

MASTER

Hippocrates

a context-aware, collaboration enabling search tool

Aravanis, G.

Award date:
2014

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain



Eindhoven University of Technology
Business Information Systems
Department of Mathematics and Computer Science

Philips Research

Master Thesis

Hippocrates: A Context-aware, Collaboration Enabling Search Tool

Author:
Georgios Aravanis (0804645)

Supervisors:
TU/e:
Mykola Pechenizkiy

Philips Research:
Anca Bucur

August, 29th 2014

Abstract

Hippocrates is a search tool that contains 1.6 million of biomedical references from Pubmed, related to the Oncology domain. It makes use of the data model incorporated into the European project EURECA and is oriented in extending the current framework with biomedical searching functionality. EURECA framework is used by oncologists for multiple use cases.

Having as input of Hippocrates the workflow state of the oncologist's workflow, but also the Electronic Health Record (EHR) of the patient, the objective of the current project was defined; to improve the searching process of the oncologist by making use of the contextual information of the case. During this project, the concept of collaboration among doctors became more and more obvious that was needed and necessary to be supported in the searching process.

Different contextualization and collaboration techniques were discovered, analyzed and finally implemented as Hippocrates components. Three different contextualization approaches were implemented; the pre-filtering in which concepts included in the EHR of the patients are used by the oncologist in order to filter the documents and get relevant only results, the query expansion that expands the user's query with the patient's observations including that of the diagnosis and the re-ranking that re-ranks the top results based on a Boolean score of concepts of the EHR included in the references.

In the collaboration perspective, an algorithmic intervention based on the I-SPY equations for scoring documents rated by similar users was implemented among multiple User Interface (UI) based intervention components such as group history ranked by rating.

The experiment/evaluation of Hippocrates besides the effort and resources devoted towards that direction, was limited to three participants, one of which is an Oncologist. The results though, even not statistically confident, indicate that the pre-filtering contextualization improves the searching process, and the rest two were observed to improve it in at least one case. Limitations were discovered during the internal evaluation of the collaborative algorithm and the I-Spy equations but also prove that could at least improve the Inverse Cumulative Gain of the result set.

The Usability of the system was highly rated and several components averagely rated were accompanied with valuable feedback that could be used for future development.

Overall, Hippocrates received positive evaluation and could be further developed and tested to prove the improvement on the searching process of the oncologists.

Preface

This master thesis is the result of my studies in the MSc “Business Information Systems” at Eindhoven University of Technology. Moreover, it is also a conclusion of my internship at the Healthcare Information Management department at Philips Research.

I would like to acknowledge my supervisors who have helped me during my graduation project. I would like to thank Dr. Mykola Pechenizkiy and my external supervisor Dr. Anca Bucur for their guidance and continuous support on difficult decisions in the thrilling scientific field of web engineering. I would also like to thank all the people working on the EURECA project who were always willing to help me especially Ahmed Ibrahim and Jasper van Leeuwen. I would like to acknowledge Dr. Norbert Graf who contributed to this project with his valuable remarks and evaluation and Kosmas Hatzidimitris as well.

Finally, I want to thank my family and friends for their support during the course of my studies. My studies would not have been completed without the financial and moral support by the G. & P. Kougiouli foundation.

Eindhoven, the Netherlands
June 2014
Georgios Aravanis

Contents

Abstract.....	2
Preface	3
Contents.....	4
Table of figures	6
1 Introduction	7
1.1 Motivation.....	7
1.2 Objective	8
1.3 Methodology.....	8
1.4 Thesis structure.....	9
2 Background and Related Work	10
2.1 Contextual information.....	10
2.1.1 Workflow state as contextual information	10
2.1.2 The Electronic Health Record (EHR) and the Common Data Model.....	13
2.2 Pubmed literature references.....	15
2.3 Web Search model.....	16
2.4 The BM25 scoring function	18
2.5 Collaborative Information Search	19
2.6 Context aware systems	23
2.7 Re-finding and revisiting the berry-picking model.....	26
3 Hippocrates, algorithms and implementation description.....	28
3.1 System description.....	28
3.2 Data import, contextual expansion and contextual query formulator.....	30
3.3 Solr configuration.....	32
3.3.1 Pubmed References Preprocessing and Indexing.....	33
3.3.2 Matching and Scoring	34
3.3.3 The More Like This component	34
3.3.4 Autocomplete	35
3.3.5 Text Highlighting	35
3.3.6 Faceted Search.....	36
3.4 Contextual Search	37

3.4.1	Query Expansion	37
3.4.2	Contextual Pre filtering –Guided Query formulation	38
3.4.3	Boolean Based Score for Contextual Re-ranking	38
3.5	Collaborative Search	38
3.5.1	Capturing information.....	38
3.5.2	Collaborative filtering re-ranking	39
3.5.3	UI enabled collaborative search modules.....	41
3.6	Other UI modules.....	43
3.6.1	Visual Notification of updated results	43
3.6.2	Re finder	43
4	Evaluation	45
4.1	First evaluation by medical expert.....	45
4.2	Experimental Design/Usability Evaluation.....	46
4.3	Results	48
5	Conclusions and Future Work.....	54
5.1	Conclusions	54
5.2	Future work.....	55
6	Works Cited.....	57
	Appendices.....	60

Table of figures

Figure 1 Iterative development	8
Figure 2 Patient's Treatment Workflow State-Diagram	11
Figure 3 HL7 RIM based Common Data Model.....	13
Figure 4 Classic Information Retrieval Model (Robertson, 1977).....	16
Figure 5 A search engine model (Smyth, Coyle, & Briggs, 2011).....	17
Figure 6 The berry picking model (Bates, 1989)	17
Figure 7 SMART notation for tf-idf variants (Manning, Raghavan, & Schütze, 2008)	18
Figure 8 Contextual depiction of Collaborative Information Seeking (CIS) (Shah, 2012)	20
Figure 9 Depiction of CIS and related topics using the dimensions of human–system and explicit-implicit collaboration (Shah, 2012).....	20
Figure 10 Architecture overview of I-SPY (Smyth, Coyle, & Briggs, 2011).....	22
Figure 11 CDAPubMed Architecture (Perez-Rey, Jimenez-Castellanos, Garcia-Remesa, Crespo, & Maojo, 2012)	24
Figure 12 Patient-based literature retrieval Architecture (Jimenez-Castellanos, Fernandez, Perez-Rey, & Viejo, 2013).....	25
Figure 13 Berry-Picking model revisited	27
Figure 14 High level system architecture of Hippocrates	29
Figure 15 Use of Diagnosis as Initial Query with alternative labels retrieved from SNOMED CT.....	31
Figure 16 Contextual expansion of the Aclarubicin Antineoplastic Agent	31
Figure 17 From the EHR to the Contextual query formulator	32
Figure 18 The MLT component	35
Figure 19 The Autocomplete feature.....	35
Figure 20 Text highlighting for the query "cancer".....	36
Figure 21 Faceted Search menus	36
Figure 22 Approaches for contextualizes search	37
Figure 23 Proposed approach in including SD in the recommendation models (Kwon, 2008).....	40
Figure 24 Case based collaboration module.....	41
Figure 25 Recommender based on the collaborative re-ranking algorithm.....	42
Figure 26 Context aware-collaborative Query Recommendations	42
Figure 27 Visual Notification of updated results module	43
Figure 28 Refinder.....	44
Figure 29 Cumulative Gain Oncologist.....	48
Figure 30 Precision @ 9 Oncologist	49
Figure 31 Cumulative Gain Informaticians.....	49
Figure 32 Discounted Cumulative Gain Informaticians	50
Figure 33 Precision@18 Informaticians	50
Figure 34 Overall usability evaluation of Hippocrates.....	52
Figure 35 Modules evaluation	52

1 Introduction

In this thesis the problem of literature search is studied. Different techniques and approaches that facilitate and improve the searching tasks were identified and evaluated and some of them implemented into the search tool Hippocrates. This search tool includes a Collaboration Searching algorithm, User Interface mediation for collaboration modules and methods of contextualized search, based on the Electronic Health Record of a patient.

