

## SNOMED CT - Terminologische Fundierung für Medizinische Wissensbasen

Workshop „Wissensbasierte Systemkomponenten  
in Krankenversorgung und Lehre  
conhIT „Satellitenveranstaltung“

Berlin, 19. April 2010

---



**Josef Ingenerf**

Institut für Medizinische Informatik  
Universität zu Lübeck

### Übersicht

- n Motivation für Klassifikationen und Terminologien
- n SNOMED CT: Einführung
- n SNOMED CT: Beschreibungslogik
- n SNOMED CT und Wissensverarbeitung

## Motivation für Klassifikationen und Terminologien?

Mangelnde maschinelle Verarbeitbarkeit natürlicher medizinischer Sprache!



Magen Befund

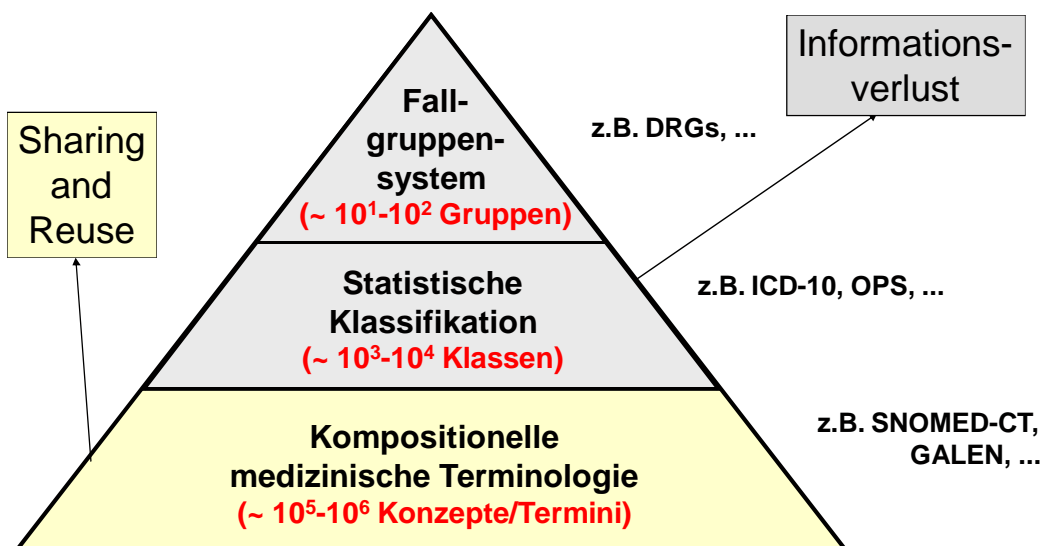
Im Antrum, hinterwandig, ist ein mitteltiefes, glatt begrenztes Ulkus im abheilenden Stadium mit einem Durchmesser von ca. 3 mm zu finden. Die Läsion blutet spritzend arteriell aus sichtbarem Gefäß (Forrest Ia). Im Magen Injektionstherapie mit 22 ml Aethoxysklerol 0,5% zur Blutstillung.

### Phänomene:

- sprachliche Variabilität (Synonymie, Homonymie, ...)  
z.B. Niere, Kidney, Nephritis, Arteria renalis
- Begriffshierarchien, (weitere) Begriffsrelationen  
z.B. Niere ist ein Organ, Nephritis ist eine Nephropathie

Ingenierf. Lübeck / 3

## Standardisierung auf verschiedenen Abstraktionsniveaus

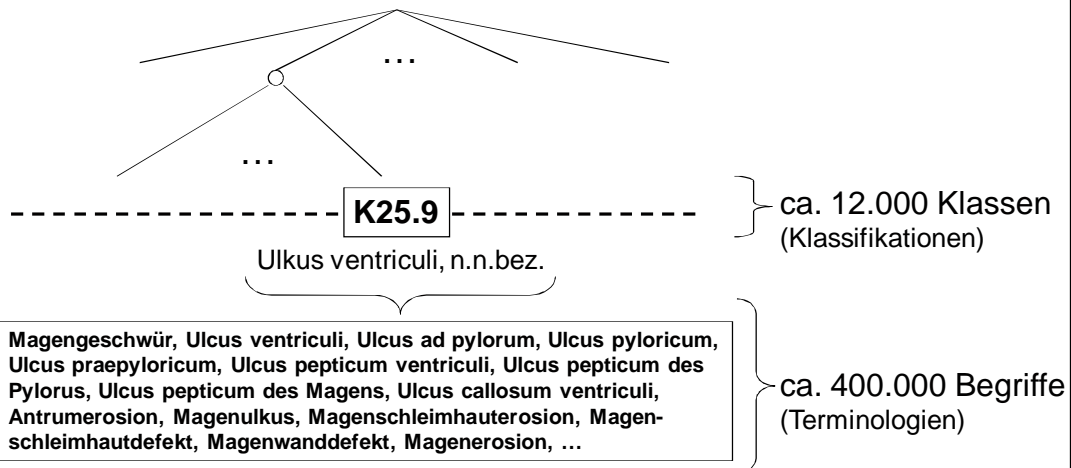


## Medizinische Fachsprache bzw. Terminologie

Ingenierf. Lübeck (4)

