Large Scale Study on Health Information Retrieval for layperson

COMP8755 – Initial Presentation

Supervisor: Assoc/Prof Hanna Suominen & Mr. Chenchen Xu
Zhengyang Liu, u6022937
The Australian National University
02/04/2020
Motivation

Clipart Library, n.d.. Retrieved from: http://clipart-library.com/clipart/gTeEeAEenc.htm
CLEF eHealth, 2020, Retrieved from: https://clefehealth.imag.fr/
Introduction

Initial Dataset -> Preprocessing -> Clean document collection -> ElasticSearch Index

Queries -> Query Expansion -> Extended Queries

Retrieval -> Ranked Docs
Dataset

- 2ndchance.info
- 03center2018:technology
- 3dprintingindustry.com
- 04center2018:technology
- 123helpme.com
- 164andmore.com
- 1800theeagle.com
- aaaa.org
- aacp.com
- aademicmeeting.org
- aafp.org
- aafa.org
- aafp.org
- aais.ca
- aap.org
- aapd.org
- aar.org
- aasm.org
- aou.edu
- abim.org
- aboutkidshealth.ca
- acaai.org
- acad.org
- academia.edu
- academic.oup.com
- academicon.com
- acatoday.org
- accesslabs.com
- accessdata.fda.gov
- accidentawillinois.com
- accconn.com
- acefitness.org

- anemia diet changes
- what products are available to help giving up smoking
- what foods contain omega 3 fatty acids and what are the benefits of eating them
- what are the common errors in prescribing drugs
- how do you cope with anxieties
- what are the benefits of spirulina
- is there a relationship between headaches and allergies
- what is ketamine what are the problems associated with it
- head and neck cancer
- HIV vaccine trials
- blood test normal values
- causes of withdrawal
- Rheumatoid arthritis causes
- Breast lift operations
- what can cause involuntary trembling
- drug food interactions
- effect of breakfast on school performance
- use voltaran emulgel 1% for a sore ankle
- treatment for pelvic inflammatory disease
- skipping breakfast school performance diet recommendations
- drug addiction organisations in ireland
- insulin resistance
- diabetes support in usa
- apnea syndrome treatment
- marijuana memory effects
- sore muscles after exercise
- percent of water in celery
- prostate cancer antiandrogen therapy
- prostate cancer toxins chemicals
- hearing loss stem cell treatment
- high blood pressure and smoking
- coffee blood pressure consequences
- gonorrhea long term effects
- high blood pressure food activities
- canada private health care
- liver function blood results
- diabetes melitus body fat mass
- ionised water therapy

html documents and queries
Background & Approach

• Information Retrieval models: Boolean Queries, Vector Space Model, Okapi BM25

• Query expansion: Reinforcement learning
Evaluation

Roadmap

• Literature review. (done)

• Document Indexing. (done)

• Query Expansion. (April)

• Evaluation. (May)

• 2 CLEF eHealth lab overview papers. (June)