Research Questions:

- **RQ 1:** In which **geographical regions** is conceptual modeling research conducted?
- **RQ 2:** How has the **quantity of conceptual modeling papers evolved over time**, taking into account specific regions and the communities of specific outlets?
- **RQ 3:** **How many authors** are active in conceptual modeling and what is their **typical number of papers**?
- **RQ 4:** **What are the major conceptual modeling topics** in terms of published research, how do they **differ per outlet** and how did they **evolve over time**?
- **RQ 5:** With the topical evolution, **where is conceptual modeling positioned now** and are there **indications of research gaps or opportunities**?
- **RQ 6:** Is there an indication of prospective topics, application areas, or domains **where conceptual modeling might be used in the future**?

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset



 Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Selection of nine English-language conference and journal outlets:

### Conferences

---

**BMSD** International Symposium on Business Modeling and Software Design

**BPMDS/EMMSAD** Joint publication of Business Process Modeling, Development and Support in conjunction with Evaluation and Modelling Methods for Systems Analysis and Development

**ER** International Conference on Conceptual Modeling

**MoDELS** International Conference on Model Driven Engineering Languages and System

**PoEM** Practice of Enterprise Modeling

### Journals

---

**CISMQ** Complex Systems Informatics and Modeling Quarterly

**EMISAJ** Enterprise Modelling and Information Systems Architectures - International Journal of Conceptual Modeling

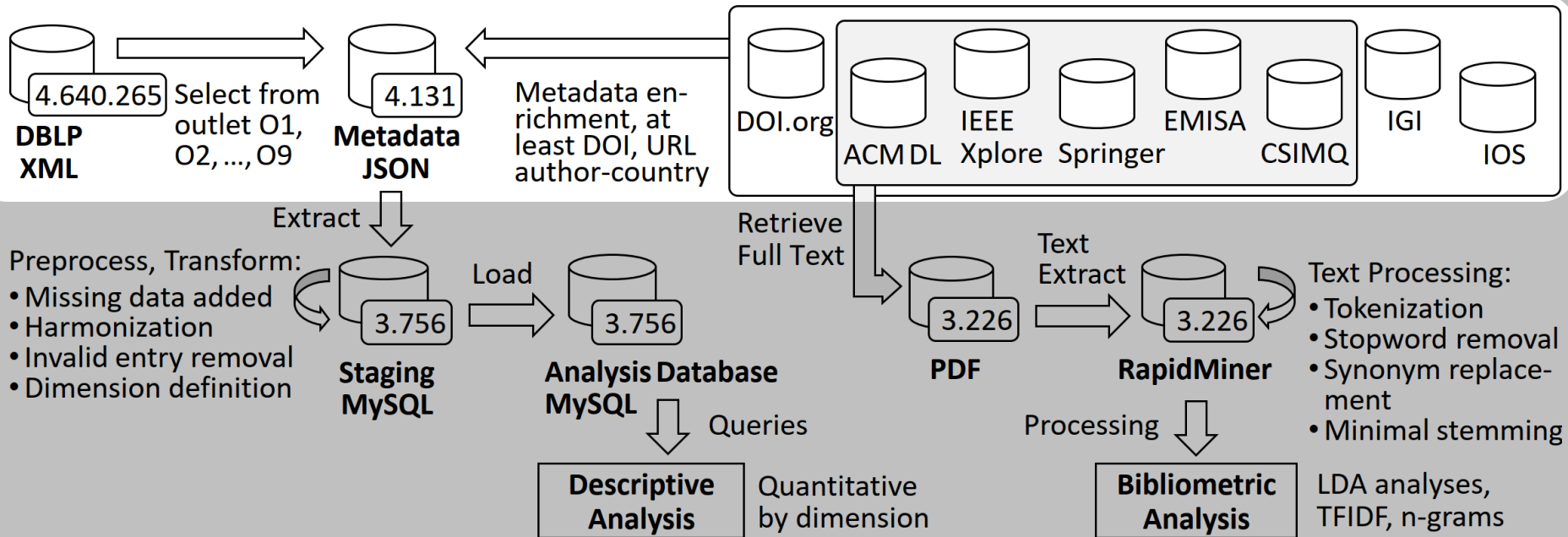
**IJISMD** International Journal of Information System Modeling and Design