## Ausdrucksschwäche von Klassifikationen

z.B. Krankheiten



Ingenierf. Lübeck / 5

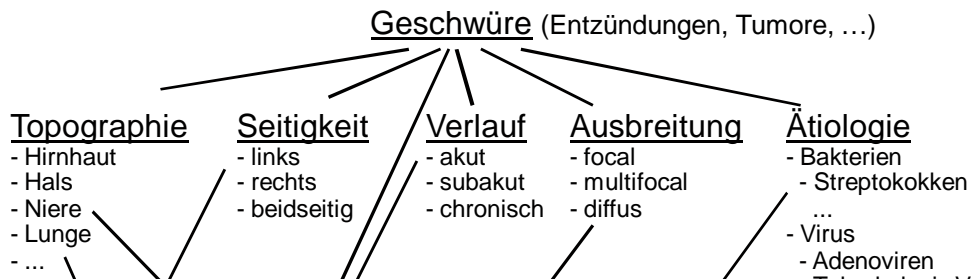
## Klassifikation und Terminologien

Klassen-ID (ICD-10)	Text	...	Term-ID (Alpha-ID)	Begriff-ID (SNOMED CT)
K25.0	Ulcus ventriculi, akut mit Blutung		-	-
	Akutes Ulcus ventriculi mit Blutung		I5497	89748001
	Akutes Ulcus ventriculi mit Hämorrhagie		I101506	
	Akutes Magengeschwür mit Blutung		I97444	
	Dieulafoy-Krankheit		I75618	(89748001)
	Dieulafoy-Ulkus		I78281	<u>WEGEN</u>
	Exulceratio simplex Dieulafoy		I31915	109558001)
K25.9	Ulcus ventriculi, weder als akut noch als chronisch bezeichnet, ohne Blutung oder Perforation		-	-
	Magengeschwür		I5052	397825006
	Ulcus ventriculi		I5051	
	Ulcus pyloricum		I11881	39204006
	Ulcus präpyloricum		I5055	22620000
	Infektion durch Helicobacter pylori bei Ulcus ventriculi (+ Sekundärkode B96.81!)		I32930	89662003

erf. Lübeck / 6

## Ausdrucksstärke von Terminologien

Terminologien zerlegen Ausdrücke in ihre Bedeutungsbestandteile und rekonstruieren die Gesamtbedeutung mit einer Grammatik.



### Erforderlich:

**Kompositioneller Ansatz** (Postkoordination)

statt **Enumeration und Hierarchisierung** (Präkoordination)

## Koordination von Klassifikationen und Terminologien

"Content Model for the ICD-11 Revision":

[bmir.stanford.edu/file\\_asset/index.php/1522/BMIR-2010-1405.pdf](http://bmir.stanford.edu/file_asset/index.php/1522/BMIR-2010-1405.pdf)

**ICD-11 Revision**

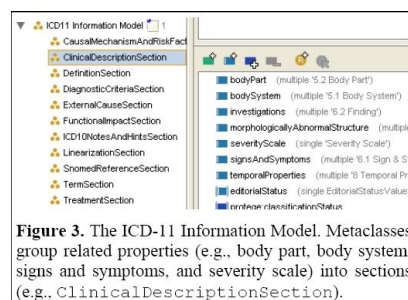
Home > iCAT - Initial ICD-11 Collaborative Authoring Tool!

This platform is used for the initial authoring of the ICD-11 alpha draft.

- The iCAT production server is available at: <http://icat.stanford.edu/>
- The demo and training iCAT platform is available on: <http://icademo.stanford.edu/>

The content of ICD-11 categories can be edited by going to the and selecting the category in the hierarchy.

<https://sites.google.com/site/icd11revision/home/icat>



**Figure 3.** The ICD-11 Information Model. Metaclasses group related properties (e.g., body part, body system, signs and symptoms, and severity scale) into sections (e.g., ClinicalDescriptionSection).

This informal model is implemented in a three-layer model documented in UML:<sup>4</sup> a) The *Foundation layer* divided into (1a) the *Ontology layer* that is intended to be aligned with a subset of SNOMED, and (1b) the *Category layer* that contains the description of each ICD category; (2) the *Linearizations layer*—a generalization of the traditional ICD classifications that provides the backwards compatibility (including their inclusions, exclusions, and residual categories) and supports new use cases.

Ingenier, Lübeck (8)

## Übersicht

- n Motivation für Klassifikationen und Terminologien
- n SNOMED CT: Einführung
- n SNOMED CT: Beschreibungslogik
- n SNOMED CT und Wissensverarbeitung

Ingenerf. Lübeck (9)

## SNOMED: Historie, siehe „<http://www.ihtsdo.org>“

1965 **SNOP**: Systematized Nomenclature of Pathology

1974 **SNOMED**: Systematized Nomenclature of Medicine,  
N = 44.587 terms

1979 **SNOMEDII**: Erweiterung

1984 **SNOMED II**: Deutsche Übersetzung und Erweiterung  
N = 80.000 terms (Prof. Wingert, jetzt ID GmbH)

1993 **SNOMED III**: SNOMED International,  
N = 130.580 terms

1998 **SNOMED® RT™**: SNOMED Reference Terminology  
(Release 1.0 Nov 2000, Release 1.1 July 2001)

2002 **SNOMED®CT™**: SNOMED Clinical Terms,  
Merger of SNOMED RT and Clinical Terms V.3 (the Read Codes)  
N = 350.000 Concepts, 800,000 Terms, >1 Mio. Relationships