1.1 Motivation

Knowledge workers devote a significant amount of time in searching for relevant to their profession information. This applies to the doctors as well, by searching for biomedical literature for recent developments in their expertise or relevant to their cases information.

Pubmed is one of the prime search engines that the doctors use. But until very recently supported only Boolean queries on a vocabulary used for document representation sorted by the date of publication, limiting the users to a complex and inefficient way of searching. This is indicated by the appearance of multiple guides on how to search PubMed. PubMed documents are accessible and free for use after an agreement with terms and conditions. This has led to multiple research projects containing Pubmed references in the index.

There are numerous attempts in optimizing the searching-retrieving-sorting mechanisms by providing alternative web tools based on the Pubmed documents. There is far less effort though in providing the right information to the doctor on the time needed.

The Scenario that the development of the tool is based on is that of the workflow of the oncologist, during which many sub-processes occur. The selections of patients for a trial or the support of an oncologist by a clinical decision support system are such examples, supported by the EURECA Framework (an EU project) in which Philips Research is contributing. In that context, an oncologist may review the patient profile and in any case an information need might arise.

The motivation for developing Hippocrates was to meet those information needs to a solution, integration ready, that will contextualize the search, making use of the context; the workflow state and the patient's Electronic Health Record without limiting to the use of the document representation of the document by using MeSH terms that are described in the relevant work.

Motivation for the collaboration part of Hippocrates was the suspected need for communicating interested references inside a medical team. This was explicitly verified by a clinician and further development was followed.

1.2 Objective

The objective of the project is to provide an open-source, extendable tool that will measurably improve the searching process of medical references of the oncologists. The project was oriented into two different approaches towards that direction; the contextualized search and the collaborative search.

In the contextualized search there is the objective to analyze, develop, implement and evaluate different approaches in including contextual information provided by the Electronic Health Record of the patient and the workflow state of the process in order to provide to the Oncologist means to reach to more relevant results faster.

In the collaborative search there is the objective to support the collaboration among the team of oncologists on the searching process by either algorithmically or through UI components.

1.3 Methodology

The methodology followed in this project was an iterative one. As shown in Figure 1, the same cycle is iterated over and over until the product, in this case Hippocrates is finalized.

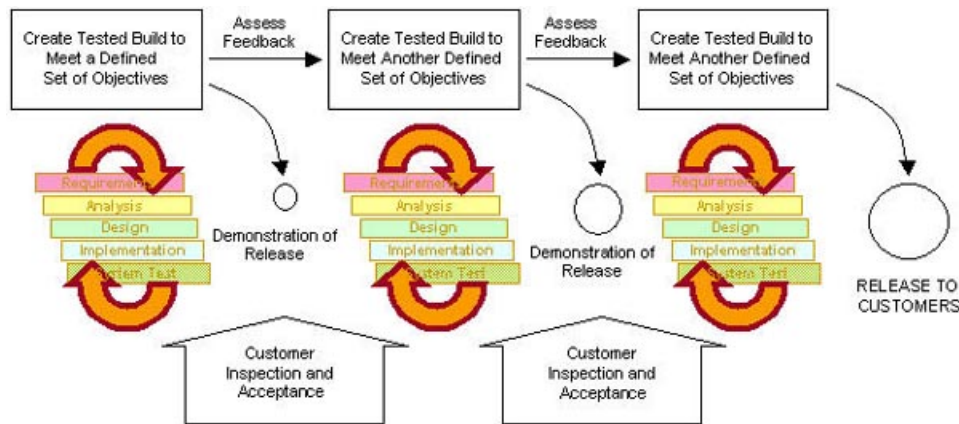


Figure 1 Iterative development¹

The project started with some basic requirements. The first version of Hippocrates included the configuration of the Solr Search Engine, the inclusion of contextual information as pre-filtering and some concepts of the collaborative UI components.

After the assessment of the valuable first feedback from an oncologist described in 4.1 there was a second cycle of Requirements, Analysis, Design, Implementation, and Testing. After that phase an internal assessment took place and after a second assessment the contextual algorithms were implemented in the second phase.

¹ <http://www.ibm.com/developerworks/rational/library/may05/bittner-spence/>

This thesis presents the results of the second external evaluation which corresponds to the end of the third cycle of the iterative development.

1.4 Thesis structure

This thesis is separated into five chapters. The Chapter 1 introduces the reader to the problem, the objective and the methodology followed. In Chapter 2, there is an overview of the input used by the system. Different information retrieval models are presented among with how collaboration and contextualization affect the searching process. Such systems that realize the above approaches are also included in the second chapter of this thesis. In Chapter 3, there is an effort to present the complete picture of the implementation of the developed tool and to link it with the approaches presented in Chapter 2 or to motivate the choices for less critical features implemented. Chapter 4 includes the first evaluation of the system by an oncologist, the experimental design and the results of the second evaluation cycle. The Chapter 5 concludes the thesis with a discussion of the evaluation results and with proposals for future work.

2 Background and Related Work

In this chapter there is an introduction to the Contextual Information used as input in the system and for each type, background information is presented. The second part is an overview of the Web Search models along with an efficient scoring algorithm of the probabilistic search branch. The third part analyses the related, to the core directions of this thesis, work regarding the effect of collaboration and contextualization in searching. Collaborative systems and context aware systems are analyzed in the present chapter.

2.1 Contextual information

As presented later in section 2.6 there are a lot of tools that make use of contextual information in order to improve the results provided to the users. Usually this contextual information includes but is not limited to the previous search logs, the accessed pages or the physical location of the user, the time or the day of the week. From the EURECA framework, there are two types of contextual information available that are used into Hippocrates, namely, the Workflow State of the treatment process and the Electronic Health Record of the patient.

2.1.1 Workflow state as contextual information

The workflow state is one of the inputs of Hippocrates. A detailed description about the use of this contextual information is presented in Chapter 3. The workflow state includes the state in which the workflow is at the moment of a literature search. The states of the workflow are:

- Admission
- Diagnosis
- Planning
- Treatment
- Follow-up monitoring and treatment
- Adverse event

As it can be derived from the Figure 2 the transition from one state to the next, signifies a progress in time. This is used as demonstrated in Chapter 3 in order to provide to the users relevant results at the time the information is needed.

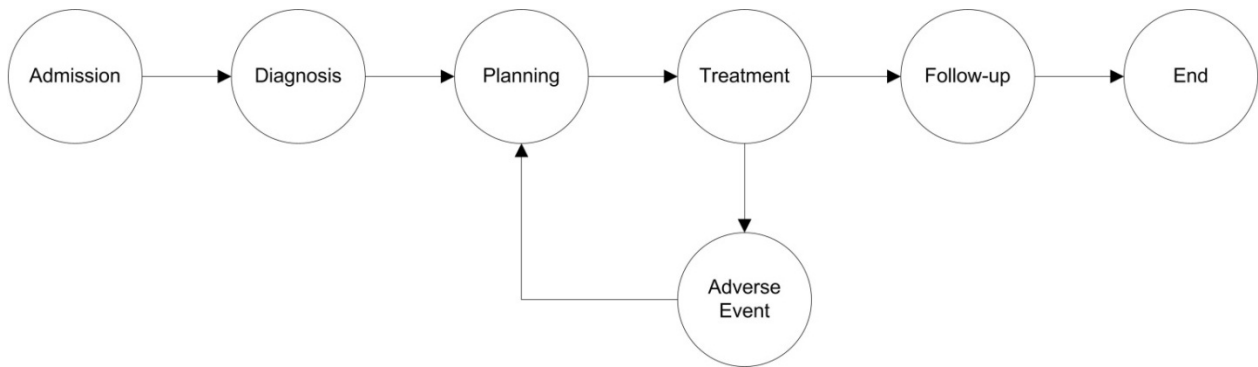


Figure 2 Patient's Treatment Workflow State-Diagram

Workflow description

In order to present the treatment process in this section, a short textual description of the process is presented.

