

Übung Netzbasierte Informationssysteme

Termin 9: Ontology Engineering

Prof. Dr. Adrian Paschke

Arbeitsgruppe Corporate Semantic Web (AG-CSW)
Institut für Informatik, Freie Universität Berlin
paschke@inf.fu-berlin.de
<http://www.inf.fu-berlin.de/groups/ag-csw/>



- **Lernziele des 8. Übungsblattes waren:**
 - Grundlage Text Mining und Web Mining Technologien
- **Wie war`s?**
 - Fehlerquellen
 - Probleme?

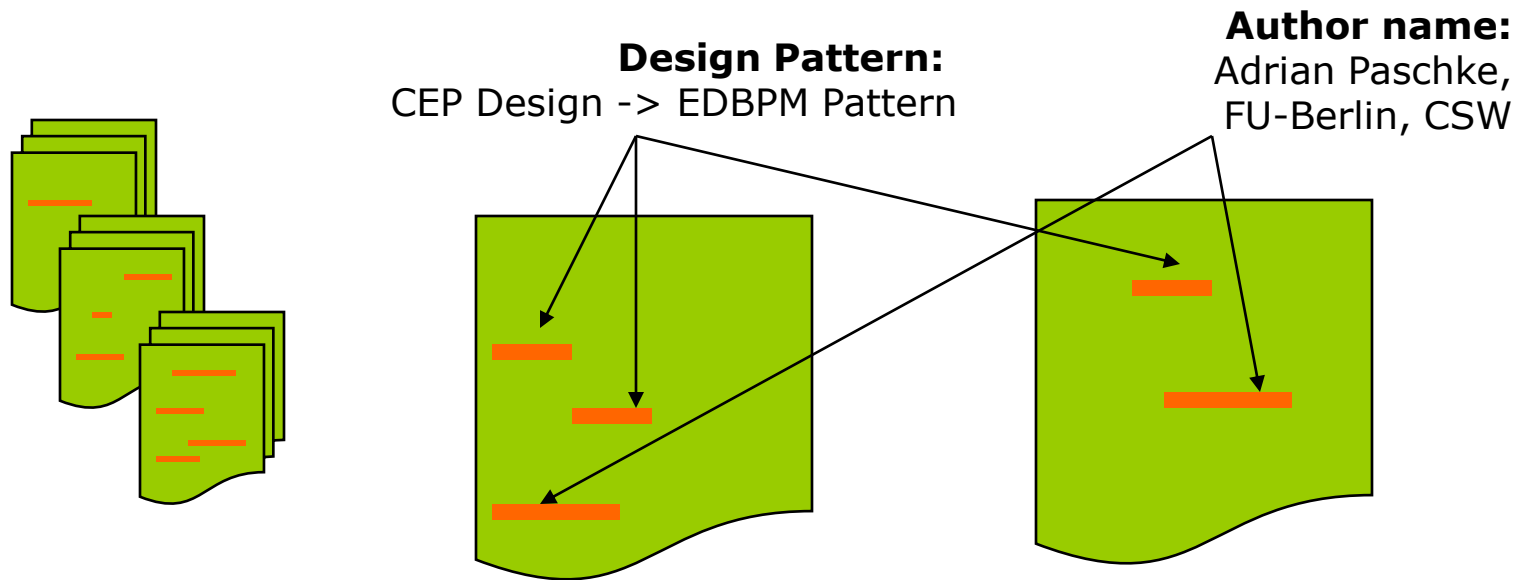
- Um Domänenannahmen explizit zu machen
 - Leichter die Annahmen in einer Domäne zu ändern
 - Leichter bestehende Daten zu verstehen und zu ändern

"liver" as an term of human anatomy
"keyboard" as term for computer equipment
Enables data integration at different granularities

- Um Domänenwissen von operationalem Wissen zu trennen (seperation of concerns)
 - Getrennte Wiederverwendung von Domänen- und Anwendungswissen

domain: *"liver" is_a "organ"*
domain: *"caspase 8" is_a "protein"*
domain: *"caspase 8" involved_in "programmed cell death"*
operational: *"proteins" have sequence data*
operational: *"protein sequences" are input for EBI-BLAST*

- Eine Community Referenz für Anwendungen
- Um ein konistentes Verständnis darüber was Information bedeutet zu teilen



Design Prinzipien für Ontologien

Ontologies should be an agreement with the truths of basic science

“caspase-8” is involved in “apoptosis”

“menopause” is part of “death”

[Example from the Gene Ontology 2005]

Use Aristotelean definitions

An A is a B which C's

- A researcher is a person who systematically investigates and studies materials and sources to establish facts and reach conclusions.
[www.spaceforspecies.ca/glossary/r.htm]



Ontologies are made to capture domain knowledge

e.g. anatomy (FMA), chemical compounds (CHEBI), etc.



