

Pizza mit Schlagobers? Ontologien und nachhaltiges Wissensmanagement

Stefan Schulz

Institut für Medizinische Informatik, Statistik und Dokumentation

Medizinische Universität Graz

stefan.schulz@medunigraz.at

Wissen und Nachhaltigkeit



natürliche Ressourcen





Wissen und Nachhaltigkeit

- Einführung
- Wissensexplosion
- Wissensarten
- Ontologisches Wissen















Wissen

- Demo: Ontologie
- Gesellschaftliche Dimension
- Diskussion

Wissen und Nachhaltigkeit?





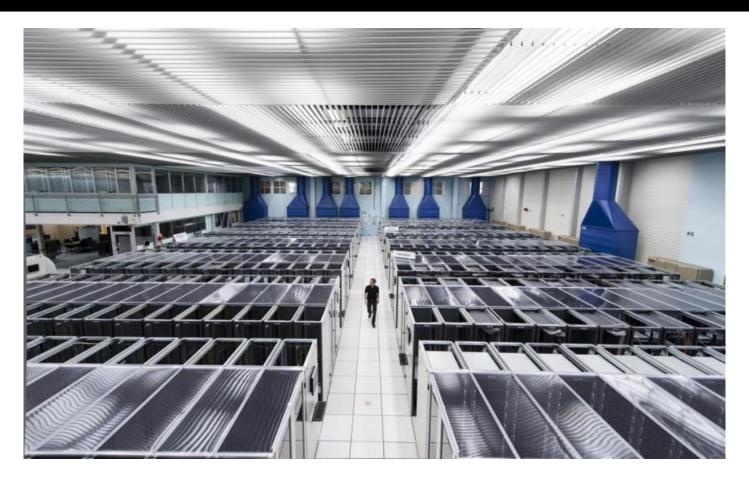




http://update2.blog.de/2009/01/08/

http://www.dw.de/ausstellung-in-der-anna-amalia-bibliothek/a-17890230 http://www.schule.rorschach.ch/potatoe/5MU023/5MU023.nim

