



Foto: Thomas Josek

Datenbanktechnologie

SoSe 2022, Woche 10: 20. Juni.

Institut für Digital Humanities, Historisch-Kulturwissenschaftliche Informationsverarbeitung | Prof. Dr. Øyvind Eide

NoSQL

- Nicht nur SQL
 - Not only SQL
- Nicht-relationaler Ansatz
- Keine festgelegten Tabellenschemata
- Alternativen zu Relationsdatenbanken
 - Relationsdatenbanken noch dominierend
- In den letzten 25 Jahren
 - aber: Netzwerksdatenbanken
 - ursprüngliche Meinung: no SQL
 - seit 2009: not only SQL

NoSQL-Typologie

- Dokumentenorientierte Datenbanken
 - Dokumente als Grundeinheit zur Speicherung der Daten
 - Z. B. XML-Datenbanken
- Graphdatenbanken
 - Beispiel folgt
- Schlüssel-Werte-Datenbank
 - Key-Value-Datenbanken
 - Hashtabelle
 - Schnell
 - Nicht für komplexe Datenstrukturen geeignet

NoSQL-Typologie

- Objektdatenbank
 - Objektorientiert
 - Klassenhierarchie
 - Kapselung der Objekte
- Spaltenorientierte Datenbanken
 - Data-Warehouse
 - Aggregate über große Zahlen ähnlicher Elemente
- Datenstrom
 - Data Stream Management System
 - Verwaltung von kontinuierlichen Datenströmen