Avoid modelling negative knowledge

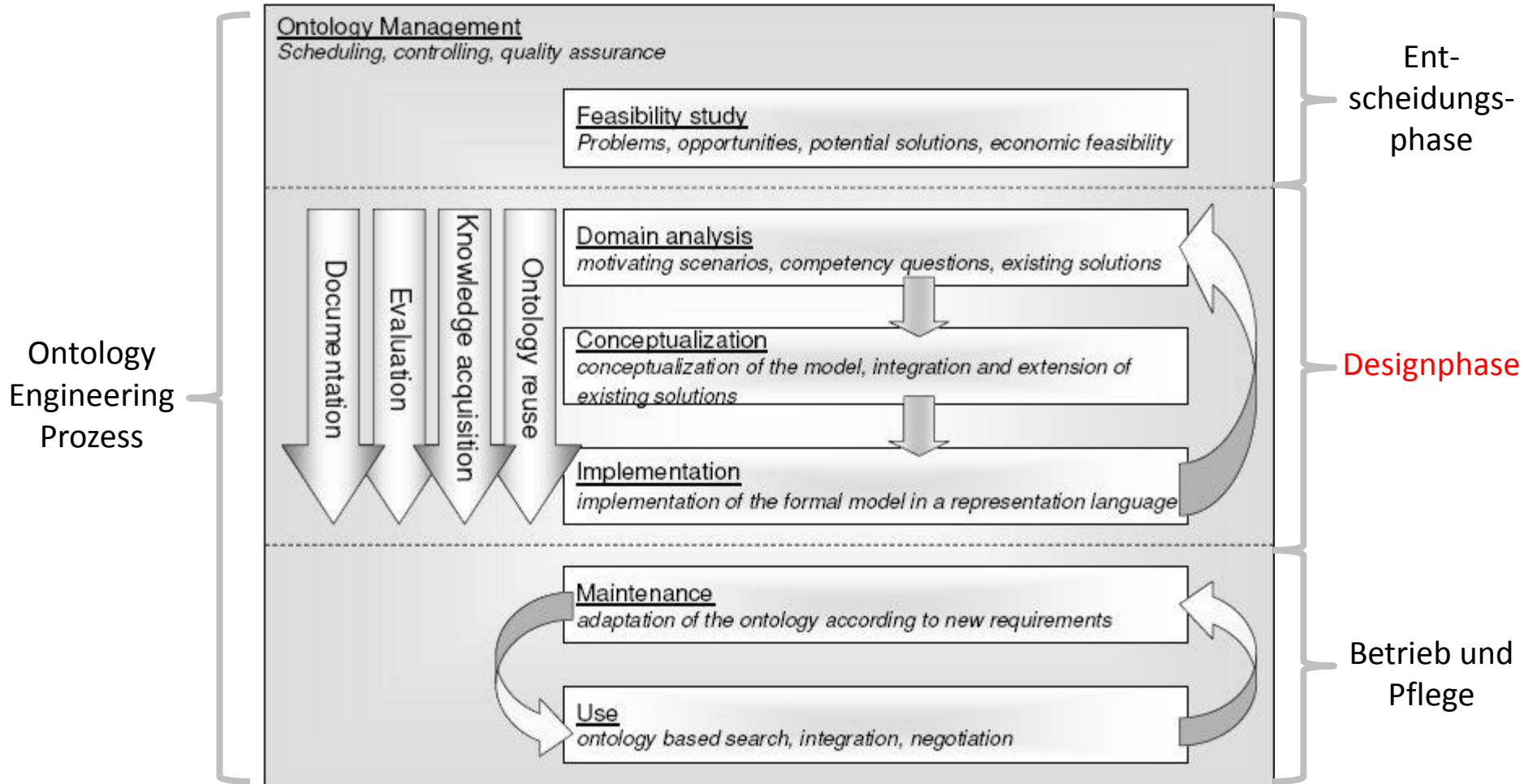
Ontology terms should be used in literature
(in written language)

Otherwise they are hardly reusable!!

- Zu Beginn individuelle „Kunst“ einzelner Experten
 - kannten die Theorie & Standards
 - konnten mit Werkzeugen umgehen
 - Aber: Wissen ist nur repräsentativ, wenn es ein Konsens vieler Personen ist
 - Teilnehmer bei Ontologie-Entwicklung haben unterschiedlich ausgeprägtes Wissen
 - der Ontologie-Entwickler ist in den seltensten Fällen auch Domänen-Experte
 - der Prozess ist kollaborativ, verteilt und asynchron
- grundsätzliche Prinzipien vom Software Engineering adoptiert
- Neueste Ansätze definieren Übersetzungsleitfäden (z.B. spezielles OE Glossar)

Ontology Engineering beschreibt den Ontologie-Entwicklungsprozess, den Ontologie-Lebenszyklus, die Methoden und die Methodiken um Ontologien zu erzeugen sowie die Werkzeuge und Sprachen, die dieses unterstützen.