Admission

The patient is admitted to the oncology clinic after a referral from a GP or from another department of the hospital and a secretary registers her. Then an appointment for consultation by an oncologist is scheduled.

Diagnosis

In this step a nurse or the oncologist records the medical history of the patient if that is not already available. The diagnosis of the patient is verified by the oncologist. The diagnosis is usually determined by external to the oncological clinic healthcare unit, and in most of the cases all the lab tests that leads to the diagnosis have been done to that external unit². The oncologist reviews those results and checks whether or not the diagnosis are based on valid evidences.

Planning

In this step the oncologist develops a treatment planning for her patient. In this step, a consultation session with the patient takes place.

Treatment

In this step the patient follows the treatment and there are some periodic meetings with the oncologist for monitoring and adjustments of the treatment if it concerns a radiation or drugs administration.

Adverse event

During the treatment and because of the exposure in radiation or medication or any other treatment solution chosen, there might appear adverse events.

² This depends on the organizational structure of a hospital, on the health care system of each country and on each case. This description is a high level overview of the process described by an oncologist specialist

Follow-up care

In the follow up care, there is a periodic consultation and monitoring of the patient in order to early detect the return of a cancer to the patient, appearance to another organ. Multiple types of tests might be applied in order to detect any changes in the condition of the patient.

This short description proves there are different types of information needs in each workflow state. This assumption is verified by a specialist who provided some of the information needs that need to be covered in each state.

Information needs per state

By the input received by a specialist (section 4.1), two of the workflow states appear to be of higher importance; the Planning of Treatment state and the adverse event state. In some cases there are information needs regarding the refinement of the diagnosis. These information needs supports the need of inclusion of this contextual information into Hippocrates. The explicit feedback received on the information needs for each state is as follows:

Diagnosis

In this workflow state the oncologist might be interested in the availability or recent developments in genetic tests and biomarkers available that corresponds to the patient's diagnosis. This is used in order to further refine the diagnosis and to improve in the next step the treatment planning.

Planning

In this state the oncologist is interested into the availability of a trial in which a patient is a match³. The treatment options and their short and long term impact on the patient are of great importance. Radiotherapy necessity, choice of drastic substances, and the corresponding dosages or surgery options are what the oncologists need to be informed about on recent developments. These treatment options need to be evaluated in terms of outcome in the treatment of the patient such as the survivability rates after administering the treatment but also in terms of quality of life. Another aspect which is of interest regarding the treatment option is the prevention of potential adverse events, in other words, which actions need to be taken to prevent them.

Adverse event

In this state the oncologist is interested into options to deal with the adverse events that have appeared to a patient. Potential information needs in this state are the alternative treatments options, dosage of radiation or dosage of the drastic substance administration adjustment.

³ This information need is planned to be covered by another module of the EURECA framework, Yakobo.

2.1.2 The Electronic Health Record (EHR) and the Common Data Model

Another input of the system is the EHR of the patient. The structure of the EHR follows a reduced data model of the HL7 Reference Information Model (RIM). The HL7 is a widely used standard for healthcare interoperability and the RIM was designed in order to be a more generic data model that will cover the needs of the domain. The RIM was used in order to simplify the standard into the newest version 3.

The EURECA project makes use of an adapted common model of another EU project, Integrate which enables interoperability of the different modules of the project and an accurate representation of clinical data of breast cancer clinical trials. This adapted common model is presented in Figure 3.

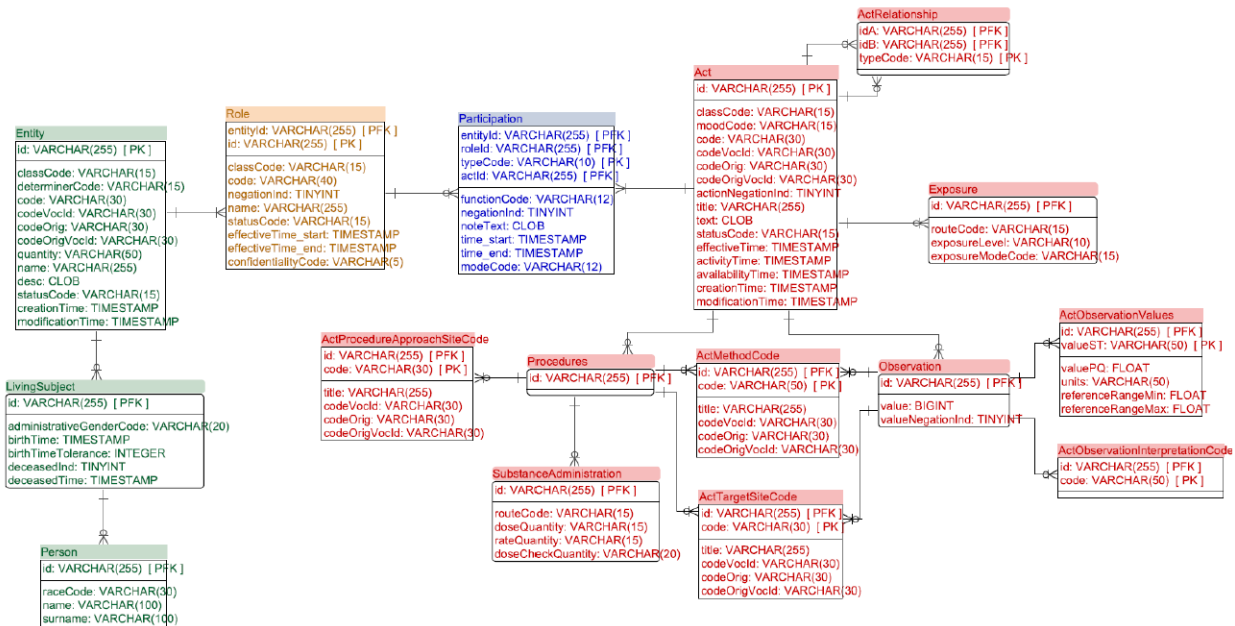


Figure 3 HL7 RIM based Common Data Model⁴

The EHR records imported into Hippocrates (as a MySQL dump) consists of 80 fictitious cases generated for the development of the different modules of the EURECA project.

There are two types of information and three main subtypes namely the personal information and the main case information including Observations, Substance administration and Procedures.

The *Living Subject* table includes:

- The id of the patient [fictitious1-fictitious80]
- The administrativeGenderCode, which is the SNOMED CT code of the ontology depending on the sex of the patient (in our context) which is 248152002 for the female and 248153007 for the male
- The birthTime, which is the timestamp of the time of birth (of the patient in our context)

⁴ INTEGRATE WP 3 D 3.5, version 1.0

The *Entity* and the *Role* tables in our dataset do not provide any additional information other than each subject is a human person and a patient respectively.

The *Participation* table links every patient to an *actID*.

An *act* is of one of the following three classes; Observation, Procedure and Substance Administration. The ontological code of each act and the corresponding title from the corresponding ontology are included in the table. The *effectiveTime* is one a column of the table of high interest, since it contains information about the time that an act took place.

An act is related to a site of the patient e.g. “breast” of the patient, whether it concerns an observation or a procedure. This information is registered on the table *ActTargetSiteCode* and each act might have multiple target sites e.g. breast, left, with laterality. Besides the title of the site, the ontology code is included as well.

For the *Observation* class acts, there is a table that adds extra information about the observation, the *ActObservationValues*. This table includes a value of an observation e.g. 0.22, the units e.g. mg/dL, and the reference range with the corresponding minimum and maximum values that the observation should be within in normal cases.

