Medical Information Retrieval

Lorraine Goeuriot
University of Grenoble Alpes
LIG, Grenoble - France
lorraine.goeuriot@imag.fr

December 5, 2019

Introduction

Purpose of the course:

1. What are the challenges for Information retrieval in the medical domain?
2. Search context/environment:
   - what are the tasks?
   - What are the information needs?
   - What data is used?
3. Which information retrieval model suits the tasks?
4. How can these models be evaluated?

Outline

Objectives

Challenges in Medical Information Search

Varying stakeholders: Patients, next-of-kins, caregivers, physicians, clinicians, researchers

Varying medical knowledge :

Among patients : short-term vs long-term disease

Among medical professionals : from medical students to specialized practitioners

Varying language skills : literacy, cross-lingual search...

Search tasks and challenges:

- For medical practitioners: Evidence-based medicine, need for precise information in daily care
- For patients: vocabulary gap, cybercondria [White and Horvitz, 2009]
- For clinicians and researchers: need for up-to-date information, systematic reviews, patients cohorts for clinical trials...

Outline

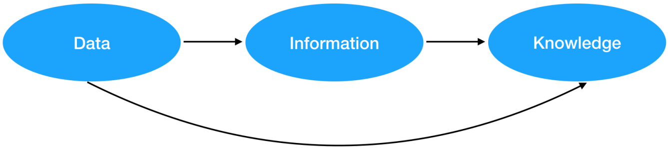
Medical Information

What is information?

Information $\neq$ data $\neq$ knowledge

- Data: observations and measurements made about the world
- Information: data brought together in aggregate to demonstrate facts
- Knowledge: what is learned from the data and information, that can be applied in new situations to understand the world

[Blum, 1984] cited in [Hersh, 2010]



Medical Information

Scientific Information Properties

Properties of scientific texts [Hersh, 2010]:

- **Growth**: The amount of scientific publications is growing exponentially
- **Obsolescence**: scientific advances, constant update of the state-of-the-art and changes in society make information quickly obsolete
- **Fragmentation**: text published often reflects only one part of a problem or situation
- **Links and citations**: strong property of scientific text, links and references allow to generate networks among works and communities
- **Propagation**: simplicity of information flow

Medical Information

A Classification of Textual Health Information

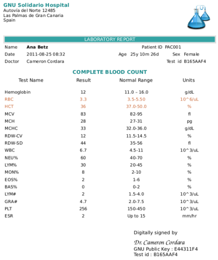
[Hersh, 2010] distinguishes two main categories of textual health documents:

- **Patient-specific information**: applies to individual patients. Tells healthcare providers, administrators and researchers about the health and disease of a patient.
  - Structured: laboratory results, vital signs
  - Narrative: history and physical, progress notes, radiology report
- **Knowledge-based information**: has been derived and organized from observational or experimental research. Usually provided in books, journals or *computerized media*.
  - Primary: original research (in journals, books, reports, etc.)
  - Secondary: summaries of research (in review articles, books, practice guidelines, etc.)

With the emergence of Web2.0, one could also consider **User-generated Content** as another category:

- Collaborative writing: wikipedia, blogs
- Social media: discussion forums, Facebook, Twitter, PatientsLikeMe

Medical Information



GNU Solidario Hospital
Autovía del Norte 12485
Las Palmas de Gran Canaria
Spain

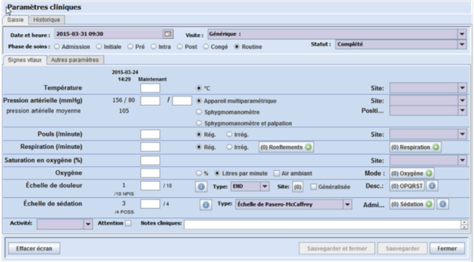
LABORATORY REPORT

Name **Ana Betz** Patient ID PAC001
Date 2011-08-25 08:32 Age 25y 10m 26d Sex Female
Doctor Cameron Cordara Test id B165AAF4

COMPLETE BLOOD COUNT

Test Name	Result	Normal Range	Units
Hemoglobin	12	11.0 – 16.0	g/dL
RBC	3.3	3.5-5.50	10^6/uL
HCT	36	37.0-50.0	%
MCV	83	82-95	fl
MCH	28	27-31	pg
MCHC	33	32.0-36.0	g/dL
RDW-CV	12	11.5-14.5	%
RDW-SD	44	35-56	fl
WBC	6.7	4.5-11	10^3/uL
NEU%	60	40-70	%
LYM%	30	20-45	%
MON%	8	2-10	%
EOS%	2	1-6	%
BAS%	0	0-2	%
LYM#	2	1.5-4.0	10^3/uL
GRA#	4.7	2.0-7.5	10^3/uL
PLT	256	150-450	10^3/uL
ESR	2	Up to 15	mm/hr

Digitally signed by
Dr. Cameron Cordara
GNU Public Key : E44311F4
Test id : B165AAF4



HEMATOLOGIE

NUMERATION ET FORMULE LEUCOCYTAIRE

	Résultats	Normes
HEMATIES	5.320.000 /mm3	[illegible]
HEMOGLOBINE.	16,5 g/100 ml	[illegible]
HEMATOCRITE.	48,7 %	[illegible]
LEUCOCYTES	6.000 /mm3	[illegible]
PLAQUETTES	226.000 /mm3	[illegible]

VITESSE DE SEDIMENTATION

1ère heure	8 mm	[illegible]

CHIMIE DU SANG

	Résultats	Normes
aspect du sérum.	normal	
GLYCEMIE	1,08 g/l 5,99 mmol/l	[illegible]
UREE	0,33 g/l 5,48 mmol/l	[illegible]
CREATININEMIE.	10 mg/l 88 umol/l	[illegible]
EXPLORATION LIPIDIQUE		
CHOLESTEROL TOTAL.	2,86 g/l 7,38 mmol/l	[illegible]
H.D.L.	0,53 g/l 1,37 mmol/l	[illegible]
TRIGLYCERIDES.	1,54 g/l 1,76 mmol/l	[illegible]
LDL CHOLESTEROL.	2,02 g/l 5,23 mmol/l	[illegible]
PROTEINE C-REACTIVE.	inf à 3 mg/l	[illegible]

Medical Information

Narrative Patient Specific Information

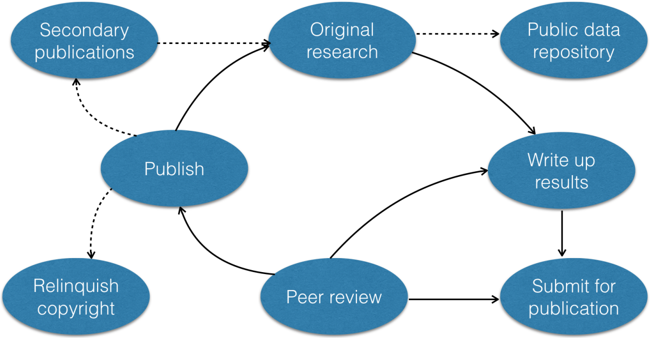
```
Admission Date:  [**2015-03-17**]        Discharge Date:  [**2015-03-24**]
Date of Birth:   [**1974-10-03**]        Sex:  F
Service: Neurosurgery
HISTORY OF PRESENT ILLNESS:  The patient is a 40-year-old female with complaints
of headache and dizziness.  In [**2015-01-14**], the patient had headache with
neck stiffness and was unable to walk for 45 minutes.   [...]
PAST MEDICAL HISTORY:  Hypothyroidism.
ALLERGIES:  Penicillin and Bactrim which causes a rash.
MEDICATIONS:  Levoxyl 1.75 mg.
PHYSICAL EXAMINATION:  On physical examination, her blood pressure was 104/73,
pulse 79.  In general, she was a woman in no acute distress.  HEENT:  Nonicteric.
Pupils are equal, round, and reactive to light.  Extraocular movements are full.
[...]

On postoperative day #1, the patient was taken to arteriogram, where she
underwent a cerebral angiogram to evaluate clipping of the aneurysm.  []
DISCHARGE MEDICATIONS:
1. Hydromorphone 2-6 mg po q4h prn.
2. Synthroid 175 mcg po q day.[...]

CONDITION ON DISCHARGE:  Stable.

FOLLOW-UP INSTRUCTIONS:  She will follow up in 10 days for staple removal with
Dr. [**Last Name (STitle) 570**].

(End of Report)
```

Discharge summary extracted from the MIMIC II dataset https://physionet.org/mimic2/.

Medical Information

Primary Knowledge-based Documents

- Contain reports of research results: discoveries, observations, description of related work and position of the report, conclusions.
- Has never been published before
- Published in books, journals or conference proceedings
- Usually a small number of documents have the highest impact



Medical Information

Primary Knowledge-based Documents

Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search

RYEN W. WHITE and ERIC HORVITZ
Microsoft Research

The World Wide Web provides an abundant source of medical information. This information can assist people who are not healthcare professionals to better understand health and illness, and to provide them with feasible explanations for symptoms. However, the Web has the potential to increase the anxieties of people who have little or no medical training, especially when Web search is employed as a diagnostic procedure. We use the term *cyberchondria* to refer to the unfounded *escalation* of concerns about common symptomatology, based on the review of search results and literature on the Web. We performed a large-scale, longitudinal, log-based study of how people search for medical information online, supported by a survey of 515 individuals' health-related search experiences. We focused on the extent to which common, likely innocuous symptoms can escalate into the review of content on serious, rare conditions that are linked to the common symptoms. Our results show that Web search engines have the potential to escalate medical concerns. We show that escalation is associated with the amount and distribution of medical content viewed by users, the presence of escalatory terminology in pages visited, and a user's predisposition to escalate versus to seek more reasonable explanations for ailments. We also demonstrate the persistence of postsession anxiety following escalations and the effect that such anxieties can have on interrupting user's activities across multiple sessions. Our findings underscore the potential costs and challenges of cyberchondria and suggest actionable design implications that hold opportunity for improving the search and navigation experience for people turning to the Web to interpret common symptoms.

