

# Mapping provenance in ontologies

František Jahoda

FNSPE

Czech Technical University

Institute of Computer Science  
Academy of Sciences of the Czech Republic

# What is Provenance?

- from the French provenir, "to come from"
- confirm the time, place, and if appropriate the person responsible, for the creation, production or discovery of the object.
- artefact's complete documented history
- no clear semantics...

# Provenance

- Is used in
  - Archaeology, Paleontology
  - Arts, Archives, Science
- Is used for
  - to discover a forgery, looted, reproduction
  - to prove ownership
  - can set the price of antiques
  - to prevent evidence tampering - Chain of custody
  - To prove ecology behaviour (wood, endangered species)

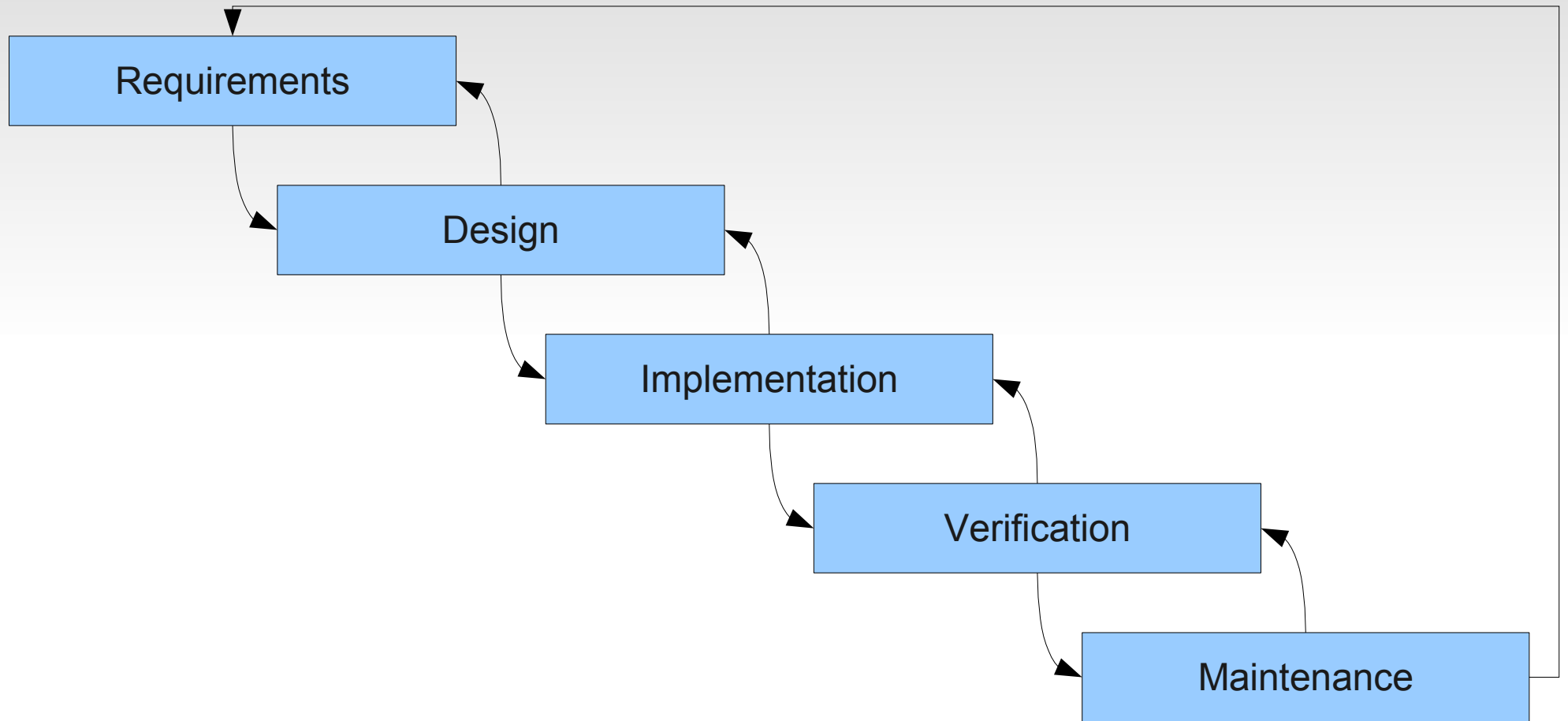
# Data provenance

- Is used for
  - Managing trustworthiness of data sources
  - Provide provability of conclusions
  - Recording history (not only actual snapshot)
- Enables learning from history, sharing data (data rights management)
- To prevent conflicts between data from two different sources