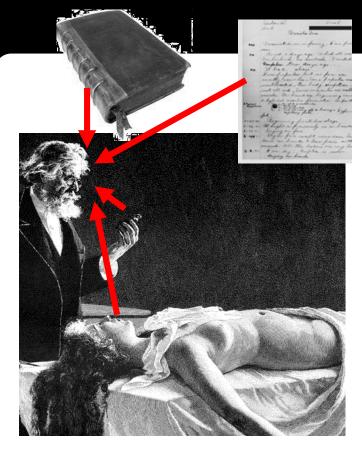
Datenexplosion: Experimentalphysik



CERN: 30 000 000 000 000 000 Byte = 30 Petabytes pro Jahr

http://home.web.cern.ch/about/computing

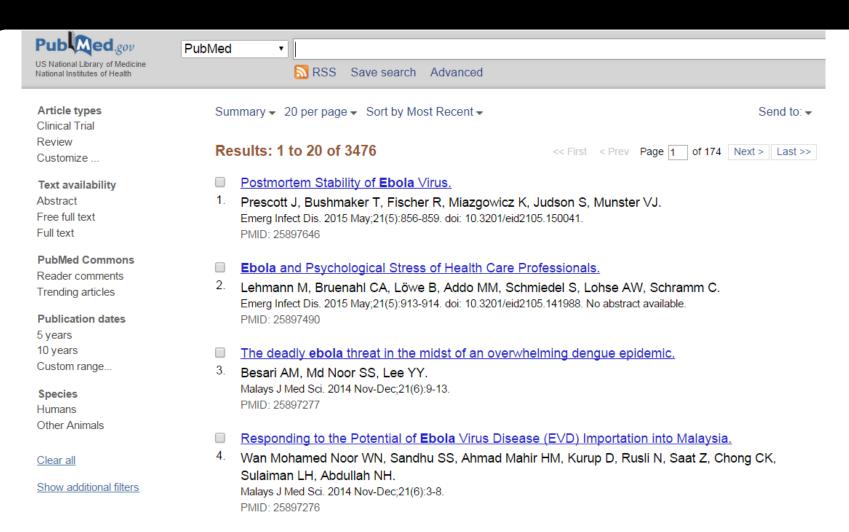
Informations explosion: Medizin





80 Megabytes pro Patient und Jahr (Beth Israel Deaconess Medical Center, 2011)

Wissensexplosion Fachliteratur



Datenbank MEDLINE: 2,5 Millionen Publikationen pro Jahr

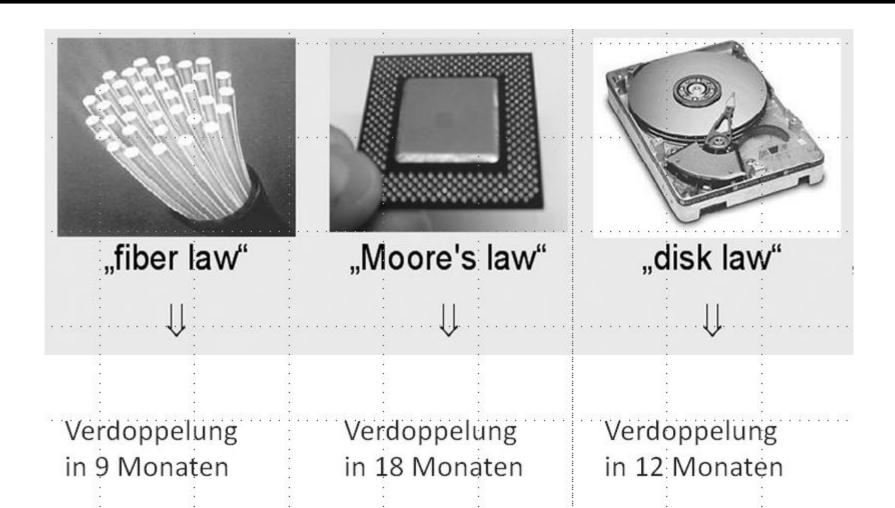
http://www.ncbi.nlm.nih.gov/

Wissensexplosion Biochemie

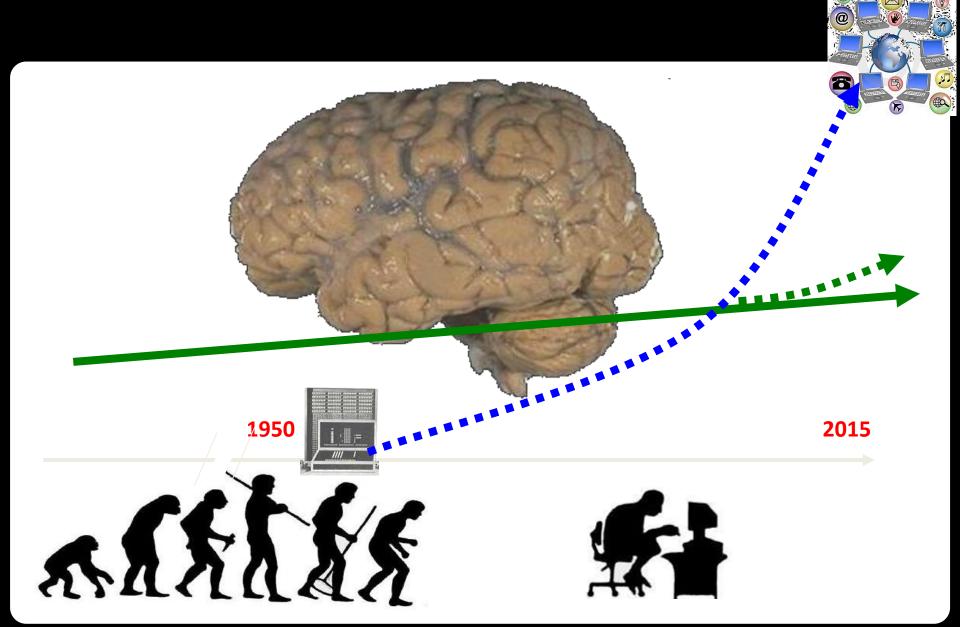
Entry 🗢	Entry name 🕏		Protein names 🗣 💟	Gene names 🕏	Organism ♦	Length 🕏
P04626	ERBB2_HUMAN	≜	Receptor tyrosine- protein kinase er	ERBB2, HER2, MLN19, NEU, NGL	Homo sapiens (Human)	1,255
Q03557	GATA_YEAST	*	Glutamyl- tRNA(Gln) amidotransferase	HER2, GEP6, LRC6, YMR293C	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	464
Q90464	Q90464_DANRE		HER-2 protein	her2, HER-2	Danio rerio (Zebrafish) (Brachydanio rerio)	108
P34708	TRA1_CAEEL	<u>F</u>	Sex-determining transformer protein	tra-1 , her-2, Y47D3A.6	Caenorhabditis elegans	1,110
Q75D84	GATA_ASHGO	A	Glutamyl- tRNA(Gln) amidotransferase	HER2, ABR140C	Ashbya gossypii (strain ATCC 10895 / CBS 109.51 / FGSC 9923 / NRRL Y-1056) (Yeast) (Eremothecium gossypii)	463
Q5AK64	GATA_CANAL	*	Glutamyl- tRNA(Gln) amidotransferase	HER2 , CaO19.11438, CaO19.3956	Candida albicans (strain SC5314 / ATCC MYA-2876) (Yeast)	450
C4YRY0	GATA_CANAW	À	Glutamyl- tRNA(Gln) amidotransferase	HER2, CAWG_04838	Candida albicans (strain WO-1) (Yeast)	450
Q6C0M4	GATA_YARLI	<u>F</u>	Glutamyl- tRNA(Gln) amidotransferase	HER2, YALI0F23441g	Yarrowia lipolytica (strain CLIB 122 / E 150) (Yeast) (Candida lipolytica)	459
Q6CXA9	Q6CXA9_KLULA		Glutamyl- tRNA(Gln) amidotransferase	HER2 , KLLA0A09845g, KLLA0_A09845g	Kluyveromyces lactis (strain ATCC 8585 / CBS 2359 / DSM 70799 / NBRC 1267 / NRRL Y-1140 / WM37) (Yeast) (Candida sphaerica)	469

UniProtKB: 47 Millionen Datensätze http://www.uniprot.org/

Maschinelle Performanz



Menschliche Performanz



Wissensexplosion Biochemie

UniProtKB

Swiss-Prot (548,208)

Manually annotated and reviewed.

TrEMBL (46,714,516)

Automatically annotated and not reviewed.

_	_		A [53]	_		
Entry 🕏	Entry name 🗣		Protein names 🕈 💟	Gene names ₹	Organism 🗣	Length ♦
P04626	ERBB2_HUMAN		Receptor tyrosine- protein kinase er	ERBB2, HER2, MLN19, NEU, NGL	Homo sapiens (Human)	1,255
Q03557	GATA_YEAST	F	Glutamyl- tRNA(Gln) amidotransferase	HER2, GEP6, LRC6, YMR293C	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	464
Q90464	Q90464_DANRE	B	HER-2 protein	her2, HER-2	Danio rerio (Zebrafish) (Brachydanio rerio)	108
P34708	TRA1_CAEEL	<u></u>	Sex-determining transformer protein	tra-1 , her-2, Y47D3A.6	Caenorhabditis elegans	1,110
Q75D84	GATA_ASHGO	<u>~</u>	Glutamyl- tRNA(Gln) amidotransferase	HER2, ABR140C	Ashbya gossypii (strain ATCC 10895 / CBS 109.51 / FGSC 9923 / NRRL Y-1056) (Yeast) (Eremothecium gossypii)	463
Q5AK64	GATA_CANAL	<u></u>	Glutamyl- tRNA(Gln) amidotransferase	HER2, CaO19.11438, CaO19.3956	Candida albicans (strain SC5314 / ATCC MYA-2876) (Yeast)	450
C4YRY0	GATA_CANAW	<u>*</u>	Glutamyl- tRNA(Gln) amidotransferase	HER2, CAWG_04838	Candida albicans (strain WO-1) (Yeast)	450
Q6C0M4	GATA_YARLI	P	Glutamyl- tRNA(Gln) amidotransferase	HER2, YALI0F23441g	Yarrowia lipolytica (strain CLIB 122 / E 150) (Yeast) (Candida lipolytica)	459
Q6CXA9	Q6CXA9_KLULA	L	Glutamyl- tRNA(Gln) amidotransferase	HER2 , KLLA0A09845g, KLLA0_A09845g	Kluyveromyces lactis (strain ATCC 8585 / CBS 2359 / DSM 70799 / NBRC 1267 / NRRL Y-1140 / WM37) (Yeast) (Candida sphaerica)	469

UniProtKB: 47 Millionen Datensätze http://www.uniprot.org/

Was ist Wissen?



Was ist Wissen?



Aspekte des Wissens



"dog" "perro", "chien", "canis", Hund" "Marley"

Denotation

Symbole

Individuen

C. K. Ogden and I. A. Richards (1923) The Meaning of Meaning

Aspekte des Wissens

Hunde sind Wirbeltiere

Universalien

Hunde können Tollwut übertragen

"Hund" und
"perro" sind
Synonyme

"Hund" ist ein Substantiv

3

Denotation

Symbole

Instantifierup

Marley ist ein Hund

2

Marley lebt in Florida

1

Individuen

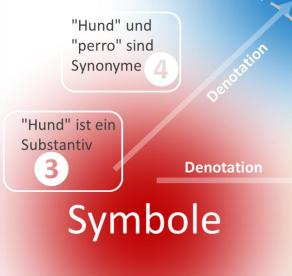
Ontologisches Wissen:

Was ausnahmslos als wahr betrachtet wird

Kontingentes Wissen:

Was typisch, wahrscheinlich oder möglich ist







Sprachliches Wissen:

Aussagen über Aufbau und Bedeutung sprachlicher Zeichen

Faktenwissen:

Aussagen über konkrete Dinge und ihre Beziehungen



1 Relation zwischen Individuen

Datum	Heim-Mannschaft	Gast-Mannschaft	Ergebnis
16. September 2014	AS Monaco	Bayer 04 Leverkusen	1:0
16. September 2014	Benfica Lissabon	Zenit St. Petersburg	0:2
01. Oktober 2014	Zenit St. Petersburg	AS Monaco	0:0
01. Oktober 2014	Bayer 04 Leverkusen	Benfica Lissabon	3:1
22. Oktober 2014	Bayer 04 Leverkusen	Zenit St. Petersburg	
22. Oktober 2014	AS Monaco	Benfica Lissabon	
04. November 2014	Zenit St. Petersburg	Bayer 04 Leverkusen	
04. November 2014	Benfica Lissabon	AS Monaco	
26. November 2014	Zenit St. Petersburg	Benfica Lissabon	
26. November 2014	Bayer 04 Leverkusen	AS Monaco	
09. Dezember 2014	AS Monaco	Zenit St. Petersburg	

Hunde sind 5 Wirbeltiere

Universalien

otation

nstantii

Marley ist ein Hund

2

Marley lebt in Florida

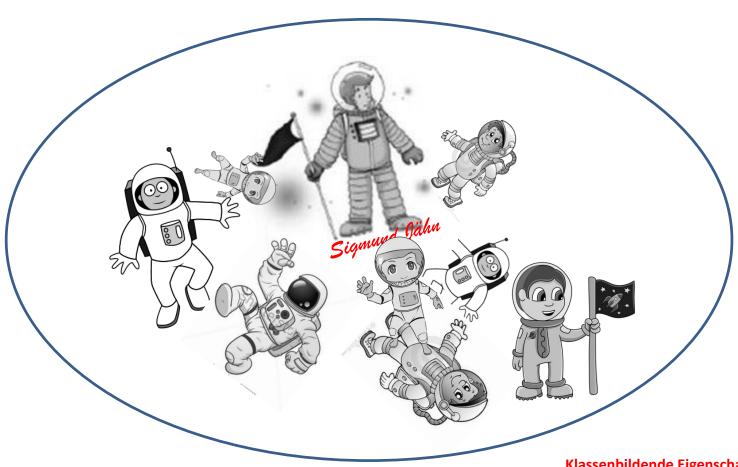


Individuen

2 Relation Individuum - Universale



2 Relation Individuum - Universale



Klasse: Astronaut

Klassenbildende Eigenschaften

- Lebender Mensch
- Technische Ausbildung
- Mindestens ein Aufenthalt im Weltraum

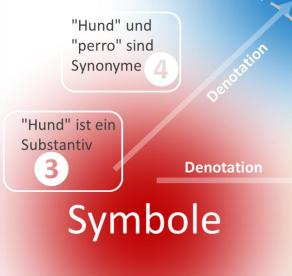
Ontologisches Wissen:

Was ausnahmslos als wahr betrachtet wird

Kontingentes Wissen:

Was typisch, wahrscheinlich oder möglich ist





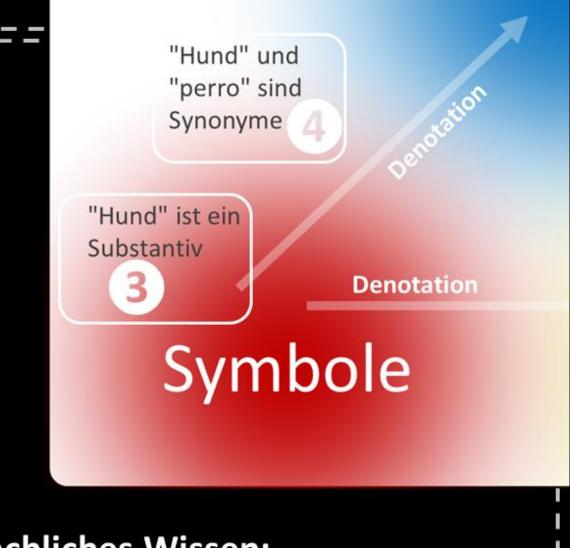


Sprachliches Wissen:

Aussagen über Aufbau und Bedeutung sprachlicher Zeichen

Faktenwissen:

Aussagen über konkrete Dinge und ihre Beziehungen



Sprachliches Wissen:

Aussagen über Aufbau und Bedeutung sprachlicher Zeichen

3 Sprachliches Wissen:

WordNet Search - 3.1

http://wordnetweb.princeton.edu/

- WordNet home page - Glossary - Help

Word to search for: dog	Search WordNet				
Display Options: (Select option to change) ▼	Change				
Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations					
Display options for sense: (gloss) "an example sentence"					

Noun

- S: (n) dog, domestic dog, Canis familiaris (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) frump, dog (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) dog (informal term for a man) "you lucky dog"
- S: (n) cad, bounder, blackguard, dog, hound, heel (someone who is morally reprehensible) "you dirty dog"
- S: (n) frank, frankfurter, hotdog, hot dog, dog, wiener, wienerwurst, weenie (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) pawl, detent, click, dog (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) andiron, firedog, dog, dog-iron (metal supports for logs in a fireplace) "the andirons were too hot to touch"

Verb

 S: (v) chase, chase after, trail, tail, tag, give chase, dog, go after, track (go after with the intent to catch) "The policeman chased the mugger down the alley"; "the dog chased the rabbit"

4 Thesaurus



building services

NT1 air conditioning

NT1 building insulation

RT insulator [6811]

NT2 sound insulation

RT insulator [6811]

RT noise [5216]

RT noise protection [5206]

NT2 thermal insulation

RT energy saving [6606]

RT insulator [6811]

RT thermal energy [6626]

SPRACHLICHE ÄQUIVALENTE

BG изолатор

ES aislante

CS izolátor

DA isolerende stof

DE Isolierstoff

ET isolaator

EL μονωτικό

EN insulator

FR isolant

HR izolator

IT isolante

LV izolācijas materiāls

LT izoliacinė medžiaga

HU szigetelőanyag

MT iżolant

NL isolerende stof

PL materiał izolacyjny

PT isolante

RO izolant

SK izolant

SL izolator

FI eriste

SV isolator

SR изолатор

МК изолатор

SQ izolant

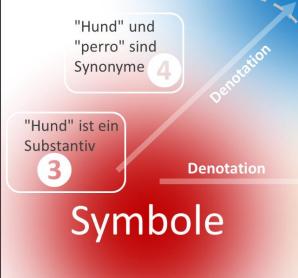
Ontologisches Wissen:

Was ausnahmslos als wahr betrachtet wird

Kontingentes Wissen:

Was typisch, wahrscheinlich oder möglich ist





Marley ist ein Hund

Marley lebt in Florida

1

Individuen

Sprachliches Wissen:

Aussagen über Aufbau und Bedeutung sprachlicher Zeichen

Faktenwissen:

Aussagen über konkrete Dinge und ihre Beziehungen



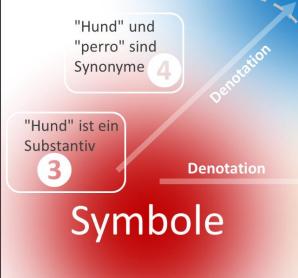
Ontologisches Wissen:

Was ausnahmslos als wahr betrachtet wird

Kontingentes Wissen:

Was typisch, wahrscheinlich oder möglich ist





Marley ist ein Hund

Marley lebt in Florida

1

Individuen

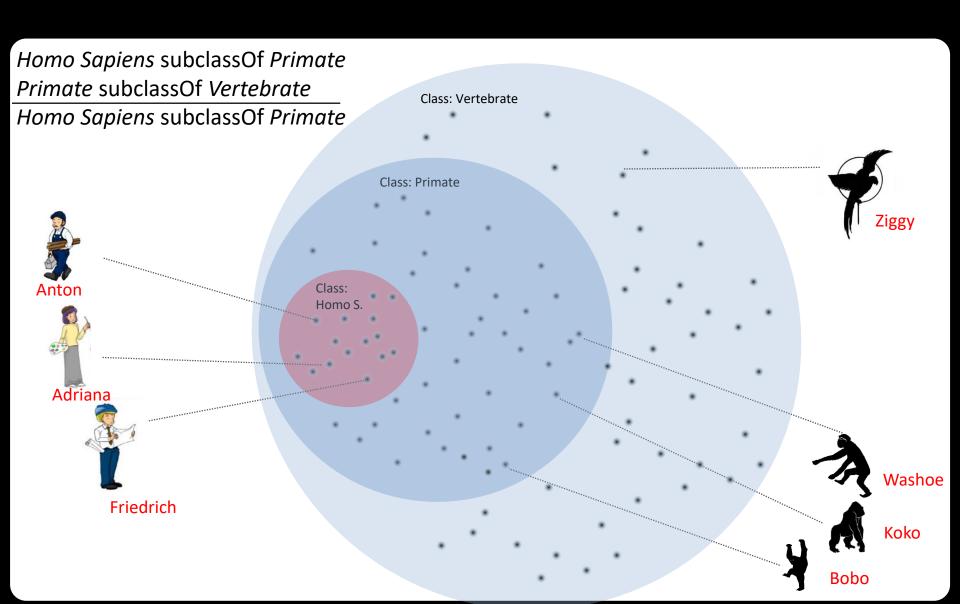
Sprachliches Wissen:

Aussagen über Aufbau und Bedeutung sprachlicher Zeichen

Faktenwissen:

Aussagen über konkrete Dinge und ihre Beziehungen

5 Taxonomien



Übung: Taxonomie

- Clinical Medicine
- - Oncology
- --- Cancer
- ---- Lung
- --- Breast
- --- Prostate
- --- Colon

- Geography
- - Countries
- --- BRIC Countries
- --- Brazil
- --- Russia
- --- India
- --- China

Übung: Korrektur

- Clinical Medicine
- - Oncology
- --- Cancer
- ---- Lung
- --- Breast
- --- Prostate
- --- Colon

- Clinical Disease
- - Oncologic Disease
- --- Cancer
- --- Lung Cancer
- --- Breast Cancer
- --- Prostate Cancer
- --- Colon Cancer

Übung: Übung: Korrektur

- Geography
- - Countries
- --- BRIC Countries
- --- Brazil
- --- Russia
- --- India
- --- China

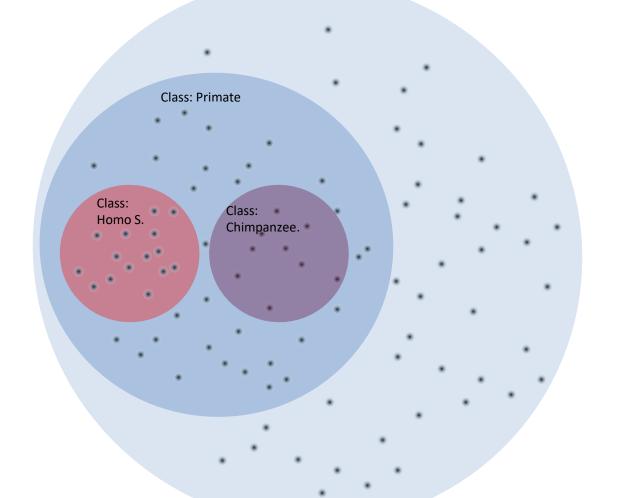
- Geographical Entity
- - Country
- --- BRIC Country

(Members: Brazil,

Russia, India, China)

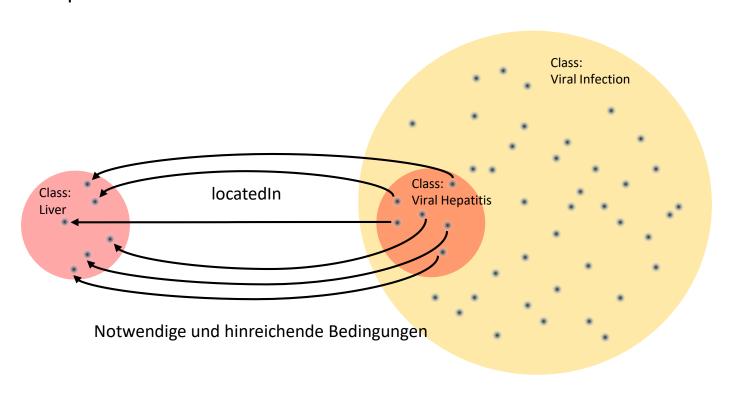
5 Partitionen

Homo Sapiens subclassOf not (Chimpanzee)



5 Aristotelische Definitionen

Viral Hepatitis equivalentTo Viral Infection and locatedIn some Liver



Kontingentes Wissen: Was typisch, wahrscheinlich oder möglich ist

Hunde können Tollwut übertragen

6 Kontingentes Wissen



- Vioxx® lindert Schmerzen
- Vioxx® verursacht Herzinfarkt
- Vioxx[®] ist Mittel der Wahl bei rheumatischer Arthritis (2002)
- Vioxx® ist als Arzneimittel verboten (2005)
- Vögel können fliegen
- Pinguine sind Vögel
- Pinguine können nicht fliegen





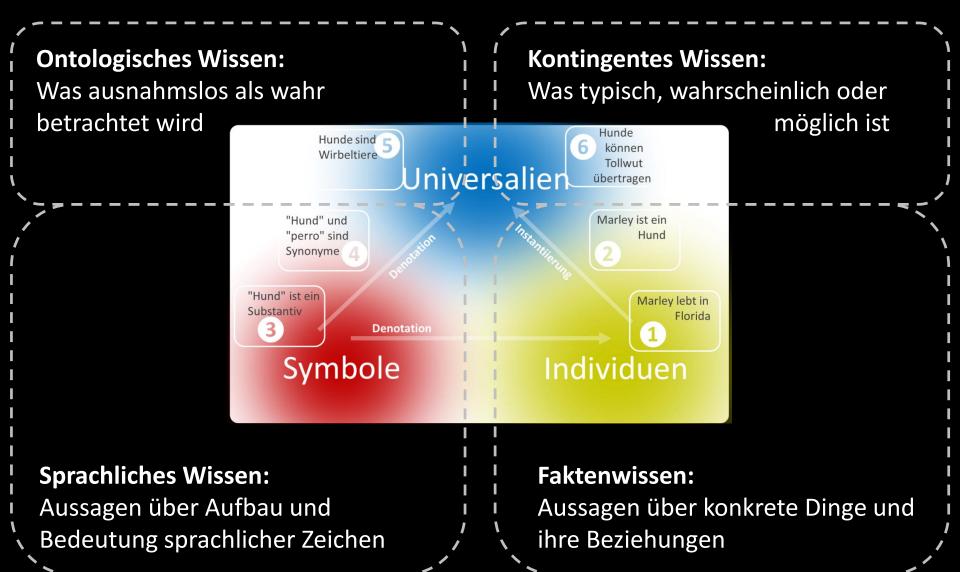
- Österreich ist kein Erdbebengebiet
- 1972 schweres Erdbeben in Seebenstein
- Seebenstein liegt in NÖ
- NÖ liegt in Österreich