Categories and Subject Descriptors: H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Search process; query formulation*

General Terms: Human Factors, Experimentation

Additional Key Words and Phrases: Cyberchondria

ACM Reference Format:

White, R. W. and Horvitz, E. 2009. Cyberchondria: Studies of the escalation of medical concerns in Web search. ACM Trans. Inf. Syst. 27, 4, Article 23 (November 2009), 37
DOI = 10.1145/1629096.1629101 http://doi.acm.org/10.1145/1629096.1629101

1. INTRODUCTION

The World Wide Web has the potential to provide valuable medical information to people, where Web sites such as WebMD (http://www.webmd.com) and MSN

Authors' addresses: R. W. White and E. Horvitz, Microsoft Research, One Microsoft Way, Redmond, WA 98052; email: {ryenw, horvitz}@microsoft.com.

DOI 10.1145/1629096.1629101 http://doi.acm.org/10.1145/1629096.1629101

ACM Transactions on Information Systems, Vol. 27, No. 4, Article 23, Publication date: November 2009.

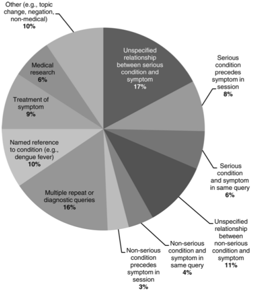


Fig. 1. Distribution of labels assigned to set of hand-labeled no-change sessions.

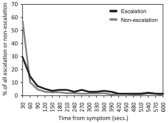


Fig. 2. Temporal distance from initial input of symptom (within session)

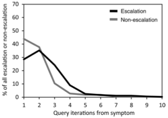


Fig. 3. Query distance from initial input of symptom (within session).

REFERENCES

Agichtein, E., Brill, E., and Dumais, S. 2006. Improving Web search ranking by incorporating user behavior information. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. 19–26.

American Psychiatric Association. 1994. *Diagnostic and Statistical Manual of Mental Disorders* 4th Ed. APA, Washington, DC.

Asmundson, J. G., Taylor, S., and Cox, B. J. 2001. *Health Anxiety: Clinical and Research Perspectives on Hypochondriasis and Related Conditions*. Wiley.

Ayers, S. and Kronenfeld, J. 2007. Chronic illness and health-seeking information on the Internet. *Health 11*, 3, 327–347.

Baker, L., Wagner, T. H., Singer, S., and Bundorf, M. K. 2003. Use of the Internet and e-mail for health care information. *J. Amer. Med. Assoc. 289*, 18, 2401–2406.

Barsky, A. J. and Ahern, D. K. 2004. Cognitive behavioral therapy for hypochondriasis. *J. Amer. Med. Assoc. 291*, 12, 1464–1470.

Barsky, A. J. and Klerman, G. L. 1983. Overview: Hypochondriasis, bodily complaints, and somatic styles. *Amer. J. Psychiatry 140*, 273–283.

Benigeri, M. and Pluye, P. 2003. Shortcomings of health-related information on the Internet. *Health Promot. Int. 18*, 4, 381–387.

Berland, G. K., Elliott, M. N., Morales, L. S., Algazy, J. I., Kravitz, R. L., Broder, M. S., Kanouse,

Medical Information

Secondary Knowledge-based documents

- All medical professionals are not researchers: primary resources need to be rephrased, summarized, synthetized
- Summary and reviews of primary resources are published in scientific journals
- Quality issue: the editorial process is not the same for secundary than primary resources
- Other category: clinical practice guidelines (many publications, very little control)

Specific case: Systematic Reviews and Meta-Analysis

- Fragmentation of the scientific literature → difficult to identify all the relevant papers on a topic
- In particular with clinical trials, large amount of publications on a similar condition or treatment
- Systematic reviews tackle a precise question, and describe the complete set of related work and factual approaches
- Meta-analysis compare results at the systematic review scale
- Topics: treatment (63%), causality and security (29%), diagnosis (4,4%), prognosis (2.1%) [Montori et al., 2004]
- Cochrane is a non-profit, non-governmental organization formed to organize medical research findings so as to facilitate evidence-based choices about health interventions `http://www.cochranelibrary.com/`

Medical Information

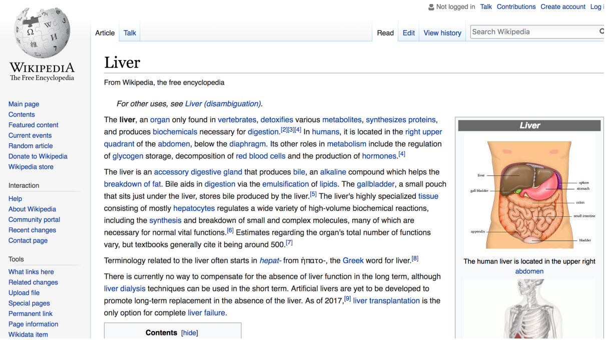
User Generated Content

Collaborative writing websites allow users to edit collaboratively documents. It can have some sort of editorial control. It includes:

- **Wikis** such as wikipedia (collective writing and control of the content)

[Blackman, 2006] showed that information contained on wikipedia wasn't erroneous (comparison on 42 topics with the Britannica Encyclopaedia)

- **Blogs**: discussion or informational website published on the Web consisting of discrete, often informal diary-style text entries ("posts").



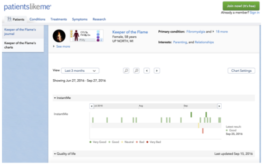
Medical Information

User Generated Content

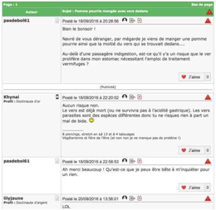
Health topics can be covered on all types of social media:

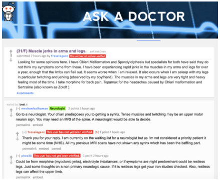
- General social media such as facebook, twitter:

- Medical social media such as PatientsLikeMe:



- Discussion forums: where all kinds of users (patients, doctors, students, nurses...) can discuss health topics





Medical Information

Certification

How can the quality of health information online be guaranteed?

The organization Health On the Net (HON) certifies the quality and validity of medical websites.
HON manually certifies website according to the following principles:

- Principle 1 : Authority - Give qualifications of authors
- Principle 2 : Complementarity - Information to support, not replace
- Principle 3 : Confidentiality - Respect the privacy of site users
- Principle 4 : Attribution - Cite the sources and dates of medical information
- Principle 5 : Justifiability - Justification of claims / balanced and objective claims
- Principle 6 : Transparency - Accessibility, provide valid contact details
- Principle 7 : Financial disclosure - Provide details of funding
- Principle 8 : Advertising - Clearly distinguish advertising from editorial content

https://www.hon.ch/HONcode/Guidelines/guidelines.html

OUTLINE

Medical Search Tasks

Information need

Information needs [Hersh, 2010]:

- *Retrospective information needs*:
  - The need for help in solving a certain problem or making a decision
  - The need for background information on a topic
- *Current awareness information needs*:
  - The need to keep up with information in a given subject area

Amount of information needed [Lancaster and Warner, 1993]

- A single fact
- One or more documents
- A comprehensive search of the literature

Types of information needs [Wilkinson and Fuller, 1996]

- Fact-finding
- Learning
- Gathering
- Exploring

States of information need [Gorman, 1995]

- Unrecognized need
- Recognized need
- Pursued need
- Satisfied need

Medical Search Queries

Typology

The types of queries that are the most widely studied are:

- Classical keyword-based queries (physician vs patients)
- Boolean queries (systematic reviews)
- Structured queries (PICO)
- Multimodal queries (text + concepts e.g. pubmed search tools)

General classification of search queries from [Broder, 2002]:

- Navigational
- Transactional
- Informational

Classification of search queries for semantic search [Bast et al., 2016]:

- Structured
- Keyword-based
- Natural language-based

Medical Search Queries

Physician Queries

- Study by [Ely et al., 1999] on family doctors questions in their daily practise.
- Observation of 100 doctors from Iowa (US)

Taxonomy of generic questions:

- What is the cause of symptom X?
- What is the dose of drug Y?
- How should I manage disease or finding X?
- How should I treat finding or disease X?
- What is the cause of physical finding X?
- What is the cause of test finding X?
- Could this patient have disease or condition X?
- Is test X indicated in situation Y?
- What is the drug of choice for condition X?
- Is drug X indicated in situation Y?

