

Morphosaurus



Morphosaurus

Statusreport Mai 2005

Grundidee

- Semantische Indexierung von medizinischen Dokumenten
- Verbesserung des Dokumentenretrievals
- Unterstützung sprachübergreifender Dokumentenrecherche
- Geeignetes Granularitätsniveau: “Subwörter”

Indexierung durch Subwörter

AG Gastroenterologie - Microsoft Internet Explorer

Datei Bearbeiten Ansicht Favoriten Extras ?

Zurück Suchen Favoriten Verlauf Wechseln zu

Adresse http://www.med.uni-heidelberg.de/chir/chirall/AG_Gastroenterologie.htm

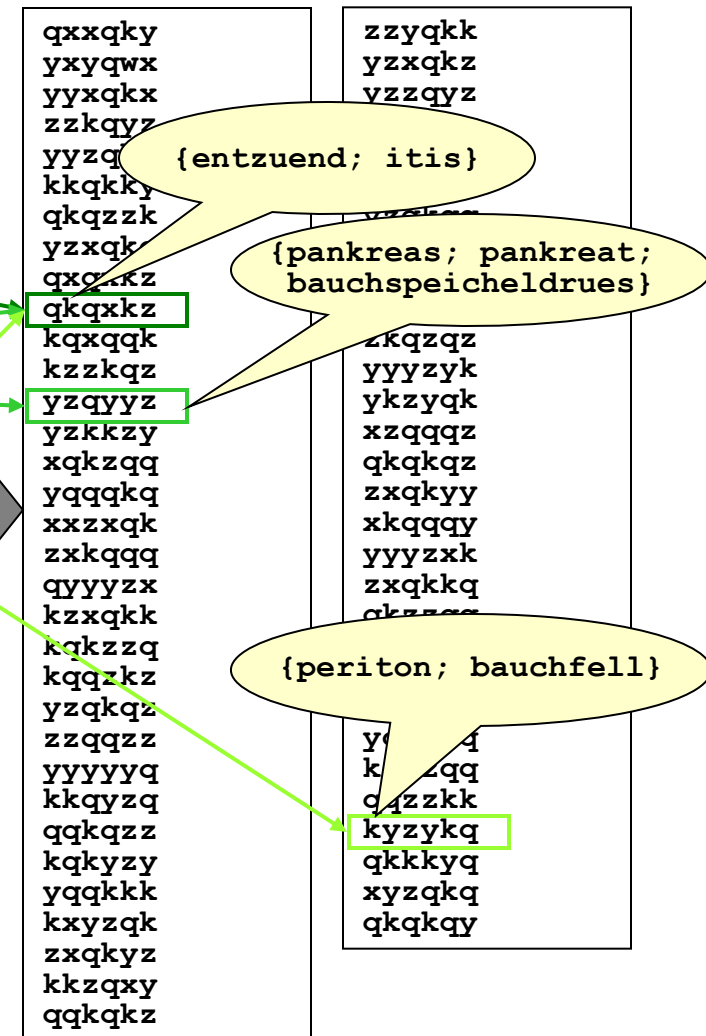
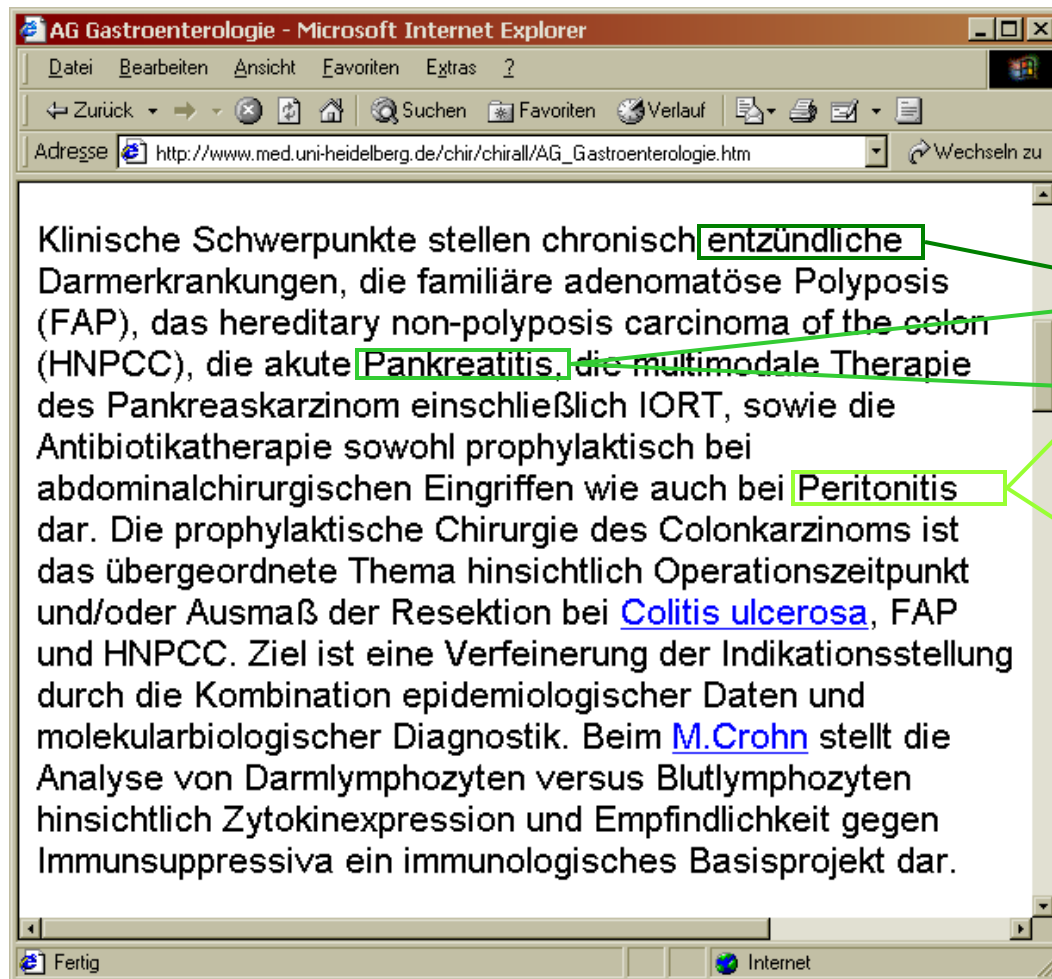
Klinische Schwerpunkte stellen chronisch entzündliche Darmerkrankungen, die familiäre adenomatöse Polyposis (FAP), das hereditary non-polyposis carcinoma of the colon (HNPCC), die akute Pankreatitis, die multimodale Therapie des Pankreaskarzinom einschließlich IORT, sowie die Antibiotikatherapie sowohl prophylaktisch bei abdominalchirurgischen Eingriffen wie auch bei Peritonitis dar. Die prophylaktische Chirurgie des Colonkarzinoms ist das übergeordnete Thema hinsichtlich Operationszeitpunkt und/oder Ausmaß der Resektion bei Colitis ulcerosa, FAP und HNPCC. Ziel ist eine Verfeinerung der Indikationsstellung durch die Kombination epidemiologischer Daten und molekularbiologischer Diagnostik. Beim M.Crohn stellt die Analyse von Darmlymphozyten versus Blutlymphozyten hinsichtlich Zytokinexpression und Empfindlichkeit gegen Immunsuppressiva ein immunologisches Basisprojekt dar.

