



Biomedical and Health Informatics
From Foundations to Applications to Policy
November 8-12, 2008 Washington, DC



Evaluation of a Document Search Engine in a Clinical Department System

Stefan Schulz¹, Philipp Daumke², Pascal Fischer³, Marcel Müller³

¹Institute of Medical Biometry und Medical Informatics, University Medical Center Freiburg

²AVERBIS GmbH, Freiburg, Germany

³Department of Dermatology, University Medical Center Freiburg, Germany

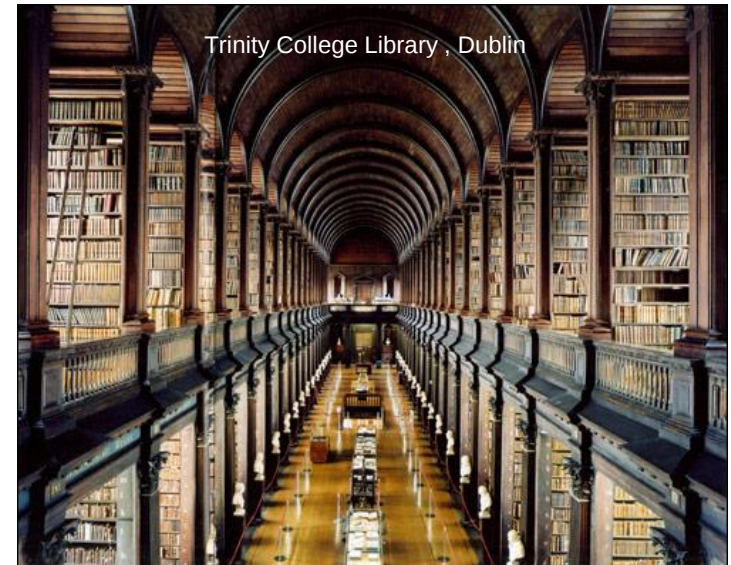
Navigation paradigms

Vertical Paradigm

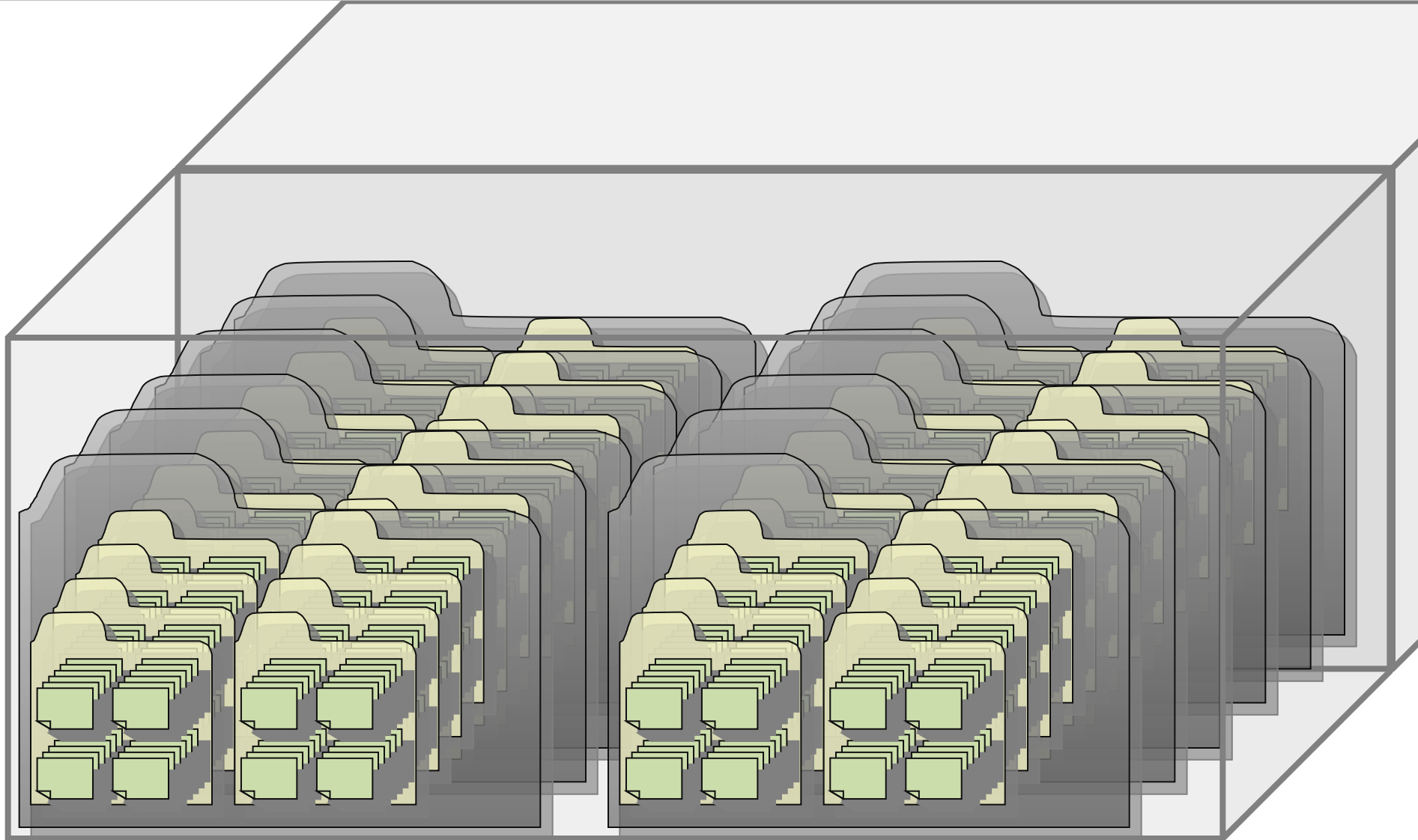


Vertical Paradigm

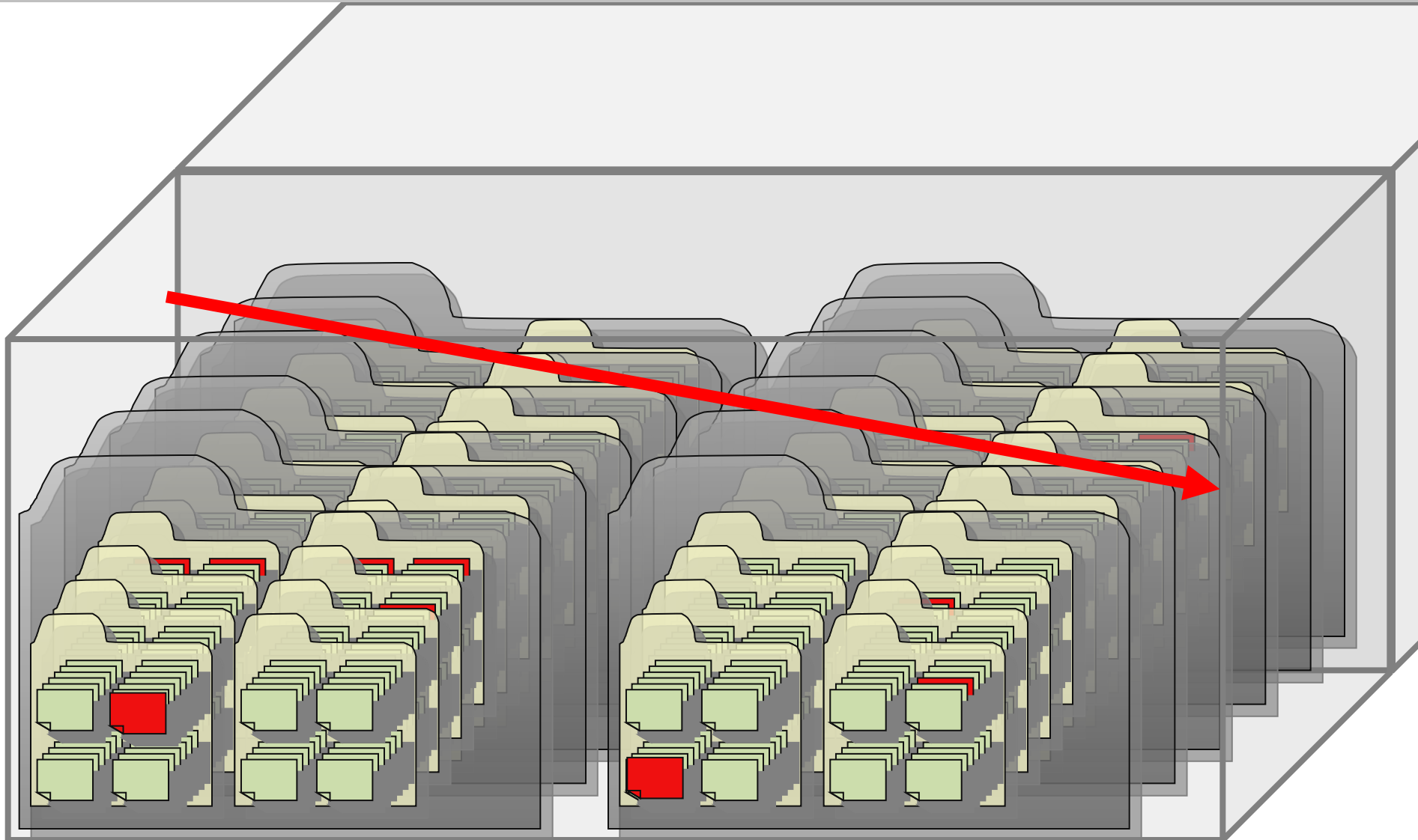
- Vertical data access:
 - hierarchical
 - ordered
 - systematic
 - complete