Medical Search Queries

Physician Queries

- Study by [Ely et al., 1999] on family doctors questions in their daily practise.
- Observation of 100 doctors from Iowa (US)

Taxonomy of generic questions:

- What is the cause of symptom X?
- What is the dose of drug Y?
- How should I manage disease or finding X?
- How should I treat finding or disease X?
- What is the cause of physical finding X?
- What is the cause of test finding X?
- Could this patient have disease or condition X?
- Is test X indicated in situation Y?
- What is the drug of choice for condition X?
- Is drug X indicated in situation Y?

- These are questions and not queries - 64% were not pursued
- In 1999 Internet was not the primary source of information

Medical Search Queries

Clinical Queries

Analysis of search queries in an EHR search utility [Natarajan et al., 2010]

- **Navigational queries (14.5%)**: were mostly aiming at retrieving a specific EHR (e.g. using the record number)
- **Transactional queries (0.4%)**: were representing an action (e.g. adding a new note)
- **Information queries (85.1%)**: the most frequent, especially among clinicians and researchers.

Top 5 semantic types of searches

Semantic type	%	Semantic type	%
Laboratory or test result	29.2	Pharmacologic substance	7.5
Disease or syndrome	21.7	Diagnostic procedure	6.2
Body part, organ or organ component	8.1		

Top 10 most frequent queries

Query	%	Query	%
class	9.8	nephrogenic	1.8
nyha	4.5	hysterectomy	1.5
hodgkins	2.9	cva	1.1
iii	2.4	ef	1.0
iv	2.3	hf	0.9

- Very short queries (1.2 term(s) on average in the corpus)
- Many acronyms (*NYHA*) and abbreviations (*tach* for tachycardia)
- Ambiguous (*class*)

Medical Search Queries

Layperson queries

Particularities and challenges [Zhang et al., 2012]

- *Conceptual level*: layperson have their own understandings and hypotheses about a particular condition.
- *Terminological level*: layperson's vocabulary doesn't match medical terminologies
- *Lexical level*: queries contain mispelling, partial words, etc.
- Short text (on average less than 3 words), ambiguous

Medical Search Queries
PICO Queries

Designed to answer Evidence-based Medicine problems, PICO stands for:

- **P**atient / **P**roblem / **P**opulation
- **I**ntervention
- **C**omparison / **C**ontrol
- **O**utcome

The formulation of a focused clinical question containing well-articulated PICO elements is widely believed to be **the key to efficiently finding high-quality evidence** and also **the key to evidence-based decisions** [Huang et al., 2006].

Example (from [Boudin et al., 2010]):

`"children with pain and fever how does paracetamol compared with ibuprofen affect levels of pain and fever?`

→

Patient/Problem: `children/pain and fever`
Intervention: `paracetamol`
Comparison: `ibuprofen`
Outcome: `levels of pain and fever`

Medical Search Queries

Clinical Search Queries (Genomics)

[Hersh and Voorhees, 2009] categorized clinical queries into several Generic Topic Types:

Generic Topic Type	Example Topic
Find articles describing standard methods or protocols for doing some sort of experiment or procedure	Method or protocol: GST fusion protein expression in Sf9 insect cells
Find articles describing the role of a gene involved in a given disease	Gene: DRD4 Disease: alcoholism
Find articles describing the role of a gene in a specific biological process	Gene: Insulin receptor gene Biological process: Signaling tumorigenesis
Find articles describing interactions (e.g. promote, suppress, inhibit, etc.) between two or more genes in the function of an organ or in a disease	Genes: HMG and HMGB1 Disease: Hepatitis
Find articles describing one or more mutations of a given gene and its biological impact	Gene with mutation: Ret Biological impact: Thyroid function

Medical Search Queries

Systematic review Queries

- Systematic reviews use boolean queries on specific databases such as the Cochrane library to retrieve all the possible relevant documents on a topic.
- Example (topic extracted from CLEF eHealth Technologically assisted reviews task [Kanoulas et al., 2017]):

```
Topic: CD009551
Title: Polymerase chain reaction blood tests for the diagnosis of
       invasive aspergillosis in immunocompromised people

Query:
exp Aspergillosis/
exp Pulmonary Aspergillosis/
exp Aspergillus/
(aspergillosis or aspergillus or aspergilloma or "A.fumigatus" or
"A. flavus" or "A. clavatus" or "A. terreus" or "A. niger").ti,ab.
or/1-4
exp Nucleic Acid Amplification Techniques/
pcr.ti,ab.
"polymerase chain reaction*".ti,ab.
or/6-8
5 and 9
exp Animals/ not Humans/
10 not 11

Pmid's:
    25815649
    26065322
    ...
```

SUMMARY

Medical information retrieval =

- Various stakeholders
- Various information needs and search tasks
- Various information sources

→ Medical IR can take as many forms as you can imagine search scenarios

Towards **semantic information retrieval**!

What makes the difference with adhoc IR:

- Very well defined search tasks
- Users willing to use enriched format
- Very rich and maintained knowledge source
- Allows richer search

Outline

Semantic Search in the Medical Domain

There are many cases in medical information search where simple term matching is not enough:

- Patient cohort search
- Evidence-based medicine
- Systematic reviews
- Low-literacy users search

Examples of queries

- Patients taking atypical antipsychotics without a diagnosis schizophrenia or bipolar depression
- Patients with Diabetes exhibiting good Hemoglobin A1c Control (<8.0%)

Example of data

- Hydromorphone 2-6 mg po q4h prn.
- On physical examination, her blood pressure was 104/73, pulse 79. In general, she was a woman in no acute distress. HEENT: Nonicteric.

WHAT IS SEMANTIC SEARCH?

[Bast et al., 2016]

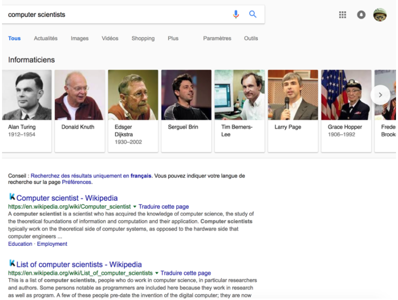
"In a nutshell, semantic search is 'search with meaning'. This 'meaning' can refer to various parts of the search process: *understanding the query* [...], *understanding the data* [...], or *representing knowledge in a way suitable for meaningful retrieval*"

- Understanding the query: instead of matching its terms to the data, extract its meaningful content
- Understanding the data: instead of just searching for term/stem matches, match meaningful entities
- Representing knowledge: define models representing knowledge in ways suitable to retrieve information

"université Grenoble" →

What is Semantic Search?

"Computer scientists"



"Female computer scientists working on semantic search"

Working on

Female

?

Semantic search

Computer scientists

What is Semantic Search?

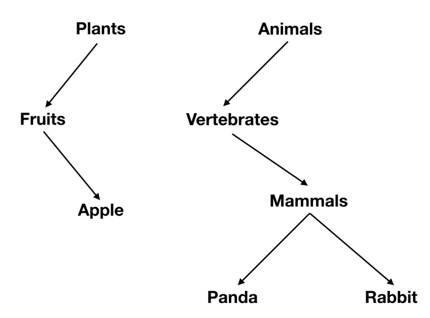
Semantics?

Explicit vs Implicit Semantics

The knowledge used in semantic search can be found or created under 2 main forms:

Explicit Semantics

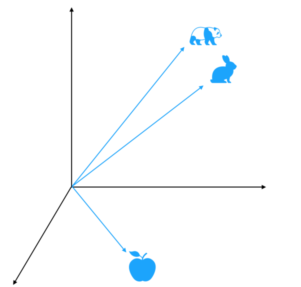
Human representation of the world and its concepts:



Implicit Semantics

"A word is characterized by the company it keeps" [Firth, 1957]

Purpose: represent words as vectors, based on their neighbours



Semantically close words will have similar vectors.

OUTLINE

Semantic Resources
Definitions

- Lexical and semantic resources are used **in many domains**
- They can be named differently
- We give here definitions usually used in Information Retrieval and Information Extraction
- Definitions are extracted from [Hersh, 2010] and [Bast et al., 2016]

A concept

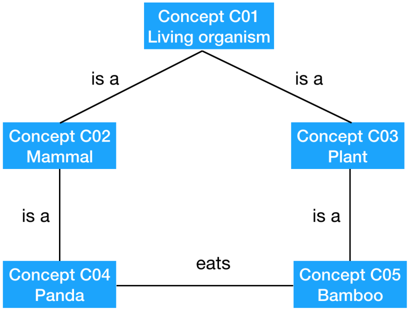
Idea or object that occurs in the world (e.g. *the condition under which human blood pressure is elevated*

A term

String of one or more words that represents a concept (e.g. *hypertension* or *high blood pressure*)

A relationship

Link between 2 concepts (e.g. the *liver* is an *organ*) or terms (e.g. *hypertension* and *high blood pressure* are synonyms)



Medical Knowledge Resources

Existing Medical thesauri (in English)

1. Medical Subject Headings (MeSH)

- Created by the National Library of Medicine *to index medical documents*
- 28,000 descriptors (concepts) with over 90,000 entry terms
- 3 types of relationships: hierarchical, synonymous, related