*2004 included in 2004AA UMLS-Metathesaurus.*

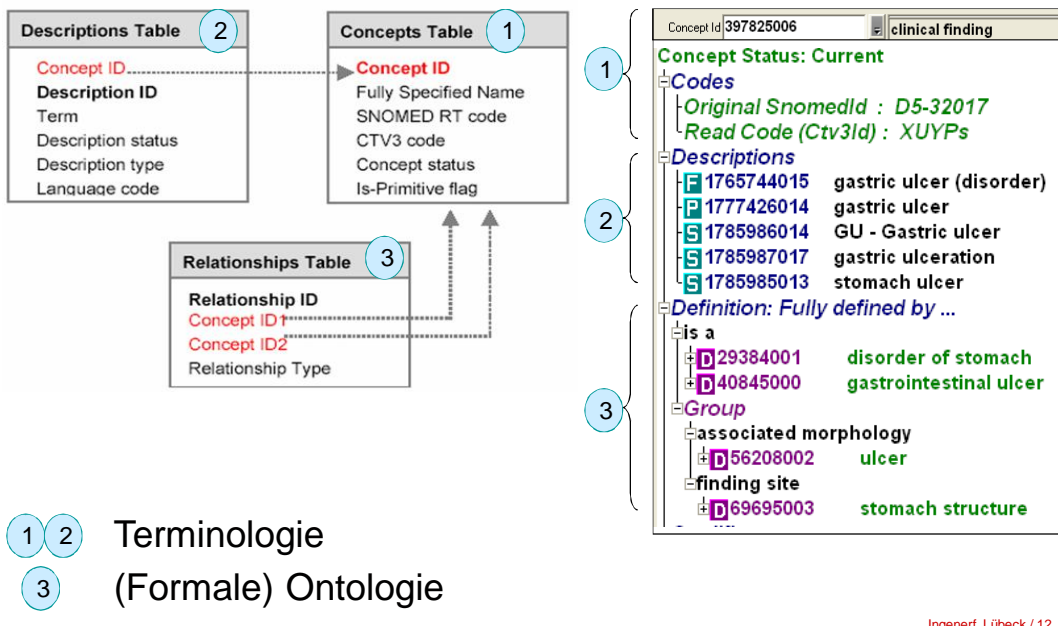
Ingenerf. Lübeck (10)

## IHTSDO: Int. Health Terminology Standard Dev. Org.

- n Internationale Non-profit-Organisation nach dänischem Recht, Sitz Kopenhagen
- n Gegründet 2006
- n Mitglieder: Australien, Kanada, Dänemark, Litauen, Niederlande, Neuseeland, Schweden, Singapur, Spanien, Vereinigtes Königreich, Vereinigte Staaten
- n Corporate Affiliates
- n Hält die Rechte an SNOMED CT seit 2007
- n CEO: Jennifer Zelmer (Kanada)
- n Chief Terminology Officer: Kent Spackman (USA)
- n <http://www.ihtsdo.org>