**SoSyM** Software and Systems Modeling



# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset

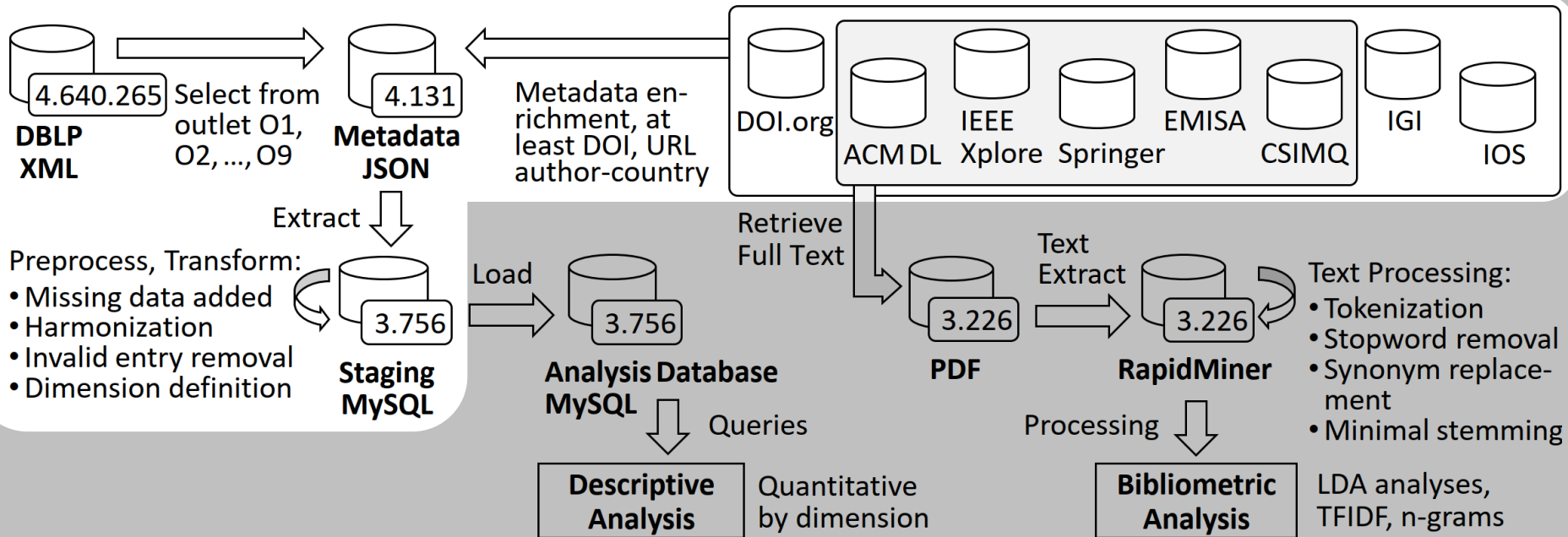



Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset

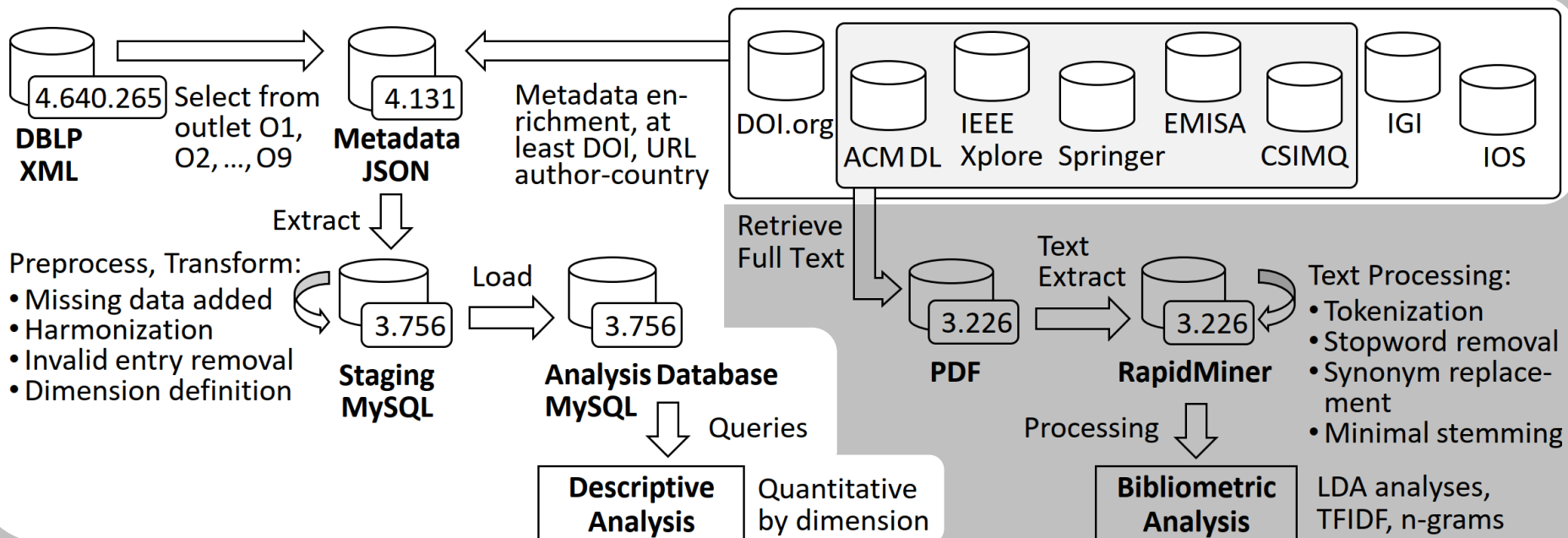


 Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset

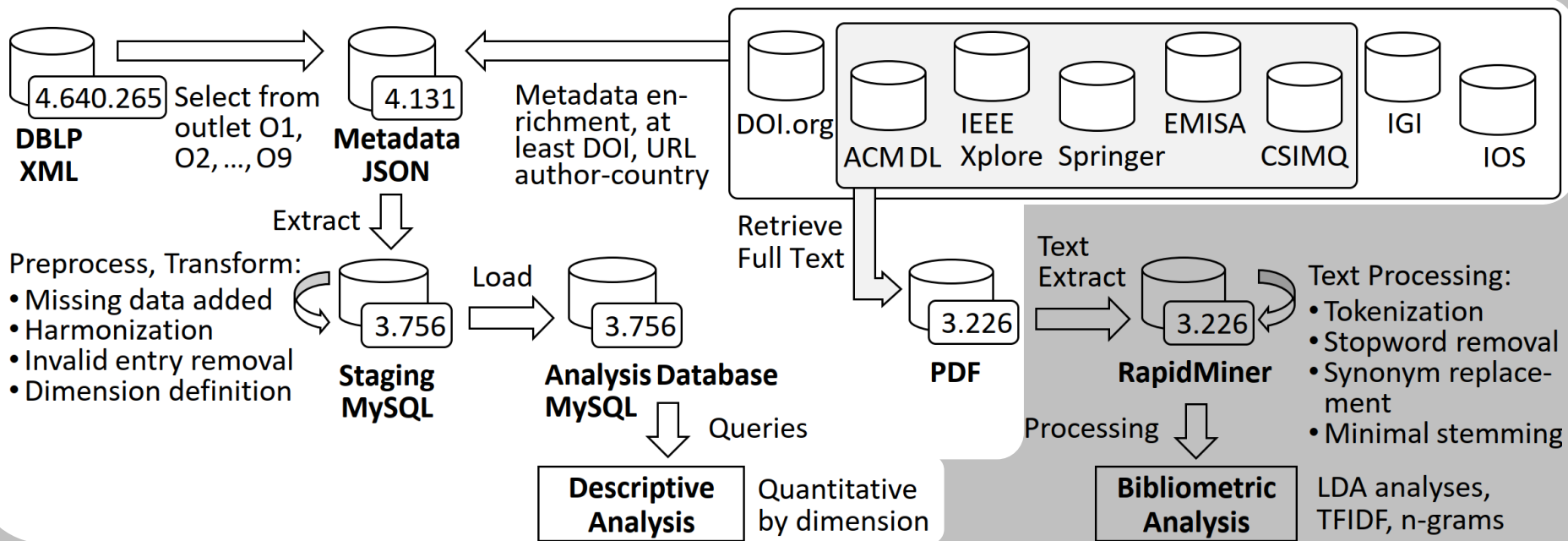


 Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset

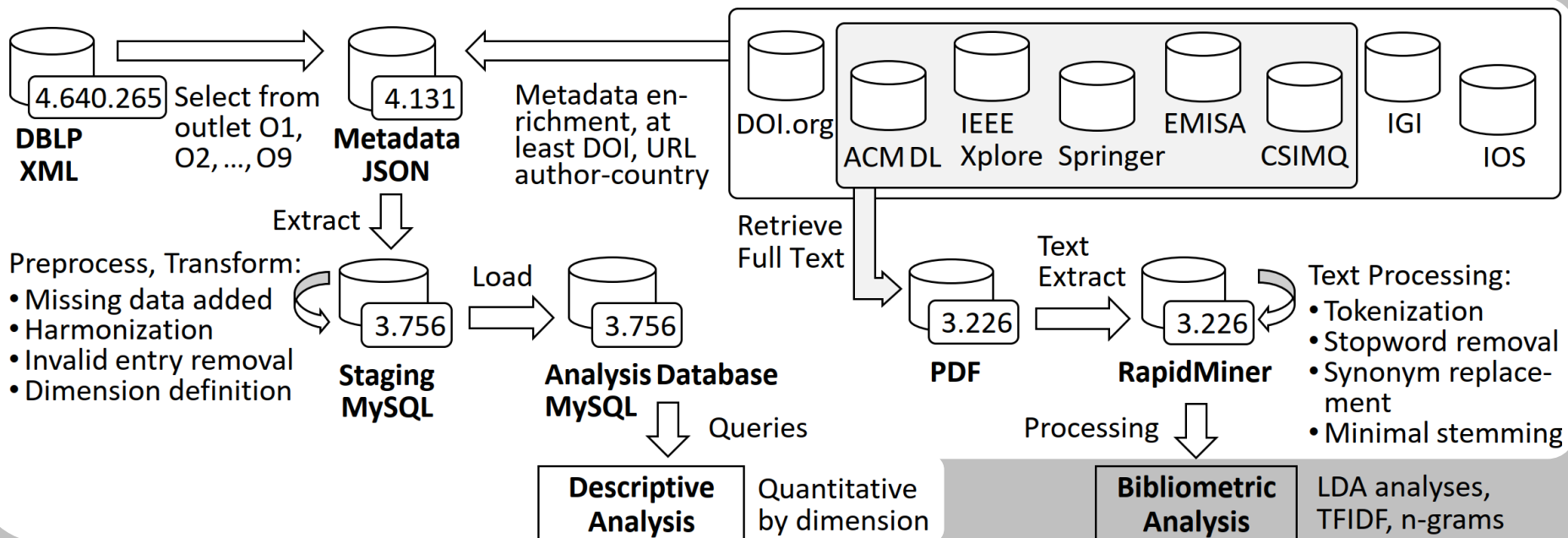


Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset

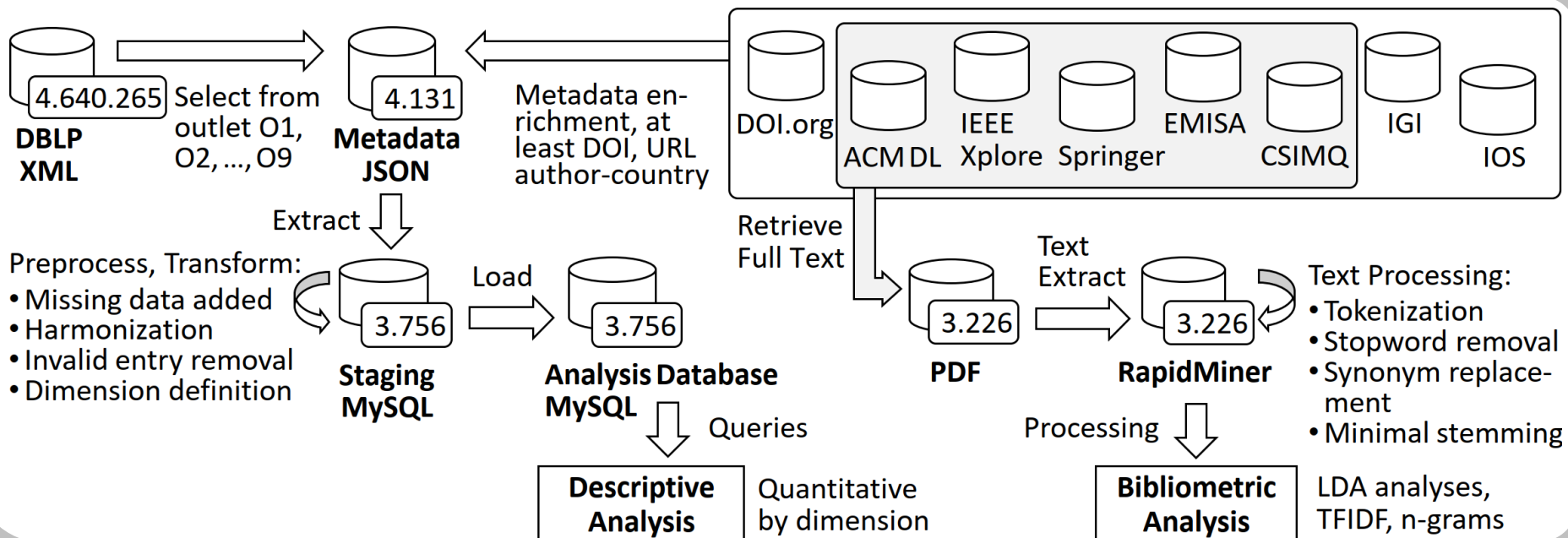


 Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# DATA COLLECTION