Horizontal Paradigm

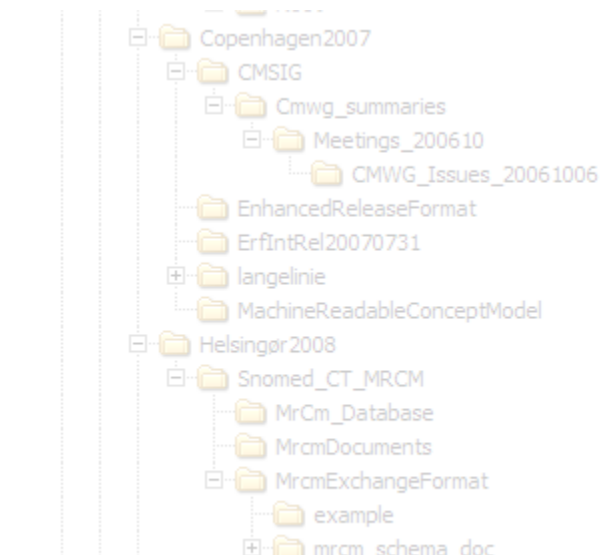


Horizontal Paradigm



Horizontal Paradigm

- Vertical data access:
 - hierarchical
 - ordered
 - systematic
 - complete
- Horizontal data access:
 - unsystematic
 - anarchic
 - incomplete



Google [Erweiterte Suche](#) [Einstellungen](#)

Web

Tipp: [Suchen nur nach Ergebnissen auf Deutsch](#). Sie können Ihre bevorzugten Spracheinstellungen in [Einstellungen](#) angeben.

[IHTSDO: International Health Terminology Standards Development ...](#) - [[Diese Seite übersetzen](#)]
 IHTSDO develops and promotes use of **SNOMED CT**, a comprehensive health and clinical terminology, so as to support safe and effective health information ...
[Karte von 325 WAUKEGAN RD, NORTHFIELD, IL 60093 einblenden](#)
[www.ihstso.org/](#) - 11k - [Im Cache](#) - [Ähnliche Seiten](#)

[IHTSDO: SNOMED CT](#) - [[Diese Seite übersetzen](#)]
 SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) is considered ...
 SNOMED CT aims to improve patient care through the development of systems ...
[www.ihstso.org/snomed-ct/](#) - 8k - [Im Cache](#) - [Ähnliche Seiten](#)

[SNOMED Clinical Terms® \(SNOMED CT®\)](#) - [[Diese Seite übersetzen](#)]
 SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) is a comprehensive clinical terminology, originally created by the College of American ...
[www.nlm.nih.gov/snomed/](#) - 12k - [Im Cache](#) - [Ähnliche Seiten](#)