A typical example of an EHR used as an input in the Hippocrates is the following of **fictitious 1**:

-livingSubject:

administrativeGenderCode: 248152002
birthTime: 1965-03-01 00:00:00

-Act:

class: OBS,
title: Infiltrating duct carcinoma
effectiveTime: 2006-02-02 00:00:00

class: OBS,
title: Tumor size
effectiveTime: 2006-02-02 00:00:00

class: PROC,
title: Surgical procedure
effectiveTime: 2006-02-02 00:00:00

class: SBADM,
title: Aclarubicin
effectiveTime: 0000-00-00 00:00:00

class: SBADM,
title: Aclarubicin
effectiveTime: 0000-00-00 00:00:00

-ActTargetSiteCode (for the Infiltrating duct carcinoma act)

code: 76752008
title: Breast structure (body structure)

code: 7771000
title: left

code: 78615007
title: with laterality

-ActObservationValue (for the tumor size)

valueST: 52
units: mm

Some of the tables were not used for the generation of the fictitious cases and thus there is a potential to expand the contextual information with valuable data such as the interpretation of an observation (normal-abnormal). There is no notation of the Act that is the current diagnosis of the patient, which is the reason that an external service is used as described in paragraph 3.2. Currently in the EURECA project there is a development of a Natural Language Processing module that will transform an EHR to the corresponding format of the CDM. For the scope of Hippocrates only the tables and the columns used to produce the fictitious data were used and further processed as described in 3.2.

2.2 Pubmed literature references

The index used in the Hippocrates system is a subset of the Medline biomedical literature. Currently, Pubmed contains more than 23 million references from Medline, life sciences journals and online books. The dataset of the references is available to anyone accepting the corresponding Terms & Conditions and filling a memorandum of understanding for the use of the data. Pubmed is operated by the National Center for Biotechnology Information of the United States Department of Health and Human Sciences.

For the Hippocrates system, 1.626.689 biomedical literature references were imported from the document set of Medline. The filtering of the articles was done by selecting as a MeSH term the word *neoplasms* which was the most generic term including all forms of tumors and cancers that would be of interest by the oncologists. The MeSH term is a vocabulary developed by Pubmed and functions as a document representation for searching the references. Until the October 2013, that was the only way of searching Pubmed, through a match between a query and a MeSH term. Though, through an internal evaluation the MeSH terms assigned to documents did not include many of the concepts appearing on

the document and the manual process used by Pubmed to assign terms of the vocabulary to each document was considered a drawback and thus was not further used by Hippocrates.

The references of Pubmed were received in XML format⁵. Selected fields were used and indexed in the search engine as described in section 3.3.1.

2.3 Web Search model

The purpose of this section is to present the web search models on which, the differentiated approaches analyzed in the next sections, are based.

The classic Information Retrieval model in Figure 4 was described by Robertson (Robertson, 1977). From the right side of the model, the user has an information need. That need is represented by a query, a set of words which is most commonly referred as a bug of words. This bug of words is entered to an IR system. On the other side of the system there are documents from which in many cases an index of terms or specific vocabularies are created. The Information Retrieval is the scientific area that used that model in order to optimize the relevancy of the matched results.

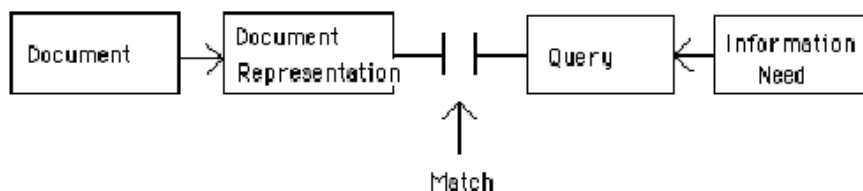


Figure 4 Classic Information Retrieval Model (Robertson, 1977)

The classic Information Retrieval model could be used in order to represent the function of a search engine as seen in the Figure 5. The offline part or crawler accesses some seed pages (initial set of pages) which are parsed and indexed. The seed pages contain links to other pages which the crawler accesses next. The online part of the Search engine model is a different representation of the Classic Information Retrieval Model previously presented.

⁵The XML files conform to the following DTDs:
http://www.nlm.nih.gov/databases/dtd/nlmmedlinecitationset_130101.dtd
http://www.ncbi.nlm.nih.gov/corehtml/query/DTD/pubmed_130501.dtd

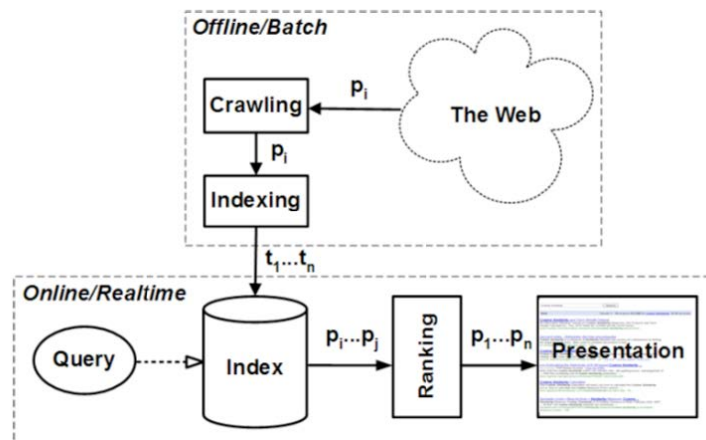


Figure 5 A search engine model (Smyth, Coyle, & Briggs, 2011)

The previous two models include only a limited view of the searching process that takes place in reality. The well cited article "THE DESIGN OF BROWSING AND BERRYPICKING TECHNIQUES FOR THE ONLINE SEARCH INTERFACE" (Bates, 1989) proposes the berry picking model.

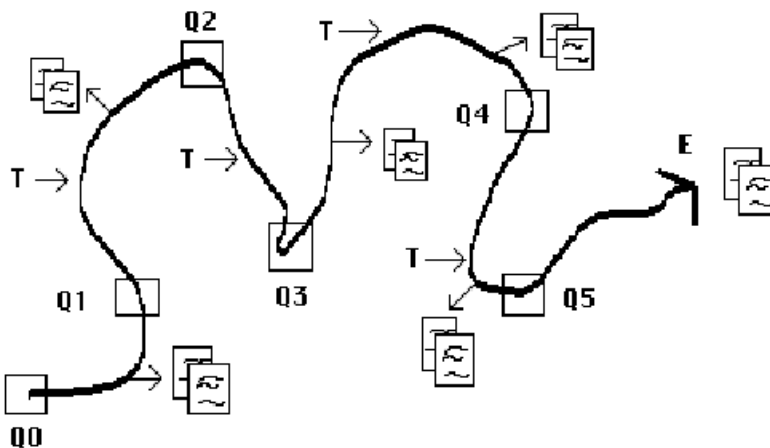


Figure 6 The berry picking model (Bates, 1989)

There are quite some differences between the classic model and the berry picking model. The first assumes that a search starts and is finalized in one action. The second proposes a repetition of the task but not in an identical manner. The information need might become more specific or broader or it could be altered to a large extent. In each step the user might find useful information and references that after the evaluation by the user, might affect and evolve the searching process. In Figure 6, Q represents the query variations from the initial Q_0 , T is the thinking of the user, and E is the exit of the searching process. It is necessary to mention that this is a path, limited to the universe of interest of the user or as will be presented in Section 2.6 the context of the search which is in short terms the universe of interest of the user or in long terms is a subarea of it.

2.4 The BM25 scoring function

The BM25 retrieval function has been considered the state of the art probabilistic scoring function for two decades. In this section, there is an introduction on the basis of the function which is the Tf-idf weighing function, and examples of usage of the system.