Hypertension MeSH Descriptor Data 2018

Details | Qualifiers | MeSH Tree Structures | Concepts

MeSH Heading
Hypertension
Tree Number(s)
C14.907.489
Unique ID
D006973
Annotation
not for intracranial or intraocular pressure; relation to BLOOD PRESSURE: Manual 23.27; Goldblatt kidney is HYPERTENSION, GOLDBLATT see HYPERTENSION, RENOVASCULAR; hypertension with kidney disease is probably HYPERTENSION, RENAL, not HYPERTENSION; venous hypertension: index under VENOUS PRESSURE (IM) & do not coordinate with HYPERTENSION; PREHYPERTENSION is also available
Scope Note
Persistently high systemic arterial BLOOD PRESSURE. Based on multiple readings (BLOOD PRESSURE DETERMINATION), hypertension is currently defined as when SYSTOLIC PRESSURE is consistently greater than 140 mm Hg or when DIASTOLIC PRESSURE is consistently 90 mm Hg or more.
Entry Term(s)
Blood Pressure, High
NLM Classification #
WG 340
See Also
Antihypertensive Agents
Vascular Resistance
Date Established
1966/01/01
Date of Entry
1999/01/01
Revision Date
2010/06/25

The 16 trees in MeSH

1. Anatomy
2. Organisms
3. Diseases
4. Chemicals and Drugs
5. Analytical, Diagnostic and Therapeutic Techniques and Equipment
6. Psychiatry and Psychology
7. Biological Sciences
8. Natural Sciences
9. Anthropology, Education, Sociology and Social Phenomena
10. Technology, Industry, Agriculture
11. Humanities
12. Information Science
13. Named Groups
14. Health care
15. Publication Characteristics
16. Geographicals

Medical Knowledge Resources

Existing Medical thesauri (in English)

2 International Classification of Medicine (ICD)

- International statistical classification of diseases and health problems
- Coded medical classification including a wide variety of signs, symptoms, trauma, etc.
- Published by the WHO
- Internationaly used to register morbidity and causes and morbidity

ICD-10 Version:2016
- I Certain infectious and parasitic diseases
- II Neoplasms
- III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
- IV Endocrine, nutritional and metabolic diseases
  - E00-E07 Disorders of thyroid gland
  - E10-E14 Diabetes mellitus
    - E10 Type 1 diabetes mellitus
    - E11 Type 2 diabetes mellitus
    - E12 Malnutrition-related diabetes mellitus
    - E13 Other specified diabetes mellitus
    - E14 Unspecified diabetes mellitus
  - E15-E16 Other disorders of glucose regulation and pancreatic internal secretion
  - E20-E35 Disorders of other endocrine glands
  - E40-E46 Malnutrition
  - E50-E64 Other nutritional deficiencies
  - E65-E68 Obesity and other hyperalimentation
  - E70-E90 Metabolic disorders
- V Mental and behavioural disorders
- VI Diseases of the nervous system
- VII Diseases of the eye and adnexa
- VIII Diseases of the ear and mastoid process
- IX Diseases of the circulatory system
- X Diseases of the respiratory system
- XI Diseases of the digestive system
- XII Diseases of the skin and subcutaneous tissue
- XIII Diseases of the musculoskeletal system and connective

.9 Without complications

E10 Type 1 diabetes mellitus

[See before E10 for subdivisions]

Incl.: diabetes (mellitus):
- brittle
- juvenile-onset
- ketosis-prone

Excl.: diabetes mellitus (in):
- malnutrition-related (E12.-)
- neonatal (P70.2)
- pregnancy, childbirth and the puerperium (O24.-)

glycosuria:
- NOS (R81)
- renal (E74.8)

impaired glucose tolerance (R73.0)
postsurgical hypoinsulinaemia (E89.1)

E11 Type 2 diabetes mellitus

[See before E10 for subdivisions]

Incl.: diabetes (mellitus)(nonobese)(obese):
- adult-onset
- maturity-onset
- nonketotic
- stable

non-insulin-dependent diabetes of the young

Excl.: diabetes mellitus (in):
- malnutrition-related (E12.-)
- neonatal (P70.2)
- pregnancy, childbirth and the puerperium (O24.-)

glycosuria:
- NOS (R81)
- renal (E74.8)

impaired glucose tolerance (R73.0)

ICD Classification

1 Certain infectious and parasitic diseases
2 Neoplasms
3 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
4 Endocrine, nutritional and metabolic diseases
5 Mental and behavioural disorders
6 Diseases of the nervous system
7 Diseases of the eye and adnexa
8 Diseases of the ear and mastoid process
9 Diseases of the circulatory system
10 Diseases of the respiratory system
11 Diseases of the digestive system
12 Diseases of the skin and subcutaneous tissue
13 Diseases of the musculoskeletal system and connective tissue
14 Diseases of the genitourinary system
15 Pregnancy, childbirth and the puerperium
16 Certain conditions originating in the perinatal period
17 Congenital malformations, deformations and chromosomal abnormalities
18 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
19 ...

Medical Knowledge Resources

Existing Medical thesauri (in English)

3. Systematized Nomenclature of Medicine (SNOMED): thesaurus designed to process clinical data
4. Cumulative Index to Nursing and Allied Health Literature (CINAHL): classical medical concepts + domain-specific ones
5. EMTREE: European MeSH, used to index EMBASE
6. PsycINFO: psychology and psychiatry thesaurus
7. Gene Ontology: description of biomolecular biology (molecular functions, biological processes, cellular components) - designed to structure the knowledge rather than index content
8. National Cancer Institute (NCI) thesaurus: knowledge model enabling cross-disciplinary communication and collaboration

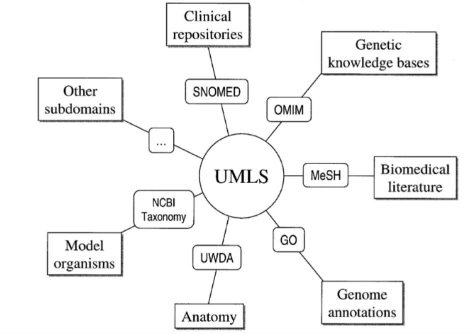
Many thesauri are also available in many well-endowed languages.

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

- Purpose: provide a mecanism to link existing medical thesaurus and controlled vocabularies
- Initiated in 1986 and maintained by the National Library of Medicine
- Contains: a metathesaurus, a semantic network, NLP tools
- Gathers more than 100 thesauri/vocabulary



Bodenreider, O. (2004) The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research, 32, D267-D270.*

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

Basic View | Report View | Raw View

- Concept: [C0004238] Atrial Fibrillation
- Semantic Type
- Definition
- Synonyms (96)
- Relations (1672) REL | RELA | RSAB| String | CUI

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

Basic View | Report View | Raw View

⊟ Concept: [C0004238] Atrial Fibrillation

DA Date 1990-09-30 05:00:00.000000000

MR Major Revision Date 2017-09-14 06:00:00.000000000

ST Status R

⊟ Semantic Type

Disease or Syndrome [T047]

⊞ Definition

⊞ Synonyms (96)

⊞ Relations (1672) REL | RELA | RSAB| String | CUI

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

Basic View | Report View | Raw View

Concept: [C0004238] Atrial Fibrillation

Semantic Type

Definition

CHV/null - rapid tremor and shake of upper chambers of the heart

CSP/null - disorder of cardiac rhythm characterized by rapid, irregular atrial impulses and ineffective atrial contractions.

HPO/null - An atrial arrhythmia characterized by disorganized atrial activity without discrete P waves on the surface EKG, but instead by an undulating baseline or more sharply circumscribed atrial deflections of varying amplitude an frequency ranging from 350 to 600 per minute. [HPO:probinson]

MEDLINEPLUS/null -

An arrhythmia is a problem with the speed or rhythm of the heartbeat. Atrial fibrillation (AF) is the most common type of arrhythmia. The cause is a disorder in the heart's electrical system.

Often, people who have AF may not even feel symptoms. But you may feel

- Palpitations -- an abnormal rapid heartbeat
- Shortness of breath
- Weakness or difficulty exercising
- Chest pain
- Dizziness or fainting
- Fatigue
- Confusion

AF can lead to an increased risk of stroke. In many patients, it can also cause chest pain, heart attack, or heart failure.

Doctors diagnose AF using family and medical history, a physical exam, and a test called an electrocardiogram (EKG), which looks at the electrical waves your heart makes. Treatments include medicines and procedures to restore normal rhythm.