Quelle: Asunción Gómez-Pérez, Mariano Fernández-López, Oscar Corcho (2004).
Ontological Engineering: With Examples from the Areas of Knowledge Management, E-commerce and the Semantic Web. Springer, 2004.



- Anforderungsanalyse
 - funktionale Anforderungen, Performanz-Anforderungen, Entwurfs-Anforderungen (z.B. mögliche Sprachen...) etc.
- Kompetenzfragen (Competency Questions)
 - Liste von Fragen, die eine Knowledge Base beantworten können muss
 - *Kann ein Student v zu Zeitpunkt w in Raum x sein, in dem Kurs y gehalten wird, für den Kurs z Voraussetzung ist?*
- Konzeptualisierung
 - Identifikation von Kerntermen aus den CQ
 - Festlegung als Konzepte in Form von ersten Ontologieklassen

- Funktionale Anforderungen für die Anwendung
- Nicht-funktionale Eigenschaften für die Entwicklung
 - individualizability/customizability
 - ability to make adaptations and differentiations of ontology
 - composability
 - ability to compose, import and integrate ontology hierarchies which are composed from modules and further submodules
 - interoperability
 - ability to cooperate with existing external tools, data sources and functionalities
 - declarative implementability, modifyability and evolvability
 - ability to declaratively implement, add or modify (unspecified) future functionality
 - reusability and interchangeability
 - ability to interchange and (re)-use ontologymodules in different target environments and by different entities (collaborative humans, agents, services)

- Nicht-funktionale Eigenschaften zur Laufzeit
 - active adaptability and user-defined configurability
 - ability to adapt to changed situations
 - usability
 - ability for the users (humans or machines) to easily use the ontology
 - understandability and explanation
 - ability of the users (humans or machines) to easily understand produced results from the ontology
 - efficiency
 - adequated performance
 - scalability
 - ability of the ontology system to grow in size
 - correctness
 - absolute ability to derive the correct conclusions
 - robustness
 - "reasonable" behavior in unforeseen circumstances and incomplete knowledge
 - safety
 - fault and conflict tolerance, information hiding (need-to-know principle) etc.
 - verifiability and validation
 - ability to analyze the correctness

Mindmap Definition

 Was ist Mindmapping?

eine Methode aus der Gehirnforschung heraus entwickelt

berücksichtigt die Arbeitsweise des Gehirns...

zwei Gehirnhälften

nicht lineares sondern vernetztes Denken

Ganzheitlichkeit

Erfassung Strukturen

Erstellung von "Gedanken-Landkarten" bzw. "Wissenslandkarten"

einfache Regeln, sofort anwendbar

as keyword

related concepts

a central concept

defined as:

Mindmap

useful for



communicating

planning



collaborating



studying



Definition

„Ein Entwurfsmuster beschreibt ein bestimmtes, in einem gegebenen Kontext immer wiederkehrendes Entwurfsproblem, sowie ein vorgegebenes Schema zu seiner Lösung“
(E. Gamma)

„A pattern is the abstraction from a concrete form which keeps recurring in specific non-arbitrary contexts.“

(D. Riehle and H. Zullighoven: *“Understanding and Using Patterns in Software Development”*)

“A design pattern addresses a recurring design problem that arises in specific design situations and presents a solution to it”

(Buschmann, *et. al.* 1996)

- Beschreiben eine vorgeschlagene Lösung eines wiederkehrenden Entwurfsproblems
- Verbreitetes Mittel um Wissen über erfolgreiche Designs explizit zu erfassen und zu transferieren.
- Semi-formalisierte Spezifikation von Lösungen zu bestimmten Problemklassen

Pattern

Name

A name used for identification

Problem

A repeating problem that occurs in a domain

Solution

Best practice solution to that problem

Consequences

Advantages and disadvantages of the recommended solution

Examples

A few examples where the recommended solution has already been applied

Anti-Pattern

Name

A succinct name to convey the essence of the anti-pattern

Problem / Bad solution

The commonly occurring mistake or bad solution that relates to the anti-pattern

Symptoms

The indications or signs of the problem

Consequences

The results of applying this anti-pattern

Root cause

This provides the context for the anti-pattern, that is, where a pattern was applied incorrectly and resulting in a problem or failed solution

Suggested solution(s)

Refactored solution that solves the problem and ensures more benefits

Adapted from Gamma, E., et al., *Design Patterns - Elements of Reusable Software*. 1995: Addison-Wesley.

Name**Intent**

Describes the underlying problem and the basic principle of the design solution.

Also Known As

Other names of the pattern, if existing.