A first scoring approach is the appearance of a term in a document (Boolean score) or in different parts of a document (zones). The score is 1 on the appearance of the term in the text or 0 otherwise. A next logical step is to consider the term frequency (tf) into a scoring function. The motivation for the term frequency is that the more frequently a term appears in a document the more relevant that document is to a query containing that particular word. There are multiple scoring functions that could be used based on the term frequency concept (Manning, Raghavan, & Schütze, 2008) shown in Figure 7. Besides the natural and the Boolean already mentioned the logarithmic, the augmented and the logarithmic average have been proposed as scoring functions.

Term frequency		Document frequency		Normalization	
n (natural)	$tf_{t,d}$	n (no)	1	n (none)	1
l (logarithm)	$1 + \log(tf_{t,d})$	t (idf)	$\log \frac{N}{df_t}$	c (cosine)	$\frac{1}{\sqrt{w_1^2 + w_2^2 + \dots + w_M^2}}$
a (augmented)	$0.5 + \frac{0.5 \times tf_{t,d}}{\max_t(tf_{t,d})}$	p (prob idf)	$\max\{0, \log \frac{N-df_t}{df_t}\}$	u (pivoted unique)	$1/u$
b (boolean)	$\begin{cases} 1 & \text{if } tf_{t,d} > 0 \\ 0 & \text{otherwise} \end{cases}$			b (byte size)	$1/CharLength^\alpha, \alpha < 1$
L (log ave)	$\frac{1 + \log(tf_{t,d})}{1 + \log(\text{ave}_{t \in d}(tf_{t,d}))}$				

Figure 7 SMART notation for tf-idf variants (Manning, Raghavan, & Schütze, 2008)

The term frequency scoring algorithms have a problem regarding the relevance of different terms of a query to the retrieved documents. All terms are affecting the relevance by the same proportion. An example is the stop words or very common words. In case of Hippocrates, the word cancer is included in the scoring function and contributes to the score in the same way as the word breast which might be rarer to the index. The idf *inverse document frequency* concept introduced by Jones as term specificity (Jones, 1972), can be used in order to weight each term depending on its scarcity in the index. The rarer a term appears in the document set, the higher the value of the idf gets. The idf of a term is defined as the logarithm of the number of the documents in the index, to the number of documents containing that term.

Having the short introduction of some basic scoring and weighting functions, next we introduce the BM25 scoring function which is based on the previous concepts.

BM25 scoring function

$$BM25 = \ln\left(1 + \frac{Numdocs - docfreq + 0,5}{docfreq + 0,5}\right) \times \frac{termfrequency(k_1 + 1)}{k_1\left(1 - b + b\frac{doclength}{avgdoclength}\right)}$$

There are two parts in the above formula for scoring a document on a given term of a query. The first is based on the idf concept introduced. The same concept was developed based on the probabilistic theory. The probability estimates without relevance feedback gives the first logarithmic part of the scoring function. An improvement of the probabilistic idf was proposed by Jones (Walker, Jones, & Robertson, 2000), by factoring to a weight that includes, the term frequency and a simple normalization factor document length to the average document length. The normalization is necessary to score documents with different lengths. In our case a small title and abstract over a citation with 3 paragraphs of abstract. After experiments with the TREC tests, the constant b was found to perform well around the value 0.75. The constant b controls how much the length normalization will affect the scoring. The constant k_1 , if set to 0, affects the tf scoring and transforms it to a Boolean score. If the value assigned to k_1 is large then the weight in total will behave linearly as tf . From the TREC tests the constant k_1 was found to perform well with values between 1.2 and 2.

The BM25 scoring function is widely used for many different cases and is considered one of the state of the art functions from the probabilistic Information Retrieval. The performance of the function in the various TREC tests contributed in the publicity received. In the system presented in Chapter 3 the scoring function selected is an expanded version of the BM25 presented in this section.

2.5 Collaborative Information Search

In this section, there is an effort to present the classification of Collaborative Information Seeking (CIS) systems and to present some of the main approaches and systems developed. Microsoft has systematically studied the way that employees collaborate through the information seeking process (Morri, 2013), (Morris, 2008). These studies took place in a different context than the medical domain and the oncologists. The email was the most common communication tool for communicating interesting findings by 46,8% followed by SMS 30,3% and talking over the phone in 27,5% of the cases. Similar results are seen in a study in Collaborative Information Seeking (Böhm, Klas, & Hemmje, 2014). The use of email for collaboration by oncologists was confirmed during the first evaluation for Hippocrates.

A first conclusion is that collaboration in information seeking is taking place among knowledge workers. A question that needs to be answered is how the collaboration information seeking relates to the information seeking model that was presented in the previous sections. In the book Collaborative

Information Seeking (Shah, 2012), such an explanation is included in order to give the contextual depiction of the subject.

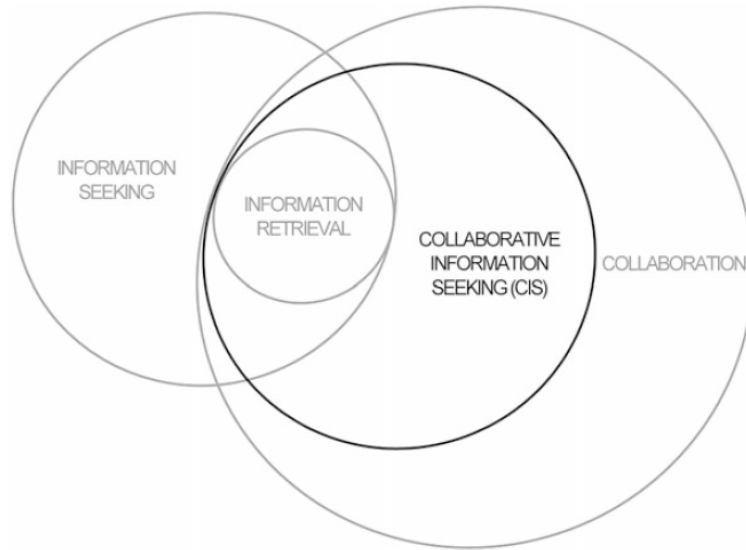


Figure 8 Contextual depiction of Collaborative Information Seeking (CIS) (Shah, 2012)

In Figure 8 information seeking is defined as the set of the atomic tasks towards finding information such as searching and retrieving (Information Retrieval) but also assessing the retrieved documents synthesizing and further processing them. On the other hand, collaboration includes the contribution of effort from the individuals in a team towards a common goal and to raise the cognition level of the group through a dialogue. So Collaborative Information Seeking includes the atomic tasks of information seeking and retrieval in a collaborative context.

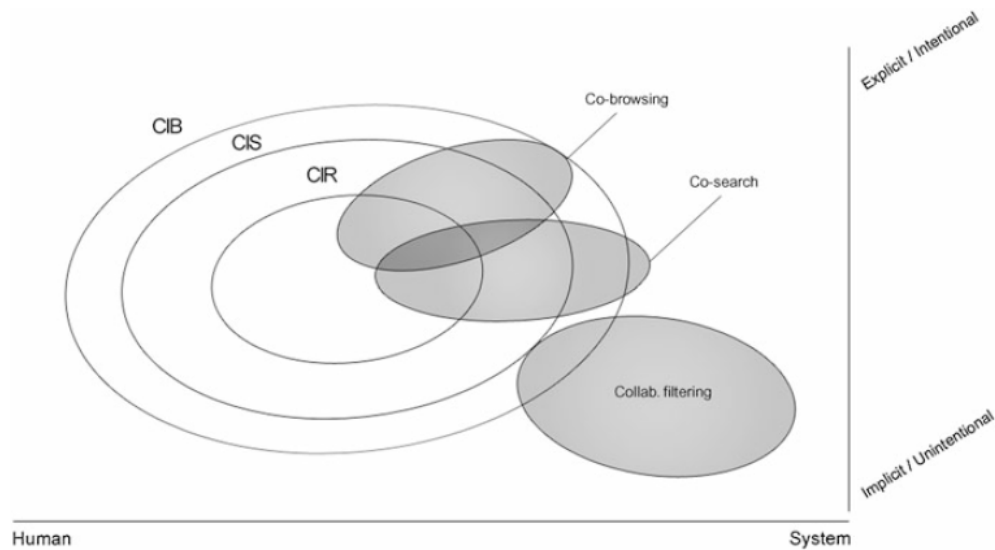


Figure 9 Depiction of CIS and related topics using the dimensions of human–system and explicit-implicit collaboration (Shah, 2012)

In the same book (Shah, 2012) the different keywords found in studies and publication for CIS related systems, were classified in the dimensions seen in Figure 9. The horizontal axe represents the human participation in the system where Collaborative Filtering presents the lowest participation. The vertical axe represents how the human input is acquired by the system in terms of intention. The three different cycles represent the Collaborative Information Behavior (CIB) which is a superset including the Collaborative Information Seeking which includes Collaborative Information Retrieval systems.