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

Basic View | Report View | Raw View

- Concept: [C0004238] Atrial Fibrillation
- Semantic Type
- Definition
- Synonyms (96)
  - ACFA (arythmie complète par fibrillation auriculaire)
  - AF
  - AF - Atrial fibrillation
  - AFib
  - ATRIAL FIBRILLATION
  - ATRIJ, FIBRILACIJA
  - AURICULAR FIBRILLATION
  - AURICULAR, FIBRILACION
  - Afib
  - Atrial Fibrillation
  - Atrial Fibrillation [Disease/Finding]
  - Atrial Fibrillations
  - Atrial fibrillation
  - Atrial fibrillation (disorder)
  - Atrieflimmer
  - Atriumfibrillatie
  - Auricular Fibrillation
  - Auricular Fibrillations

Medical Knowledge Resources

Existing Medical thesauri (in English)

The Unified Medical Language System (UMLS)

Basic View | Report View | Raw View

⊞ Concept: [C0004238] Atrial Fibrillation
⊞ Semantic Type
⊞ Definition
⊞ Synonyms (96)
⊟ Relations (1672) REL | RELA | RSAB| String | CUI

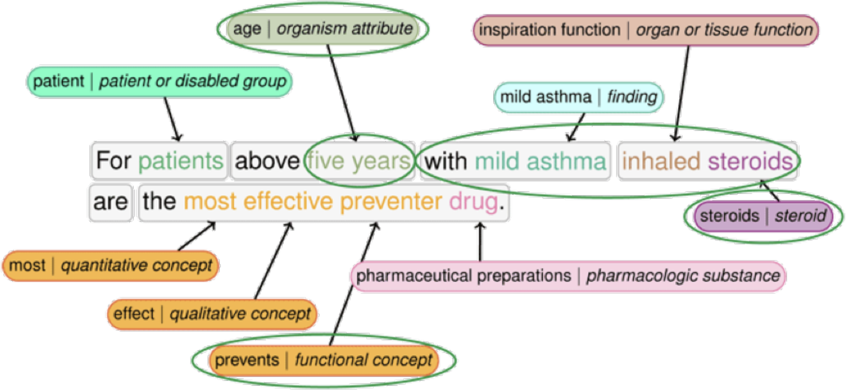
[: 1 - 10 : »]

AQ | | MSH | In Blood | C0005768
AQ | | MSH | In Cerebrospinal Fluid | C0007807
AQ | | MSH | chemically induced | C0007994
AQ | | MSH | Taxonomic | C0008903
AQ | | MSH | Congenital MeSH qualifier | C0009678
AQ | | MSH | nutritional management | C0012160
AQ | | MSH | pharmacotherapeutic | C0013217
AQ | | MSH | Economic | C0013557
AQ | | MSH | embryologic | C0013943
AQ | | MSH | enzymology | C0014445

Medical Knowledge Resources

Semantic Annotation

Annotated sentence:

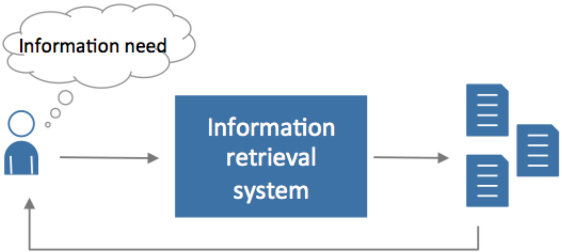


http://ieg.ifs.tuwien.ac.at/~gschwand/mapface/project_page/img/corrections.gif

OUTLINE

Keyword search on texts

The IR view



- Traditional IR data: queries and documents
- Traditional IR models: lexical representation and matching [Robertson et al., 1981, Salton, 1971]
- Traditional IR relevance assumption: topical relevance [Borlund, 2003]

Lexical matching Vs. Semantic matching

On the semantic gap in IR

She takes **just like a woman**, yeah she does. She makes love **just like a woman**, yeah she does And she aches **just like a woman** But she breaks **just** like a little girl

Just like a woman

One of many Dylan songs with an unclear subject. It's often thought tobe about fellow folk-singer Joan Baez, with whom Dylan had arelationship. Edie Sedgwick, an actress affiliated with Andy Warhol, isalso thought to have inspired the song.

- Understand broad language: what's behind the surface of strings?
  - **Semantic representation rather than string representation**
  - Disambiguation of **entities**, **concepts** and roles
  - Reasoning and inference of relations
- Understand broad relevance: what's behind the surface of matching?
  - Semantic matching rather than string matching
  - Relevance matching vs. semantic matching [Guo et al., 2016]

Outline

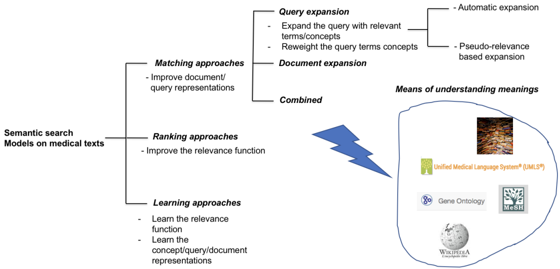
Roadmap

- Traditional IR
  - Q: bag of words
  - D: bag of words
  - RSV(Q,D): Alignment of Q and D

- Semantic (medical) IR
  - Q:
    - Bag of words
    - Bag of words **and concepts/entities**
    - **Embeddings**
  - D:
    - Bag of words
    - Bag of words **and concepts/entities**
    - **Embeddings**
  - RSV(Q,D): **Semantic inference**

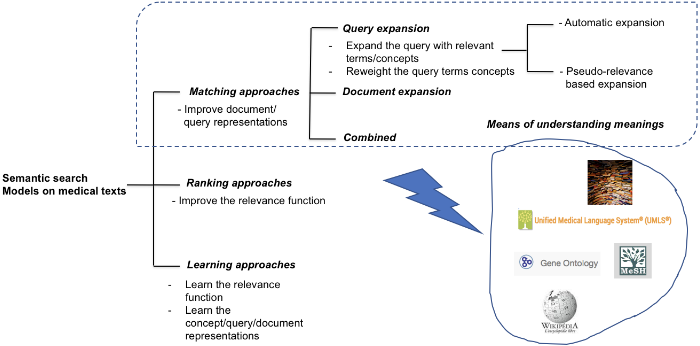
Roadmap

- Traditional IR
  - Q: bag of words
  - D: bag of words
  - RSV(Q,D): Alignment of Q and D
- Semantic (medical) IR
  - Q:
    - Bag of words
    - Bag of words **and concepts/entities**
    - **Embeddings**
  - D:
    - Bag of words
    - Bag of words **and concepts/entities**
    - **Embeddings**
  - RSV(Q,D): **Semantic inference**



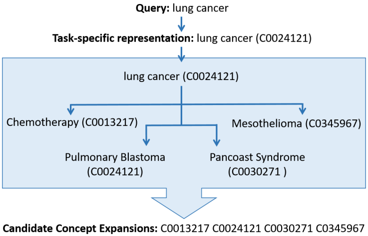
OUTLINE

QUERY/DOCUMENT EXPANSION



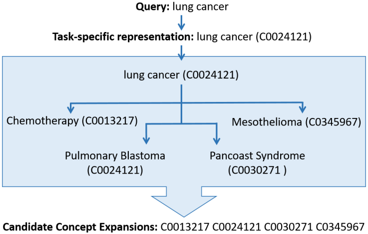
QUERY/DOCUMENT EXPANSION

- Query/document expansion
  - Enhance the Query/Document using:
    - evidence from related words/terms in semantic resources;
    - relevance feedback signals

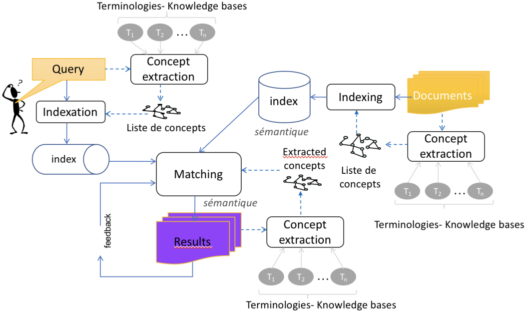


Query/document expansion

- Query/document expansion
  - Enhance the Query/Document using:
    - evidence from related words/terms in semantic resources;
    - relevance feedback signals

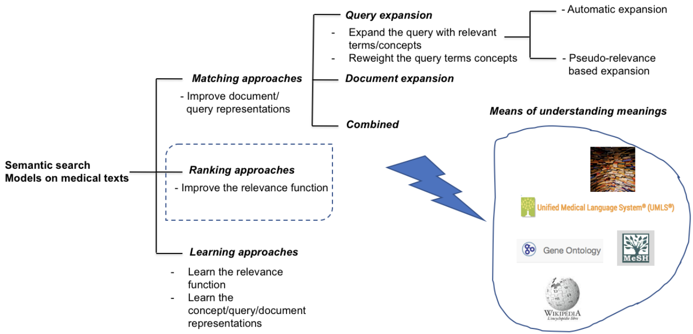


  - The full framework



Outline

Document Ranking



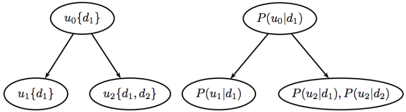
Document ranking

- How to incorporate semantics in the document relevance estimation?
  - Ranking as a semantic inference [Goodwin and Harabagiu, 2016, Koopman et al., 2016, Cao et al., 2011]
  - Ranking as learning the discriminant relevant (semantic) features [Balaneshin-kordan and Kotov, 2016, Xiong and Callan, 2015, Soldaini and Goharian, 2017]

Document ranking

Ranking as a semantic inference: a graph-based approach [Koopman et al., 2016]

- **Key model components**
  - Graph-based representation of the documents
  - Document ranking as an inference process over related concepts in the graph
  - Knowledge resources with directed relationships between concepts
  - Different types of relationships
- **Key inference rationale:** tune the inference mechanism according to semantic gap issues: lexical mismatch, granularity mismatch, conceptual mismatch
  - Lexical mismatch (eg., *hypertension* vs. *high blood pressure*): association and deductive inference
  - Granularity mismatch (eg., *antipsychotic* and *Diazepman*): introduce uncertainty in the taxonomic (hierarchical eg., IS A) relationships
  - Conceptual mismatch (eg., *treatments* $\rightarrow$ *disease*): deductive inference and logical deduction
- The Graph-based corpus representation



(a) Basic node-document representation. (b) Representation with initial probabilities assigned to node.

