

Web Search and Mining (F033583)



The Search System

realselfSEARCH

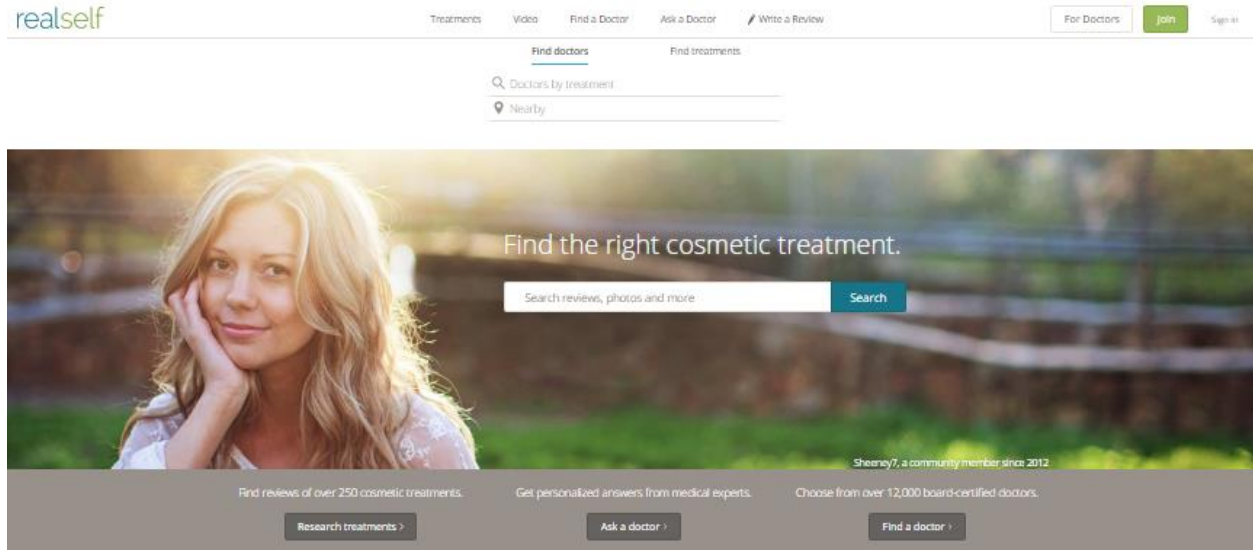
Charlie Kruger J11603099010

Daniel Aldana 116030990081

Introduction

Realself is the website where you can find a veridic information about any kind of treatment, where will be performed and under which hands will be made, the website consists on a large database of doctors, surgeons and specialists of the health, the website as well has a large section of questions where the patients ask about any concerns and the doctors across the world may answer.

Every doctor will be graded according to 1 to 5-star qualification system for every treatment that have performed, the main idea of realself is to offer serious and trusty reviews about doctors, treatments and patient inquiries.



Main Workflow of realself search system

The crawled website it is <https://www.realself.com/>, it will be crawled two specific sections: the doctor section which is under `/xmlsitemap-Dr{}.xml` and questions section, contained in `/xmlsitemap-Question{}.xml` both of these sections consists of the titles of the questions or doctors names, the heading which is the answer title of the question and the title of the reviews that a doctor is answering along with the content of the question and doctor information

The system consists of the following parts:

1. The Web Crawler which crawls the follow site and obtain the frequently asked questions and the doctor list.
2. The Parser obtains extracts specific information of the crawled files in bulk and gives a first structure.
3. The Intermediate Part creates from the extracted information preliminary files per word and connects it to the webpages where those words occur.
- 4.1 The PageRanker calculates the popularity of a webpage based on the number of links which point to it from other webpages.
- 4.2 The QualityScorer assigns a quality to a webpage. The quality is determined by the number of answers on a question website and by the amount of reviews in case of a doctor's page.
- 4.3 Term Frequency-Inverse Dictionary Frequency(TF-IDF) is the matrix for tf-idf scoring will sort all the documents by the respective relevance.
5. The Indexer finally creates the final dictionary using the combined score of PageRanker, Quality Score and TF-IDF.
6. This can then be used via a Graphical User Interface (GUI) that consists of a search button and a textbox for the input of the keywords.
7. The Query Parser converts the keywords to an appropriate format which can be matched against the dictionary.
- 8.The returned Results are presented on the graphical interface in a ranked fashioned and with links.

