

Wirtschaftsinformatik II – Meilicke/Stuckenschmidt

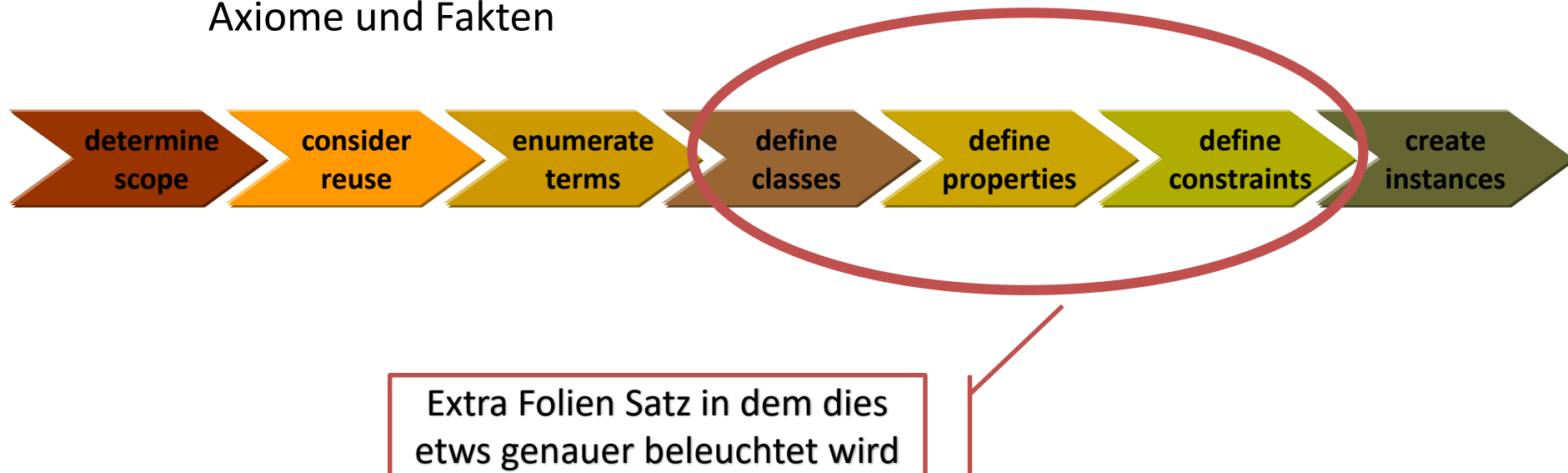
Wie man Ontologien entwickelt  
Der Ablauf des Gesamtprozesses

# ONTOLOGIEN

## ENTWICKLUNGSPROZESS

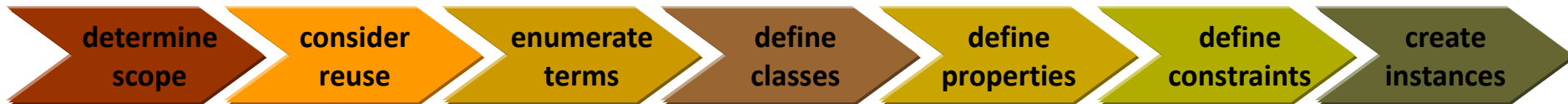
# Ontology Engineering

- Die systematische Erstellung einer Ontologie für eine bestimmte Domäne
  - Vom Abgrenzen der Domäne bis zum konkreten Modellieren der Axiome und Fakten