Document ranking

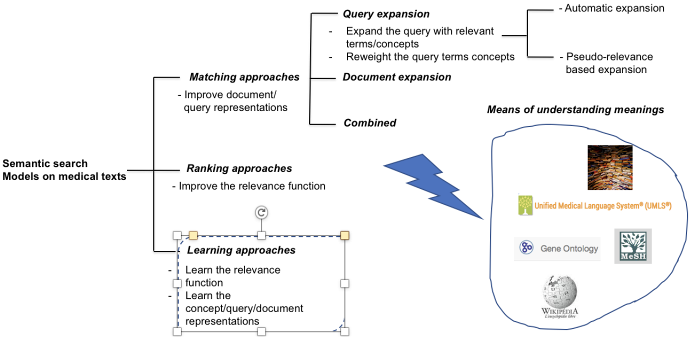
Discussion

- **A few work addressed the semantic search at the relevance function level**
  - Identify logical matching between words and concepts
  - Identify relevant semantic features that connect words to concepts, queries to documents

- **Findings: the general trend**
  - High-level inference yiels to high computational complexity
  - The good balance between lexical matching and semantic matching is difficult to tune
  - Robustness to concept annotation quality is important

Outline

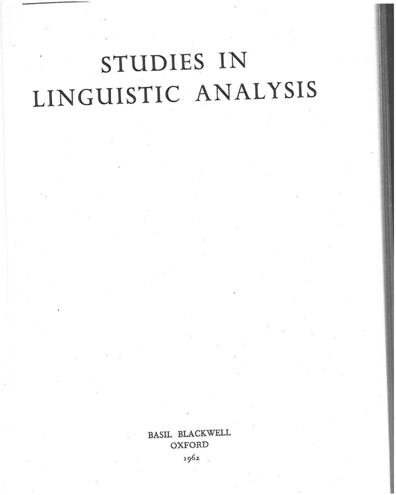
LEARNING


Semantic search
Models on medical texts
Matching approaches
- Improve document/
query representations
Query expansion
- Expand the query with relevant terms/concepts
- Reweight the query terms concepts
- Automatic expansion
- Pseudo-relevance based expansion
Document expansion
Combined
Ranking approaches
- Improve the relevance function
Learning approaches
- Learn the relevance function
- Learn the concept/query/document representations
Means of understanding meanings
Unified Medical Language System® (UMLS®)
Gene Ontology
MeSH
WIKIPEDIA

Fundamentals

Distributional semantics

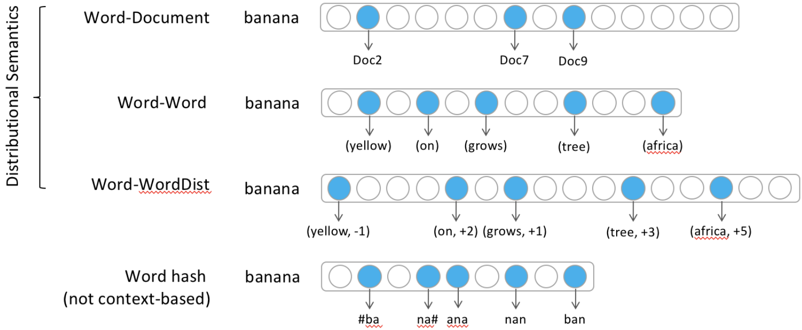
You shall know a word by the company it keeps


STUDIES IN
LINGUISTIC ANALYSIS

BASIL BLACKWELL
OXFORD
1962

Fundamentals

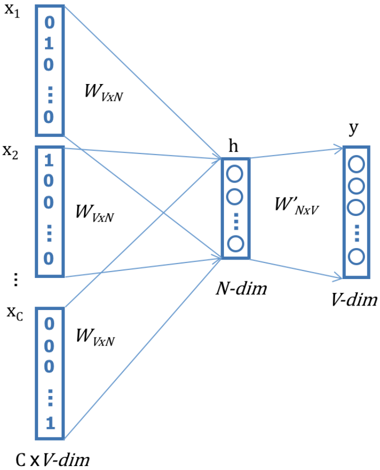
Distributed representations of words



Fundamentals

Distributed representations of words: seminal work [Mikolov et al., 2013]

- **The CBOW model**



$$h = \frac{1}{C} W^T \cdot \left(\sum_{i=1}^{C} x_i \right)$$

$$y = W'^T \cdot h$$

$$P\left(y_j | \{x_1, x_2, \ldots, x_C\}\right) = \frac{exp(y_j)}{\sum_{j'=1}^{V} exp(y_{j'})}$$

Objective function:

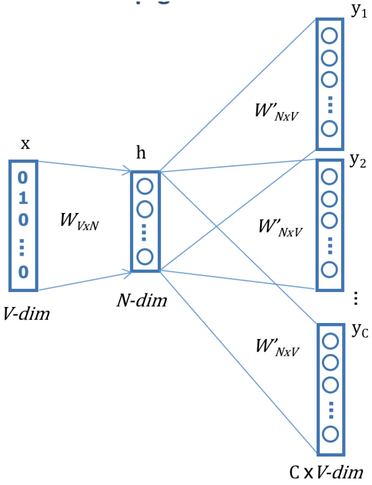
$$E = -\log P(y_j | \{x_1, x_2, \ldots, x_C\})$$

$$E = -y_j + \log \sum_{j'=1}^{V} \exp(y_{j'})$$

Fundamentals

Distributed representations of words: seminal work [Mikolov et al., 2013]

- **The Skip-Gram model**



$$h = W^T \cdot x$$

$$y_c = W'^T \cdot h$$

$$P\left(y_{c,j}|x\right) = \frac{exp(y_{c,j})}{\sum_{j'=1}^{V} exp(y_{j'})}$$

Objective function:

$$E = -\log P(y_1, y_2, \ldots, y_c|x)$$

$$= -log \prod_{c=1}^{C} \frac{exp\left(y_{c,j}\right)}{\sum_{j'=1}^{V} exp\left(y_{j'}\right)}$$

$$= -\sum_{j=1}^{C} y_j + C \cdot log \sum_{j'=1}^{V} exp(y_{j'})$$

Representation learning for medical search

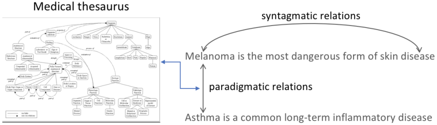
Overview of early research

- What do the models learn?
  - *Word, concept embeddings*: bridge the gap between explicit semantics driven by knowledge resources and implicit semantics driven by the corpus [De Vine et al., 2014, Limsopatham and Collier, 2016, Liu et al., 2016, Ghosh et al., 2017]
  - *Word, concept and document embeddings*: ...to improve semantic document representations [JA et al., 2014, Nguyen et al., 2017, Loza Mencía et al., 2016, Peng et al., 2016, Choi Y, 2016]
  - *Medical objects of interest*: care events/episodes, disease [Ghosh et al., 2016, Moen et al., 2015, Choi et al., 2016], patient representations [Baytas et al., 2017, Ni et al., 2017, Zhu et al., 2016]
- For which search tasks?
  - Relevance matching (eg., document retrieval, case-episode retrieval)
  - Semantic matching (eg., patient similarity)

Representation learning for medical search

Learning word, concept representations

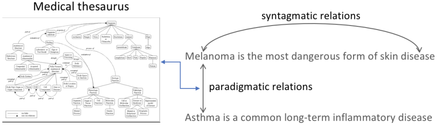
- Different purposes yield to different objective functions
  - Learn readable concept representations from raw texts: driven by syntactic and paradigmatic relations provided in knowledge-bases
  - Learn concept representations from annotated texts: valid through concept similarity provided by knowledge bases
  - Learn concept and associated poly-senses



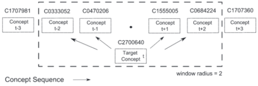
Representation learning for medical search

Learning word, concept representations

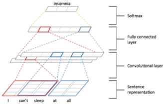
- Different purposes yield to different objective functions
  - Learn readable concept representations from raw texts: driven by syntactic and paradigmatic relations provided in knowledge-bases
  - Learn concept representations from annotated texts: valid through concept similarity provided by knowledge bases
  - Learn concept and associated poly-senses



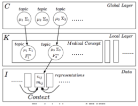
- Different neural architectures
  - Extension of the CBOW and Skip-Gram models
  - Deep architectures (CNN, RNN, ...)



Devine et al. 2014



Limsopatham et al. 2016



Representation learning for medical search

Discussion

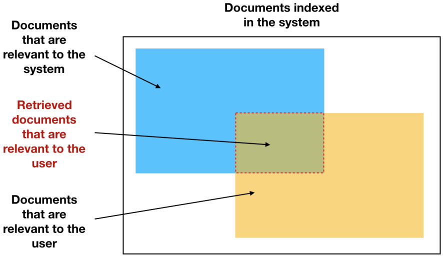
- In summary
  - Recent trend toward the use of neural models in medical search: early stage, not yet mature work but seem promising
  - Learned representations reusable in a wide range of search tasks and prediction tasks
  - Background knowledge (eg., Knowledge-base, expert's assessments) driven representations increases the readability of the representations
- Pending issues
  - What are the impacting factors? What works vs. fails in the black box?
  - Non availability of a hight amount of labeled data (eg., patient similarity, IR tasks)
  - Sensitivity to a large size of network parameters, hyper-parameters and models parameters

OUTLINE

Challenges in Evaluating Medical Information Retrieval

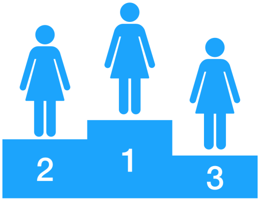
Evaluation at the system level

It's all a matter of precision and recall...



$$Precision = \frac{|P \cap R|}{|R|}, Recall = \frac{|P \cap R|}{|P|}$$

... And of rank!