## Data collection and analysis process based on the DBLP XML dataset



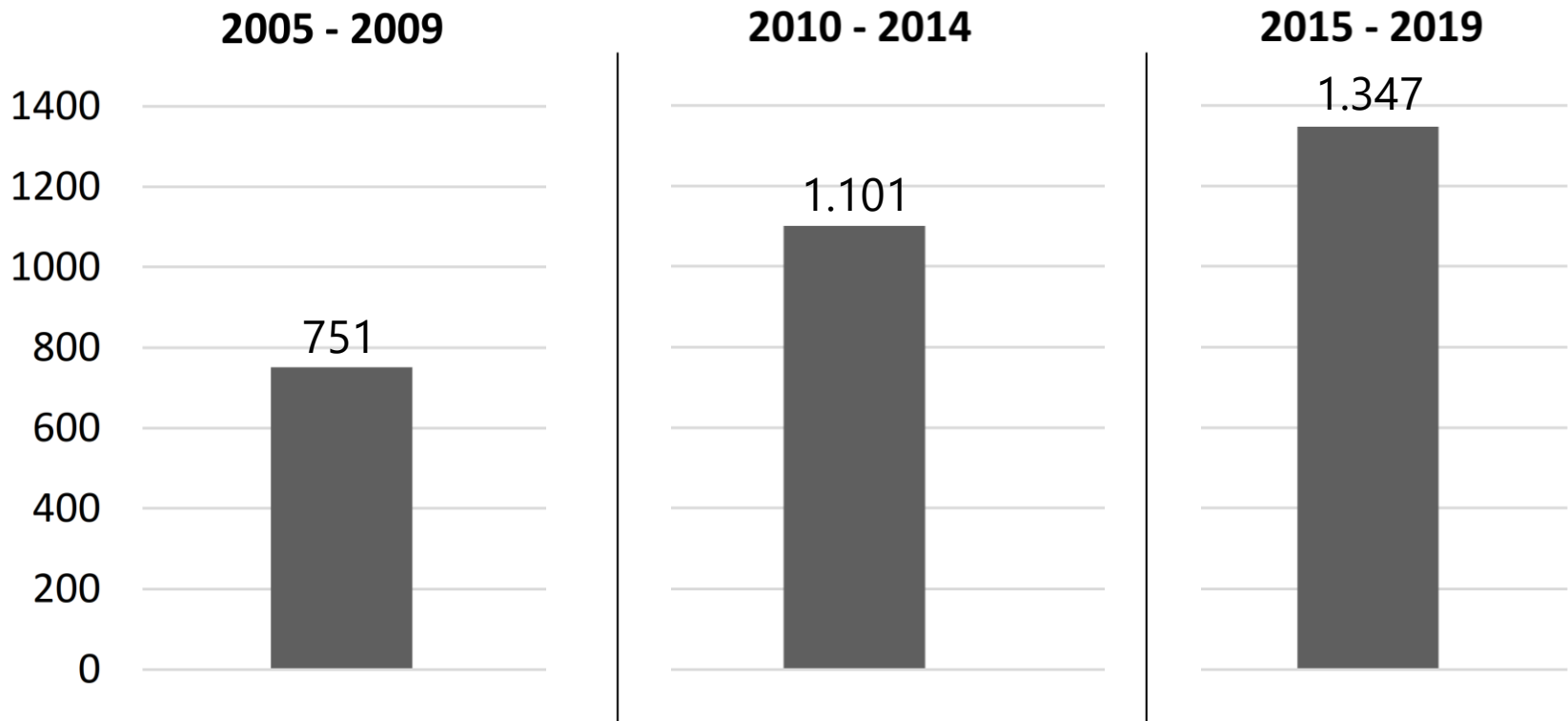
 Dataset containing n elements

O1: BMSD, O2: BPMDS/EMMSAD, O3: ER, O4: MoDELS, O5: PoEM, O6: CSIMQ, O7: EMISAJ, O8: IJISMD, O9: SoSyM

# RESULTS – DESCRIPTIVE ANALYSIS

## Quantity of conceptual modeling papers over time (RQ 1, 2)

### ➤ Number