

# Ontologien in der klinischen Forschung

Frank Loebe, Ronald Speer  
IMISE, Universität Leipzig

TMF Workshop „Terminologien und Ontologien in der  
klinischen Forschung und Versorgung“  
29.01.2007



# Hintergrund

- Zunehmende Vernetzung der medizinischen Forschung
- Ziele:
  - Expertenwissen bündeln
  - Behandlungsqualität verbessern
  - Kosten senken
  - Transfer von Forschungsergebnissen in die Versorgung erleichtern
- Grundproblematik
  - einheitliche Begrifflichkeiten
  - „semantischer“ Datenaustausch

im IMISE/KKSL:  
**Klinische Studien**

# Begriffe in klinischen Studien (I)

- Mehrfachverwendung von Begriffen und Definitionen, z.B. in
  - Studienprotokollen, SOPs, Dokumentationsbögen, ...
  - Studiendatenbanken, Software-Anwendungen
- ⇒ Inkonsistenzen
- ⇒ Mehraufwand



# Begriffe in klinischen Studien (II)

- Uneinheitliche Definitionen
    - z.B. unterschiedliche Maßeinheiten von Laborparametern
  - Ungenaue Definitionen, z.B.
    - keine präzise Beschreibung zur Durchführung von Verfahren (z.B. Messverfahren)
- ⇒ Unterschiedliche Interpretationen möglich
- ⇒ Schlechte Nachvollziehbarkeit
- ⇒ Meta-Analysen werden erschwert



# Begriffe in klinischen Studien (III)

- Übliche Lösung: Definition eines Standards
- **Aber:** einheitliche Definitionen nicht immer möglich
- Gründe: Unterschiede im/in
  - Abstraktionsgrad z.B. allgemeinmedizinisch, studienspezifisch
  - Krankheitsbild z.B. Bewertung als Remission, Progress, ...
  - Expertenmeinungen aufgrund mangelnder Beweislage
  - Sichten beteiligter Personen z.B. Pathologe, Biometriker, Informatiker

# Werkzeug Onto-Builder V1

- Aufbau und Verwaltung von Data Dictionaries
- Dezentrale Erstellung der Definitionen
  - alternative Definitionen zu einem Begriff sammeln
  - vergleichend analysieren
  - Konsensfindung unterstützen aber nicht erzwingen
- Domänenunabhängigkeit



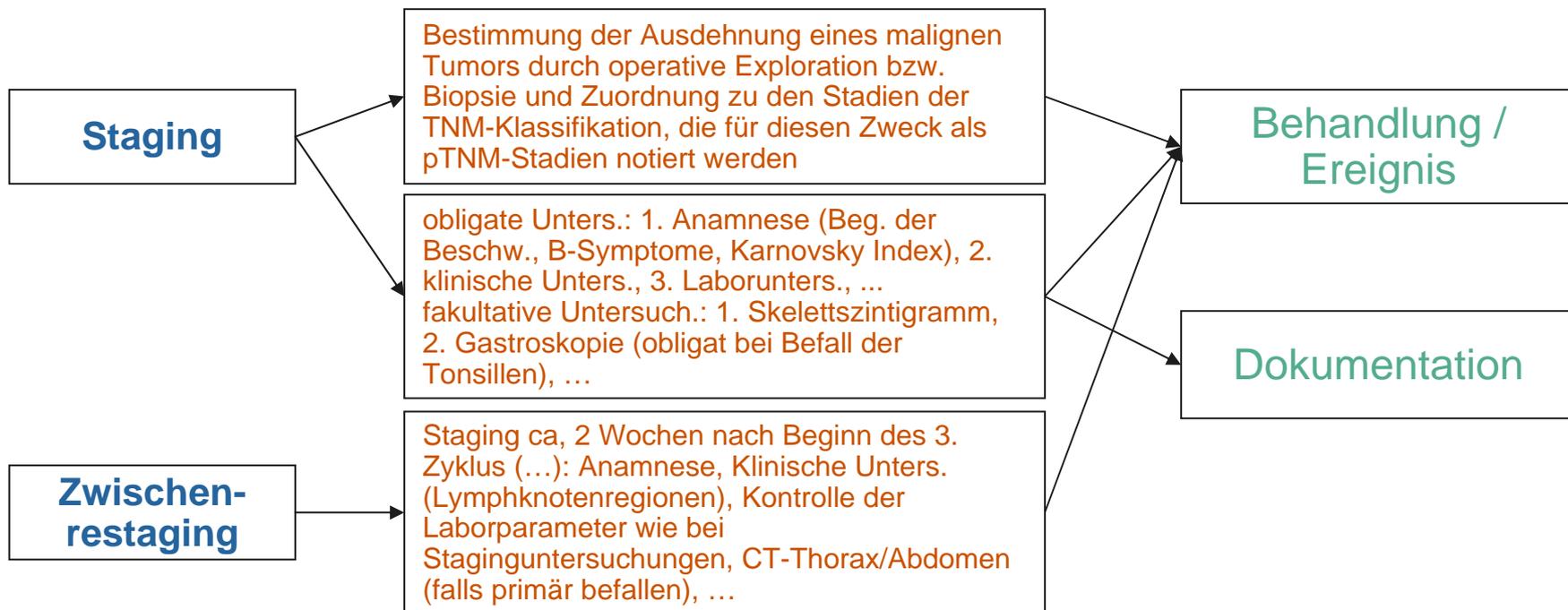
# Modell Onto-Builder V1

n Terme

m Definitionen

k Kontexte

Beispiel (Ausschnitt):