Other classifications of such systems include the location of the team members, Collocated-Remote located, the data synchronization, Synchronously-Asynchronously (Pickens, Golovchinsky, Shah, Qvarfordt, & Back , 2008) and the type of system intervention User Interface-Algorithmically.

Collaboration through algorithmic intervention

The main focus of the present study is towards the Collaborative Information Retrieval and the systems using algorithmic intervention.

The I-SPY collaborative filtering algorithm

The core assumption on which the I-SPY algorithm was based on is that the documents that are accessed more often are those that are more relevant to a group of users issuing the same or a similar query (Smyth, Coyle, & Briggs, 2011) (SMYTH, BALFE, FREYNE, BRIGGS, COYLE, & BOYDELL, 2004) (Balfe & Smyth, 2007). The group of users is referred to users of a certain community that either belongs to a certain company or working team etc. with similar information needs. An example given to motivate the development of the I-Spy system is that 65% of the queries in a software company share more than 50% of the keywords with 5 other queries.

The basis of the implementation of the I-Spy system is the registration of a log file with the access of each web page on a given query in a Hit Matrix as seen in Figure 10 Architecture overview of I-SPY (Smyth, Coyle, & Briggs, 2011)Figure 10.

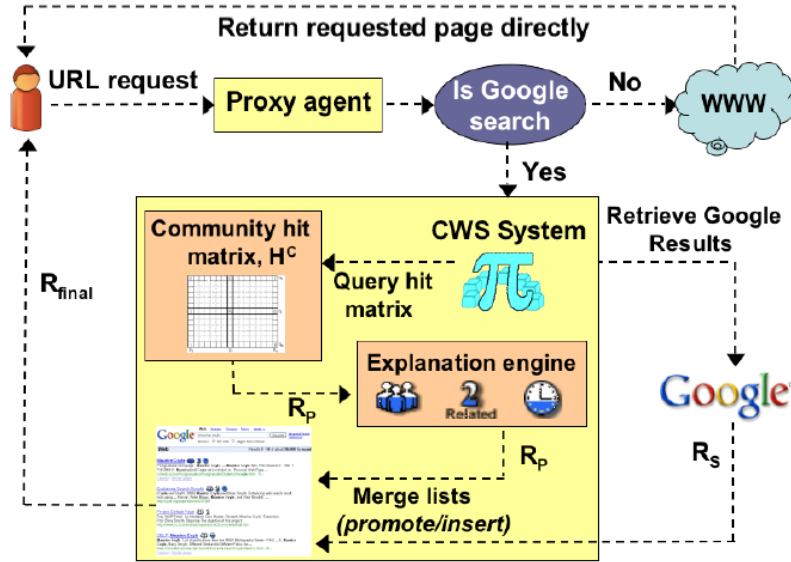


Figure 10 Architecture overview of I-SPY (Smyth, Coyle, & Briggs, 2011)

The Hitmatrix contains the queries as rows and the documents as columns. The value of the cells is the number of times that each document p_j was selected for a particular query q_i . Then a relevance criteria is introduced as the number of times a page is selected to the total pages accessed for a particular query as seen in equation 1. In general, a threshold of minimum similarity could be set in order to scope down the number of documents to calculate the score for a given query. Finally, in 3 the score of relevancy of a certain page could be calculated.

$$Relevance(p_j, q_i) = \frac{H_{ij}}{\sum_{\forall j} H_{ij}} \quad (1)$$

$$Sim(q, q') = \frac{|q \cap q'|}{|q \cup q'|} \quad (2)$$

$$WRel(p_j, q_T, q_1, \dots, q_n) = \frac{\sum_{i=1 \dots n} Relevance(p_j, q_i) \cdot Sim(q_T, q_i)}{\sum_{i=1 \dots n} Exists(p_j, q_i) \cdot Sim(q_T, q_i)} \quad (3)$$

The top 3 results in case of the i-spy implementation are used as promoted pages. A relative benefit of 30% is claimed by judging as successful, the searching sessions with at least one page click.

The Haystack system

The Haystack system (Smyth, Coyle, & Briggs, 2011) uses defined tasks as staks where people can join and rate pages as relevant or not relevant and tag results with keywords. The backend system of Haystack includes a recommender system that promotes relevant result on top of results from a search engine. The recommendation algorithm is not published. Though, the signals used to score the documents and to produce the recommendations include the query terms used and tag terms used to rate or share the document, the counts of positive and negative ratings and the number of times shared and tagged.

The same strategy is applied as in I-Spy regarding the promotions. When there are documents fulfilling a minimum threshold, are projected in the top 3 places of the result set.

Example of Synchronous Collaboration through algorithmic intervention system

Until this point an overview of the I-Spy and Haystack systems is given. Both systems share some characteristics, one of which is the asynchronous nature of the collaboration. In the CIS book (Shah, 2012) multiple systems are presented that promotes collaboration in a synchronous way such as Cerchiamo (Golovchinsky, Adcock, & Pickens, 2008) where the members of a team have specific roles, namely Prospector, Miner, domain expert etc. The prospector performs the queries, and the miner rates the results and based on the algorithmic part of the system, recommendations are presented in a common UI of all the users.

2.6 Context aware systems

In this section there is a presentation of the related work regarding the use of contextual information in the medical literature domain, but also to other approaches followed in context aware systems.

Use of the EHR in the literature search

As mentioned in a research study about context aware retrieval (Brown & Jones, 2001) there are two ways that context could contribute in the Retrieval process; in the IR systems to provide an improved mechanism of weighting terms and for the IF systems to develop improved user profiles. In this section several Context aware IR systems are presented.

One of the first approaches in the use of contextual information is the Medline button (Cimino, Johnson, Aguirre, Roderer, & Clayton, 1993). The approach starts with the selection of the terms that are relevant to the information need raised to a doctor. Next the system proposes a set of potential relevant questions which is based on the analysis of user questions posed to librarians. The doctor selects the question that describes the information need the most. The system matches the relevant terms to the MeSH terms. The ICD-9CM standard used in the coding of the EHR was not performing well in the direct matching to the MeSH term, nor the use of UMLS. For improving the procedure they extended the UMLS mappings. Then the system provides the references from the Medline. This pattern of manual selection of relevant terms, matching to the MeSH terms and retrieving the results is seen in other research efforts with some modifications.

In the Evidence Based Practice (EBM) the PICO framework is used in order the doctor to formulate focused questions and to search for relevant evidences. PICO is an acronym for Patient problem, Intervention, Comparison and Outcome (Richarson, Wilson, Nishikawa, & Hayward, 1995). This framework was utilized into two different methods and compared to a standard Pubmed interface

(Schardt, Adams, Owens, Keitz, & Fontelo, 2007). A precision score comparing the retrieval of golden standards articles to the total set of the retrieved articles was measured in the three different interfaces. The first one, besides the PICO entries, requested the age group of the patient, the gender and the Publication Type that the user was interested into. The second protocol filtered further the results requesting input from the user regarding the type of question needed to be answered. The result of this research was not statistically significant since the 95% of confidence interval of the precision scores overlapped. The precision scores though were higher for both the context aware searches. Furthermore, there is some consideration in the suitability of the framework in representing clinical information needs (Huang Xiaoli, Lin Jimmy, & Demner-Fushman , 2006).