of publications over time



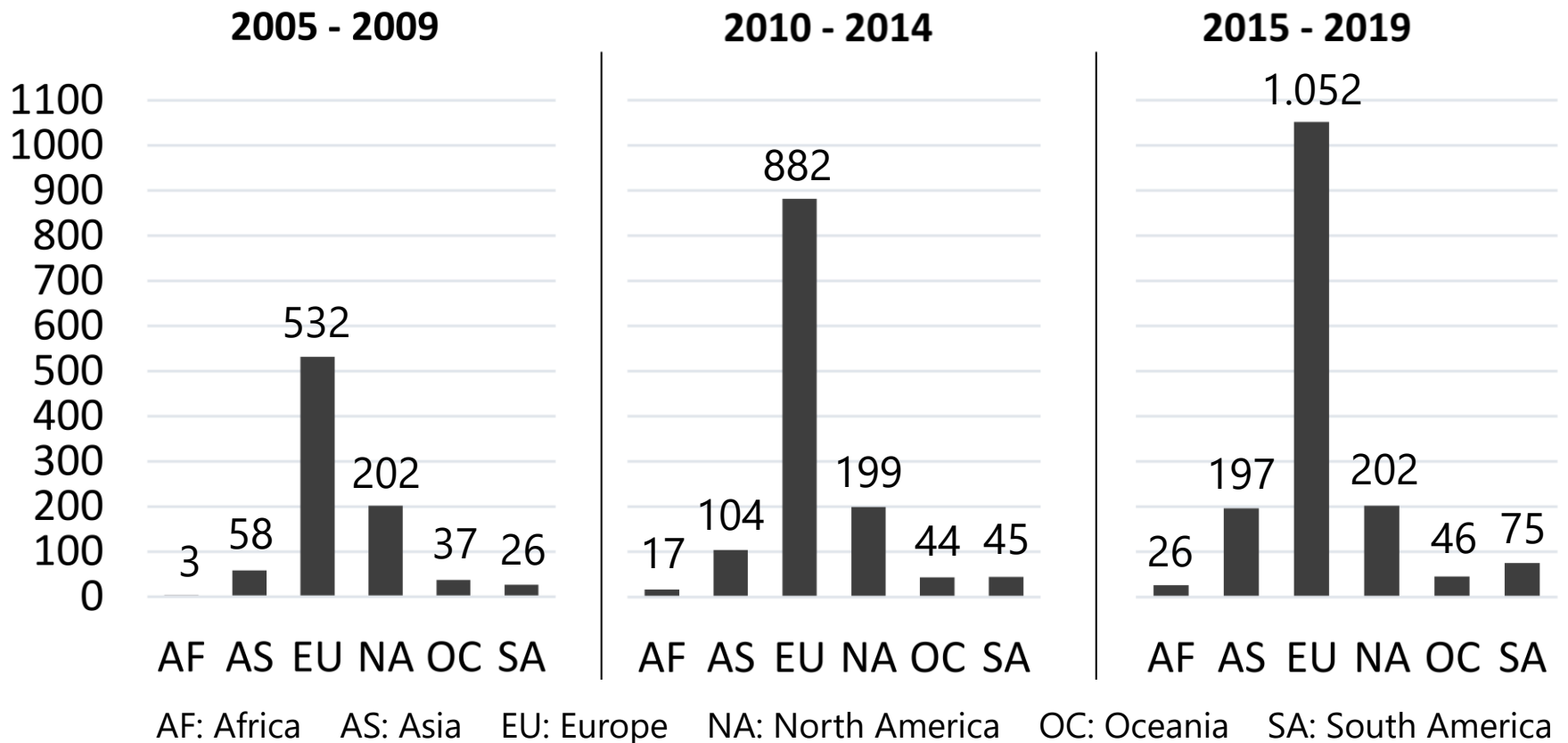
2005 – 2019: n=3.199

*see Table 3 for details*

# RESULTS – DESCRIPTIVE ANALYSIS

## Geographical regions of conceptual modeling research (RQ 1, 2)

- Number of publications originating from authors affiliated with institutions in the following continents