Detailed Description of the Search System Elements

Web Crawler:

The .html files are stored in the user's directory under the folder `tsg/raw` once there the folders will be created per the crawled files downloaded, those categories are listed as following

- Questions
- Doctors

Item Names

Crawled Site: SITE = '<https://www.realself.com>'

Questions items:

Questions Site: `/xmlsitemap-Question{}.xml`

Questions Files: `faq_question_toronto-70-years-marionette-lines-feel-older-recommend.html`

```
C:\Users\Daniel\DanPycharmProjects\py_src>python 0_crawl faq
INFO:root:Downloading URL https://www.realself.com/xmlsitemap-Question0.xml
INFO:root:Downloading URL site https://www.realself.com/question/toronto-70-years-marionette-lines-feel-older-recommend
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_toronto-70-years-marionette-lines-feel-older-recommend.html
INFO:root:Downloading URL site https://www.realself.com/question/charlotte-nc-42-yrs-time-facelift
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_charlotte-nc-42-yrs-time-facelift.html
INFO:root:Downloading URL site https://www.realself.com/question/toronto-fillers-lips-tear-troughs
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_toronto-fillers-lips-tear-troughs.html
INFO:root:Downloading URL site https://www.realself.com/question/bahrain-rf-micro-prp-after-35-tca-chemical-peel-sensitive-skin-skin-irritated
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_bahrain-rf-micro-prp-after-35-tca-chemical-peel-sensitive-skin-skin-irritated.html
INFO:root:Downloading URL site https://www.realself.com/question/fort-worth-tx-laser-lipo
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_fort-worth-tx-laser-lipo.html
INFO:root:Downloading URL site https://www.realself.com/question/orange-county-ca-lost-breast-profile
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_orange-county-ca-lost-breast-profile.html
INFO:root:Downloading URL site https://www.realself.com/question/bn15-9hj-i-exercised
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_bn15-9hj-i-exercised.html
INFO:root:Downloading URL site https://www.realself.com/question/chicago-il-24y-kids-breast-lift-breast-lift
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_chicago-il-24y-kids-breast-lift-breast-lift.html
INFO:root:Downloading URL site https://www.realself.com/question/miami-fl-changing-implants-change-rippling
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_miami-fl-changing-implants-change-rippling.html
INFO:root:Downloading URL site https://www.realself.com/question/san-jose-ca-long-after-bbl-start-putting-scar-cream-scars
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_san-jose-ca-long-after-bbl-start-putting-scar-cream-scars.html
INFO:root:Downloading URL site https://www.realself.com/question/toronto-long-wait-after-jaw-surgery-before-booking-consultation-combined
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/faq_question_toronto-long-wait-after-jaw-surgery-before-booking-consultation-combined.html
```

Doctor items:

Doctor Site: `/xmlsitemap-Dr{}.xml`

Doctor Files `doctor_Anesthesiologist_Alex-Roher.html`

```
C:\Users\Daniel\DanPycharmProjects\py_src>python 0_crawl doctor
INFO:root:Downloading URL https://www.realself.com/xmlsitemap-Dr0.xml
INFO:root:Downloading URL site https://www.realself.com/find/Florida/Ocala/General-Surgeon/Juan-Gutierrez
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Florida.Ocala.General-Surgeon_Juan-Gutierrez.html
INFO:root:Downloading URL site https://www.realself.com/find/Alabama/Birmingham/Plastic-Surgeon/Robert-Oliver-Jr.
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Alabama.Birmingham.Plastic-Surgeon_Robert-Oliver-Jr..html
INFO:root:Downloading URL site https://www.realself.com/find/New-York/New-York/Dentist/Thomas-Connelly
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_New-York.New-York.Dentist_Thomas-Connelly.html
INFO:root:Downloading URL site https://www.realself.com/find/North-Carolina/Charlotte/Dentist/Mitesh-Dhulab
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_North-Carolina.Charlotte.Dentist_Mitesh-Dhulab.html
INFO:root:Downloading URL site https://www.realself.com/find/North-Carolina/Raleigh/Plastic-Surgeon/Benjamin-Wood
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_North-Carolina.Raleigh.Plastic-Surgeon_Benjamin-Wood.html
INFO:root:Downloading URL site https://www.realself.com/find/Texas/Austin/Physician/Gabriel-Salinas
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Texas.Austin.Physician_Gabriel-Salinas.html
INFO:root:Downloading URL site https://www.realself.com/find/Texas/Houston/Physician/Mark-Hennessey
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Texas.Houston.Physician_Mark-Hennessey.html
INFO:root:Downloading URL site https://www.realself.com/find/Pennsylvania/Philadelphia/Physician/Kathleen-Brennan
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Pennsylvania.Philadelphia.Physician_Kathleen-Brennan.html
INFO:root:Downloading URL site https://www.realself.com/find/Florida/Key-West/Plastic-Surgeon/Scott-Loessin
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Florida.Key-West.Plastic-Surgeon_Scott-Loessin.html
INFO:root:Downloading URL site https://www.realself.com/find/Washington/Kirkland/Plastic-Surgeon/Aysel-Sanderson
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Washington.Kirkland.Plastic-Surgeon_Aysel-Sanderson.html
INFO:root:Downloading URL site https://www.realself.com/find/India/India/Plastic-Surgeon/Dushyanth-Kalva
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_India.India.Plastic-Surgeon_Dushyanth-Kalva.html
INFO:root:Downloading URL site https://www.realself.com/find/New-Jersey/Shrewsbury/Oral-&Maxillofacial-Surgeon/Manolis-Manolakis
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_New-Jersey.Shrewsbury.Oral-&Maxillofacial-Surgeon_Manolis-Manolakis.html
INFO:root:Downloading URL site https://www.realself.com/find/Ontario/Toronto/Plastic-Surgeon/Cory-Goldberg
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Ontario.Toronto.Plastic-Surgeon_Cory-Goldberg.html
INFO:root:Downloading URL site https://www.realself.com/find/Georgia/Atlanta/Plastic-Surgeon/Kristin-Ann-Boehm
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Georgia.Atlanta.Plastic-Surgeon_Kristin-Ann-Boehm.html
INFO:root:Downloading URL site https://www.realself.com/find/Florida/Kissimmee/Plastic-Surgeon/Donovan-Rosas
INFO:root:File at C:\Users\Daniel\tsgdataRS/raw/doctor_Florida.Kissimmee.Plastic-Surgeon_Donovan-Rosas.html
```

we crawled 222,229 files in total, which represents 30Gb of information, the resulting dictionary is magnitudes smaller.

Parser:

The parser read all the crawled html files and converts into a json file with a more structured information with parsed contents.

The output files contain the following list of attributes:

- type - the category of a website; either a question or a doctor
- title – the name question page or the name of the doctor
- heading: (answer titles(questions), review titles(doctor))
- content – the entire text content of a website
- uuid - the identification of the file, how to find unique IDs by hash or content)
- link (the entire link of the html file on the realself.com site)
- listing

The content section is a parsed string list of the content separated by spaces and converted into lowercases.

OUTPUT

The folder parsed will be created inside the raw directory, the names of the files will be UUID.json

Doctor Files

```
C:\Users\Daniel\DanPycharmProjects\py_src>python 1_parse
C:\Users\Daniel\DanPycharmProjects\py_src\tsg\parser\base.py:24: FutureWarning: The behavior of this method will change in future versions. Use specific 'len(elem)' or 'elem is not None' test
.
if not tree.getroot():
INFO:root:Parsed file C:\Users\Daniel\tsg\data\raw\doctor_Alabama.Birmingham.Plastic-Surgeon_Robert-Oliver-Jr..html with uuid doctor_Alabama.Birmingham.Plastic-Surgeon_Robert-Oliver-Jr.
INFO:root:Parsed file C:\Users\Daniel\tsg\data\raw\doctor_Florida.Key-West.Plastic-Surgeon_Scott-Loessin.html with uuid doctor_Florida.Key-West.Plastic-Surgeon_Scott-Loessin
INFO:root:Parsed file C:\Users\Daniel\tsg\data\raw\doctor_Florida.Kissimmee.Plastic-Surgeon_Donovan-Rosas.html with uuid doctor_Florida.Kissimmee.Plastic-Surgeon_Donovan-Rosas
```