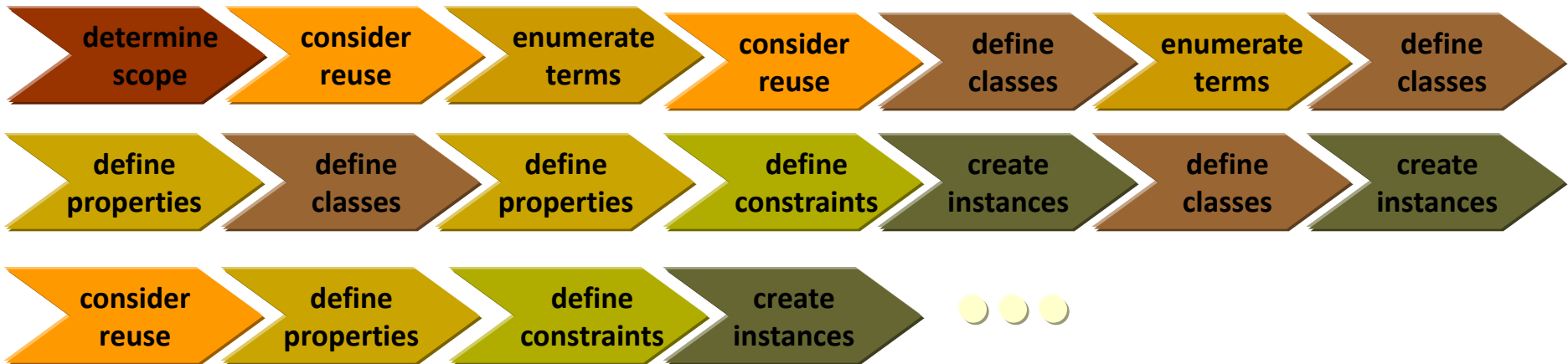


# Theorie und Praxis

In der Theorie:



In der Praxis:



# Domäne festlegen



- Welches Anwendungsgebiet soll abgedeckt werden ?
  - Meist wichtiger: Welche Gebiete sollen NICHT abgedeckt werden
- Für welchen Zweck wird die Ontologie entwickelt?
  - Web-Interface für jederman vs. Unternehmensinterne Anwendung
  - Darstellung statischen Wissens vs. Wissensbasis kollaborativ
- Insbesondere:
  - Welche Fragen über das Anwendungsgebiet soll die Ontologie beantworten?

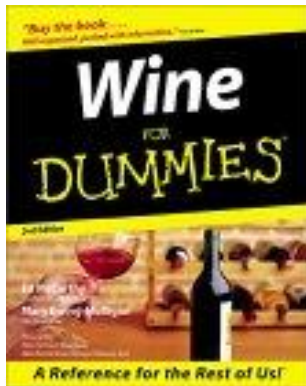
# Beispiel



**Ontologie über  
Wein und Essen**

**Französische  
Weine und  
Regionen**

**Kalifornische  
Weine und  
Regionen**



# Zweck der Ontologie

- Webseite auf der Nutzer sich über Weine informieren können
  - Kategorien von Weinen sollen vorgegeben sein
  - Nutzer stellt Weine ein und kann Weine bewerten
  - Weinhersteller stellt Wein ein
  - Weinhersteller können Werbung schalten
  - Nutzer können Speisen empfehlen, die zu den Weinen passen
  - Nutzer kann sich informieren, welche Speisen zu welchem Wein passen
  - Nutzer kann Wein online kaufen über ein Interface zu einem Drittanbieter

# Kompetenzfragen

- Zum Beispiel
  - Welche Charakteristika sollten bei der Auswahl eines Tischweins berücksichtigt werden?
  - Ist Chianti ein Rotwein oder ein Weisswein?
  - Passt ein Cabernet Sauvignon zu Meeresfrüchten ?
  - Welche Art von Wein passt am besten zu dunklem Fleisch?
  - Ändern sich die Eigenschaften eines Weines mit dem Jahrgang?
  - Welches sind gute Jahrgänge für einen Napa Zinfandel?
  - ...