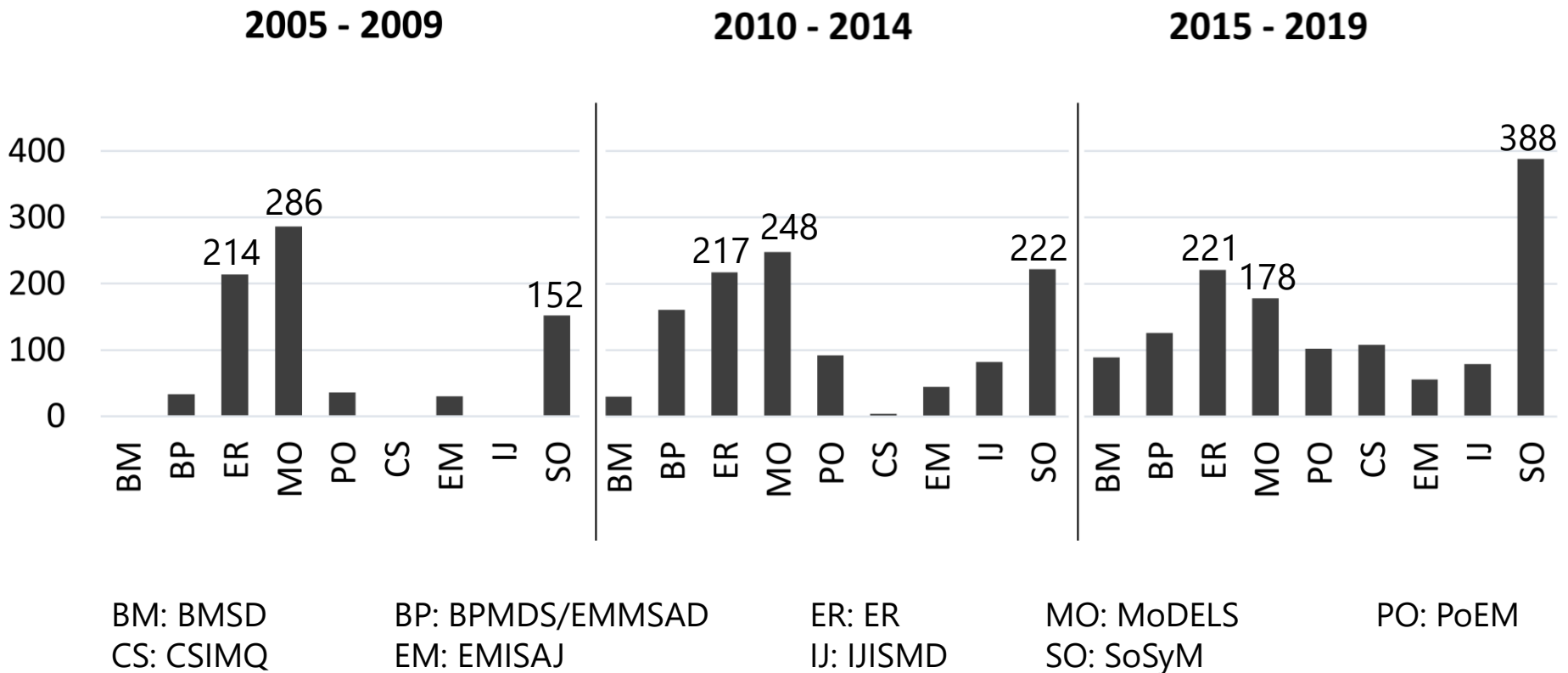
*see Table 2 for details*



# RESULTS – DESCRIPTIVE ANALYSIS

## Outlets in conceptual modeling research over time (RQ 1, 2)

➤ Number of publications for the top 3 outlets by quantity:



see Table 3 for details

# RESULTS – DESCRIPTIVE ANALYSIS

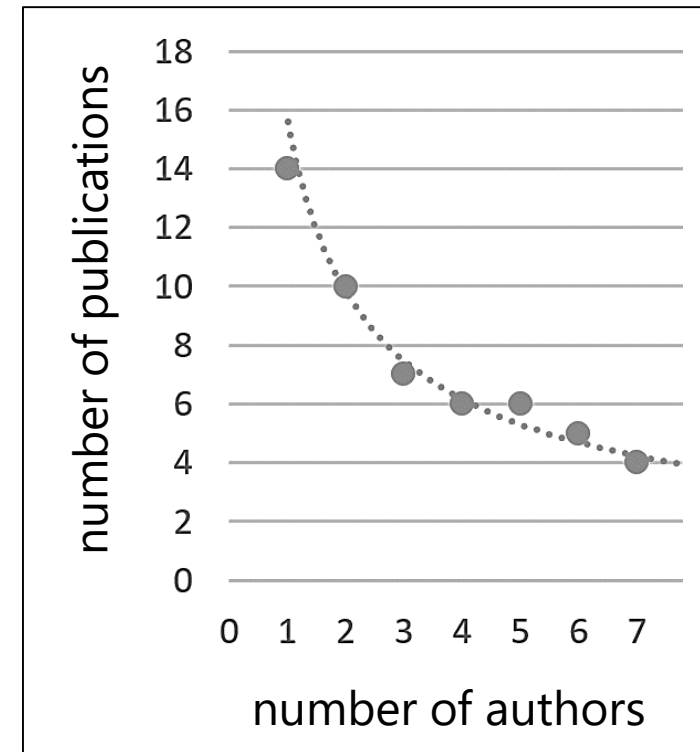
## Authors in conceptual modeling and their number of papers (RQ 3)

### ➤ Number of authors

- $n = 5.141$  in the time frame 2005 – 2019 (inclusive)
- 69% of authors have exactly one publication
- 31% have between 2 and 61 publications

### ➤ Percentile measures

- Skew observed: 80% of authors have less than  $P_{80} = 2$  publications
- Steep increase after  $P_{90} = 4$  and  $P_{95} = 6$  indicating a power law distribution



*see Table 1 for the number of authors per outlet*

# RESULTS – DESCRIPTIVE ANALYSIS

## Discussion

- Quantitatively, a steady growth overall and for most outlets over the time frames 2005 – 2009 and 2010 – 2014 and 2015 – 2019
- The community has a relatively small core of authors
- Many of the 5.141 authors are not part of it, possibly from collaborating fields