Altkanzler Helmut Schmidt bunkert 38.000 Menthol-Zigaretten



- Rauchen verursacht Lungenkrebs Helmut Schmidt raucht seit 85 Jahren
- Helmut Schmidt hat keinen Lungenkrebs

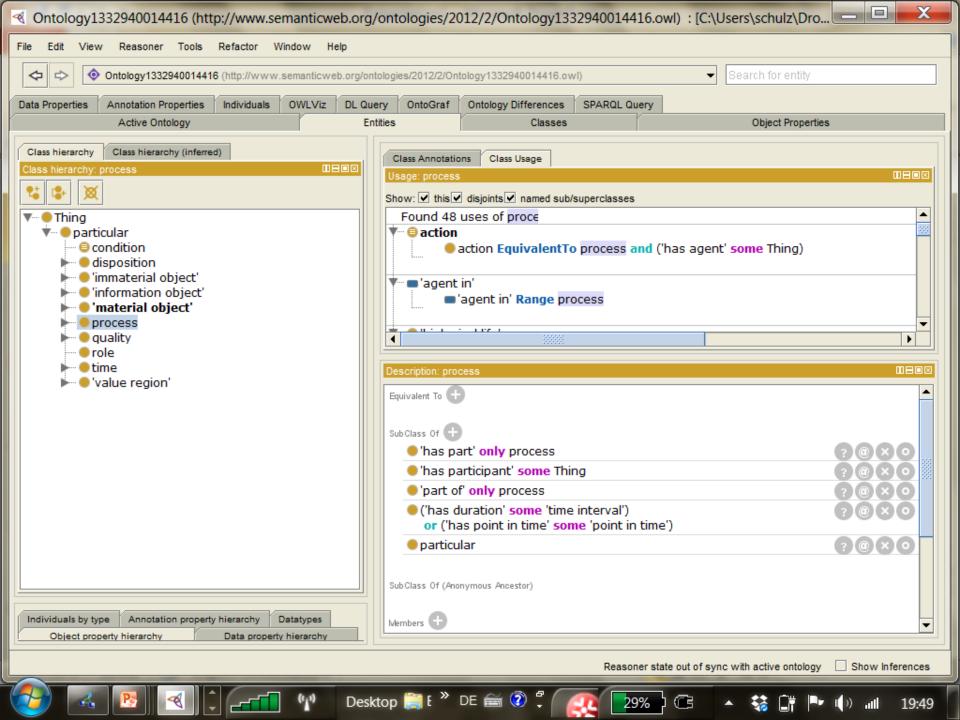
Welches Wissen ist nachhaltig?