# Wiederverwendung



- Warum sollte man andere Ontologien wiederverwenden?
  - Aufwand sparen
  - Einheitliche Sicht auf eine Anwendungsdomäne
  - Fehler vermeiden
  - Eigenes Expertenwissen fehlt (Beispiel: Dozent und Weindomäne)
  - Von etablierten und erfolgreich verwendeten Ontologien profitieren



# Wo gibts Ontologien zum Wiederverwenden?

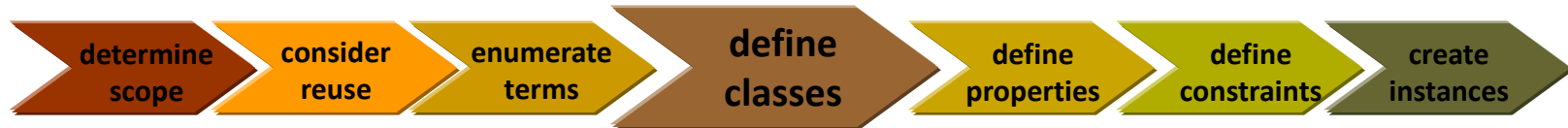
- Ontologie-Bibliotheken, z.B.
  - DAML ontology library (<http://www.daml.org/ontologies>)
  - Protégé ontology library (<http://protege.stanford.edu/plugins.html>)
- Suchmaschinen für Ontologien
  - Swoogle (<http://swoogle.umbc.edu/>)
  - Watson (<http://watson.kmi.open.ac.uk/>)
- Schemata und Klassifikationen
  - Insbesondere: RDF Schemas
  - Für Produkte gibt es verschiedenste Klassifikationen, z.b: eCl@ss

# Enumerate Terms



- Was sind wichtige Begriffe im Anwendungsgebiet?
  - Begriffe aus Competency Questions
  - Begriffe aus Texten über die Domäne
  - Begriffe aus Use Cases
  - ...

# Konzepte definieren



- Welche Begriffe bezeichnen Konzepte?  
(Mengen von Objekten mit gleichen Eigenschaften)
  - Die Klasse aller Weine
  - Die Klasse aller Weingüter
  - Die Klasse aller Rotweine
- Konzepte werden in einer Begriffshierarchie organisiert
- Kriterium: Teilmengenbeziehung zwischen Objekten
  - Jedes Objekt des Unterkonzeptes ist auch ein Objekt des Oberkonzeptes

# Beispiel aus der Weindomäne

- Positive Beispiele
  - Apple ist eine Unterklasse von Fruit  
*Every apple is a fruit*
  - Red wines ist eine Unterklasse von Wine  
*Every red wine is a wine*
  - Chianti wine ist eine Unterklasse von Red wine  
*Every Chianti wine is a red wine*
- Negative Beispiele
  - Bordeaux ist keine Unterklasse von France
  - Winery ist keine Unterklasse von Wine
  - ...

# Rollen und Attribute definieren



- Rollen (und Attribute) beschreiben Eigenschaften, die alle Objekte einer Klasse besitzen.
  - Eigenschaften, die alle gemeinsam haben
  - Eigenschaften, die Objekte unterscheiden
- Welche Begriffe in der Domäne beschreiben diese Eigenschaften ?

*Each wine will have color, sugar content, producer, etc.*

# Rollen vs. Attribute

- Rollen bezeichnen Eigenschaften, die sich als Verbindung zwischen Instanzen darstellen lassen
  - ... hasColor ...
  - Wenn man Wine und Color als Konzepte eingeführt hat, dann wäre hasColor eine Rolle
- Attribute bezeichnen Eigenschaften, die sich als Verbindung zwischen einer Instanz und einem Datenwert darstellen lassen
  - ... hasColor ...
  - Wenn man Color nicht als Konzept eingeführt hat, sondern mögliche Werte “red” und “white” usw. sind, dann handelt es sich um ein Attribut