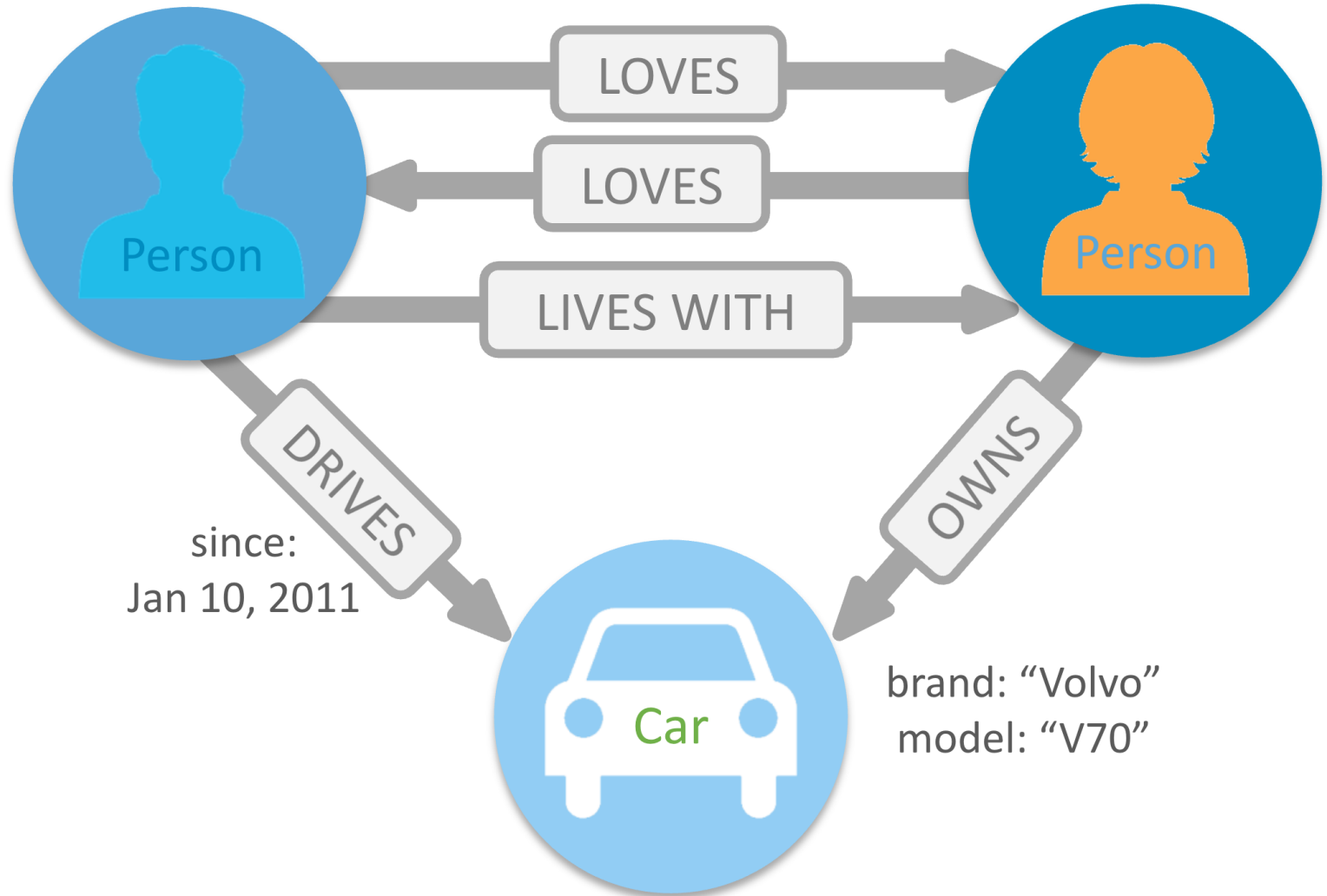
Neo4j

- Beispiel einer NoSQL-Datenbank
- Speicherung:
 - Kanten
 - Knoten
 - Attributen
- Programmiersprache: Cypher
 - SQL-basiert
 - Abfragen von Graphdaten
 - Muster und Relationen
 - Patterns and relationships

Neo4j

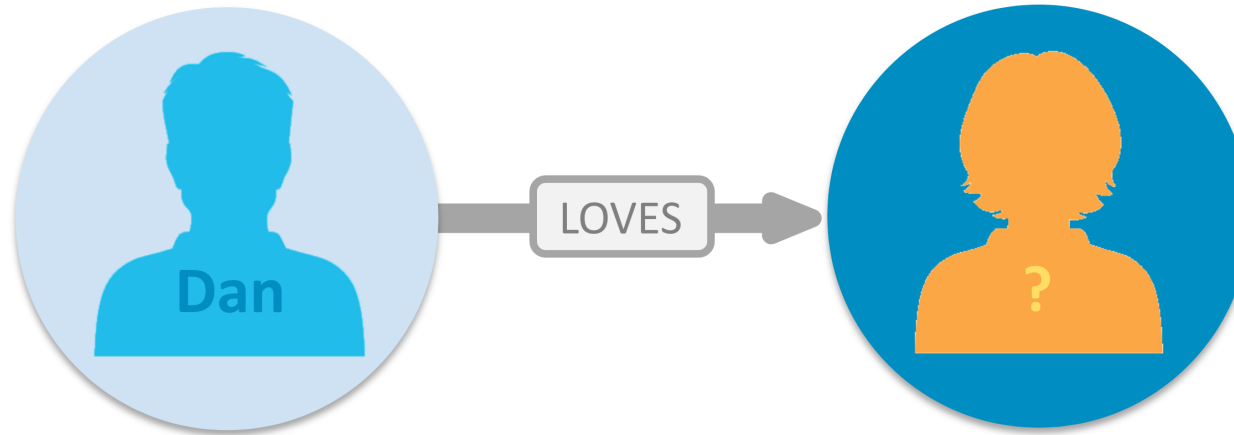
name: "Dan"
born: May 29, 1970
twitter: "@dan"