Fertig Internet

abdomin
adenom
akut
analys
antibiot
ausmass
basis
biolog
blut
chirurg
chroni
darm
daten
diagnost
eingriff
empfindlich
entzuend
epidemiolog
express
famili
fap
fein
heredit
hinsichtlich
hnpcc
immun
indik
iort
itis
karzin
klin
kolitis
kolon

kombin
krank
krohn
lymph
modal
molekul
multi
non
operation
ordn
osis
pankreas
pankreat
periton
polyp
projekt
prophylakt
punkt
resekt
schwerpunkt
stell
suppress
thema
therap
ueber
ulzer
versus
zeit
ziel
zyt
zytokin

Indexierung durch Subwort – Synonymklassen-IDs



Sprachübergreifende Indexierung

High TSH values suggest the diagnosis of primary hypothyroidism ...

Erhöhte TSH-Werte erlauben die Diagnose einer primären Hypothyreose ...

Original

Interlingua

#up	tsh	#value	#suggest
#diagnost		#primar	#small
#thyre			

#up	tsh	#value	#permit
#diagnost		#primar	#small
#thyre			

Orthografische Normalisierung

Orthografische Regeln

high tsh values suggest the diagnosis of primary hypothyroidism ...

erhoehte tsh-werte erlauben die diagnose einer primaeren hypothyreose ...

**Zerlegungsalgorithmus
Subwort-Lexikon**

Semantische Normalisierung

Subwort-Thesaurus

high tsh value s suggest the diagnos is of primar y hypo thyroid ism

er hoeh te tsh wert e erlaub en die diagnos e einer primaer en hypo thyre ose

Morphosaurus - Komponenten

Kernsystem

- Subwort – Lexikon
- Präprozessor
- Segmentierer / Indexer
- Suchmaschine

Werkzeuge

- Lexikon-Editor (Morphoedit)
- Validierungstools

MorphoEdit WEB - Microsoft Internet Explorer

File Bearbeiten Ansicht Favoriten Extras ?