Another tool that makes use of the contextual information is the CDAPubMed (Perez-Rey, Jimenez-Castellanos, Garcia-Remesa, Crespo, & Maojo, 2012). This browser extension extracts relevant sections from HL7 CDA formatted EHRs, using a natural language processing module matches the documents to the MeSH vocabulary, and uses those terms for a guided query formulation on the Pubmed. The query produced by the CDAPubMED is the intersection of the MeSH terms of the publication with the union of the MeSH terms selected from the EHR and the diagnosis of the patient.

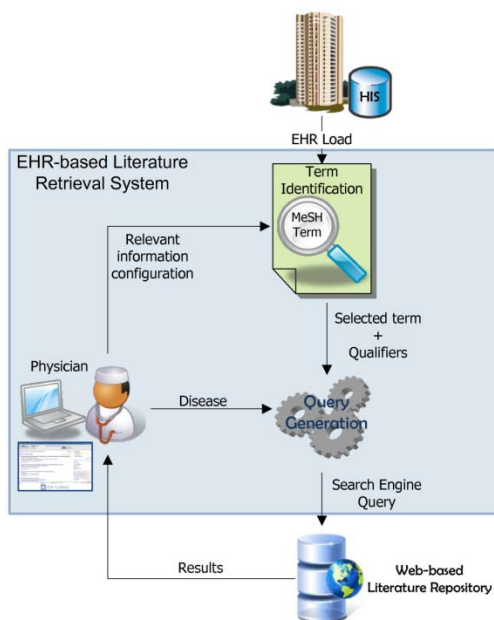


Figure 11 CDAPubMed Architecture (Perez-Rey, Jimenez-Castellanos, Garcia-Remesa, Crespo, & Maojo, 2012)

The use of contextual information from both sensors and EHRs has been proposed (Jimenez-Castellanos, Fernandez, Perez-Rey, & Viejo, 2013) for a use case of diabetes and arterial hypertension. The first step of the proposed solution is the extraction of information from the EHR and the sensors of a patient. The information is divided into structured parameters and unstructured parameters. Depending on the use case, a set of rules for the inclusion of the structured parameters could be developed. The included parameters are then mapped and classified according to MeSH terms for the selection phase by an expert. The selected by the expert parameters are then translated into a PICO query and further

transformed to match the requirements of different medical literature sources. The final step is the merge of the results. There is a claimed precision of 90% on the retrieved results.

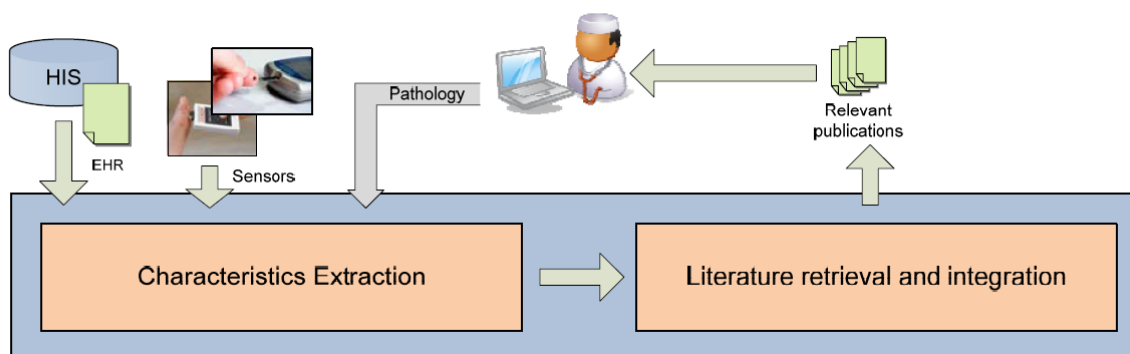


Figure 12 Patient-based literature retrieval Architecture (Jimenez-Castellanos, Fernandez, Perez-Rey, & Viejo, 2013)

Another research on the use of EHR in the searching process is that of Mendonca et al. (Mendonça, Cimino, & Johnson, Using narrative reports to support a digital library., 2001). Their methodology is based on the assumption that there is no judgment of the relevance of the terms of the EHR by the user. This is the opposite of the methodology used in the two systems presented above. Their hypothesis is that the rarer a term in the EHRs the most significant they are and could be used in searching the literature. In the system architecture, there is a data mining system that corresponds a finding with a body part. The research was initially limited into the evaluation of different techniques (TF-IDF combinations) in order to evaluate the data mining system and the recognition of the important terms of an EHR. There is a clear suggestion on evaluating the results of this study in the direct use to the literature search process. After the evaluation of the NLP system an initial study took place (Mendonça, Cimino, & Johnson, Using Patient Data to Rank Records of Literature Retrieval, 2002) showing correlation of the relevance of the results through the user's selection of relevant terms and the automated (TF-IDF) based approach.

A different approach towards the objective of improving the searching process is proposed by Yadav and Poellabauer (Yadav & Poellabauer, 2012). The system developed, contains a search engine (Solr) that has indexed web sites after a Selective Web Crawling with trusted URLs as seeds. Besides that, it makes use of an in-house developed Personal Health Record from which the present and previous conditions are extracted and imported as a query terms with the ability to apply different weights. The terms are inserted into a query and the matching and scoring is done with an extension of the Lucene tf/idf function. The most relevant results of the indexed pages are then presented to the user. There is no clear indication of the efficiency of the approach besides the occurrence of terms included in the query within the top 5 results of the search.

Other context aware systems and architectures

In the previous section, the use of contextual information contained in the EHR of a patient in search engines is presented. In this subsection, the research done to other extensions of the context aware area is examined.

Other context aware searching approaches include into the scoring function contextual information from previous queries or click through data from the logs of the user (Xiang, Jiang, Pei, Sun, Chen, & Li, 2010), (Jiang, Kenneth , & Wilfred, 2011) (Fonseca, Golgher, Possas, Ribiera-Neto, & Ziviani, 2005) that leads to search Personalization. Another approach towards that direction is the use of ontological user profiles (Sieg, Mobasher, & Burke, 2007) where user queries contribute to the creation of the user profiles.

Context awareness in most of the cases is referred to the existence of sensors that record the location of the user, the time of the day etc. In a study about the digital library of the future (Noh, 2013) several context aware services are suggested based on location sensors of the visitors and their friends for automating the services provided by the library such as recommendations but also for safety issues. These sensor oriented approaches are dominated in the context aware searching systems as presented in the study performed by Hon et al. (Hong, Suh, & Kim, 2009).

To some extent, these sensors signals are used to context aware recommending systems. These recommending systems extend the traditional recommending systems adding one additional dimension for each type of contextual information chosen to be included into the recommending model. A detailed analysis of these systems is included in Context Aware Recommending Systems (Adomavicius & Tuzhilin, 2011). The traditional recommending model follows the formula:

$$R: User \times Item \rightarrow Rating$$

In the formula above, R is the estimation of a rating for an Item to a certain User. A Recommending System recommends the top K Items to each User. In CARS (Context AwaRe Systems) there is a third aspect incorporated to the previous formula that affects the prediction of ratings, which is the chosen contextual information:

$$R: User \times Item \times \mathbf{Context} \rightarrow Rating;$$

The reference to the CARS was due to some interesting concepts that were incorporated to the analysis for Hippocrates. The three approaches into creating a CARS is the pre-filtering, the post filtering and the contextual modeling. An analogy to the searching process is the query expansion, the re-ranking and the scoring algorithm respectively which are implemented into Hippocrates as seen in Chapter 3.