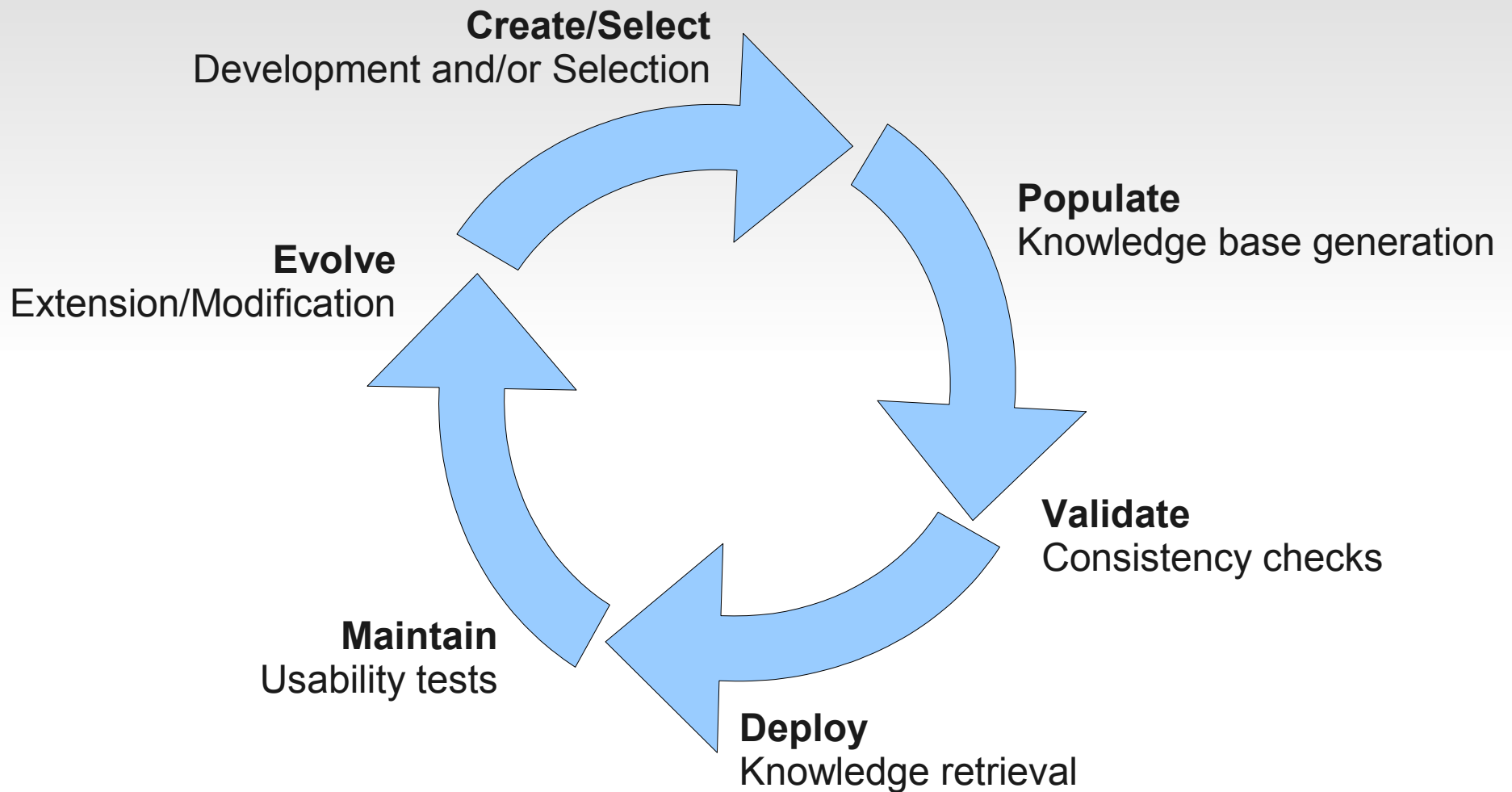
# Ontology provenance

- Ontologies have different life cycle
- Less control over ontology sources
- Ontologies are more rigid than data
- Part of the ontology creates bridge between abstract model and concrete data
- Logical nature of ontologies have to be considered
  - Derived axioms
  - Change of one axiom can change meaning of others