Links LEO MERMIG PubMed WordNet MorphoSaurus MI Elektronische Zeitschriftenbibliothek Online English to Portugese to English Dictionary

Adresse http://www.ler.pucpr.br/MEWeb/index.jsp

New Lexeme Thesaurus: Join Rel UnRel Tools: Mesh UMLS WordStat Export Exit

☒ ALL ☒ German ☒ English ☒ Portuguese ☒ Spanish ☒ French ☒ Swedish

Order 1: Lexeme Order 2: Lexeme ab Search

☐		abdom	4	Stem
☐		abdom	4	Stem
☐		abdom	4	Stem
☐		abdome	4	Stem
☐		abdomem	4	Stem
☐		abdomen	4	Stem
☐		abdomen	4	Stem
☐		abdomen	4	Stem
☐		abdomen	4	Stem
☐		abdomen	4	Stem
☐		abdomens	4	Stem
☒		abdomin	4	Stem
☐		abdomin	4	Stem
☐		abdomin	4	Stem

Order 1: Lexeme Order 2: Lexeme ab Search

	Lexeme	EqClass	Type	Total: 388
☐	ab	201138	Invariant	
☐	ab	24079	Prefix	
☐	ab	60000	Prefix	
☐	ab	60000	Prefix	
☐	ab	60000	Prefix	
☐	aba	200269	Suffix	
☐	abacat	14402	Stem	
☐	abacavir	502123	Stem	
☐	abacavir	502123	Stem	
☐	abacavir	502123	Stem	
☐	abacax	23890	Stem	
☐	abactal	25299	Stem	
☐	abactal	25299	Stem	

Eq Class 4 for indexing Unjoin

- ☐ abdomen
- ☐ abdomin
- ☐ abdomino
- ☐ baeuch
- ☐ bauch
- ☐ lapar
- ☐ unterleib
- ☐ venter
- ☐ ventr
- ☐ abdom
- ☐ abdomen
- ☐ abdomin

sense_of EqClass: 502277

- ☐ gastr
- ☐ gastr

word_part_of EqClass: 501983

- ☐ epigastr
- ☐ oberbauch

word_part_of EqClass: 28424

- ☐ enzaktin
- ☐ enzactin

word_part_of EqClass: 36759

- ☐ celiotom
- ☐ _celiotom

MorphoEdit WEB © PUC-PR Preferred Resolution: 1024x768 Internet

Morphosaurus - Komponenten

Experimentelle Komponenten:

- Akronym-Lerner
 - Disambiguierer
 - Morphoogle
 - Dynamisches Dictionary
 - Tools zur automatischen Lexikonaquisition
- ... basierend auf Ko-Okkurrenzhinformatiön aus großen mehrsprachigen Medizinkorpora

... Radiation therapy, high-dose X-rays used to kill cancer cells and shrink tumors, is often used to treat adenocarcinomas of the pancreatic head.

Stand des Lexikons

- Ca. 65 000 Einträge und 25 000 Äquivalenzklassen
- Sprachen:
 - Deutsch
 - Englisch
 - Portugiesisch
 - Spanisch
 - Schwedisch
 - Französisch

Partner

- Medizinische Informatik Freiburg:
Projektleitung, SW-Entwicklung und Pflege,
Lexikonentwicklung
- Computerlinguistik Jena
Hosting, wissenschaftliche Beratung
- PUCPR Curitiba / Brasilien
Lexikonentwicklung, Entwicklung von
Lexikonwerkzeugen
- SemanticMining-Partner
Lexikonentwicklung

Aufgabenverteilung

- Projektleitung: Stefan Schulz, Kornel Marko
- Lexikon: Philipp Daumke, Susanne Hanser, Roosewelt Andrade + 4 Hiwis (D) + 3 Hiwis (Brasilien) + 1-2 (Schweden)
- SW-Entwicklung:
Kornel Marko, Philipp Daumke, Edson Pacheco, Jan Pätzold + Hiwis
- Wissenschaft: Stefan Schulz, Kornel Marko, Philipp Daumke, Udo Hahn, Edson Pacheco
- Externe Partner / Kunden:
Kornel Marko, (Martin Romacker)

Finanzierung

- EU: NoE SemanticMining (bis 12/06)
- DFG: KoMoDoRe (bis 03/06)
- BMBF: Brasilien-Austausch (bis 12 / 06)

Externe Partner

Machbarkeitsstudien (im Gange):

- ZbMED, Köln: Indexierung der Literaturdatenbank CCMed
- Rhön-Kliniken (+ Siemens): “EPA-Google”
- IS-GBE (Stat. Bundesamt / RKI) :
Retrievalinterface für Materialien zur Gesundheitsberichterstattung

(geplant)

- HCPA (Porto Alegre, Brasilien): Recherche und Wissensextraktion aus elektronischer Patientenakte

Potentielle Partner

Termine mit:

- TEMIS group (Text Mining)
- Novartis (Martin Romacker in Text / Knowledge Management – Gruppe)
- DIMDI

SWOT - Analyse

- Strengths
- Weaknesses
- Opportunities
- Threats