name: "Ann"
born: Dec 5, 1975



<https://neo4j.com/developer/graph-database/>

Neo4j und Cypher



NODE

Relationship

NODE

```
MATCH (:Person { name:"Dan" } ) -[:LOVES]-> ( whom ) RETURN whom
```

LABEL

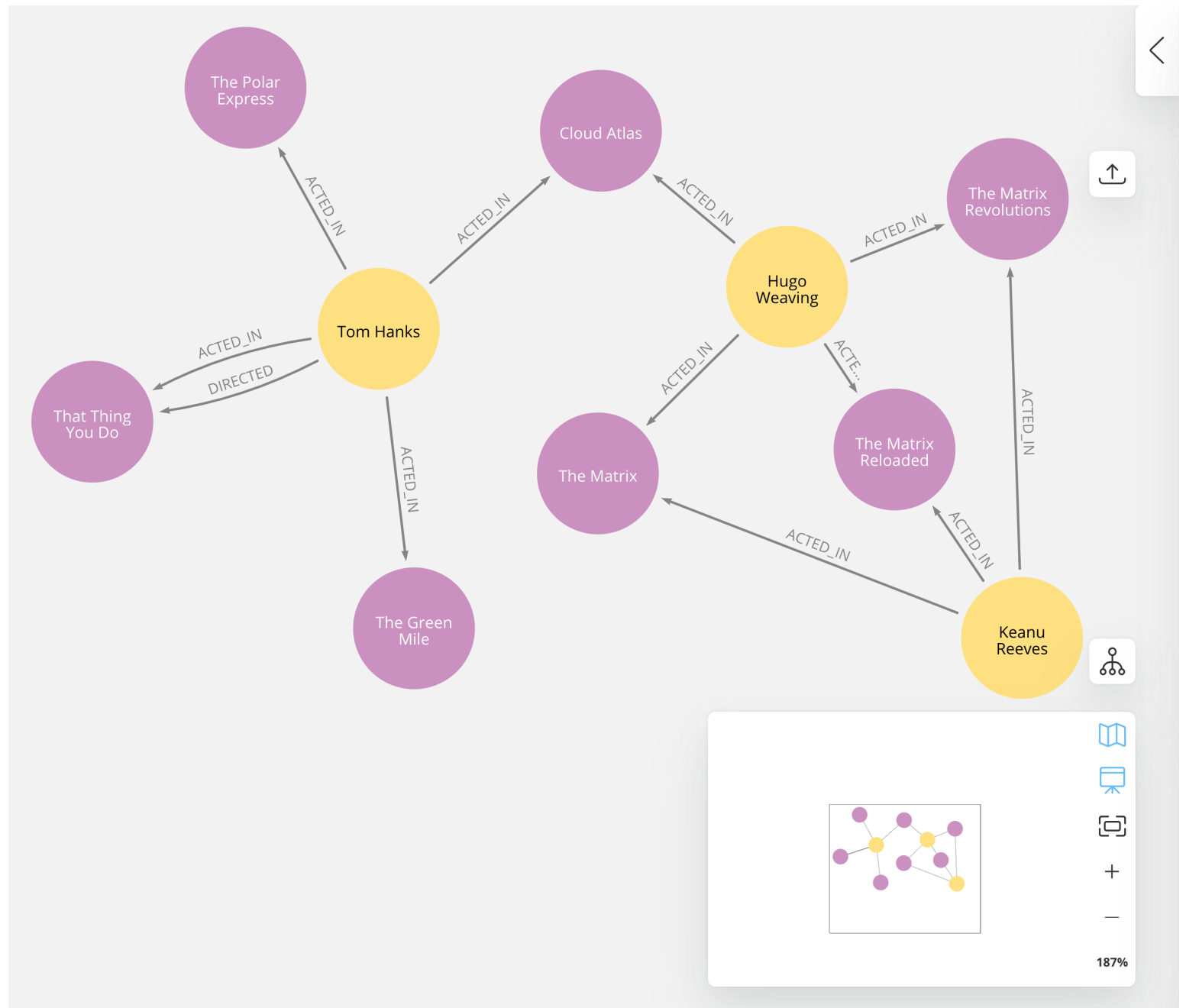
PROPERTY

VARIABLE

```
CREATE (ee:Person {name: "Emil", from: "Sweden"})
```

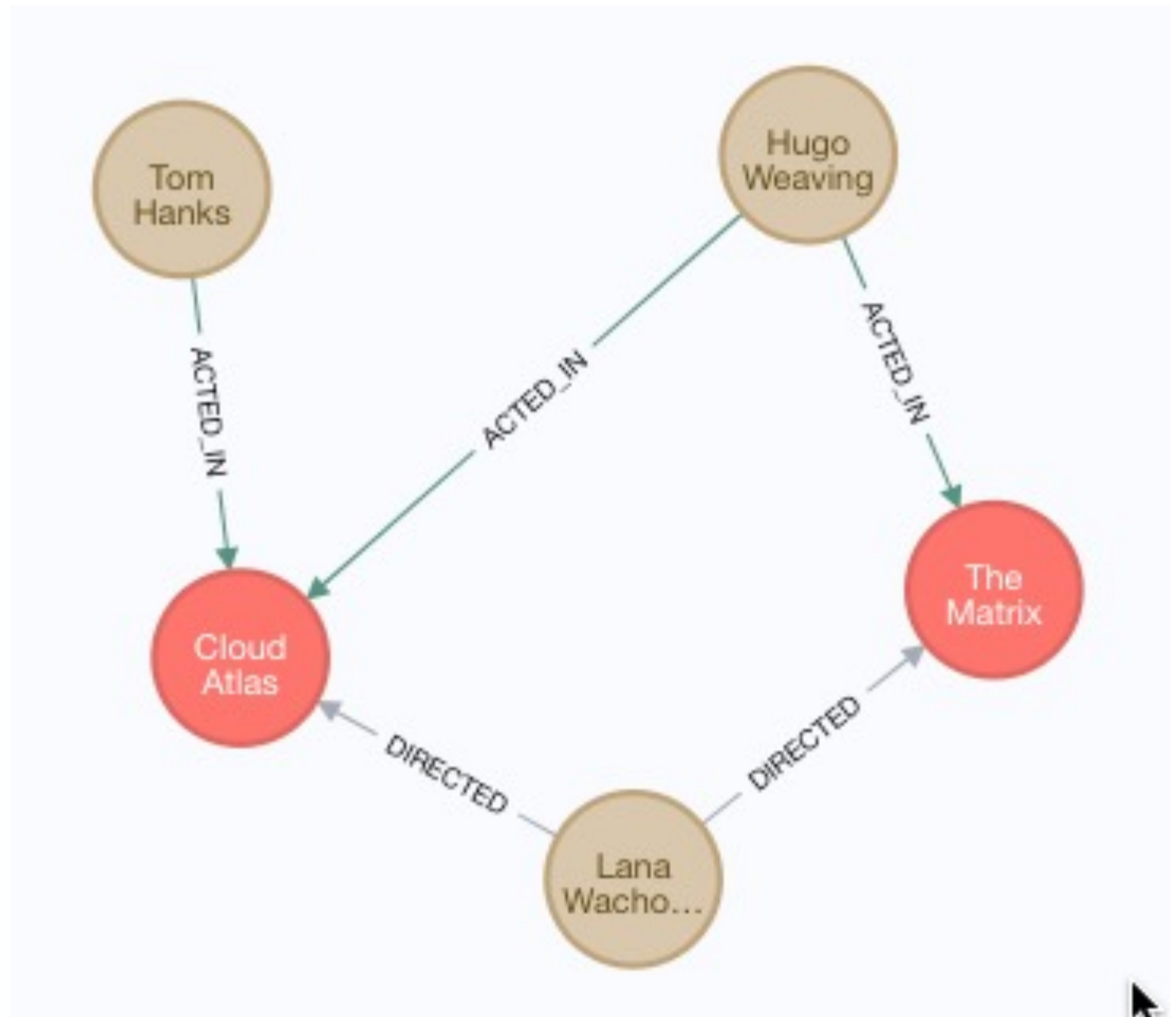
<https://neo4j.com/developer/graph-database/>