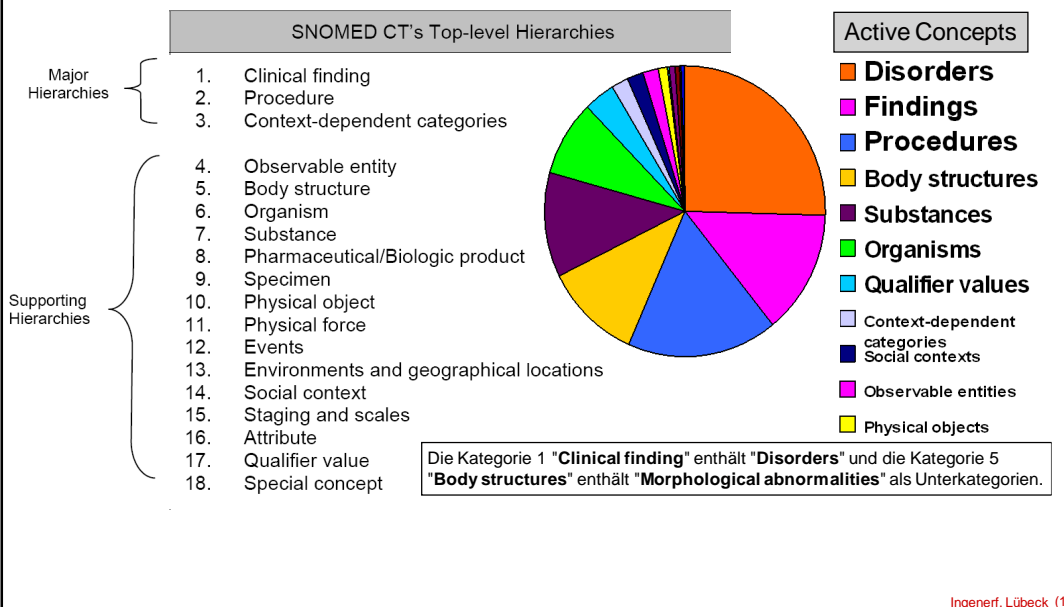
Ingenierf. Lübeck / 11

## SNOMED CT: Core table format

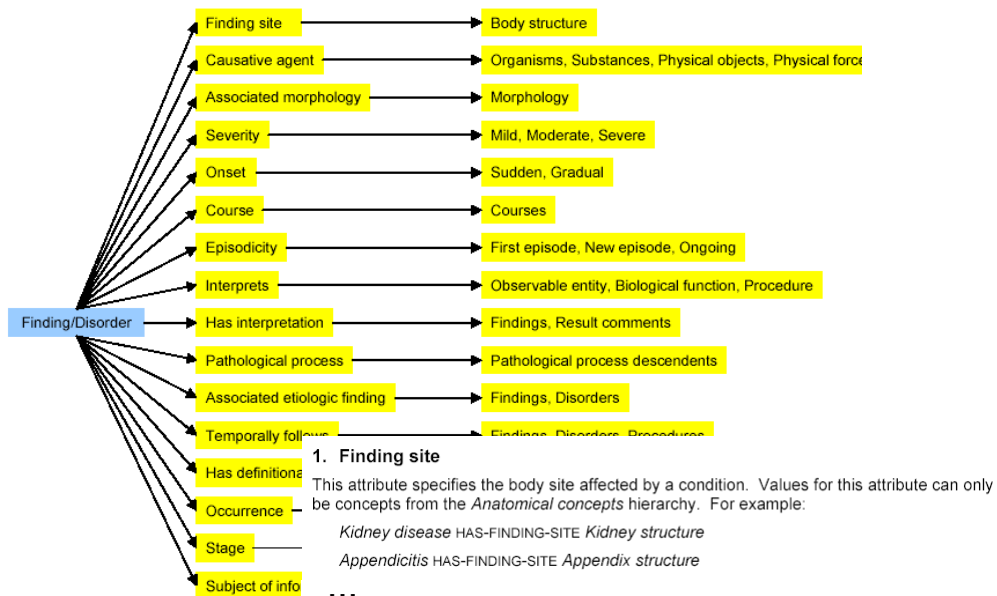


Ingenierf. Lübeck / 12

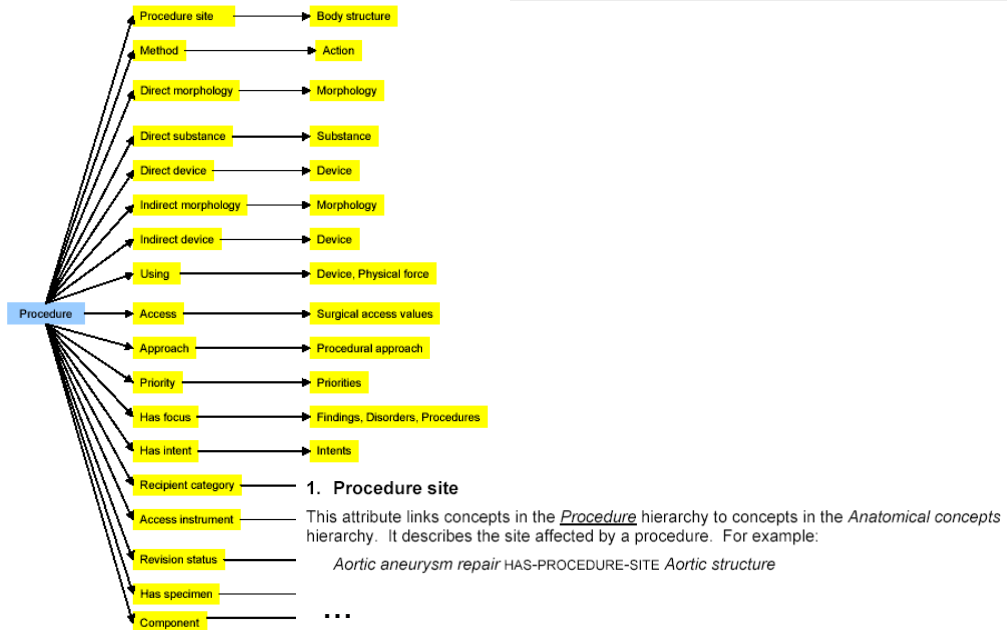
# SNOMED CT: Übersicht



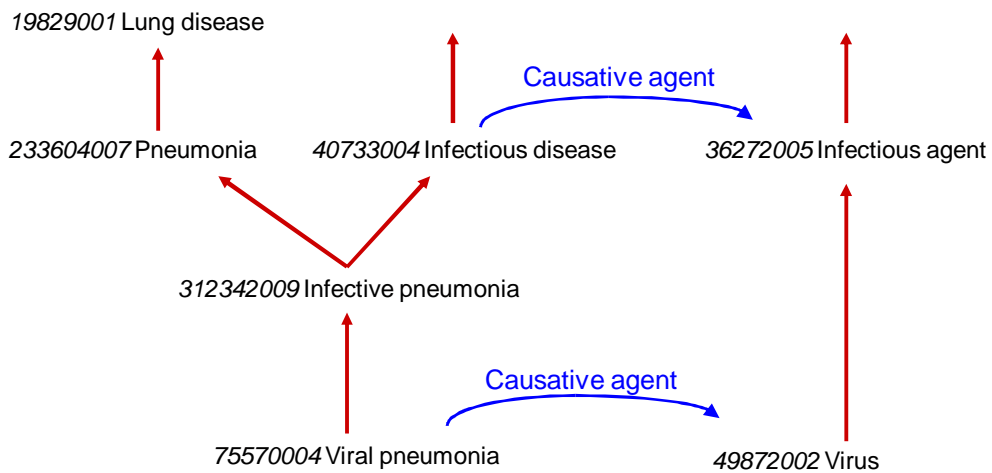
# SNOMED attributes used for finding and disorder concepts



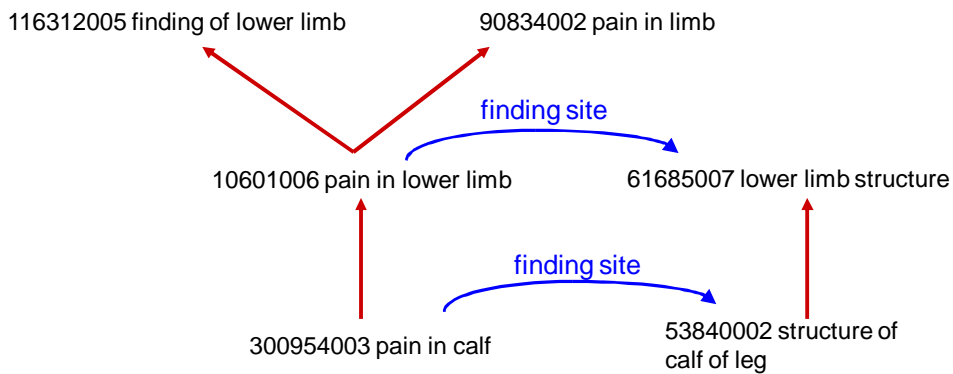
## SNOMED attributes used for procedure concepts



## SNOMED CT: isa- and attribute-relations



## SNOMED CT: isa- and attribute-relations



Ingenierf. Lübeck / 17

## Formale Terminologie, z.B. SNOMED CT

**CONCEPT:** 397825006 Gastric ulcer

**DESCRIPTORS (TERMS):**

Gastric ulcer (1777426014, preferred), Stomach ulcer (1785985013),  
GU - Gastric ulcer (1785986014), Gastric ulceration (1785987017)

**DEFINITION:**

Fully defined by ...