* Parts of the content section were removed, otherwise the follow image could not be displayed correctly.

```
doctor_Alabama.Birmingham.Plastic-Surgeon_Robert-Oliver-Jr..json
1  {"listings_count": 19.5, "text": "selecting dr oliver was one of the best decisions i have ever made he went the extra mile and i am 100 satisfied with my procedure i would highly recommend him to anyone and i have actually been doing so dr oliver performed my lower body lift breast lift and breast augmentation i am beyond happy with the treatment and care that ive received from him his nurse mandi and his office staff not only are they extremely professional but they are super responsive and attentive the thin incision lines and overall results are more than i could have ever imagined its obvious that he cares dr robert oliver will be performing my procedure i met him in 2013 and was very impressed i spoke with 4 plastic surgeons in an attempt to find the best fit for myself dr oliver is a 3rd generation surgeon whom truly cares about the well being of his patients he is very thorough and has exceptional follow up skills he actually emails his patients directly and is not offended by questions his goal is to ensure complete understanding of the procedure and to resolve concerns in addition to his professionalism and beautiful outcomes he is also a man of faith which makes me extra happyhis nurse mandi is also a jewel ill disagree with my colleagues somewhat and point out that its a matter of tradeoffs it is the general thought by most plastic surgeons read more i would consider these 2 machines functionally very similar in what they produce we looked at both and went with the apollo as it 1 doesnt require read more i would endorse the use of a product called mepitac which is a kind of inexpensive silicone tape you can buy a 9 foot roll of this product for read more generally speaking the more complex the design of an implant the more likely they are to failrupture this has been demonstrated repeatedly over 30 read more 1 you got implants waaaaaay to big to be supported by your tissues2 you will need a complex revision to improve this and would likely need to use read more these type of lateral brow surgeries are actually very effective for lateral brow hooding which is usually what people come in complaining about read more there is no reason to need to separate your lift and implant procedure per se thats usually only really needed for more complex mastopexies where read more the femma s is a decidedly silly acronym for a horizontal breast reduction that type of incision is kind of finicky and works best for read more as noted we all have our biases for incisions for these types of surgery we uses multiple types t scar true vertical j shaped vertical read more the literature suggests that there is no significant clinical affects on result from compression garments for breast or body contouring surgery at read more", "title": "robert oliver jr md", "url": "https://www.realself.com/find/Alabama/Birmingham/Plastic-Surgeon/Robert-Oliver-Jr", "content": "treatments video find a doctor ask a doctor write a review photos blog realself trends for doctors claim your profile advertise realself top doctors realself fellowship realself university medibeauty today my account join now sign in help 2017 realself inc all rights reserved realself is a registered trademark of realself inc 2017 realself inc all rights reservedrealself is a registered trademark of realself inc terms of service privacy policy", "uuid": "doctor_Alabama.Birmingham.Plastic-Surgeon_Robert-Oliver-Jr.", "heading": "hands down amazing experience great experience birmingham al tummy tuck birmingham al", "type": "raw\\doctor"}
```

Question Files

INFO:root:Parsed file C:\Users\Daniel\tsgdataRS\raw\faq_question_bn15-9hj-i-exercised.html with uuid faq_question_bn15-9hj-i-exercised
INFO:root:Parsed file C:\Users\Daniel\tsgdataRS\raw\faq_question_charlotte-nc-42-yrs-time-facelift.html with uuid faq_question_charlotte-nc-42-yrs-time-facelift