# Erfahrungen aus Version 1

- Unzulänglichkeiten Terminologien/Ontologien
    - Keine / nur ungenaue Definitionen
    - Keine Alternativen, keine Sichten
  - Modellerweiterungen für Onto-Builder V2
    - Einführung von Konzepten, Kategorien, Relationen
    - strukturierte Definitionen → z.B. Items
    - Ansatz für Kontexte und Sichten
  - „Organisatorische“ Probleme
    - Anreiz zur Mitarbeit unzureichend – Mehrwert nicht deutlich
    - Kritische Masse von Inhalten und Nutzern nötig
- v.a. Fokussierung auf Mehrwert für Nutzer**

# Ausschnitt Basisdatensatz KN HI

## Modul K: Körperliche Untersuchung

Item - ID	Item - Name	Item - Format	Item Herkunft	Exp.	Codelist-ID	Valide Werte (Range)	Consistency Check (Prüfregel → Fehlernachricht)
KGROESSE	Körpergröße	Zahl (5.2)	BCD	x		130 – 220 cm	<ul style="list-style-type: none"> <li>Angaben außerhalb des Prüfbereiches → [Körpergröße] liegt außerhalb des Prüfbereiches von 130 – 220 cm</li> </ul>
KGEWICHT	Gewicht	Zahl (5.2)	BCD	x		35 – 150 kg	<ul style="list-style-type: none"> <li>Angaben außerhalb des Prüfbereiches → [Gewicht] liegt außerhalb des Prüfbereiches von 35 – 150 kg</li> </ul>
KRRSYS / KRRDIA	Blutdruck (sitzend)	Zahl (3.0)	BCD	x		Systolisch: 70 - 250 mmHg Diasolisch: 40 - 150 mmHg	<ul style="list-style-type: none"> <li>Angaben außerhalb des Prüfbereiches → [Blutdruck systolisch / diastolisch] liegt außerhalb des Prüfbereiches von 70 - 250 mmHg / 40 - 150 mmHg (das jeweils Zutreffende angeben)</li> <li>KRRSYS &lt; KRRDIA → [Blutdruck systolisch] kleiner als [Blutdruck diastolisch]</li> </ul>
KHFREQ	Herzfrequenz	Zahl (3.0)	BCD	x		30 – 180 / min	<ul style="list-style-type: none"> <li>Angaben außerhalb des Prüfbereiches → [Herzfrequenz] liegt außerhalb des Prüfbereiches von 30 – 180 / min</li> </ul>

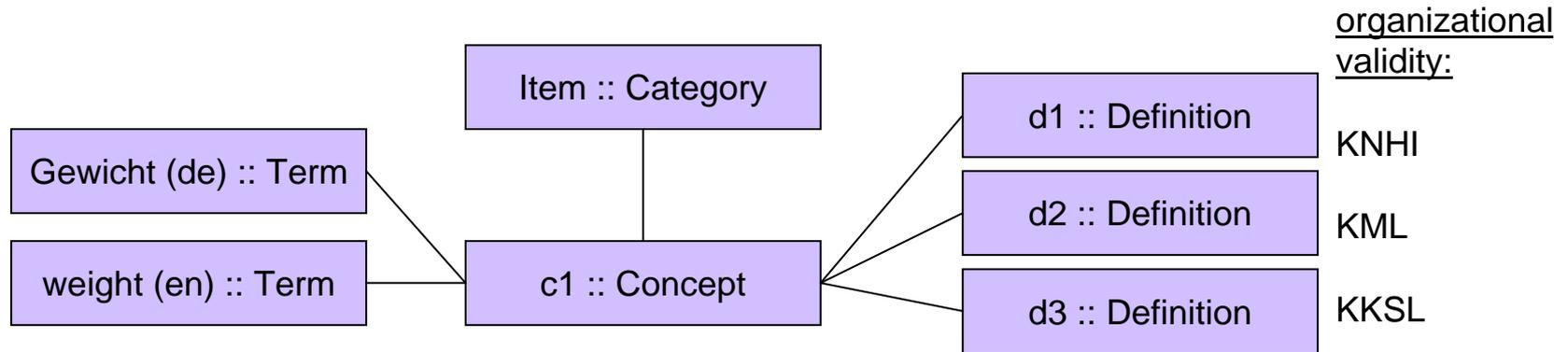
Zahlenformat: x.x = Stellen **gesamt** (ohne Komma) . Stellen **nach** dem Komma     Item-Herkunft: BCD = BCD-Item / HD = Harmonisierte Dokumentation / ST = Studienspezifisch