- *Is a* 64572001 Disease (disorder)
- Group
  - *Associated morphology* 56208002 Ulcer
  - *Finding site* 69695003 Stomach structure

§ Semantik-freie, identifizierende Kodes

§ Explizite Relationen

siehe Cimino-Kriterien für Terminologische Systeme!

Ingenierf. Lübeck (18)

## SNOMED CT: Beispiel

**CONCEPT:** 397825006 Gastric ulcer

**DESCRIPTORS (TERMS):**

Gastric ulcer (1777426014, preferred), Stomach ulcer (1785985013),  
GU - Gastric ulcer (1785986014), Gastric ulceration (1785987017)

**Defining concepts instead of creating hierarchies manually:**

**DEFINITION:**

Fully defined by ...

- *Is a* 64572001 Disease (disorder)
- Group
  - *Associated morphology* 56208002 Ulcer
  - *Finding site* 69695003 Stomach structure

} using description  
logic (DL) formalism.

**Creating subsumption-hierarchy automatically by a classifier:**

**INFERRED SUPERORDINATE CONCEPTS** (by a classifier)

- 29384001 Disorder of stomach
- 40845000 Gastrointestinal ulcer

Ingenierf. Lübeck (19)

## Übersicht

- n Motivation für Klassifikationen und Terminologien
- n SNOMED CT: Einführung
- n SNOMED CT: Beschreibungslogik
- n SNOMED CT und Wissensverarbeitung

Ingenierf. Lübeck (20)

## Mögliche Beschreibungssprachen

### n Natürliche Sprache

Jede Hepatitis ist eine Entzündung, die in einer Leber lokalisiert ist.  
Jede Entzündung in einer Leber ist eine Hepatitis.

### n Prädikatenlogik

**"x"**  $instanceOf(x, Hepatitis) \dot{\cup} instanceOf(x, Inflammation) \dot{\cup}$   
 $\$y: instanceOf(y, Liver) \dot{\cup} hasLocation(x,y)$

### n Beschreibungslogik

$Hepatitis \equiv Inflammation \dot{\cup} \$hasLocation.Liver$

**Formale Sprache: „Rechnen“**

„Stefan Schulz, HL7-Jahrestagung, 2009“

Ingenieur, Lübeck (21)

## Beschreibungslogiken

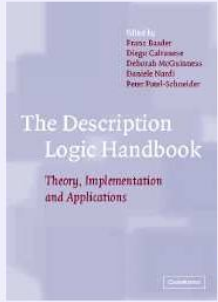
### n Abwägung: Performanz gegen Sprachumfang

### n OWL-DL:

- u Standardisierte Sprachspezifikation (W3C)
- u Bewährte Editoren (Protégé)
- u angepasste Reasoner (classifier) -  
bei Verwendung des vollen Sprachumfangs nur  
bedingt skalierbar

Ingenieur, Lübeck (22)

## Beschreibungslogiken



**The Description Logic Handbook**  
Theory, Implementation and Applications

Edited by Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, Peter Patel-Schneider

Published January 2003

574 pages 14 tables 53 figures

Hardback | ISBN: 0521781760

For price and ordering options, inspection copy requests, and reading lists please select:  
[UK](#) | [North America](#) | [Australia & New Zealand](#)

- 3. Complexity of reasoning F. M. Donini
- 4. Relationships with other formalisms U. Sattler, D. Calvanese and R. Molitor
- 5. Expressive description logics D. Calvanese and G. De Giacomo
- 6. Extensions to description logics F. Baader, R. Küsters and F. Wolter
- Part II. Implementation: 7. From description logic provers to knowledge representation systems D. L. McGuinness and P. F. Patel-Schneider
- 8. Description logics systems R. Möller and V. Haarslev
- 9. Implementation and optimisation techniques I. Horrocks
- Part III. Applications: 10. Conceptual modeling with description logics A. Borgida and R. J. Brachman
- 11. Software engineering C. Welty
- 12. Configuration D. L. McGuinness
- 13. Medical informatics A. Rector
- 14. Digital libraries and web-based information systems I. Horrocks, D. L. McGuinness and C. Welty
- 15. Natural language processing E. Franconi
- 16. Description logics for data bases A. Borgida, M. Lenzerini and R. Rosati

Ingenier, Lübeck / 23

## SNOMED CT: Description Logic (Baader, Dresden)

### DL-Dialekt EL:

Name	Syntax	Semantics
top	$\top$	$\Delta^I$
conjunction	$C \sqcap D$	$C^I \cap D^I$
existential restriction	$\exists r.C$	$\{x \in \Delta^I \mid \exists y \in \Delta^I : (x, y) \in r^I \wedge y \in C^I\}$

### DL-Dialekt EL+:

#### weitere Rollen-Axiome:

general concept inclusion	$C \sqsubseteq D$	$C^I \subseteq D^I$
role inclusion	$r_1 \circ \dots \circ r_n \sqsubseteq s$	$r_1^I \circ \dots \circ r_n^I \subseteq s^I$

Ingenier, Lübeck / 24

## DL-Dialekt „EL+“

### Beispiel:

Pericardium  $\sqsubseteq$  Tissue  $\sqcap \exists$ contained-in.Heart  
Pericarditis  $\sqsubseteq$  Inflammation  $\sqcap \exists$ has-location.Pericardium  
Inflammation  $\sqsubseteq$  Disease  $\sqcap \exists$ acts-on.Tissue  
Heartdisease  $\doteq$  Disease  $\sqcap \exists$ has-location.Heart  
Heartdisease  $\sqsubseteq \exists$ has-state.NeedsTreatment **GCI**  
has-location  $\circ$  contained-in  $\sqsubseteq$  has-location **Role Inclusion**

d.h. Pericarditis  $\sqsubseteq$  Heartdisease  
 $\sqsubseteq \exists$ has-state.NeedsTreatment