* Parts of the content section were removed, otherwise the follow image could not be displayed correctly.

```
faq_question_bn15-9hj-i-exercised.json
1 [{"listings_count": 4.0, "text": "push ups are not considered light exercise and should not be done at 4 weeks
i would suggest meeting with your surgeon and being checked just to make sure you are doing well please
follow your surgeons advice on exercising thank you for your question at 4 weeks you should not be doing push
ups you need to let your muscles heal and adjust to having implants in place i would make an appointment to
be evaluated by your surgeon to make sure you have not done any damage it is extremely important to follow
your post op instructions", "title": "what are the signs that ive started exercising too soon after breast
augmentation 5 weeks post op", "url": "https://www.realself.com/question/bn15-9hj-i-exercised", "content":
"treatments video find a doctor ask a doctor write a review for doctors sign in for doctors join sign in join
search treatments breast augmentation qa what are the signs that ask a doctor reviews 22933 pictures 32343 qa
29877 videos 1905 forum 1743 guides cost what are the signs that ive started exercising too soon after breast
augmentation 5 weeks post op georgia1305 bn15 9hj 20 hours ago im 5 12 weeks post op but started to do light
arm exercises at about 4 weeks and this included push ups im now worried i might have done some damage
exercise damage 1 month post op doctor answers 4 exercise june 22nd 2017 votecount 0 votecount 1 votecount 1
votecount 0 votecount 1 votecount 1 push ups are not considered light exercise and should not be done at 4
weeks i would suggest meeting with your surgeon and being checked just to make sure you are doing well please
follow your surgeons advice on exercising connie hiers md san antonio plastic surgeon 46 out of 5 stars 13
reviews wwwfindmybeautifulcom visit website june 22nd 2017 exercise votecount 0 votecount 1 votecount 1
votecount 0 votecount 1 votecount 1 push ups are not considered light exercise and should not be done at 4
weeks i would suggest meeting with your surgeon and being checked just to make sure you are doing well please
follow your surgeons advice on exercising connie hiers md san antonio plastic surgeon 46 out of 5 stars 13
reviews wwwfindmybeautifulcom visit website exercise after surgery june 22nd 2017 votecount 0 votecount 1
votecount 1 votecount 0 votecount 1 votecount 1 thank you for your question at 4 weeks you should not be
doing push ups you need to let your muscles heal and adjust to having implants in place i would make an
appointment to be evaluated by your surgeon to make sure you have not done any damage it is extremely
important to follow your post op instructions francis johns md greensburg plastic surgeon 48 out of 5 stars
73 reviews wwwartisan plasticsurgerycom visit website june 22nd 2017 exercise after surgery | ", "uuid":
"faq_question_bn15-9hj-i-exercised", "heading": "exercise exercise exercise after surgery exercise after
surgery really depends really depends exercise after breast augmentation exercise after breast augmentation",
"type": "raw\\faq"}]
```


The Intermediate:

The intermediate reads all JSONs file and generates the full list of terms with their respective links to the documents and the number of occurrences.

OUTPUT

- The files will be stored under the directory /raw/intermediate
- For every term one .CSV file will be created
- The files will have the following items:
 1. name : term.json
 2. type
 3. title
 4. content
 5. text

```
C:\Users\Daniel\DanPycharmProjects\py_src>python 2_intermediate
INFO:root:Skipping file C:\Users\Daniel\tsgdataRS\intermediate2\raw\doctor.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\thicker.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\until.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\7277.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\make.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\cause.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\silk.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\must.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\question.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\152.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\please.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\sufficient.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\couldnt.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\scott.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\5152012big.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\excess.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\work.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\keys.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\activities.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\bone.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\three.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\comfortable.csv
INFO:root:Creating new index file C:\Users\Daniel\tsgdataRS\intermediate2\didnt.csv
```

Dentistry.csv

A	B	C	D	E	F
uuid	type	title	content	heading	text
parsed\doctor_Florida.Kissimmee.Plastic-Surgeon_Donovan-Rosas	0	0	1	0	0
parsed\doctor_Georgia.Atlanta.Plastic-Surgeon_Kristin-Ann-Boehm	0	0	2	0	0
parsed\doctor_New-Jersey.Shrewsbury.Oral-&-Maxillofacial-Surgeon_Manolis-Manolakakis	0	0	1	0	0
parsed\doctor_North-Carolina.Charlotte.Dentist_Mitesh-Dhulab	0	0	1	0	0

The Indexer

The indexer creates a term dictionary and index based on all the links of the documents, the term frequencies and the frequencies per documents, the dictionaries will be created in alphabetical order.

Dictionaries

The dictionaries are structured as follows:

```
term{j} doc{i,1}:weight_doc{i,1}, doc{i,2}:weight_doc{i,2}, ...  
term{j + 1} doc{i + 1, 1}:weight_doc{i + 1, 1}, doc{i + 1, 2}:weight_doc{i + 1, 2}, ...  
term{j + 2} doc{i + 2, 1}:weight_doc{i + 2, 1}, doc{i + 2, 2}:weight_doc{i + 2, 2}, ...
```

Page Rank

The page rank will take in first place the links and will obtain the doc file string and the type of this document: doctor or question, once there the list will be unified and converts all the keys to uuid, the link dictionary will be created and sorted alphabetically.