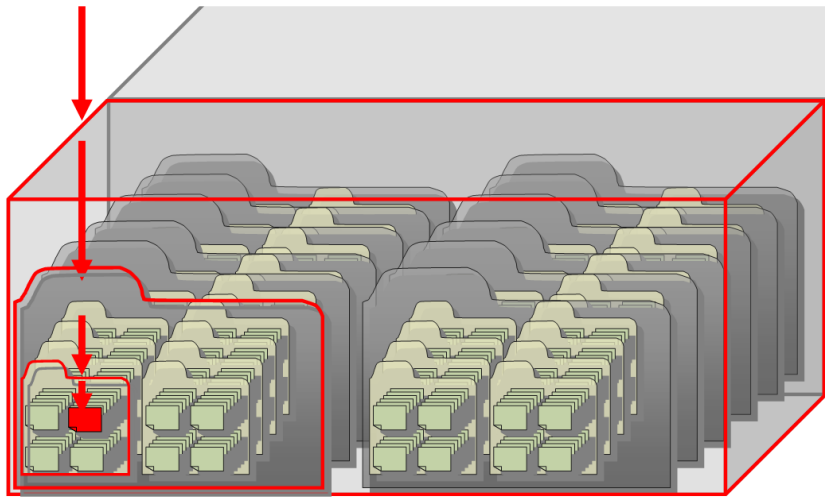
[SNOMED CT - Wikipedia, the free encyclopedia](#) - [[Diese Seite übersetzen](#)]
 SNOMED CT (Systematized Nomenclature of Medicine -- Clinical Terms), is a systematically organized computer processable collection of medical terminology ...
[en.wikipedia.org/wiki/SNOMED_CT](#) - 34k - [Im Cache](#) - [Ähnliche Seiten](#)

[Systematisierte Nomenklatur der Medizin - Wikipedia](#)
 Die fusionierte Nomenklatur wurde **SNOMED CT** (Clinical Terms) genannt. ... Seit April 2007 liegen die Rechte an **SNOMED CT** bei der International Healthcare ...
[de.wikipedia.org/wiki/SNOMED](#) - 26k - [Im Cache](#) - [Ähnliche Seiten](#)

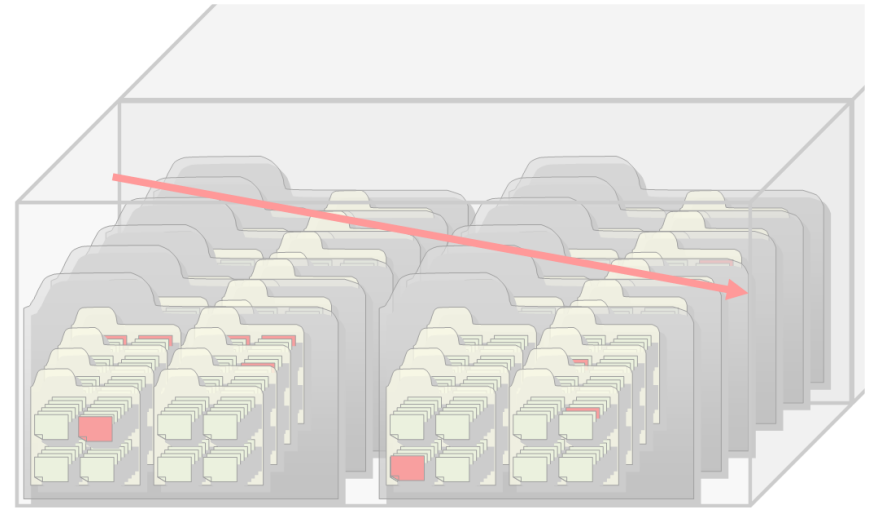
[International Health Terminology Standards Development Organisation](#) - [[Diese Seite übersetzen](#)]
 For information on **SNOMED** as related to the International Health Terminology Standards Development Organisation, please visit [www.ihstso.org](#). ...

Document access in the electronic health record

Document access in the electronic health record



vertical



horizontal

Document access in the electronic health record

- Vertical access appropriate for systematic patient / case related routines
- Horizontal access:
 - querying structured **data** across patients: for business statistics, clinical epidemiology
 - use case for querying unstructured **documents** across patients?
- Can users benefit for “googling” the EHR?