Motivation

Depicts a scenario of a concrete design problem and gives a solution.

Applicability

Itemizes the situations or conditions, in which the pattern can be applied.

- Characteristics
- Resources
- Other environmental factors

Structure**Participants**

Lists the classes and objects, which are described by the design pattern.

Consequences

Discusses the advantages and disadvantages of the pattern, and makes suggestions about variations.

Variants

Brief description of variants, extensions or specializations of the design pattern.

Known Uses

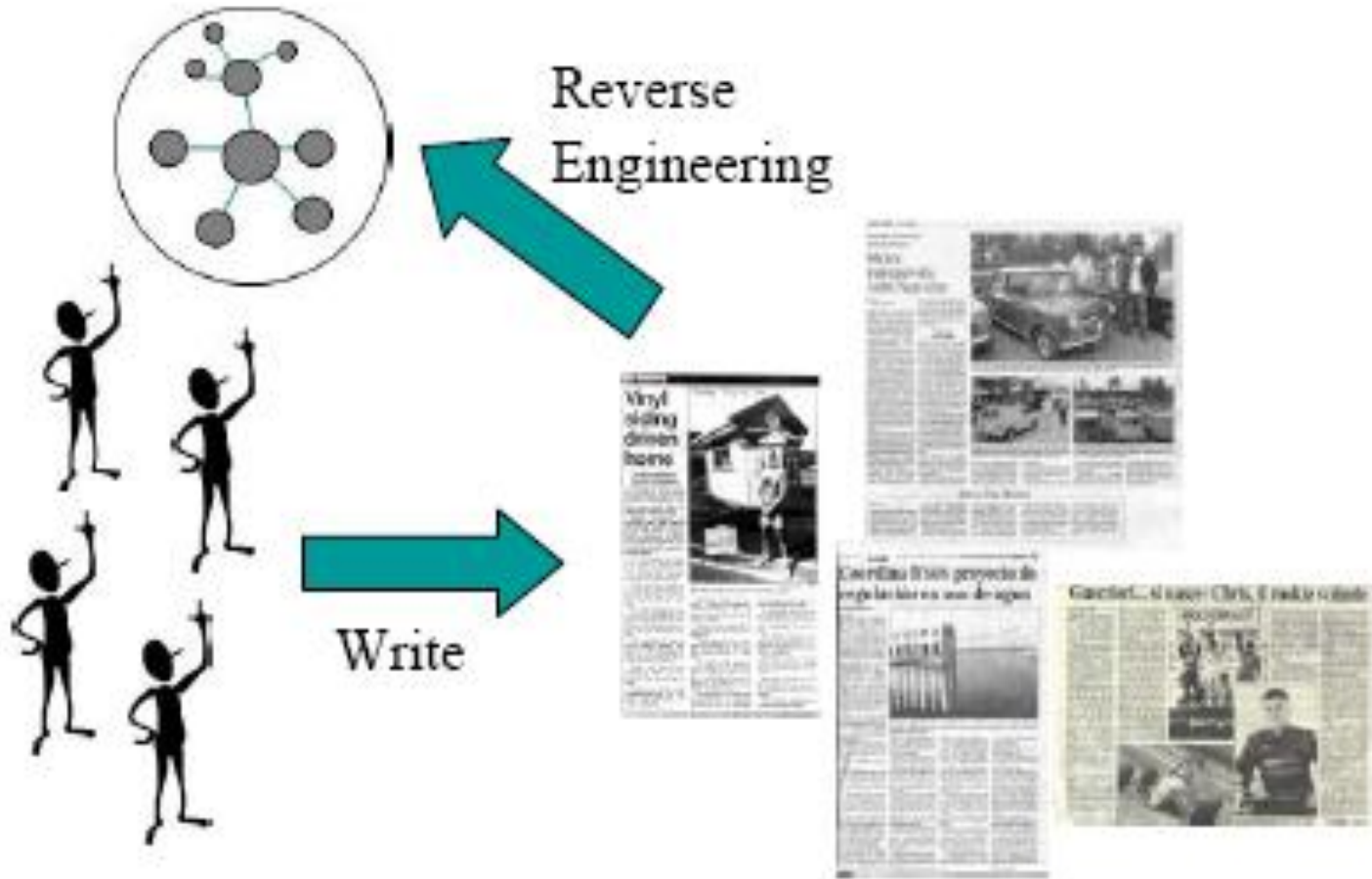
Examples of applications, in which the design pattern is used.

Related Patterns

Lists patterns, which describe a similar task, and shows the relations to other patterns.

Ontology Learning Vision : Erkennung des "Weltmodells"

Shared World Model



- **Termine**

- Ausgabe: 27.1.2009 Abgabe bis: 3.2.2009, 16:00 Uhr

- Fragen?

Viel Spaß und Erfolg !