Neo4j



<https://neo4j.com/developer/graph-visualization/>

Neo4j



<https://neo4j.com/developer/graph-visualization/>

Über Neo4j

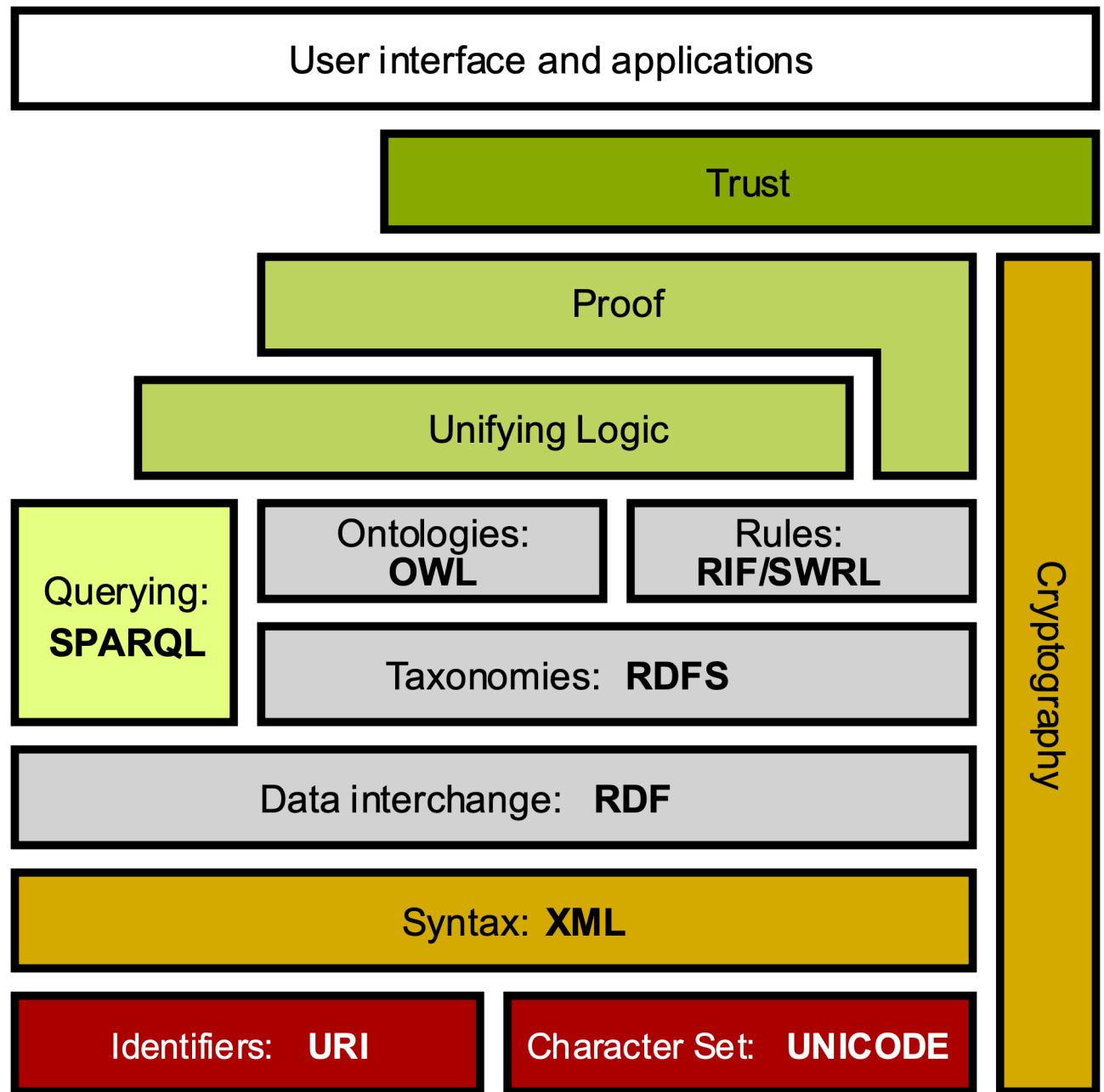
- <https://neo4j.com/>

Tutorials:

- <https://neo4j.com/developer/get-started/>

Semantic Web

Von Webseiten zu maschinenlesbaren Daten



<https://en.wikipedia.org/wiki/File:Semantic-web-stack.png>

Ontologien in der Philosophie

- Lehre des Seins
- Teil vom Metaphysik
- Fragen
 - wie Entitäten in grundlegende Kategorien eingeteilt werden
 - welche Entitäten auf der fundamentalsten Ebene existieren
- Bildung von Systeme von Kategorien
- Substanzen, Eigenschaften, Relationen, Sachverhalte, Ereignisse

Ontologien im Informatik

- Formalisierte geteilte Konzeptualisierungen

„An ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an ontology is a systematic account of Existence. For AI systems, what ‘exists’ is that which can be represented.”