The output is finally the dictionaries with the docs IDs (keys) and the scores values

Method Score Calculation

Cosine Calculation

- The query should be splitted by terms and for every term should perform:
 - a. Calculate in the query the weight of the terms.
 - b. Calculate the weight of the term in the document.
 - c. Add the product of 2 and 3 to score **[doc_id]**
 - d. Add the square of product of 2 and 3 to length **[doc_id]**
- Each score should be normalized using the following formula:

$$Score[doc_id] = \frac{score[doc_id]}{length[doc_id]}$$

The Quality Score

The quality score it is included already when the indexer is performed in the tf-idf phase.

Query parser

The Query parser takes the query input and pass to the ranker, check as well if the returned documents is lower than certainly number, in given case it will try the possible combinations of the query to fill the K docs.

In given case the query is:

” Breast reduction surgery” if < K docs, then try “breast reduction” and “reduction surgery” first, later fill K docs if don not work then fill “Breast reduction surgery” and Fill K docs.

Based on the given query, the algorithm will be executed and this will return in the indexes and get relevant information.

- tf-idf
- Ordering importance (according to the different indexes which has different weights)
 - Match title first
 - Match body later
- Semantic scoring based on query
 - Relation between the requested terms on the initial query and some specific term on the index.

Accepted Queries

Apart from the simple and pure keyword search a range of filters can be used.

The keywords can be aimed at:

- either only doctors or only question pages ('filter: {doctor, question}')
- zones on the pages, like the title (question or name of a doctor), the heading (answer titles or review titles) or the text (answers or reviews) ('filter: {title, heading, text}').
- the specialization of doctors ('specialization:< a specialization >').
- the quality / rating of the targeted source. Doctors with a review rating of more or less than x or answers to question with a certain number of upvotes ('filter: rating{<,>}<number>').

After the query got parsed those arguments are forwarded to the ranking part.

There, either an appropriate separate dictionary gets used (title, heading, text) or the results of the standard dictionary get further filtered (doctor, question, specialization, rating).

Future Work

- Terms can be further processed to stem forms to match a wider range of relevantly similar words.
- Allow combinations of query filters.

Graphical User Interface (GUI)

The GUI will consist of a textbox where the initial input will be written and one search button

Web Search Engine

realselfSEARCH

video treatments specialization:plastic-surgeon

Search

Found 520 matching documents

DOCTOR: [k roxanne grawe md](#)

1.0

...video find a doctor ask a doctor write a review for doctors sign in for doctors join sign in join search ... tx post medical school training residency plastic surgery the ohio state university medical center **treatments** offered at roxy plastic surgery acne scars treatment acne treatment african american rhinoplasty ag...**treatments video find** a doctor ask a doctor write a review for doctors sign in for doctors join sign in join search docto...ite a review for doctors sign in for doctors join sign in join search doctors ohio columbus plastic **surgeon** k roxanne grawe md 49 out of 5 stars 49 stars from 145 reviews k roxanne grawe md 49 stars from 145...

DOCTOR: [mark d epstein md](#)

0.998193837608522

...video find a doctor ask a doctor write a review for doctors sign in for doctors join sign in join search ...appt and get their free 20 units contact our office for details show less show more 3500 thermiva 3 **treatments** for 3500 no expiration date 32 used redeem now redeem now details thermiva full treatment one treat...**treatments video find** a doctor ask a doctor write a review for doctors sign in for doctors join sign in join search docto...eview for doctors sign in for doctors join sign in join search doctors new york stony brook plastic **surgeon** mark d epstein md 49 out of 5 stars 49 stars from 238 reviews mark d epstein md 49 stars from 238 r...

DOCTOR: [michele a shermak md](#)

0.9856427560158065

Technology used at the RealSelf search system.

The follow components were using during the construction of the search system

Python 3.6.1

Plugins:

- Requests: (connect to realself.com and their branches)
- Lxml: (crawling and data scrapping)
- Pandas: (data structure and data analysis)
- Validators: (different validations thought the search system)
- Scipy & Numpy: (used for mathematics, numeric calculations / PageRank)
- Flask: (Graphical User Interface)

Docker:

Container platform for run the program on any machine, regardless the OS and other factors, eliminate the need to install multiple plugins across all the machines the program may run of.