Unless they are looking for the entire set of documents, nobody goes through the entire set of results.
Ranked metrics:

- P@N
- Mean Average Precision (MAP) [Voorhees, 1998]
- Normative Discounted Cumulation Gain [Jarvelin and Kekalainen, 2000]

Challenges in Evaluating Medical Information Retrieval

Evaluation at the document level

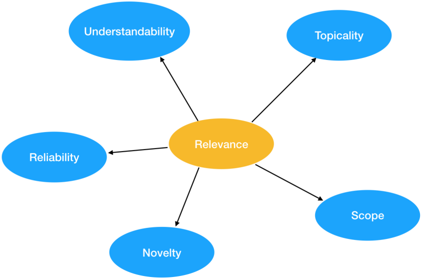
In classical IR

A relevant document contains the query's terms (topicality)

In semantic IR

A relevant document contains terms that are semantically related to the query's terms (semantic topicality)

Relevance has many other dimensions [Zhang et al., 2014]



In the medical domain:

- For patients:
  - Documents must be readable and understandable for a given user
  - The information contained in the documents should be trustworthy
- For medical professionals:
  - Documents must contain up-to-date information
  - Documents must properly cover the topic searched

Evaluation at the document level

Integration of relevance dimensions in the evaluation metrics [Zuccon, 2016]:

- Gain-Discount framework: $M = \frac{1}{N} \sum_{k=1}^{K} d(k).g(d@k), g(d@k) \propto f(P(R|d@k))$ with K the depth of the assessment, $d(k)$ the discount function and $g(d@k)$ the gain function for document d at rank k
- Integration of the relevance dimensions in this framework: $P(R|d@k) = P(D_1, ..., D_n|d@k) = \prod_{i=1}^{K} P(D_i|d@k)$
- Rank-biased precision: $RBP = (1-\rho) \sum_{k=1}^{K} \rho^{r-1} r(d@k)$, with $r(d@k)$ and estimation of $f(P(R|d@k))$, ρ^{k-1} a geometric function of the rank estimating the discount, and $1-\rho$ a normalisation component
- Adaptation of the Rank-Biased Precision measure to topicality- and understandability-based relevance :

$$uRBP = (1-\rho) \sum_{k=1}^{K} \rho^{k-1} r(d@k).u(d@k)$$

$$uRBP \propto (1-\rho) \sum_{k=1}^{K} \rho^{k-1} P(R|d@k).P(U|d@k)$$

Challenges in Evaluating Medical Information Retrieval

Each search task has its proper objectives:

- How should the retrieval and the ranking be implemented?
- How should the system be evaluated?

Examples:

- Physician adhoc search: priority given to the rank, P@10, the topicality, scope...
- Patient adhoc search: priority given to the rank, P@10, the topicality, understandability, readability...
- Clinical trials: priority given to the rank, the topicality, the scope, the novelty...
- Systematic reviews: priority given to the recall, the topicality, the scope...

OUTLINE

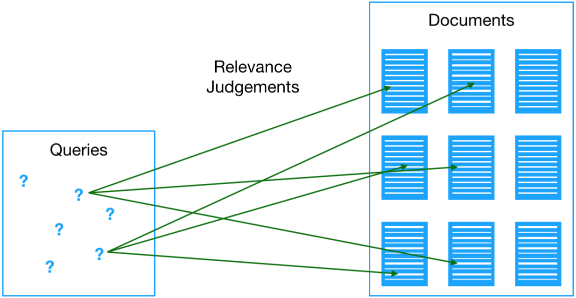
Evaluation Challenges

What is a benchmark?

- Comparing 2 search systems results on a common dataset allows to compare their effectiveness.
- These common datasets are called *benchmarks*.

An IR benchmark contains:

- A document collection that can be indexed
- A set of topics (enriched queries)
- Relevance judgements (linking queries to the relevant documents in the collection)



Evaluation Challenges

The Cranfield Paradigm

Given:

1. A test collection (T, D, R)
2. A retrieval run for the test collection : a doc-list L_t for each topic t in T

For each topic t in T

- Use a measure (e.g. P@10) to compute the quality of L_t

Combine scores:

- Mean average precision

Relevance judgement:

- For a given topic $t \in T$, a given document $d \in D$, $R(d,t)$ is the relevance score of d for topic t.
- $R(d,t)$ can be:
  - a discrete value: e.g. $\in 0, 1$ for binary assessment or $\in 0, 1, 2, 3$ for graded assessment
  - a continuous value: e.g. $\in [0, 1]$
- Assumption: if $R(d,t,u_1)$ is the judgement of assessor u_1 on topic t and document d and $R(d,t,u_2)$ the judgement of assessor u_2 on topic t and document d, $R(d,t,u_1) = R(d,t,u_2)$

Summary of the Benchmarking Activities

Venue	Task	Dataset	Activity
TREC	Genomics adhoc retrieval	Clinical information need Biomedical articles	Terminated
	Genomics passage retrieval	Clinical information need Biomedical articles	Terminated
	Medical records	Patient cohort search	Terminated
	Clinical decision support / Precision medicine	Case reports Biomedical articles	Ongoing
CLEF	ImageCLEF medical retrieval	Image and medical reports Collection of medical images	Terminated
	CLEF eHealth consumer search	Health information need Large web crawl	Ongoing
	CLEF eHealth technological assisted reviews	Boolean queries Biomedical articles	Ongoing

The majority of these datasets are still available and can be used for research!

Conclusion

A large and growing body of work on semantic search in the medical domain

- Focus on task, user profile, information need elicitation in context (time, task, user's expertise, etc.)
- Model semantic w.r.t. polyrepresentation view: document collections, knowledge bases, users, etc.
- Shift from lexical matching to sematic matching by considering domain-specific peculiarities
- Understand relevance assessment facets according to task, user (laypeople vs.expert)

References I

Balaneshin-kordan, S. and Kotov, A. (2016).
Optimization method for weighting explicit and latent concepts in clinical decision support queries.
In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, ICTIR '16, pages 241–250, New York, NY, USA. ACM.

Bast, H., Björn, B., and Haussmann, E. (2016).
Semantic search on text and knowledge bases.
Found. Trends Inf. Retr., 10(2-3):119–271.

Baytas, I. M., Xiao, C., Zhang, X., Wang, F., Jain, A. K., and Zhou, J. (2017).
Patient subtyping via time-aware lstm networks.
In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, pages 65–74.

Blackman, S. (2006).
Nature has wikipedia in its cities.
The scientist, 20(2).

Blum, B. (1984).
Information Systems for Patient Care.
Springer-Verlag.

Borlund, P. (2003).
The concept of relevance in ir.
J. Am. Soc. Inf. Sci. Technol., 54(10):913–925.

Boudin, F., Nie, J.-Y., and Dawes, M. (2010).
Clinical information retrieval using document and pico structure.
In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, HLT '10, pages 822–830, Stroudsburg, PA, USA. Association for Computational Linguistics.

References II

Broder, A. (2002).
A taxonomy of web search.
SIGIR Forum, 36(2):3–10.

Cao, Y., Liu, F., Simpson, P., Antieau, L., Bennett, A., Cimino, J. J., Ely, J., and Yu, H. (2011).
Askhermes: An online question answering system for complex clinical questions.
Journal of Biomedical Informatics, 44(2):277 – 288.

Cartright, M.-A., White, R. W., and Horvitz, E. (2011).
Intentions and attention in exploratory health search.
In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, pages 65–74, New York, NY, USA. ACM.

Choi, E., Bahadori, M. T., Searles, E., Coffey, C., Thompson, M., Bost, J., Tejedor-Sojo, J., and Sun, J. (2016).
Multi-layer representation learning for medical concepts.
In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 1495–1504.

Choi Y, Chiu CY-I, S. D. (2016).
Learning low-dimensional representations of medical concepts.
In *AMIA Summits on Translational Science Proceedings*, pages 41–50.

De Vine, L., Zuccon, G., Koopman, B., Sitbon, L., and Bruza, P. (2014).
Medical semantic similarity with a neural language model.
In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, CIKM '14, pages 1819–1822, New York, NY, USA. ACM.

Dinh, D. and Tamine, L. (2012).
Towards a context sensitive approach to searching information based on domain specific knowledge sources.
Web Semantics: Science, Services and Agents on the World Wide Web, 12(0).

References III

Dinh, D., Tamine, L., and Boubekeur, F. (2013).
Factors affecting the effectiveness of biomedical document indexing and retrieval based on terminologies.
Artificial Intelligence in Medicine, 57(2):155 – 167.