## Modul Z: Symptome und Zeichen der Herzinsuffizienz

Item - ID	Item - Name	Item - Format	Item Herkunft	Exp.	Codelist-ID	Valide Werte (Range)	Consistency Check (Prüfregel → Fehlernachricht)
ZBELASTD	Belastungsdyspnoe	radio button	BCD	x	CLJNNE		
ZRUHED	Ruhedyspnoe	radio button	BCD	x	CLJNNE		
ZOEDEME	Periphere Ödeme	radio button	BCD	x	CLJNNE		
ZHALSV	Halsvenenstauung	radio button	BCD	x	CLJNNE		
ZRASSEL	Pulmonale Rasselgeräusche	radio button	BCD	x	CLJNNE		

Zahlenformat: x.x = Stellen **gesamt** (ohne Komma) . Stellen **nach** dem Komma     Item-Herkunft: BCD = BCD-Item / HD = Harmonisierte Dokumentation / ST = Studienspezifisch

# Bsp.: Item 'Gewicht'



## Definition d1:

Name	Gewicht
Kurzname	KGEWICHT
Format	Zahl (3.2)
Herkunft	→Baseline, →Follow-up
Valide Werte	35 – 150 kg
Prüfregel / Fehler	B,F: Angaben außerhalb der angegebenen Bereiche → Kontrolle CRF → Rückfrage
Item aus BDS?	ja
Modul	→ K

# Zwischenbilanz

- bislang Prototyp, aber höhere Resonanz als bei V1
- Gründe:
  - Kritische Masse gegeben
  - Hoher Grad der Wiederverwendbarkeit
  - Harmonisierung wünschenswert:
    - höhere Qualität
    - bessere Vergleichbarkeit
    - geringerer Aufwand in Studienplanung
  - Leichtere Vergleichbarkeit gegenüber Freitextdefinitionen
- Hauptfaktoren:
  - **Erkennbarer Nutzen**
  - **Intuitiver Zugang**

# Erfahrungen in anderen Projekten



MAX-PLANCK-GESELLSCHAFT

- Biologie: günstigeres Umfeld für Ontologien
  - Gute Akzeptanz weg. „Datenflut“
  - Einheitliche Organisation der Open Biomedical Ontologies (OBO)
  - Kaum Integration mit medizinischen Term./Ont.
    - Relevanz z.B. bei Studien mit genetischen Daten
- Ontologien in der Chirurgie:
  - Ziel: konsistente Erweiterbarkeit
  - Abdeckung, Einbindung bestehender Systeme problematisch

iccas





# Terminologien / Ontologien in der klinischen Forschung

- Große Anzahl von Systemen:
  - Terminologien ICD, LOINC, ATC, ...
  - Formale Ontologien GALEN, SNOMED, ...
  - (Standards) HL7 (RIM), CDISC (ODM, CDA)
- Geringer Nutzungsgrad, wo nicht erzwungen
- Bekannte Probleme ungelöst
  - Rector: Why is Clinical Terminology so Hard? (1999)
  - Cimino: Desiderata for Clinical Terminologies (1998)
- Integrationsansätze und -systeme unzureichend
  - UMLS
  - Einzelabbildungen z.B. ICD → GALEN (2006)



# Zusammenfassung und Ausblick

- Anforderungen
  - Erkennbarer Nutzen durch Terminologien / Ontologien
    - Meta-Analysen, Vergleiche, Austausch
  - Intuitive Zugangsmöglichkeiten
    - inkl. Umgang mit Definitionsalternativen
  - Mittel zur Integration bestehender Systeme
    - z.B. mit Kernontologien
  
- Erforderliche Projekte
  - mit echtem Austauschbedarf u./o.
  - Partnern mit hohem Verteilungsgrad oder Verzahnung bei unterschiedlichen Modellen

## Vielen Dank für Ihre Aufmerksamkeit!

Kontakt: Frank Loebe  
frank.loebe@imise.uni-leipzig.de

Ronald Speer  
ronald.speer@imise.uni-leipzig.de

# Erfahrungen aus Version 1

- „Organisatorische“ Probleme
    - Anreiz zur Mitarbeit unzureichend – Mehrwert nicht deutlich
    - Kritische Masse von Inhalten und Nutzern nötig
  - Modellerweiterungen für Onto-Builder V2
    - Trennung Term – Konzept
    - Kategorien von Begriffen **z.B. Items**
    - strukturierte kategorienabhängige Definitionen
    - Relationen
    - Ansatz für Kontexte und Sichten
- v.a. Fokussierung auf Mehrwert für Nutzer**