# Software project life cycle (Waterfall)



# Ontology life cycle



# Changes in ontology

- Understanding of subject area is developing
  - e.g. three fish species can be only one species in different age
- Input data have changed meaning
- An error in ontology was found
- New requirements / extending ontology
- Speed up



# How to implement provenance?

- Encoding information about ontology
  - Information about ontology should have own ontology. (answering queries)
- Example of provenance data model
- How to encode derived axioms?

# Encoding information about ontology in OWL DL

- Statements about axioms, higher order statements are not possible
- Annotations compatible with OWL 1.0
  - Reifying axioms
  - Annotation properties
- Annotations using axioms URI

# Axioms reifying

- Axiom in original ontology
  - $\text{PERSON} \subseteq \text{MORTAL}$
- Using the meta-ontology, we can represent this ontology as follows:
  - `CLASS(Mortal)`
  - `CLASS(Person)`
  - `SUBCLASS_OF_AXIOM(axiom1)`
  - `SUBCLASS_OF_SUPERCLASS (axiom1, Mortal)`
  - `SUBCLASS_OF_SUBCLASS (axiom1, Person)`
  - `SUBCLASS_OF (Person, Mortal)`

# Axioms reifying

- Adding annotations
  - CREATOR (axiom1, Aristotle)
  - CONFIDENCE (axiom1, 0.95)
- But meta ontology is considerably bigger..