“Googling” medical narratives

- Word index is not enough
 - Term variations
 - Spelling variations
 - Synonyms
- Example: German
“Colon Cancer”
- Linguistic variations
(morphology, syntax, (morpho-
semantics) make (medical)
document retrieval difficult

Web search	# Hits	
	total	exclusive
Kolonkarzinom	2070	1780
Colonkarzinom	248	135
Colonicarcinom	111	73
Colon-Ca	203	169
Kolon-Ca	66	46
Dickdarmkrebs	4000	3610
Dickdarmkarzinom	288	175
Dickdarmcarcinom	13	10
Kolonkarzinoms	471	253
Kolonkarzinome	275	139
Kolonkarzinomen	265	166

Morphosemantic Document Indexing

Morphosemantic Document Indexing

- *“The true, significant elements of language are . . . either words, significant parts of words, or word groupings.” [Sapir 1921]*
- *Extract significant word fragments (subwords) and map them to semantic identifiers / concepts:*

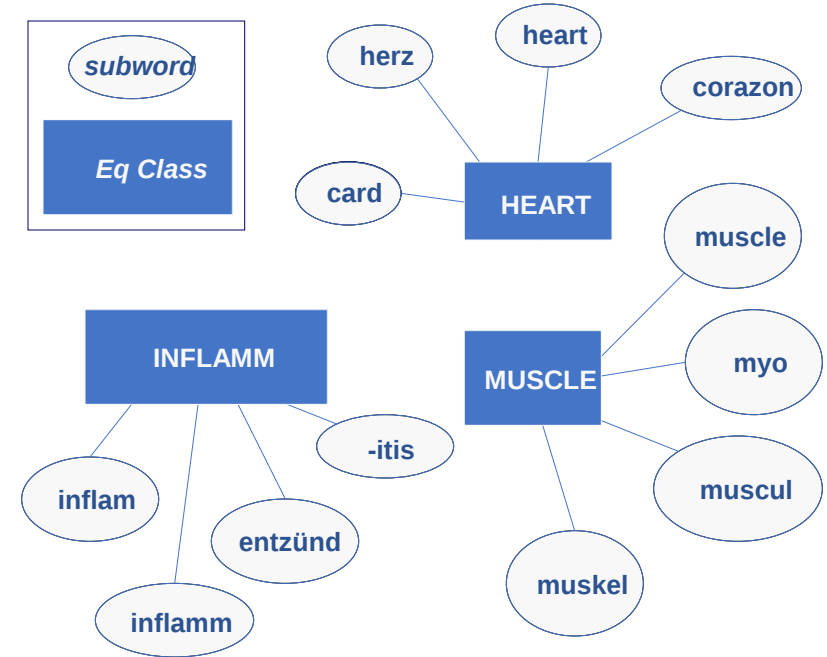


–#derm = { **derm**, **cutis**, **skin**, **haut**, **kutis**, **pele**, **cutis**, **piel**, ... }

–#inflamm = { **inflamm**, **-itic**, **-itis**, **-phlog**, **entzuend**, **-itis**,
-itisch, **inflam**, **flog**, **inflam**, **flog**, ... }

MorphoSaurus semantic indexer

- Thesaurus ~ 21.000 equivalence classes
- Lexicon entries:
 - English: ~23.000
 - German: ~24.000
 - Portuguese: ~15.000
 - Spanish : ~11.000
 - French: ~ 8.000
 - Swedish: ~10.000
 - Italian: ~ 4.000



Segmentation:

Myo | kard | itis

Herz | muskel | entzünd | ung

Inflamm | ation of the heart muscle

Indexation:

#muscle #heart #inflamm

#heart #muscle #inflamm

#inflamm #heart #muscle

Implementation

Implementation

- Department of Dermatology, University Medical Center, Freiburg
- Scope:
 - 30,000 German-language narratives (discharge letters, surgical reports and immuno-dermatologic findings)
 - Database of dermatological images
- Pilot users: 25 physicians and students

Arztbrief-Recherche Hautklinik Freiburg - suitSearch - Mozilla Firefox

DateiBearbeitenAnsichtChronikLesezeichenExtrasHilfe

Universitätsklinikum Freiburg - Abt. Medizinische Informatik & Universitäts-Hautklinik

SUITSEARCH[®]

HEALTH RECORDS

hauttumor

SucheZurücksetzen

PIZ:

Nachname:

Vorname:

Geschlecht:

Relevanz

10

14635 Ergebnisse in 24 msec gefunden. 100% Genauigkeit. --- Zeige Ergebnisse 81 bis 90: [Info](#)

Tipp: Externe Suche mit "hauttumor" in » ICD-10 » Literatur » Bilder

00000000

[Patient 1](#)

X

21.11.1922

♂

... der sich zwischen dem 13.10. und 9.11.2004 in unserer ambulanten Behandlung befand. Diagnosen: 1 **Basaliom** vom superfiziellen Typ Rücken links. 2 Basaliom vom superfiziellen Typ Rücken BWS links. 3 Basaliom vom solid-zystischen Typ Schulter links. Typ Brust rechts. Anamnese: Die Vorstellung des Patienten erfolgte wegen seit ca. einem Jahr bestehender scharf begrenzter erythemato-squamöser tumoröser **Hautveränderungen** im Bereich des Rückens und der Brust. Beurteilung/Therapie: Am 9.11.2004 erfolgte die Exzision von insgesamt vier tumorösen Hautveränderungen in oben ...
[Ersteller:](#) Dr. Andrea Schlöbe [Erstellung:](#) 24.11.2004

15186941

[Patient 2](#)

X

31.08.1944

♂

... der sich am 07.10. und 18.10.2005 in unserer Ambulanz vorstellte. Diagnose: **Keratoakanthom** Anamnese/Befund: Herr X berichtete über ein schnell wachsenden halbkugelförmigen hyperkeratotischen Tumor auf der linken Brust der sich im Verlauf der ...

... vulgaris zu denken. Der Tumor ist zu beiden Seiten in toto exzidiert. Beurteilung/Therapie: Wir exzidierten am 18.10.05 die verdächtige **Hautveränderung** in Lokalanästhesie. Die histologische Untersuchung lieferte das Bild eines epidermalen Tumors vereinbar mit dem Bild eines Keratoakanthoms. Der Tumor ist ...
[Ersteller:](#) Dr. Constance Huyke [Erstellung:](#) 11.11.2005

Stefan Schulz, Philipp Daumke, Pascal Fischer, Marcel Müller Evaluation of a Document Search Engine in a Clinical Department System

Web-Fotofinder - Universitäts-Hautklinik Freiburg - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Web-Fotofinder

Universitäts-Hautklinik Freiburg

Datum
04.12.2006 12:58:07 Uhr

Patient
Patient 1, geb. 31.08.1944 (Plz 00000000)

Benutzer
marcelm

Diese Webseite stellt Ihnen 90147 Photos der Universitäts-Hautklinik, welche in den Jahren 2000 bis 2006 erstellt und gespeichert wurden, zur Verfügung. Bitte beachten Sie, dass die Angaben zur Patientenidentifikation fehlerhaft sein können.



Datum: 30.12.2002 (Nr. 28814)
Lokalisation: KH

Datum: 19.01.2004 (Nr. 42184)
Lokalisation: KH

Evaluation Methodology

1. User satisfaction (questionnaire): Impact of horizontal EHR navigation on:
 - clinical performance
 - scientific performance
 - medical education
2. Clinical-epidemiological use case
 - select **suspected** or **manifest** syphilis cases based on information discharge summaries (coding unreliable)
 - typical terms: “*Syphilisverdacht*”, “*syphilitischer Primäraffekt*”, “*Luesserologie*”, ...
 - precision / recall analysis