Ingenierf. Lübeck / 25

## SNOMED CT: Beispiel (Definieren statt hierarchisieren)

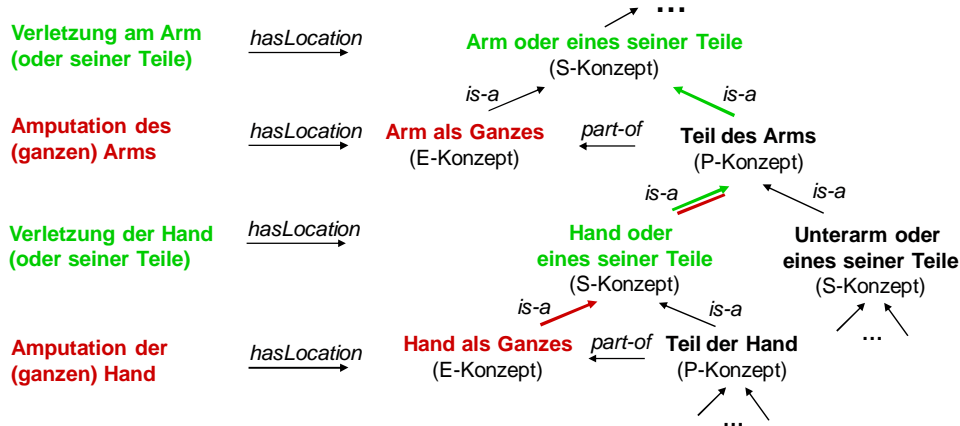
### Descriptions

- ⊖ F 2394452012 Helicobacter-assoziiertes pylorisches Ulkus (Störung)
- ⊖ P 2394453019 Helicobacter-assoziiertes pylorisches Ulkus
- ⊖ Fully defined by ...
  - ⊖ 116680003 ist ein/e
    - ⊕ D 128070006 infektiöse Krankheiten des Abdomens
    - ⊕ D 312121001 bakterielle Infektionskrankheit des Magen-Darm-Trakts
    - ⊕ D 39204006 pylorusnahes Magengeschwür
    - ⊕ D 6185008 in Verbindung mit Helicobacter stehende Krankheit
  - ⊖ 246075003 verursachendes Agens
    - ⊖ D 80774000 helicobacter pylori
  - ⊖ Group
    - ⊖ 116676008 assoziierte Morphologie
      - ⊕ D 56208002 Geschwür
    - ⊖ 363698007 Befundlokalisation
      - ⊕ D 280119005 Struktur des Magenpförtners

⊖ Die Begriffshierarchie wird rechnergestützt deduziert.

Ingenierf. Lübeck / 26

## SEP-Tripel-Ansatz für partitives Wissen



Ingenerf, Lübeck (27)

## SNOMED-CT: Anatomie

### SEP-Tripel:

- anatom. Struktur
- anatom. Ganzes (Entity)
- anatom. Teil (Part)

ConceptId	280119005	Struktur des Magenpförtners (Anatomie)
Description Id	1702394017	Anatomie
Detail	Distributed relationships	Hierarchy
ConceptStatus	Current	Subtype hierarchy
Legacy codes	SNOMED: T-D071F CTV3ID: Xa1II	
Descriptions	1702393011 Struktur des Magenpförtners (Anatomie) 1702394017 Struktur des Magenpförtners 1225725016 pylorus 1702395016 Pförtner 2000734015 Pförtnerstruktur	245415005 Region des Magens 91837002 anatomische Mündung/Öffnung 280119005 Struktur des Magenpförtners 362143006 gesamter Pförtner 70710007 Glandula pylorica im Magen 15048009 Struktur des Pförtnerteils des Magens 245416006 Teil der Magenpförtnerregion 66051006 Struktur des Antrum pyloricum 77612007 Struktur des Eingangs des Magenpförtners 19987003 Struktur des Pförtnerkanals 367602001 Struktur des präpylorischen Abschnitts
Primitive	118680003 ist ein/e 245415005 Region des Magens 91837002 anatomische Mündung/Öffnung 123005000 Teil von 181246003 gesamter Magen 118760003 gesamte Eingeweide 181244000 gesamter oberer Gastrointestinaltrakt 302553009 gesamtes Abdomen	

Annotations in the hierarchy view:

- 245415005 Region des Magens → „S-Konzept“
- 91837002 anatomische Mündung/Öffnung → „E-Konzept“
- 280119005 Struktur des Magenpförtners → „E-Konzept“
- 245416006 Teil der Magenpförtnerregion → „P-Konzept“

Lübeck (28)

## SNOMED-CT: Morphologie

<b>ConceptId</b>	56208002	<b>Geschwür (morphologische Anomalie)</b>
<b>Description Id</b>	1739741010	
Anatomie		
Detail	Distributed relationships	Hierarchy
<b>ConceptStatus Current</b> Legacy codes SNOMED: M-38000 CTV3ID: XaB0y Descriptions F 1739745018 Geschwür (morphologische Anomalie) P 1739741010 Geschwür S 1739740011 Ulcus - Läsion S 1739742015 Ulzeration S 1739743013 ulzerierende Läsion S 1739744019 Ulkus Primitive 118680003 ist ein/e D 107658001 mechanische Anomalie		U 107658001 mechanische Anomalie C 56208002 Geschwür C 47656000 Abklatschulkus C 45771005 akutes blutendes Ulkus C 40518009 akutes oberflächliches Ulkus C 26317001 akutes Ulkus C 110426005 Aphthe C 55075001 blutendes Ulkus C 37904008 chronisches blutendes Ulkus C 405719001 chronisches Ulkus C 63280002 dendritisches Ulcus C 125316002 diffuse Ulzeration C 45197007 fokales Ulkus C 413285006 Geschwür mit Kraterbildung C 25845007 heilendes Ulkus C 125315003 inkurables Ulkus

