



Medizinische Universität Graz

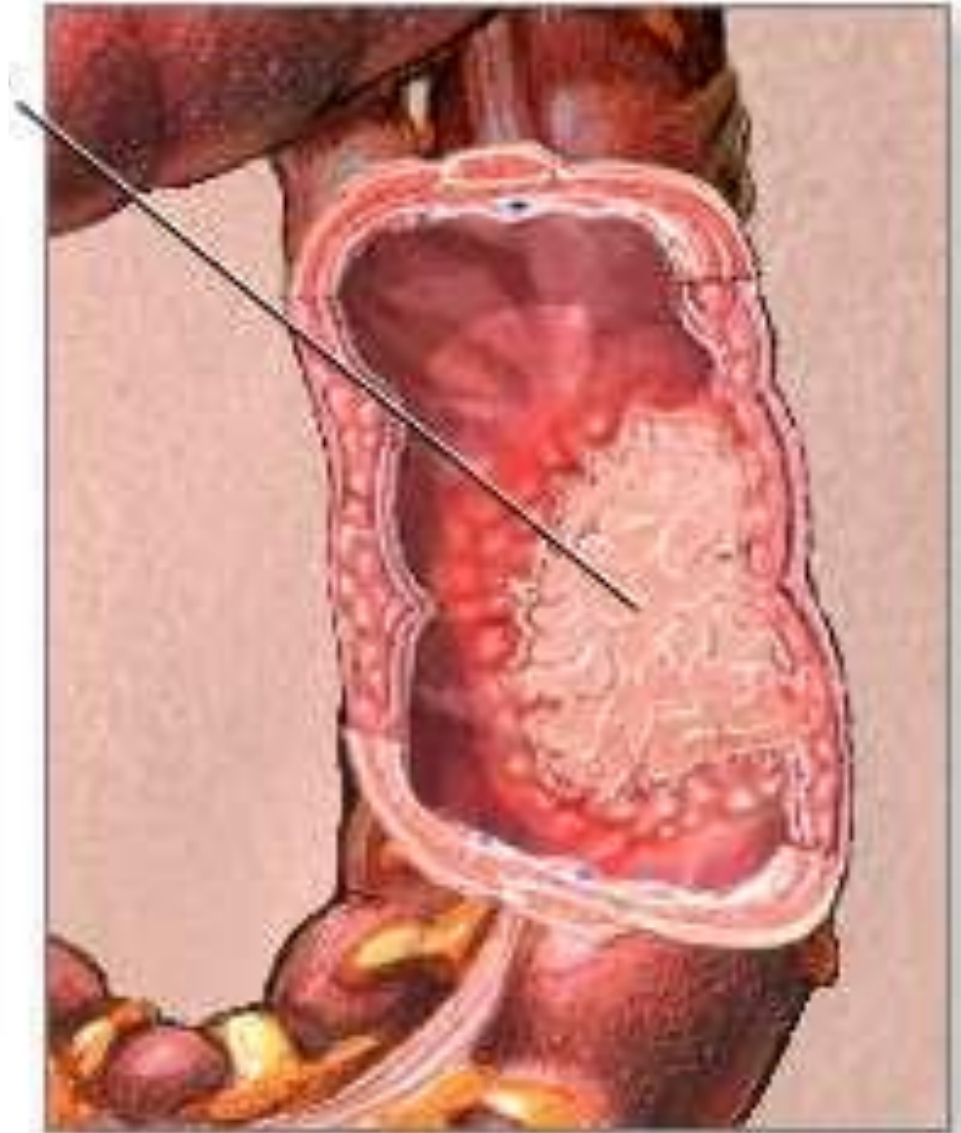
Modul 14, WS 2014 - 2015

Informationsrecherche in der Medizin

Stefan Schulz

© Institut für Medizinische Informatik, Statistik und Dokumentation

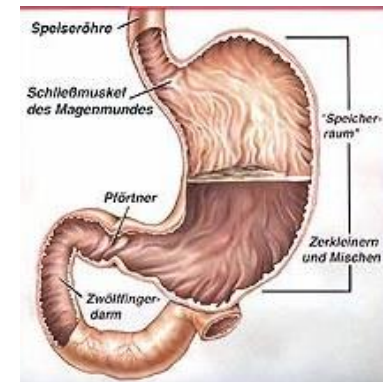
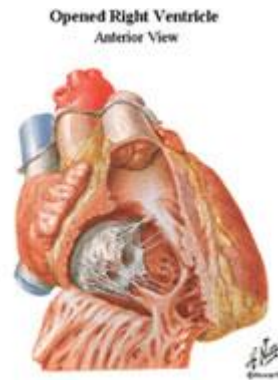
Beispiel 1



Hier wird ein medizinisches Phänomen veranschaulicht. Sie wollen hierzu nach Dokumenten recherchieren. Formulieren Sie eine Suchabfrage (ein Fachterminus genügt)

Beispiel 2

- Ein Patient kennt das Wort „Ventrikel“ nicht und gibt das Wort in eine Suchmaschine ein
- Skizzieren Sie das Problem



Hier wird ein medizinisches Phänomen veranschaulicht. Sie wollen hierzu nach Dokumenten recherchieren. Formulieren Sie eine Suchabfrage (ein Fachterminus genügt)