Ely, J. W., Osheroff, J. A., Ebell, M. H., Bergus, G. R., Levy, B. T., Chambliss, M. L., and Evans, E. R. (1999).
Analysis of questions asked by family doctors regarding patient care.
BMJ, 319:358–361.

Firth, J. R. (1957).
"a synopsis of linguistic theory 1930-1955". studies in linguistic analysis.
F.R. Palmer, ed. (1968). Selected Papers of J.R. Firth 1952-1959, pages 1–32.

Ghosh, S., Chakraborty, P., Cohn, E., Brownstein, J. S., and Ramakrishnan, N. (2016).
Characterizing diseases from unstructured text: A vocabulary driven word2vec approach.
In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 1129–1138.

Ghosh, S., Chakraborty, P., Cohn, E., Brownstein, J. S., and Ramakrishnan, N. (2017).
Joint learning of representations of medical concepts and words from ehr data.
In *Proceedings (IEEE Int Conf Bioinformatics Biomed)*, Biomed '17, pages 764–769.

Goodwin, T. R. and Harabagiu, S. M. (2016).
Medical question answering for clinical decision support.
In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 297–306, New York, NY, USA. ACM.

Gorman, P. (1995).
Inforamtion needs of physicians.
Journal of the American Society for Information Science, 46:729–736.

References IV

Guo, J., Fan, Y., Ai, Q., and Croft, W. B. (2016).
A deep relevance matching model for ad-hoc retrieval.
In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 55–64, New York, NY, USA. ACM.

Hersh, W. (2010).
Information Retrieval - A Health and Biomedical Perspective.
Springer.

Hersh, W. and Voorhees, E. (2009).
Trec genomics special issue overview.
12:1–15.

Huang, X., Lin, J., and Demner-Fushman, D. (2006).
Evaluation of pico as a knowledge representation for clinical questions.
In *AMIA Annual Symposium Proceedings*, pages 359–363.

JA, M.-G., O, M.-A., and M, S. (2014).
Exploring the application of deep learning techniques on medical text corpora.
In *Studies in health technology and in informatics*, pages 205:584–8.

Jarvelin, K. and Kekalainen, J. (2000).
IR evaluation methods for retrieving highly relevant documents.
In *Proceedings of the 23rd SIGIR conference*, pages 41–48.

Jimmy, Zuccon, G., and Koopman, B. (2018).
Choices in knowledge-base retrieval for consumer health search.
In *Proceedings of ECIR'2018*. LNCS.

References V

Kanoulas, E., Li, D., Azzopardi, L., and Spijker, R. (2017).
Clef 2017 technologically assisted reviews in empirical medicine overview.
In *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF) - Online working notes*.

Koopman, B., Zuccon, G., Bruza, P., Sitbon, L., and Lawley, M. (2016).
Information retrieval as semantic inference: A graph inference model applied to medical search.
Inf. Retr., 19(1-2):6–37.

Lancaster, F. and Warner, A. (1993).
Information Retrieval Today.

Limsopatham, N. and Collier, N. (2016).
Normalising medical concepts in social media texts by learning semantic representation.
In *Proceedings of the International Conference on the Association for Computational Linguistics*, ACL '16, pages 1014–1023.

Limsopatham, N., Macdonald, C., and Ounis, I. (2013).
Learning to handle negated language in medical records search.
In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, CIKM '13, pages 1431–1440, New York, NY, USA. ACM.

Liu, X., Nie, J.-Y., and Sordoni, A. (2016).
Constraining word embeddings by prior knowledge – application to medical information retrieval.
pages 155–167.

Loza Mencía, E., de Melo, G., and Nam, J. (2016).
Medical concept embeddings via labeled background corpora.
In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 4629–4636.

References VI

Martinez, D., Otegi, A., Soroa, A., and Agirre, E. (2014).
Improving search over electronic health records using umls-based query expansion through random walks.
Journal of Biomedical Informatics, 51:100 – 106.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013).
Efficient estimation of word representations in vector space.
CoRR, abs/1301.3781.

Moen, H., Ginter, F., Marsi, E., Peltonen, L.-M., T, S., and S., S. (2015).
Care episode retrieval: distributional semantic models for information retrieval in the clinical domain.
BMC Medical Informatics and Decision Making, 15.

Montori, V. M., Wilczynski, N. L., Morgan, D., and Haynes, R. B. (2004).
Systematic reviews: a cross-sectional study of location and citation count.
BMC Medicine, 1(2).

Natarajan, K., Stein, D., Jain, S., and Elhadad, N. (2010).
An analysis of clinical queries in an electronic health record search utility.
International Journal of Medical Informatics, 79(7):515–522.

Nguyen, G.-H., Tamine, L., Soulier, L., and Souf, N. (2017).
Learning Concept-Driven Document Embeddings for Medical Information Search.
In *Conference on Artificial Intelligence in Medicine (AIME 2017)*, Vienna, Austria.

Ni, J., Liu, J., Zhang, C., Ye, D., and Ma, Z. (2017).
Fine-grained patient similarity measuring using deep metric learning.
In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, CIKM '17, pages 1189–1198.

References VII

Oh, H.-S. and Jung, Y. (2015).
Cluster-based query expansion using external collections in medical information retrieval.
Journal of Biomedical Informatics, 58:70 – 79.

Peng, S., You, R., Wang, H., Zhai, C., Mamitsuka, H., and Shanfeng Zhu1, . (2016).
Deepmesh: deep semantic representation for improving large-scale mesh indexing.
Bioinformatics, 32.

Robertson, S. E., van Rijsbergen, C. J., and Porter, M. F. (1981).
Probabilistic models of indexing and searching.
In *Proceedings of the 3rd Annual ACM Conference on Research and Development in Information Retrieval*, SIGIR '80, pages 35–56, Kent, UK, UK. Butterworth & Co.

Salton, G. (1971).
The SMART Retrieval System—Experiments in Automatic Document Processing.
Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Soldaini, L. and Goharian, N. (2017).
Learning to rank for consumer health search: A semantic approach.
In Jose, J. M., Hauff, C., Altıngovde, I. S., Song, D., Albakour, D., Watt, S., and Tait, J., editors, *Advances in Information Retrieval*, pages 640–646, Cham. Springer International Publishing.

Soldaini, L., Yates, A., and Goharian, N. (2017).
Learning to reformulate long queries for clinical decision support.
Journal of the Association for Information Science and Technology, 68(11):2602–2619.

Soldaini, L., Yates, A., Yom-Tov, E., Frieder, O., and Goharian, N. (2016).
Enhancing web search in the medical domain via query clarification.
Inf. Retr., 19(1-2):149–173.

REFERENCES VIII

Sondhi, P., Sun, J., Zhai, C., Sorrentino, R., and Kohn, M. (2012).
Leveraging medical thesauri and physician feedback for improving medical literature retrieval for case queries.
Journal of the American Medical Informatics Association: JAMIA, 19(5):851–858.

Voorhees, E. (1998).
Variations in relevance judgments and the measurement of retrieval effectiveness.
In *Proceedings of the 21st SIGIR conference*, pages 315–323.

Wang, C. and Akella, R. (2015).
Concept-based relevance models for medical and semantic information retrieval.
In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, pages 173–182, New York, NY, USA. ACM.

White, R. W. and Horvitz, E. (2009).
Cyberchondria: studies of the escalation of medical concerns in web search.
ACM Transactions on Information Systems (TOIS), 27(4):23.

Wilkinson, R. and Fuller, M. (1996).
Integration of information retrieval and hypertext via structure.
Information Retrieval and Hypertext, pages 257–271.

Xiong, C. and Callan, J. (2015).
Esdrank: Connecting query and documents through external semi-structured data.
In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, pages 951–960, New York, NY, USA. ACM.

Zhai, C. and Lafferty, J. (2001).
Model-based feedback in the language modeling approach to information retrieval.
In *Proceedings of the Tenth International Conference on Information and Knowledge Management*, CIKM '01, pages 403–410, New York, NY, USA. ACM.

References IX

Zhang, Y., Wang, P., Heaton, A., and Winkler, H. (2012).
Health information searching behavior in medlineplus and the impact of tasks.
In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, IHI'12, pages 641–650.

Zhang, Y., Zhang, J., Lease, M., and Gwizdka, J. (2014).
Multidimensional relevance modeling via psychometrics and crowdsourcing.
In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '14, pages 435–444, New York, NY, USA. ACM.

Zhu, D. and Carterette, B. (2012).
Improving health records search using multiple query expansion collections.
In *2012 IEEE International Conference on Bioinformatics and Biomedicine*, pages 1–7.

Zhu, Z., Yin, C., Qian, B., Cheng, Y., Wei, J., and Wang, F. (2016).
Measuring patient similarities via a deep architecture with medical concept embedding.
In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 749–758.

Znaidi, E., Tamine, L., and Latiri, C. (2015).
Answering pico clinical questions: A semantic graph-based approach.
In Holmes, J. H., Bellazzi, R., Sacchi, L., and Peek, N., editors, *Artificial Intelligence in Medicine*, pages 232–237, Cham. Springer International Publishing.

Znaidi, E., Tamine, L., and Latiri, C. (2016).
Aggregating semantic information nuggets for answering clinical queries.
In *Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, April 4-8, 2016*, pages 1041–1047.

Zuccon, G. (2016).
Understandability biased evaluation for information retrieval.
In *European Conference on Information Retrieval*, pages 280–292. Springer.