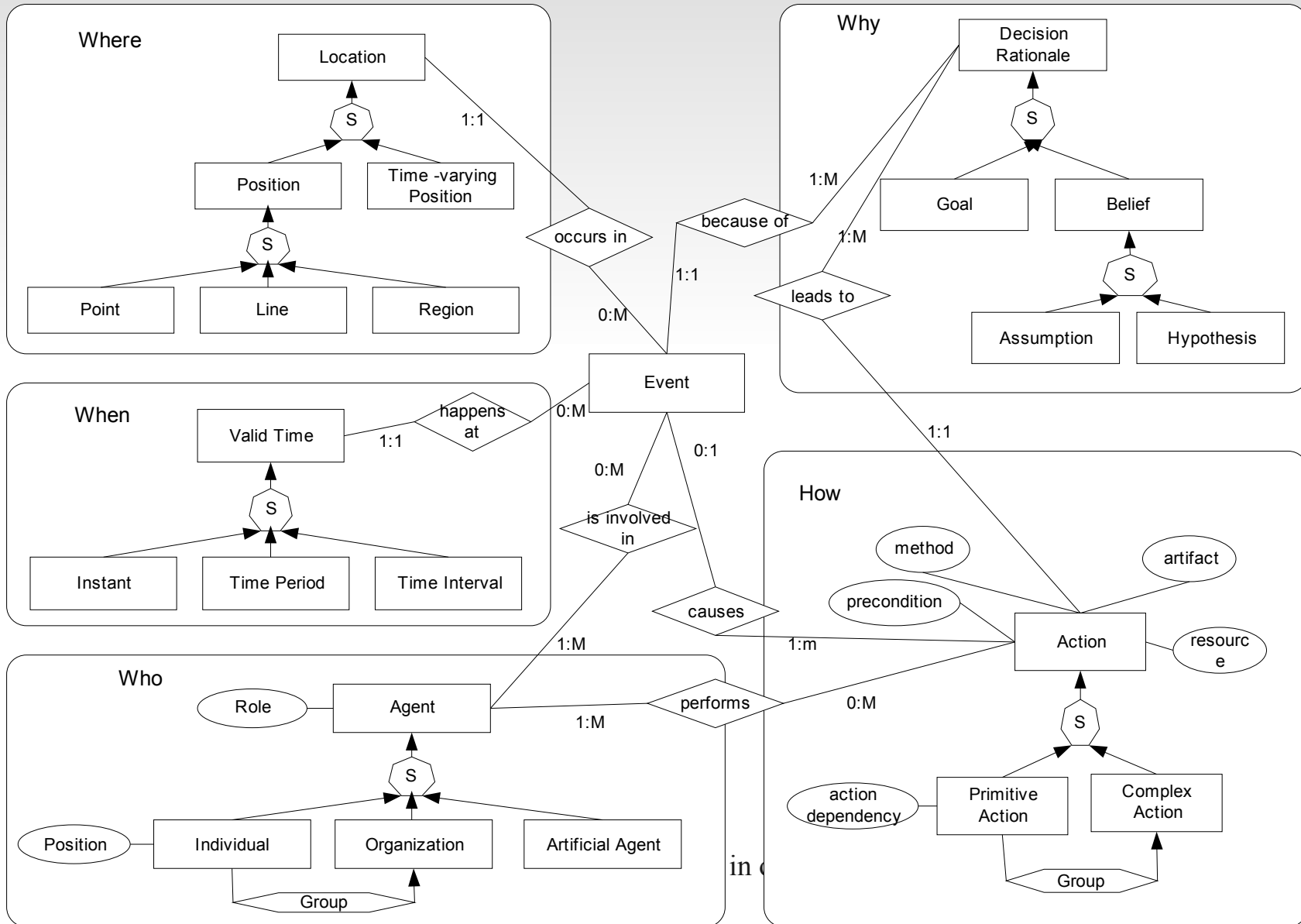
# Annotation properties

- Using Annotation properties (same ontology)
- Example
  - $\text{PERSON} \subseteq \text{MORTAL}$  (axiom1)
  - *CREATOR* (axiom1, Aristotle)
  - *CONFIDENCE* (axiom1, 0.95)
- But, provenance is treated outside the OWL logical semantics

# Annotation with meta-ontology

- Using Meta-Ontology (separate ontology)
- Extension in OWL 2.0 draft
- Example (in original ontology)
  - $\text{PERSON} \subseteq \text{MORTAL}$  (axiom1)
- In meta-ontology
  - $\text{AXIOM}(\text{axiom1})$
  - $\text{CREATOR}(\text{axiom1}, \text{Aristotle})$
  - $\text{CONFIDENCE}(\text{axiom1}, 0.95)$

# Example of provenance ontology



# Example of provenance ontology

- Representation of history as a sequence of events which cause changes of state.
- Events have properties:
  - When – event time (point, interval)
    - Temporal granularity (day, business day, hour)
  - Where – place of event (point, line, region)
    - Position can change with time



# Example of provenance ontology

- How – actions leading to the event
  - Primitive actions, complex actions
  - Preconditions, Resources
  - Methods – what was done, capture action's parameters
  - Artefacts – data manipulated by the action (input/output)
- Who – refers to people and and organisation involved
  - Agents – individual, organization, software agent
  - Role – functional responsibility (e.g. in organization)
- Why – rationale of an action
  - Goals, Beliefs – Assumptions, Hypothesis

# How to encode derived axioms?

- Each inferred axiom  $\alpha$  is based on a minimal set of base axioms. Let is union of all axioms in such sets called  $O_{AJ(\alpha)}$
- Each base axiom is connected in meta-ontology with provenance atoms (describing provenance information, N:M binding).
- Therefore, axiom  $\alpha$  should be connected with all provenance atoms related to  $O_{AJ(\alpha)}$

# References

- Denny Vrandečić , Johanna Völker, Peter Haase, Thanh Tran Duc, Philipp Cimiano: A Metamodel for Annotations of Ontology Elements in OWL DL
- Dr. Sudha Ram: the Active Conceptual Modeling of Learning Workshop, Space and Naval Warfare Systems Center , San Diego, May 9-12, 2006
- Paul Buitelaar: eLearning4NLP Workshop Prague, June 2007
- Miroslav Vacura, Vojtěch Svátek: Pattern-Based Representation and Propagation of Provenance Metadata in Ontologies, Poster Proceedings EKAW2008