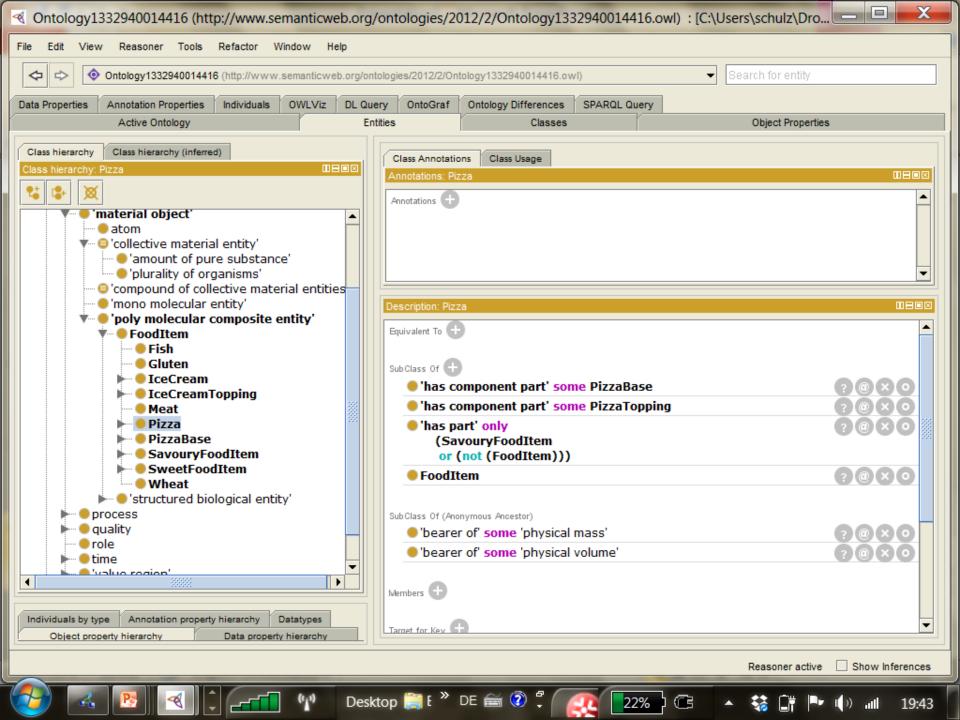


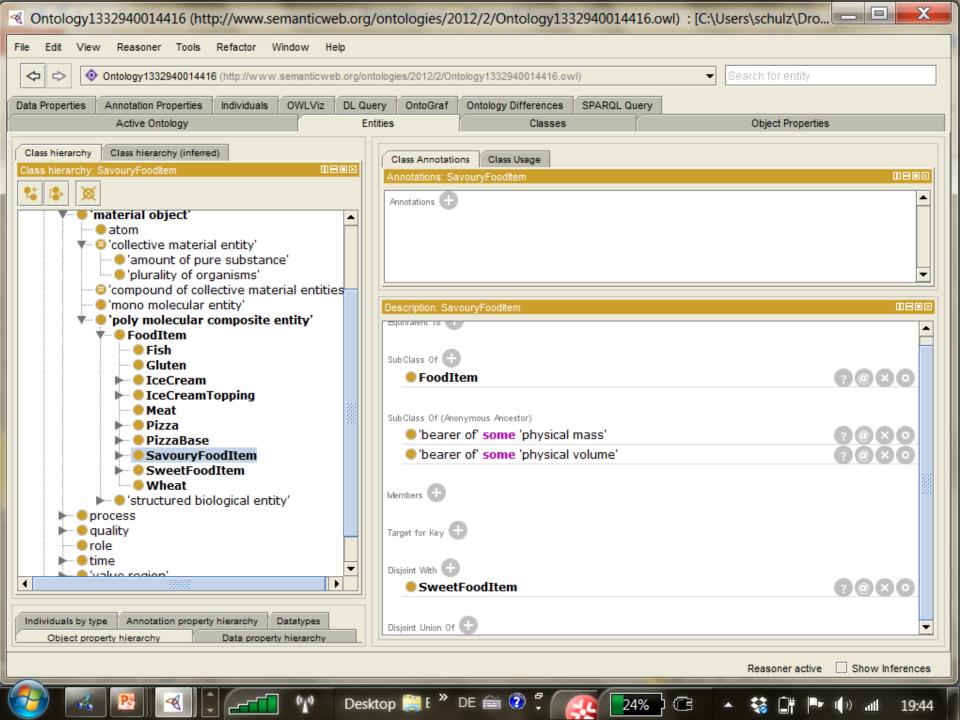


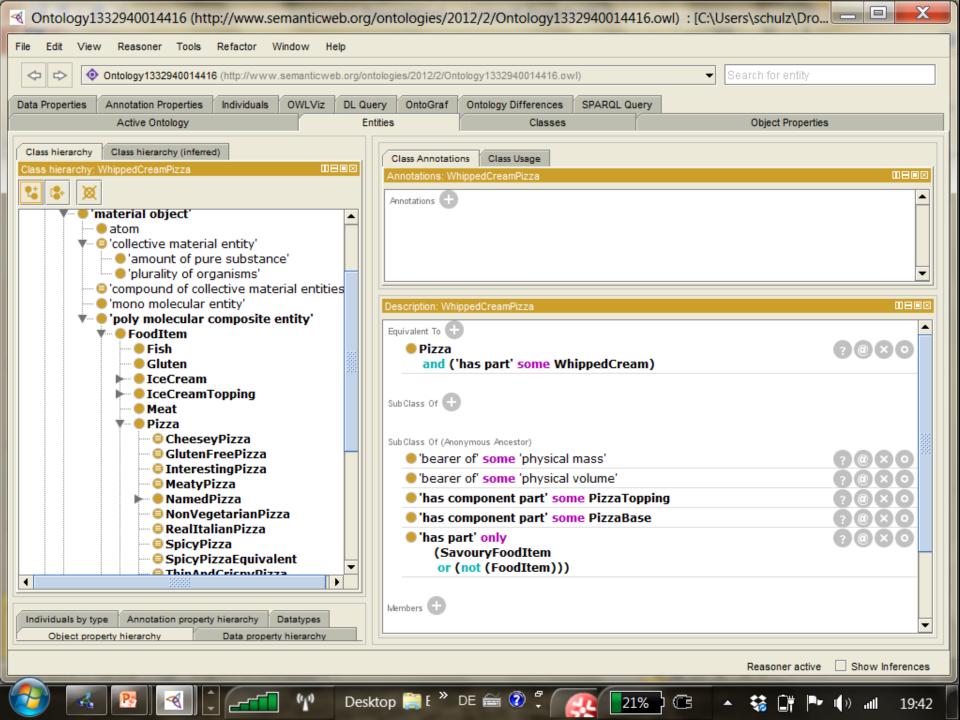
Demonstration: Pizza-Ontologie

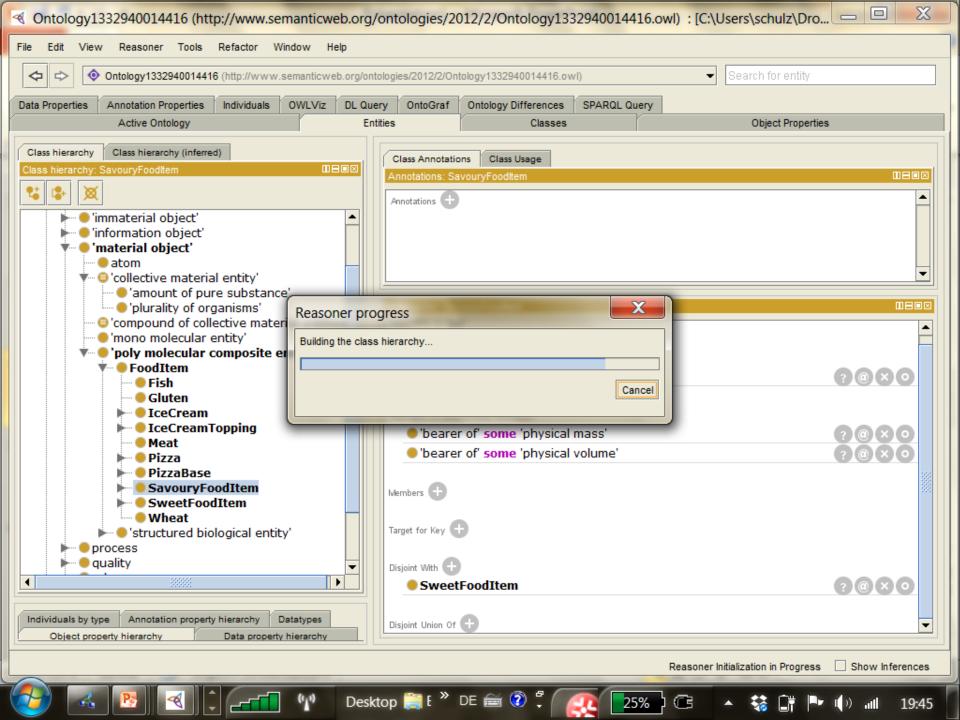
	BAUS hgehend bis 100 Uh		v.diebausatzlokale 5 Tage im Jah	
○ große Pizz○ kleine Pizz○ Minipizza AB	AB)CO)E(F)G]HILMINOOP(R)	omatensauce + Oregano + Käse (kein Analogkäse) € 7,0 + 4 Zutaten € 5,3 Jede weitere Zutat € 0,30 Bitte bei der C R E W abgeben! € 4,1		
echt. österr. Schinken (kein Pizzablock) Salami scharfe Salami ital. Prosciutto Speck Cayennepfeffer Curry Thunfisch delphinschonend gefange Meeresfrüchte Kapern	Champignons frisch Briekäse Kürbiskerne milde Pfefferoni scharfe Pfefferoni	Artischocken Ananas Oliven Feta Schafskäse Mozzarella Parmesan Blauschimmel- käse Gorgonzola Kernöl Sonnenblumen- kerne österr. Ei BIO Freiland	 ☐ Indianerbohnen ☐ Käferbohnen ☑ Schlagobers ☐ Spinat ☐ Broccoli ☐ Lauch ☐ Paprika frisch ☐ Knoblauch ☐ Tabasco ☐ Piri-Piri 	Schnecken Zucchini Öst. BIOTofu geräuchert Chili Pulver Tomatenpesto Kürbiskernpesto Basilikumpesto Nutella Rucola frisch Gummibärli