Zwei Sprachphänomene, die die Textrecherche erschweren

- Synonymie: Ein Gegenstand lässt sich durch unterschiedliche sprachliche Zeichen ausdrücken
 - Kolon = Dickdarm
 - EKG = Elektrokardiogramm
 - „die Antrumschleimhaut ist durch Lymphozyten infiltriert“ = „lymphozytäre Infiltration der Antrum mukosa“
- Homonymie (Ambiguität) : Unterschiedliche Gegenstände werden mit demselben sprachlichen Ausdruck belegt
 - „HIW“ = Hinterwandinfarkt oder Harnwegsinfekt
 - "Leitung": Nervenleitung oder Abteilungsleitung
 - "Verstopfung": Obstipation oder Gefäßverschluss
 - "Bruch": Fraktur, Hernie, Bruch'sche Membran im Auge

Arten von Retrieval

- Dokumentenretrieval
 - beantwortet Nutzeranfragen mit Listen potentiell relevanter Dokumente
 - Früher: typische Aufgabe von Dokumentations- und Bibliothekskräften, heute: "Googlen" als zentrale Kulturtechnik der Informationsgesellschaft
- Faktenretrieval
 - beantwortet eine Nutzeranfrage mit einer Auflistung potentiell relevanter Fakten
- Termretrieval
 - Auffinden von passenden Terminologie- / Klassifikationstermen Kodieren, "Verschlüsseln"

Übung

- Ein Lehrbuch enthält Krankheitsbeschreibungen zu : *Pneumonie, Pankreatitis, Hepatitis, Arthritis*, Diabetes mellitus, *Thyreoiditis*, Gonarthrose, *M. Crohn, Appendizitis, Rachitis*.
- Ziel: Selektiere alle Dokumente über entzündliche Erkrankungen
- Methode: Jeder Titel, der den Teilstring „itis“ enthält, wird als relevant betrachtet.

Wie ist die Precision, wie der Recall dieser Methode ? Gegeben eine große ($n > 1000$) Dokumentensammlung, was ist einfacher zu messen, Precision oder Recall?

Einfache Freitextsuche

Beispiel:

Es sollen alle stationären Fälle mit gesichertem Magenulkus im Jahr 2013 gezählt werden.

Probleme:

- Nicht alle Fälle mit Magenulkus (Nebendiagnose) sind mit dem ICD-Kode K25 kodiert.
 - Es kommt vor, dass ungesicherte Verdachtsdiagnosen auch mit K25 kodiert sind.
- Diagnosenkodes zur Selektion unzuverlässig

Simple Freitextsuche "Magenulkus" durch alle verfügbaren Dokument:

- Kein Treffer bei "**Magenulcus**" oder "**Magengeschwür**"
- Kein Treffer bei "Stress**ulkus** im **Corpus**bereich"
- Treffer bei "**Magenulcus** kann nicht ausgeschlossen werden"
- Treffer bei "Familienanamnese: Vater **Magenulcus**"
- Treffer bei "Gabe von Omeprazol zur Prophylaxe eines **Magenulcus**"

Semantische Suche

In Befundtext: "Stressulcus im Corpusbereich"

Anfrage: "Magenulcus"

ANNOTATIONEN

"Stressulcus im Corpus"

"Ulcus im Corpus"

"Ovarialulcus"

"Gehirnulcus"

"Uterusulcus"

"Magenulcus"

~~"Ovarialulcus"~~

~~"Gehirnulcus"~~

~~"Uterusulcus"~~

"Magenulcus"

"Corpusbereich" =_{syn}

"Corpus"

"Stressulcus" =_{syn}

"Stressulcus"

LEXIKON

"Corpus luteum"

"Corpus callosum"

"Corpus uteri"

"Corpus ventriculi"

ONTOLOGIE

└ Ulcus

└ Stressulcus

└ Druckulcus

(...)

Teil von Ovar

Teil von Gehirn

Teil von Uterus

Teil von Magen

soweit einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des Pylorus zeigt die Dünndarmwandung eine sanduhrartige Stenose. Im Magen- und Duodenallumen reichlich zähflüssiger Schleim, sanguinolent; die Schleimhaut ist insgesamt livide. Auf lamellierenden Schnitten zähfestes weißliches, teilweise nodulär konfiguriertes Gewebe, ohne das Gallengänge manifest werden

REFERENZKORPUS

Retrieval - Perspektiven

- Unterstützung beim Erwerb von "Retrieval-Kompetenz"
- Nicht jede Recherche ist gleich:
 - **Precision**-orientierte Recherche: Vollständigkeit sekundär, schneller Überblick:
 - Suche nach Kasuistiken für Lehrzwecke
 - Suche nach „ähnlichen“ Behandlungsfällen
 - setzt voraus, dass eher zu viele Dokumente vorhanden sind
 - **Recall**-orientierte Recherche: Vollständigkeit essentiell, "Stecknadel-in-Heuhaufen"
 - wissenschaftliche Auswertungen
 - Qualitätssicherungsmaßnahmen
 - Entscheidungsunterstützung

Wie ist der Einfluss auf Precision und Recall von: Synonymauflösung, Begriffshierarchien, Verwendung mehrdeutiger Terme, Ignorieren von Negationsausdrücken, ...



Diplomarbeiten in der medizinischen Informatik

Themen:

- Navigation in der elektronischen Krankenakte
- Weiterentwicklung von Kodierungs- und Klassifikationssystemen
- Verarbeitung medizinischer Sprache
- Medizinische Wissensrepräsentation
- Mobiles Computing in der Medizin

Kontakt: Stefan Schulz

stefan.schulz@medunigraz.at