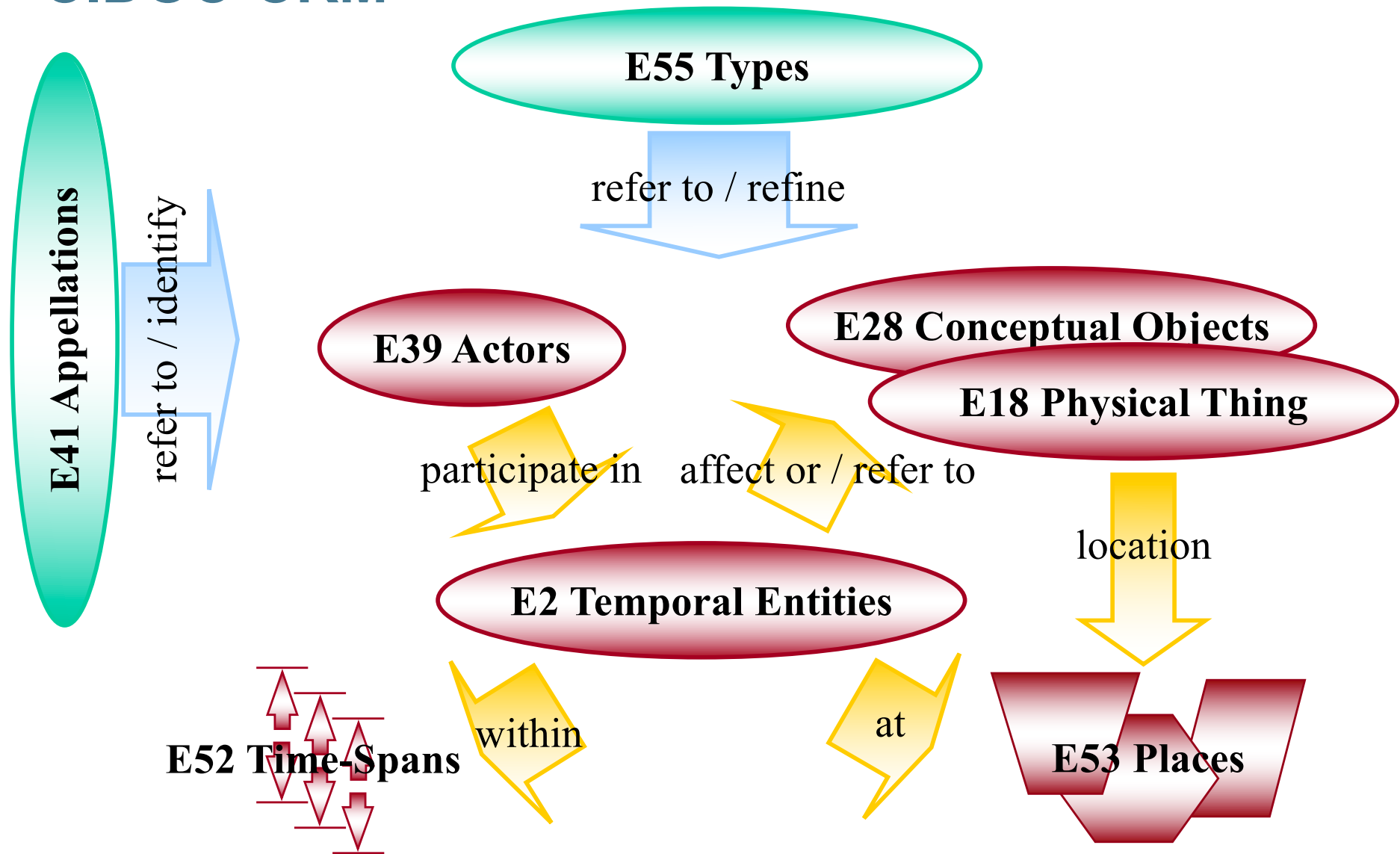
Gruber, T.R., 1993. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2), pp. 199–220.

- Eine Art von Datenmodell beschäftigt sich mit formalisierten Konzeptualisierungen statt mit Problemen mit der Implementation