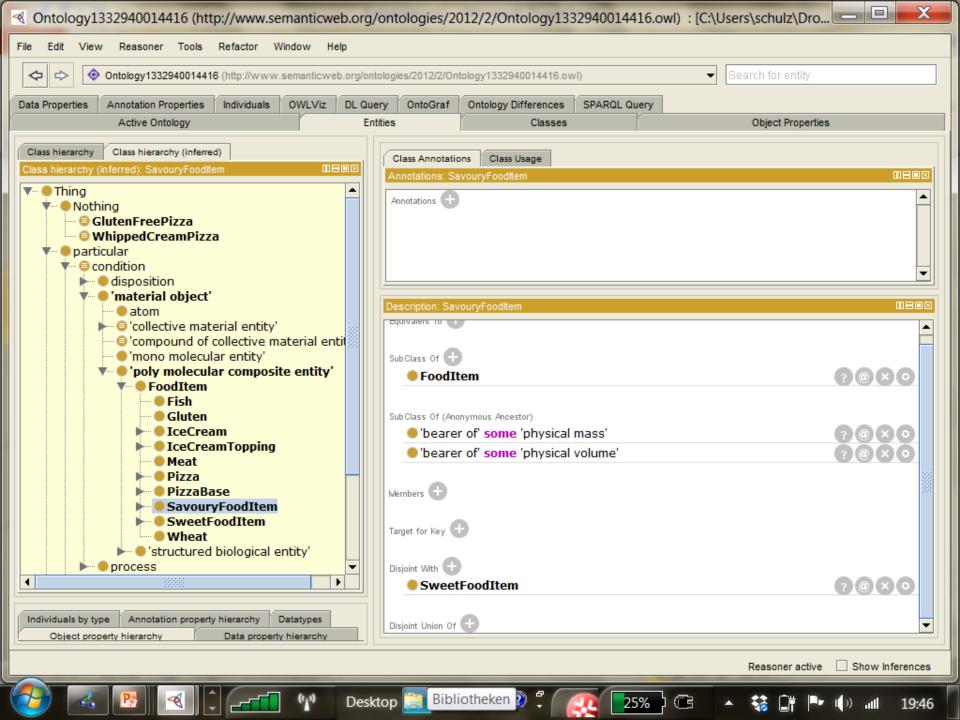


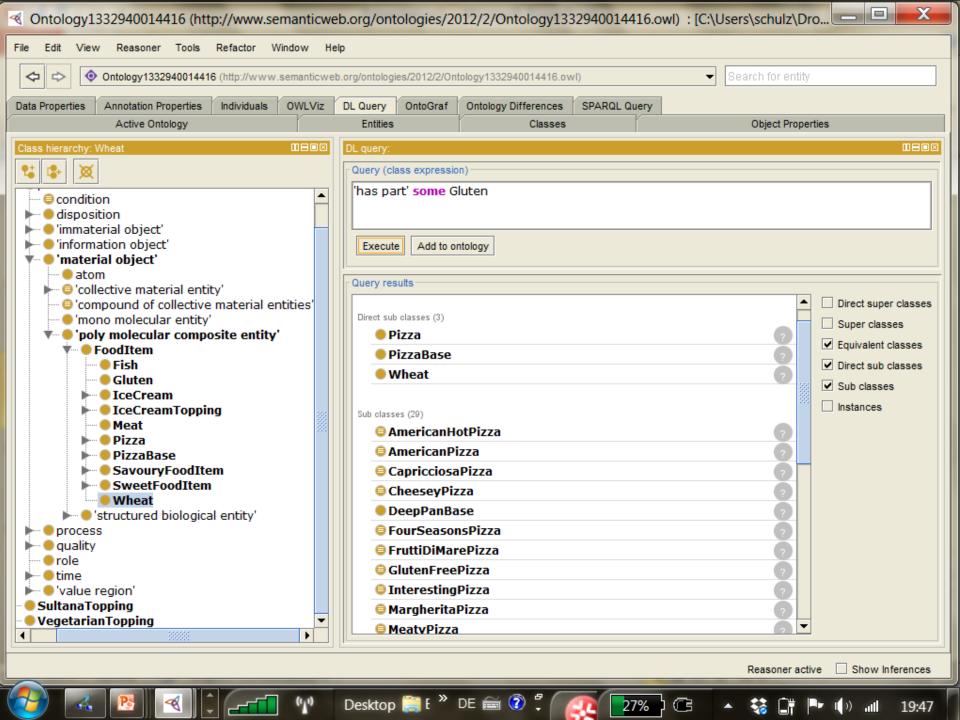


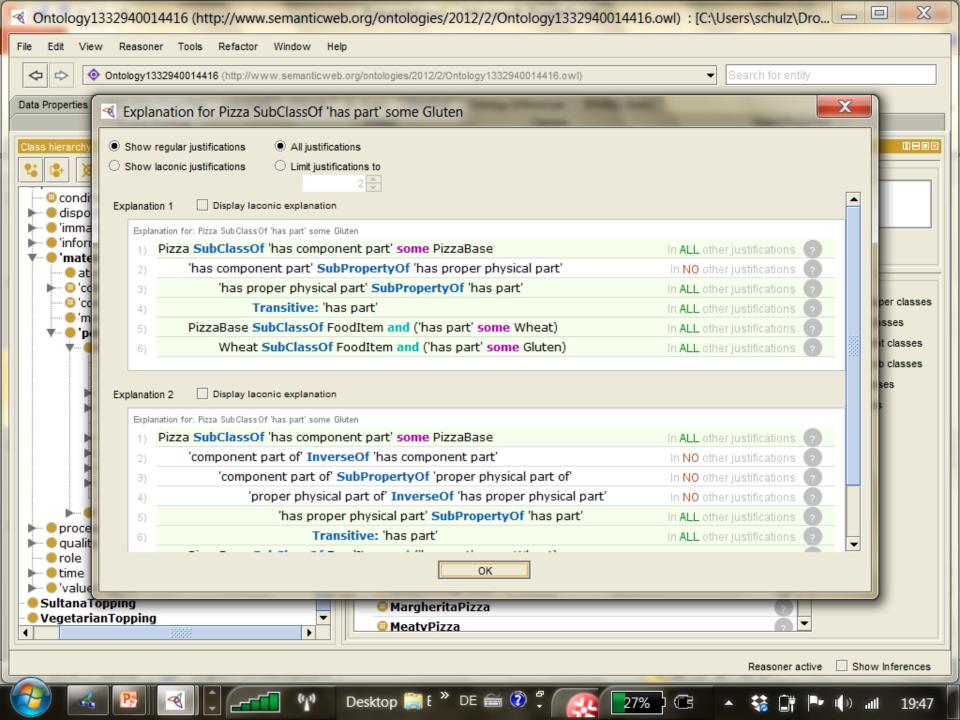












SNOMED CT

- Repräsentiert Inhalte der medizinischen Dokumentation
- Ca. 300.000 bedeutungstragende Element ("Konzepte")
- Terme in diversen Sprachen

SNOMED CT als formales System

Parent(s):

(Select a parent to make it the "Current Concept".)

Disorder of appendix (disorder)

Inflammation of large intestine (disord/

Current Concept: Appendicitis (disorder)

Child(ren):

(N=14) (Select a child to make it the "Current Concept".)

There are 5 Retired Children. Show Retired Children

Acute appendicitis (disorder)

Amebic appendicitis (disorder)

Appendicitis of a pelvic appendix (disorder)

Atypical appendicitis (disorder)

Catarrhal appendicitis (disorder)

Chronic appendicitis (disorder)

Complicated appendicitis (disorder)

Focal appendicitis (disorder)

Hierarchien:

Strikte

Spezialisierung

(is-a)

arrent Concept:

ully Specified Name: Appendicitis (disorder)

onceptId: 74400008

Defining Relationships:

Is a Disorder of appendix (disorder)

Inflammation of large intestine (disorder) Is a

Group 1

Associated morphology (attribute)

Inflammation (morphologic abnormality) Finding site (attribute) Appendix structure (body structure)

This concept is fully defined.