# Rollen/Attribute/Klassen

- Ist es besser eine Eigenschaft als Attribut, als Rolle oder als Klasse zu modellieren?
  - *Wine(wine182), hasColor(wine182, "red")*
  - *Wine(wine182), hasColor(wine182, red), Color(red)*
  - *RedWine(wine182)*
- Welche Möglichkeit ist die richtige?
  - Keine eindeutige Antwort möglich
  - Die erste Möglichkeit ist die schlechteste, da man klar zwischen *white*, *rose*, und *red* unterscheiden kann
    - Wenige, endlich viele Werte => besser keine Repräsentation als String

# Rollen/Attribute/Klassen: Alternative

$RedColor \sqsubseteq Color$   
 $WhiteColor \sqsubseteq Color$   
 $RoseColor \sqsubseteq Color$

$RedColor(darkred)$   
 $RedColor(cinnabar)$   
 $RedColor(purple)$

...

$RoseColor(lightred)$

...

$hasColor(wine182, darkred)$   
 $RedWine \equiv \exists hasColor. RedColor$

**$O \models Redwine(wine182)$**

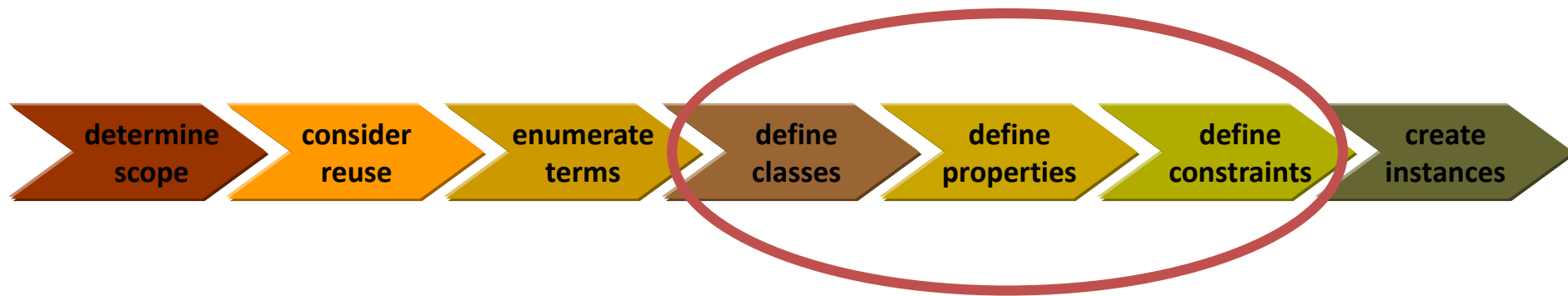


# Instanzen

- Die ABox zur TBox kann auf zwei Arten entstehen
  - Wird im Prozess des Ontology Engineering mit erstellt
    - Sollte dann auch in den Kompetenzfragen auftauchen
  - Wird durch die Nutzung des Informationssystems erstellt
- Denkbar sind auch Mischformen
  - *darkred, red, rose* werden vorgegeben
  - Konkrete Weine werden von nutzenden Weingütern eingestellt
    - Dabei werden auch Assertions wie *hasColor(weine117, red)* mit in die ABox aufgenommen
- Ob eine Instanz oder ein Konzept zum selben Zweck verwendet wird, kann von der Modellierung abhängen
  - Beispiel: *red* als Instanz vs. *RedWine* als Klasse

# Zusammenfassung Foliensatz

- Entwicklungsmodell für Ontologien



- Bereits im Detail betrachtet:
  - Wie verwende ich hierfür Beschreibungslogik
  - Wie verwende ich hierfür Protege

# Fazit: Erster Block

- Grundlagen der Modellierung
  - Statt UML- und ER- Diagrammen logische Formeln zur Beschreibung von Modellen
  - Andere (graphische) Modellsprachen sind in der Regel einfacher und können auf Logik zurückgeführt werden
- Modelltheoretische Fundierung vs. konkrete Anwendung zu Modellierungszwecken
- Aussagenlogik, Prädikatenlogik, Beschreibungslogik