# RESULTS – CONTENT-BASED ANALYSIS

## Major conceptual modeling topics in terms of published research (RQ 4, 5)

- **Topic modeling approach**

- **Method:** Latent Dirichlet Allocation (LDA) from MALLET library
- **LDA:** for all words of all documents, a sampling scheme and learning algorithm determines the mapping to a particular topic and its weight
- **Result:** fixed number of 8 topics consisting of the top 5 weighted words

# RESULTS – CONTENT-BASED ANALYSIS

## Major conceptual modeling topics in terms of published research (RQ 4, 5)

### ➤ ER publications (2005 – 2019):

#### Topic 1

model  
conceptual  
system  
design  
information

#### Topic 2

schema  
data  
node  
database  
query

#### Topic 3

process  
model  
service  
data  
event

#### Topic 4

goal  
model  
requirement  
value  
system

#### Topic 5

model  
type  
class  
relationship  
object

#### Topic 6

ontology  
model  
type  
concept  
relation

#### Topic 7

data  
user  
concept  
web  
result

#### Topic 8

set  
constraint  
relation  
data  
tuple

*see Table 5 for weights and further details*

# RESULTS – CONTENT-BASED ANALYSIS

## Major conceptual modeling topics in terms of published research (RQ 4, 5)

### ➤ ER publications (2005 – 2019) – interpretation of topics:

#### Conceptual Modeling Approaches

model  
conceptual  
system  
design  
information

#### Databases

schema  
data  
node  
database  
query

#### Process Modeling

process  
model  
service  
data  
event

#### Goals and Requirements

goal  
model  
requirement  
value  
system

#### Classes and Types

model  
type  
class  
relationship  
object

#### Ontologies

ontology  
model  
type  
concept  
relation

#### Data, User, Web

data  
user  
concept  
web  
result

#### Relational Theory

set  
constraint  
relation  
data  
tuple

*see Table 5 for weights and further details*

# RESULTS – CONTENT-BASED ANALYSIS

## Major conceptual modeling topics in terms of published research (RQ 4, 5)

### ➤ All publications (2005 – 2019):

#### Topic 1

process  
model  
business  
service  
system

#### Topic 2

model  
language  
tool  
metamodel  
uml

#### Topic 3

model  
transformation  
rule  
graph  
element

#### Topic 4

model  
system  
test  
software  
feature

#### Topic 5

model  
use  
case  
software  
tool

#### Topic 6

state  
event  
model  
transition  
system

#### Topic 7

class  
constraint  
model  
type  
object

#### Topic 8

data  
schema  
database  
query  
set

*see Table 4 for weights and further details*

# RESULTS – CONTENT-BASED ANALYSIS

## Major conceptual modeling topics in terms of published research (RQ 4, 5)

### ➤ All publications (2005 – 2019) – interpretation of topics:

#### **Process Modeling**

process  
model  
business  
service  
system

#### **Modeling Language**

model  
language  
tool  
metamodel  
uml

#### **Model Transformation**

model  
transformation  
rule  
graph  
element

#### **Software**

model  
system  
test  
software  
feature

#### **Use Cases**

model  
use  
case  
software  
tool

#### **State, Transition**

state  
event  
model  
transition  
system

#### **Classes and Types**

class  
constraint  
model  
type  
object

#### **Databases**

data  
schema  
database  
query  
set

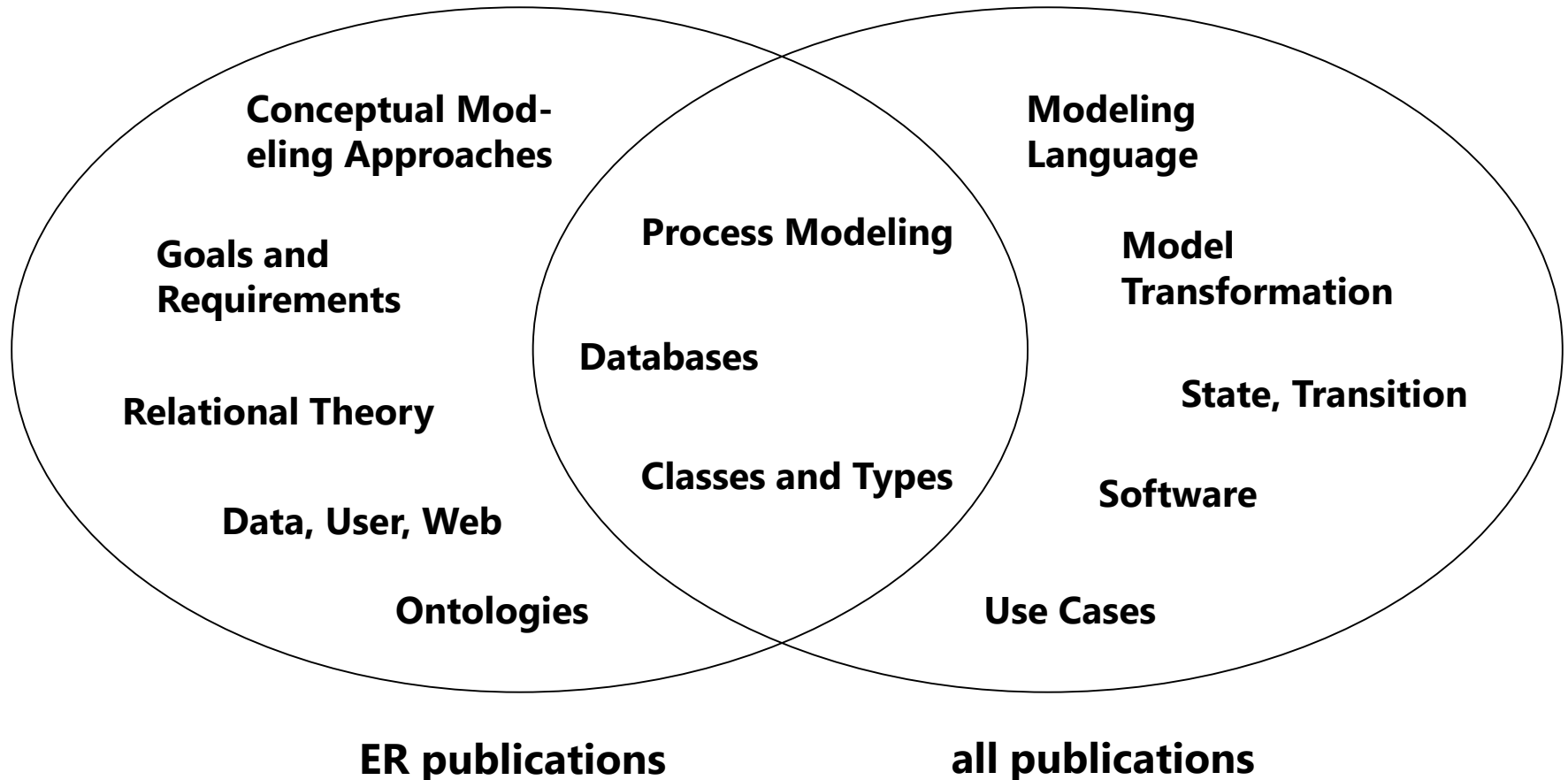
*see Table 4 for weights and further details*