Ingenerf. Lübeck (29)

## SNOMED CT: Defizite

### Biopsy Planned:

impliziert Existenz von: [Biopsy](#)

### Drug\_Abuse\_Prevention :

impliziert Existenz von : [Drug Abuse](#)

### Suspected Gallstones:

impliziert Existenz von : [Gallstones](#)

### Absence of Arm:

impliziert Existenz von : [Upper Limb Structure](#)

### Amputation of toe:

: wird klassifiziert als: [Amputation of foot](#)

### Absence of liver or gallbladder NOS :

wird klassifiziert als: [Congenital absence of liver and gallbladder](#)

### Proximal hemiphalangectomy of toe :

wird klassifiziert als: [Amputation of toe.](#)

Concept Status: **Current**

- Descriptions
  - F biopsy planned (situation)
  - P biopsy planned
- Definition: Fully defined by ...
  - is a
    - D operative procedure planned
  - Group
    - associated procedure
      - D biopsy
    - procedure context
      - D planned
    - temporal context
      - D current or specified
    - subject relationship cont
      - D subject of record

C1 – Rel – C2 zu interpretieren als:  
 " x: instanceOf(x, C1) P  
 \$y: instanceOf(C2) ∪ Rel(x,y)

„Stefan Schulz, HL7-Jahrestagung, 2009“

Ingenerf. Lübeck (30)

## Übersicht

- n Motivation für Klassifikationen und Terminologien
- n SNOMED CT: Einführung
- n SNOMED CT: Beschreibungslogik
- n SNOMED CT und Wissensverarbeitung

Ingenierf. Lübeck (31)

## Was leisten formale Ontologien?

- n Exakte, logikbasierte Beschreibungen von Typen, die durch konkrete Objekte der Welt instanziiert werden
- n Repräsentation von stabilen, kontextunabhängigen Grundannahmen
- n Verwendung von maschinellem Schließen, z.B. basierend auf Beschreibungslogiken (OWL-DL)

Ingenierf. Lübeck (32)

## Was leisten formale Ontologien nicht?

- n Repräsentation kontextabhängigen Wissens
  - u „Heuschnupfen ist die häufigste Allergie in D“
- n Repräsentation probabilistischem Wissens
  - u „5% der Hepatitiden verlaufen anikterisch“
  - u Rauchen ist ein Risikofaktor für KHK
- n Default / kanonisches Wissen
  - u „Der Mensch hat 32 Zähne“
- n Dispositionen:
  - u „Gleevec® ist indiziert bei CML“
  - u „Aspirin® greift die Magenschleimhaut an“

**Ontologie ≠ Wissensrepräsentation**

Ingenierf. Lübeck (33)

## Ontologien und Wissensbasierte Systeme

Regeln: IF QueryKonzept THEN Action

Auswertung des Query-Konzeptes:

Classifying des Query-Konzeptes in SNOMED CT:

$K \sqsubseteq$  QueryKonzept?

Ingenierf. Lübeck (34)

# Ontologien und Wissensbasierte Systeme

Colantonio et al. (2007). An Approach to Decision Support in Heart Failure

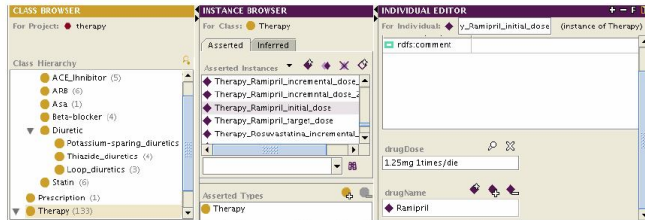
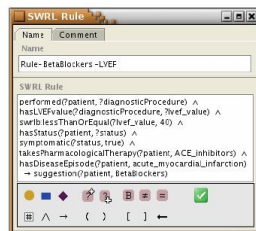


Fig. 4. A Fragment from the therapy core ontology

Ontologie (OWL)

The set of rules elicited using a logic implication format in natural language has been simply transformed into a set of SWRL rules. An example rule is shown in Fig. 5 as it has been defined in Protégé.



Rules (SWRL)

Ingenerf, Lübeck / 35

## SNOMED-Ausdrücke: Normalformen

### SNOMED Clinical Terms® Guide

August 2006 Revision (see status statement on page 2)  
Version 5

#### Transforming Expressions to Normal Forms

**Example:** SNOMED CT includes the following concepts:

- *Repair of inguinal hernia* (SCTID = 44558001)
- *Laparoscopy* (SCTID = 73632009)
- *Endoscopy with surgical procedure* (SCTID = 85876005)

SNOMED CT also includes a pre-coordinated concepts for this procedure

- *Laparoscopic repair of inguinal hernia* (SCTID = 15018006)

This makes it possible to code the procedure in (at least) two ways:

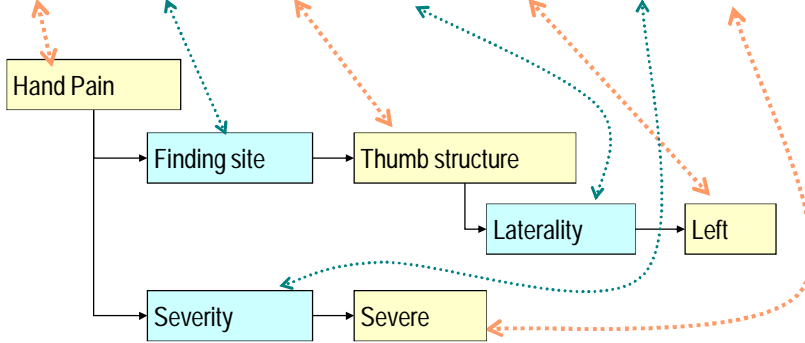
- 15018006 (pre-coordinate)
- 44558001 + 73632009 + 85876005 (post-coordinated)