Gold standard and baselines

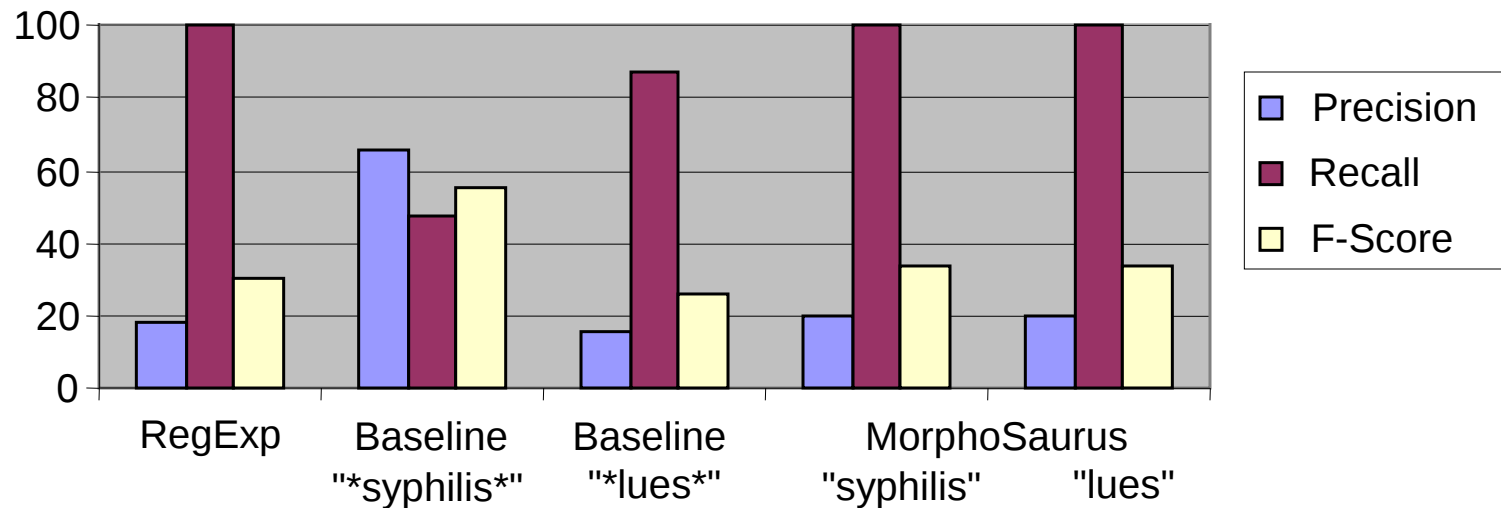
- RegExp search:
 - “**lues**”, “**luet**”, “**syphil**”, “**fta-abs**”, “**fta abs**”, “**schanker**”, “**pallid**”, “**treponem**”: *yield: 226 / 30,000 documents*
 - manual relevance check: 40 / 226 → Gold standard
- Baseline 1: substring search for “**syphilis**”
- Baseline 2: substring search for “**lues**”
- Experiment: MorphoSaurus search for “*Lues*” and “*Syphilis*”

Results

Results of retrieval experiments

- Baseline 1: substring search for “**syphilis**”
 - precision = 65.5%, recall = 47.5%, F-score: 55.1
- Baseline 2: substring search for “**lues**”
 - precision = 15.4, % recall = 87.5%, F-score: 26.2
- MorphoSaurus 1: search for “*syphilis*”
 - precision = 20.1 %, recall = 100%, F-score 33.5
- MorphoSaurus 2: search for “*lues*”
 - precision = 20.1 % recall = 100%, F-score 33.5

Comparative results



- Morphosaurus search retrieves all relevant documents
- equivalent to the laborious RegExp search
- Baseline search “*syphilis*” yields the best F-value
- Usefulness of Morphosaurus search depends on task

User satisfaction results

- User satisfaction (n = 20)
 - “system could enhance my clinical performance”:
80%
 - “have a positive impact on my scientific work”
90%
 - „has a positive impact on medical education“
50%

(reason: restriction to discharge summaries and pictures not sufficient, no good integration with the EHR so far)

Conclusion

- Horizontal access to EHR data well accepted
- New usage of medical record narratives
 - Search for related cases / similar images / cases for medical education
- Morphosemantic indexing for document filtering (under reserve: single use case / German language):
 - recall (not precision) optimized
 - robust against query formulation
 - outperforms naïve substring search in precision
 - equal to laborious search by hand-crafted regular expressions



Biomedical and Health Informatics
From Foundations to Applications to Policy
November 8-12, 2008 Washington, DC



Evaluation of a Document Search Engine in a Clinical Department System

Stefan Schulz¹, Philipp Daumke²,
Pascal Fischer³, Marcel Müller³

¹Institute of Medical Biometry und Medical Informatics, University Medical
Center Freiburg

²AVERBIS GmbH, Freiburg, Germany

³Department of Dermatology, University Medical Center Freiburg, Germany

Thank You!

Contact:

Stefan Schulz

<http://purl.org/steschu>

<http://www.averbis.de>

stschulz@uni-freiburg.de