Oualifiers:

View Qualifying Characteristics and Facts

Descriptions (Synonyms):

Preferred: Appendicitis:

Fully Specified Name: Appendicitis (disorder)

Synonym: Appendicitis, NOS

SNOMED CT als formales System

Parent(s):

(Select a parent to make it the "Current Concept".)

Disorder of appendix (disorder)

Inflammation of large intestine (disorder)

Current Concept:

Appendicitis (disorder)

Child(ren):

(N=14) (Select a child to make it the "Current Concept".)

There are 5 Retired Children. Show Retired Children

Acute appendicitis (disorder)

Amebic appendicitis (disorder)

Appendicitis of a pelvic appendix (disp

Restriktionen: auf einfacher Beschreibungslogik beruhend:

C1 - Rel - C2 zu interpretieren als:

 $\forall x$: instanceOf(x, C1) \Rightarrow

 $\exists y$: instanceOf(C2) \land Rel(x,y)

Current Concept:

Fully Specified Name: Appendicitis (disorder)

74400008

ConceptId:

Defining Relationships:

Disorder of appendix (disorder) Is a

Is a Inflammation of large intestine (disorder)

Group 1

Associated morphology (attribute)

Inflammation (morphologic abnormality) Finding site (tribute) Appendix structure (body structure)

This concept is full

Qualifiers:

View Qualifying Ch

Descriptions (S

Preferred:

Fully Specified Name

Synonym:

Relationen (Attribute): z.B.

Associated morphology

Finding site

(50 Relationstypen)

SNOMED CT als formales System

Current Concept:

Fully Specified Name: Entire upper limb (body structure)
ConceptId: 182245002

Defining Relationships:

Is a Entire limb (body structure)

Is a Upper limb structure (body structure)

This concept is primitive.

definierte vs. primitive Konzepte

Current Concept:

Fully Specified Name: Appendicitis (disorder)

ConceptId: 74400008

Defining Relationships:

Is a Disorder of appendix (disorder)

Is a Inflammation of large intestine (disorder)

Group 1

Associated morphology (attribute)

Finding site (attribute)

Inflammation (morphologic abnormality)
Appendix structure (body structure)

This concept is fully defined.

Qualifiers:

View Qualifying Characteristics and Facts

Descriptions (Synonyms):

Preferred: Appendicitis

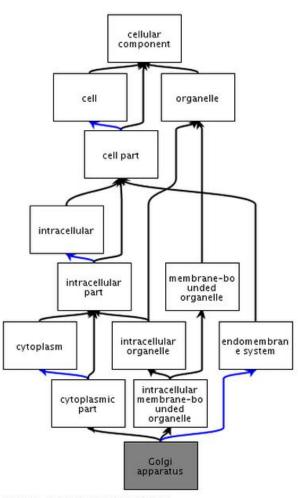
Vorteile von SNOMED CT

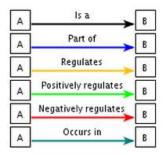
- Standardisierte Beschreibung medizinischer Sachverhalte
- Präziseres Wiederfinden von Krankheitsfällen "Diabetiker mit Parkinson, die keine Dopaminagonisten einnehmen und als Fahrzeuglenker bei einem Unfall verletzt wurden
- Automatisierte Entscheidungsunterstützung
- Automatische Wissensakquisition aus PatientInnendaten

Andere biomedizinischen Ontologien

- ChEBI: Chemische Entitäten
- GO (Gene Ontology)
 - Cell Component
 - Molecular Function
 - Biological Process
- FMA: Foundational Model of Anatomy

Beispiel Gene Ontology





QuickGO - http://www.ebi.ac.uk/QuickGO

Gesellschaftliche Aspekte

- Nachhaltige Wissensinhalte
 - Inhalte, die sich nicht verändern, und daher als stabiles Fundament für Wissenskonstruktionen dienen
 - Wissensquellen, die allgemein zugänglich sind
- Wissensinhalte, die durch Partikularinteressen abgeschirmt sind, stehen der nachhaltigen Nutzung im Weg

Beispiel terminologische / ontologische Standards

SNOMED CT:

Internationaler Standard, verwaltet von gemeinnütziger

NGO (IHTSDO)

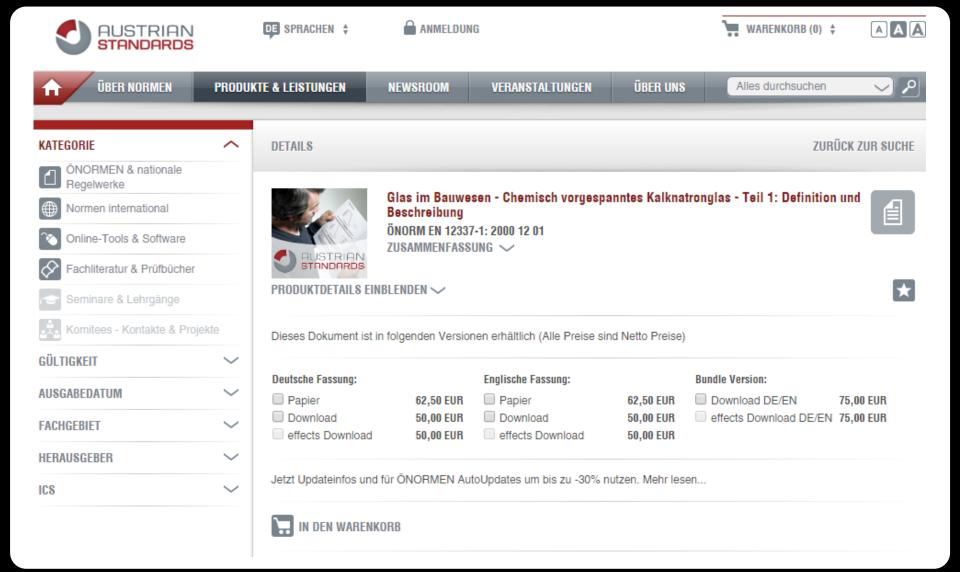
Routinenutzung lizenzpflichtig

- Affiliate Licence
- Mitgliedschaft (auf nationaler Ebene)
- Österreich ca.200.000 €/Jahr



Österreich ist bisher kein Mitglied, Einführung fraglich

Beispiel Industriestandards



Diskussion

- Vorteile genormter Ontologien und Terminologien:
 - Stabile Benennungen
 - Logische Begriffsdefinitionen, präziser als sprachliche Definitionen
 - Genaue Abgrenzung "fließender" Begriffe
 - "Schokoladenpizza": Kuchen oder Pizza?
 - 13-jähriger: Kind oder Jugendlicher
 - Interoperabilität über Disziplinen und Sprachen hinweg
 - Widersprüche können maschinell aufgedeckt werden
- Plädoyer: Genormte Ontologien und Terminologien als Gemeingut zu behandeln