**Notwendigkeit einer Normalform für (definierte) präkoordinierte und postkoordinierte Begriffe, um die Äquivalenz alternativer Begriffsausdrücke automatisch auswerten zu können.**

Ingenerf, Lübeck / 36

## Compositional grammar: „Severe pain at the left thumb“

53057004:363698007=(76505004:272741003=7771000),272141005=24484000



Serialisierung eines Konzeptausdrucks.

Ingenerf, Lübeck / 37

## SNOMED-Ausdrücke: Normalformen für Subsumptionstests

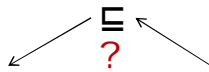
### 3.1 Candidate and predicate expressions

This paper considers transformations that can be applied to expressions, to enable effective subsumption testing. In any subsumption test there are two expressions, one of which is being tested for subsumption by the other. To distinguish these expressions the following definitions are used:

**Candidate expression** – An expression that is being tested to see if it is subsumed by another expression.

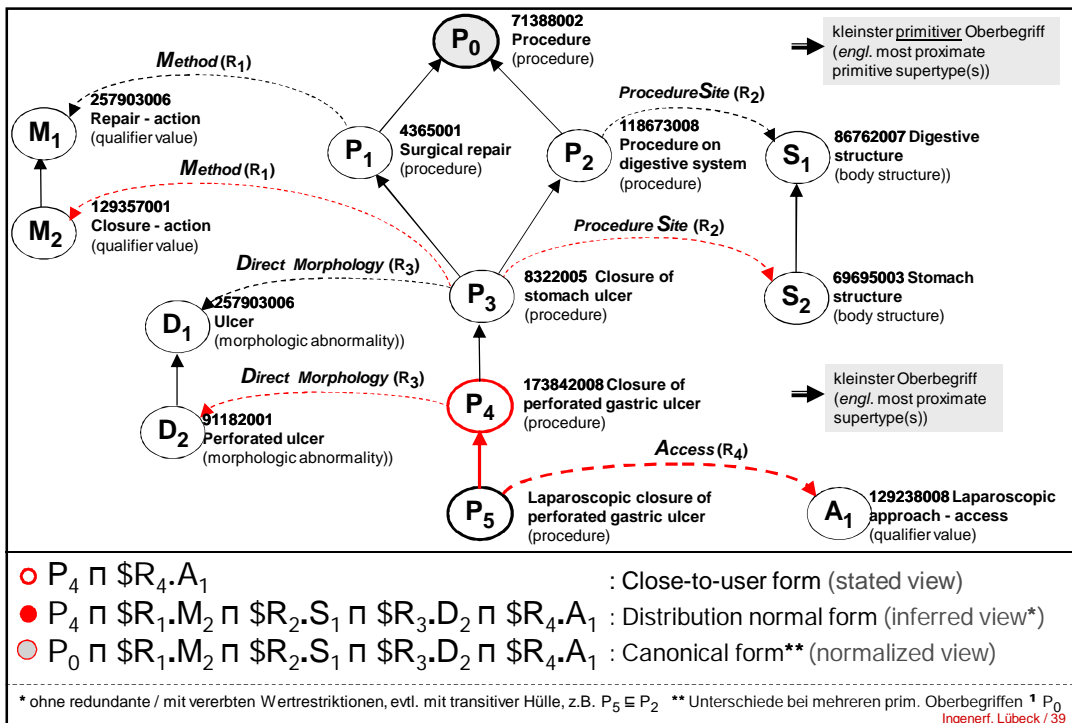
**Predicate expression** – An expression that is being tested to see if it subsumes another expression.

Examples



Predicate	Candidate	Test result
Fracture of femur	Fracture of neck of femur	True
	Fracture of bone	False
Fracture of bone	Fracture of femur	True
	Fracture of neck of femur	True
Asthma ( <i>in patient</i> )	Family history of asthma	False
	Severe asthma ( <i>in patient</i> )	True
Family history of respiratory disease	Family history of asthma	True

Ingenerf, Lübeck / 38



## Short and long normal forms of SNOMED CT expressions

### 4.5 Concept definitions in normal forms

#### 4.5.1 Long Normal Form

A form which when applied to a candidate expression allows effective computation of whether it is subsumed by a predicate expression.

Predicate Expression  $\sqsupseteq$  Candidate Expression  
 LONG-NF

#### 4.5.2 Short Normal Form

A form which when applied to a predicate expression allows effective computation of whether a candidate expression is one of its subtypes.

Predicate Expression  $\sqsupseteq$  Candidate Expression  
 SHORT-NF  
 (u.a. für Queries)

## IHTSDO Workbench

INTERNATIONAL HEALTH TERMINOLOGY  
STANDARDS DEVELOPMENT ORGANISATION



### **Press Release: “IHTSDO Launches Global Health Terminology Workbench”**

*Tools to Benefit Member Nations and Other SNOMED CT Users*

#### **Backgrounder: About the IHTSDO Workbench**

The IHTSDO Workbench includes a set of tools that will form the foundation of a multi-lingual modular workbench. The overall environment will initially host terminology browsing, authoring, subset and other reference set management, mapping, and namespace management applications. Users can either work independently or can collaborate on these tasks. IHTSDO intends to add additional modules to the Workbench over time and to encourage partners to build and share complementary tools.

Ingenierf. Lübeck (41)