# RESULTS – CONTENT-BASED ANALYSIS

Major conceptual modeling topics in terms of published research (RQ 4, 5)

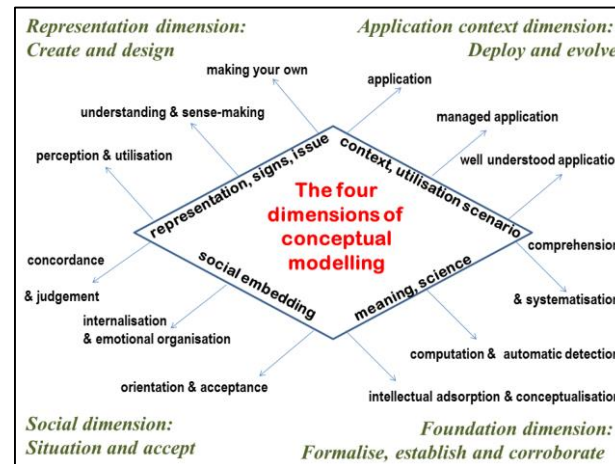
➤ Intersection of ER topics and overall conceptual modeling topics:



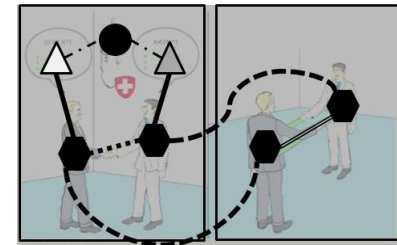
# RESULTS – CONTENT-BASED ANALYSIS

## Discussion

- **Identification of topics in the areas:**
  - Fundamental and general aspects of modeling and schemas
  - Application domain centered on process modeling
  - Specific languages and notations, e.g. state, transition
  - Technical aspects, e.g. transformation
  
- **Few topics identified for:**
  - Methodological topics
  - Application domains beside processes
  - Other domains missing, e.g. Law, Humanities, Natural Sciences



*E.g. Kropp, Thalheim (2020) on conceptual modelling and humanities [17]*



*E.g. Fill (2012) on visual law representations [11]*

# RESULTS – CONTENT-BASED ANALYSIS

## Discussion

- **Summary of indications from further LDA analyses:**
  - Observations over time
    - Model transformation has been of constant interest
    - Fewer publications on software models in recent years (2015 – 2019) compared to prior time frames
    - More publications on process and business information systems modeling in recent years (2015 – 2019) compared to prior time frames
  - Regional differences between North America and Europe
    - Modeling tools, goals and requirements stronger in North America
    - Processes, Databases and data schemas stronger in Europe
    - Software, data and transformation comparable in both regions

Note: these trends are only meaningful if the dataset is a good sample for the field of conceptual modeling.

*see Tables 6, 7, 8, 9 for topics and further details*

# RESULTS – CONTENT-BASED ANALYSIS

## Discussion of Limitations

We do believe the dataset to be a good sample, however, please note:

- Varying data availability per outlet, e.g. fewer recent MoDELS publications
- LDA interpretation
- Selection of only a subset of outlets for publishing research
  - Only including English-language publications
  - Not including workshops
  - Not including publications in domain-oriented outlets

# OUTLOOK

## Past Trends and Future Prospects

- Major conceptual modeling topics found to be well established
  - LDA indicates focus on fundamental topics, processes, information systems
- Application domains and specific fields are underrepresented
  - Broaden the topics of major conceptual modeling outlets, e.g. in regard to humanities, natural sciences, law
- Make conceptual modeling more accessible to other fields and users not working in computer science and information systems

## Next Steps

- Design and conduction of tests for the interpretations and hypotheses derived from the data
- Targeted analysis for results and identified topics
- Extension of the analysis method, e.g. using semantic approaches