Ontologie-Spezifikation

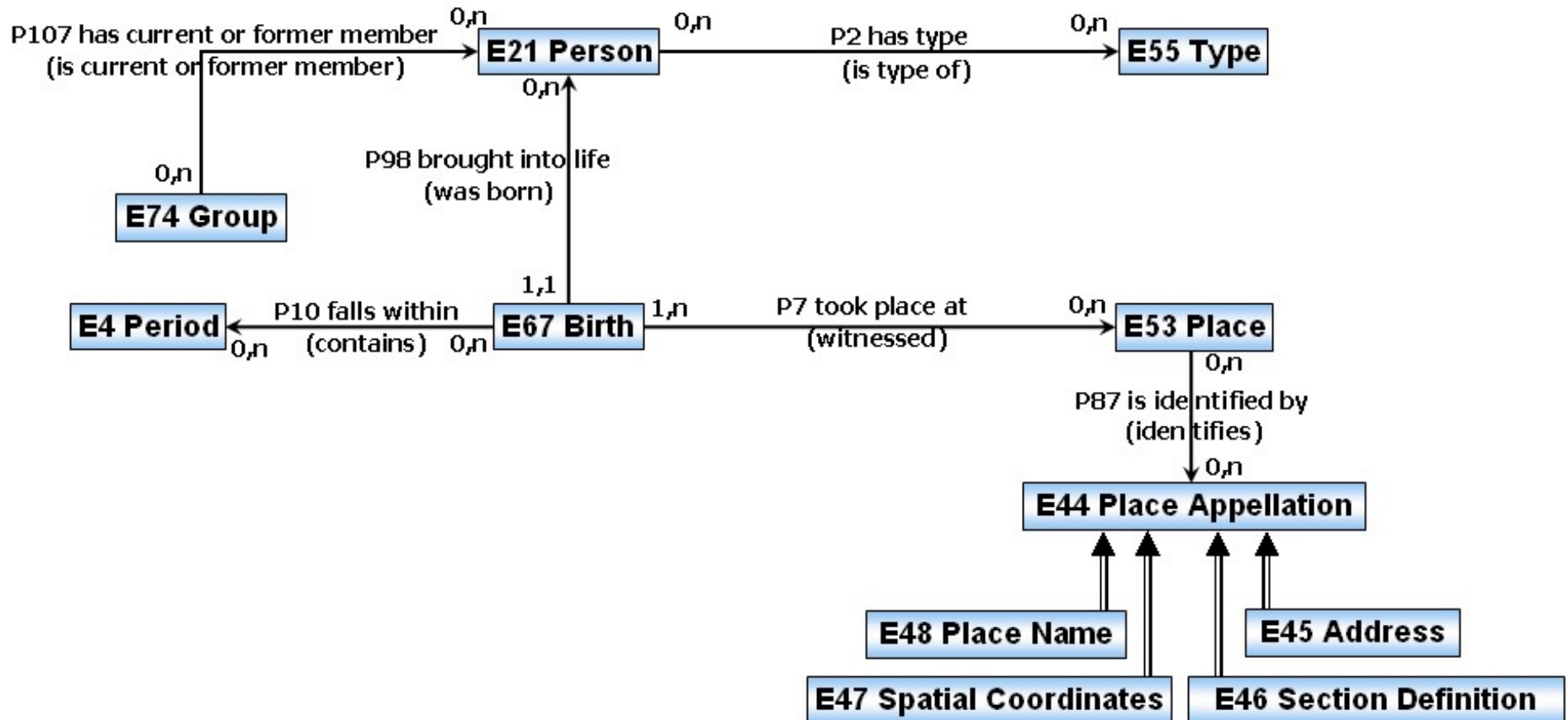
- Klassen und Eigenschaften
 - classes and properties
 - Repräsentiert Konzepte und ihre Verbindungen
- Konzepthierarchien
 - Allgemeinheit bzw. Spezialisierung
 - Superklassen bzw. Subklassen
 - Is-a-Relationen
- Instanzen
 - Objekte
- Relationen zwischen Klassen
 - und zwischen Objekte

CIDOC-CRM



CIDOC-CRM Beispiel

PERSON NATIONALITY INFORMATION



Ontologien bzw. Datenbankschema (SQL)

Datenbankschema:

- Fokus: Daten
- DB-Beschränkungen
 - für Integrität
 - auf Meinung anspielen
- Keine is-a-Hierarchien
- SQL
 - Abfragen
 - Datenintegrität

Ontologien:

- Fokus: Meinung
- Axiome der Ontologien
 - Meinungsspezifikation
 - Eventuell für Integrität
- Is-a ist Backbone
- Theorembeweiser (theorem provers)
 - Neue Informationen erschließen
 - Konsistenz sicherstellen

<https://pdfs.semanticscholar.org/b44f/a4592b69183c1965d0075dea1a3bc58dfbfe.pdf>

Protégé

- Editor zur Modellierung von Ontologien
 - Informatik
- Protégé-OWL
- WebProtégé bzw. Protégé Desktop

Über Protégé

- <https://protege.stanford.edu>

Tutorials

- <https://protege.stanford.edu/support.php>
- https://drive.google.com/file/d/1A3Y8T6nIfXQ_UQOpCAr_HFSCwpTqELeP/view

Aufgaben

- Zwei Alternativen
 - Minimum: eine von beiden Lösungen

Alternative 1:

- Entwickeln Sie eine Version der Tier-Datenbank in Neo4j
 - wenn erwünscht können Sie eine andere Graph-Datenbank verwenden
- Ein paar Daten sollten im System sein

Alternativ 2:

- Entwickeln Sie eine Tier-Ontologie in Protégé, mit
 - Is-a-Relationen
 - Klassen und Eigenschaften (classes and properties)