2.7 Re-finding and revisiting the berry-picking model

Re-finding is one of the concepts used in Hippocrates. According to a query log from Yahoo, user queries are repeated in 33% of the cases and frequently the pages accessed are the same as in the previous time (Teevan, Adar, Jones R, & Potts, 2007). In case of Hippocrates though re-finding is implemented as a separated module in order to be able to measure the usefulness of such functionality in the oncologists' daily workflow. On the contrary many studies present ways of incorporating re-finding with the finding task at the same time (Teevan, The Re:Search Engine: Simultaneous Support for Finding and Re-Finding, 2007).

To conclude this chapter, by summarizing the main concepts presented, the berry-picking model is revisited. As shown in the Figure 13 a team has a given context or a specific Universe of Interest which is a superset of the individuals' universes. The context aware systems, extracts the information from the universe and modifies the queries (Q) of the users. The systems reviewed in the section 2.6 from the medical domain are based on the Medline button and follow the pattern term extraction, MeSH term mapping, and user selection. The query in the system is used to score the documents having a minimum number of matched words based on a scoring function such as the BM25. The user might want to reach again a previously found document (re-finding R). Another group member such as user 2 might have similar information needs since she is working on the same context (case). In this case collaboration could enable a more efficient searching process by using the same or similar queries or retrieving related documents previously rated by the user 1. In section 2.5 such systems were reviewed with the most detailed described and relevant to the project that of the I-SPY system.

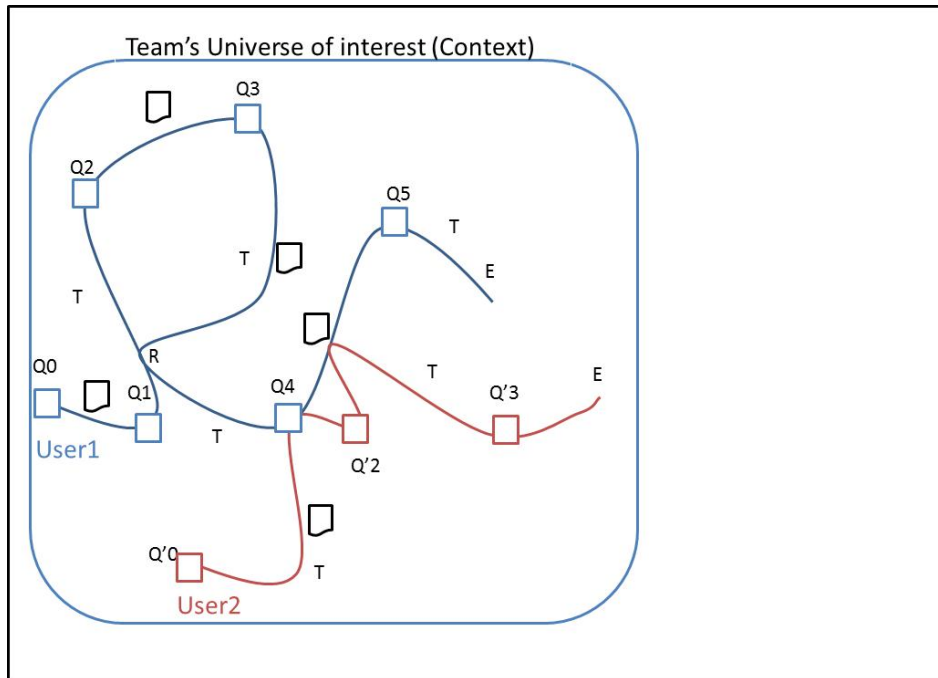


Figure 13 Berry-Picking model revisited

3 Hippocrates, algorithms and implementation description

In this chapter, there is a presentation of the implementation details of the Hippocrates search tool. From the data import till the UI, all the modules of the tool are described in the following sections.

3.1 System description

The Hippocrates tool is developed in a local WAMP installation the details of which are:

- Database MySQL v.5.6.12
- Web server Apache v.2.4.4 (Win64)
- PHP 5.4.12

The Search Engine is a configuration of the Solr v.4.5.1. The components are installed on a Windows 7 x64, Intel Corei5 @2,6GHz 2540M 4GB RAM system.

For the evaluation of the tool, a migration of the tool with selected features was necessary. The URL of the server where the tool is set up is: <http://hippocrates.ehv.campus.philips.com>.

- Database MySQL v.5.1.73
- Web server Apache v. 2.2.15
- PHP 5.3.3

The server runs with 1 vcpu of an Intel(R) Xeon(R) CPU E7- 4830 @ 2.13GHz processor with 4GB RAM on a Red Hat Enterprise Linux Server release 6.5 (Santiago).

The high level system architecture of the system is shown in Figure 14 and each component is analyzed in the following sections.

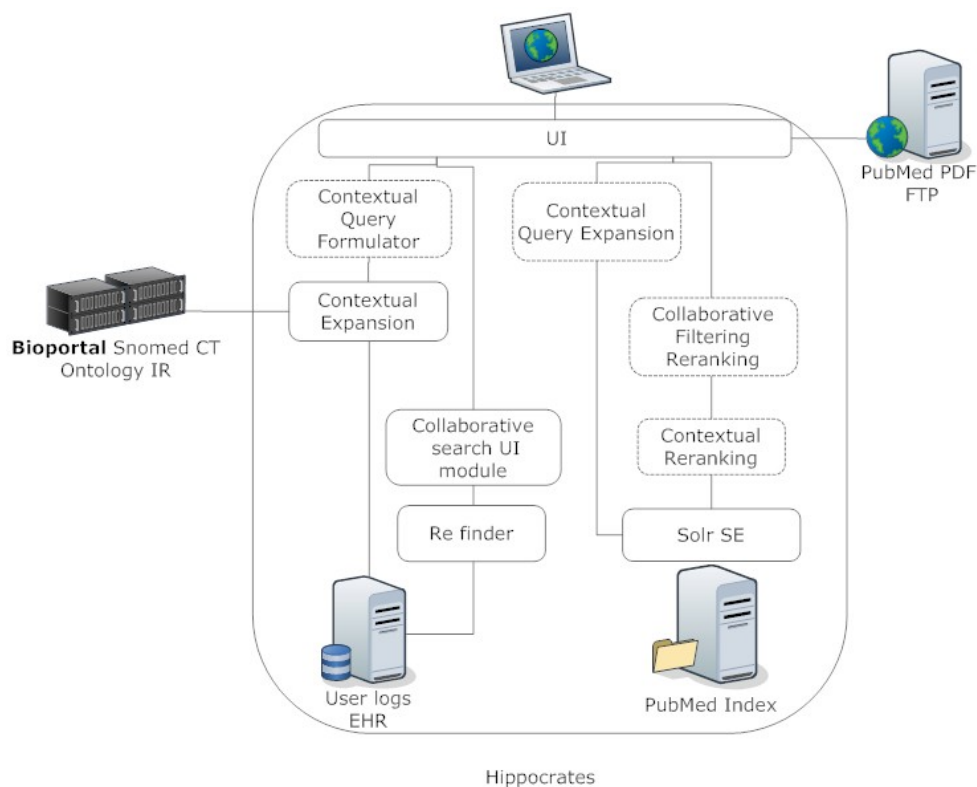


Figure 14 High level system architecture of Hippocrates

In order to complete the overview of Hippocrates, the next table presents the functionality of each module of the Architecture of Figure 14.

Module	Functionality overview
Contextual Expansion	Loads the EHR of the patient, processes it and expands it through the use of the Bioportal service
Bioportal API	Identifies the diagnosis from the EHR, retrieves alternative labels and sibling medications of the identified antineoplastic agent administered to the patient
Contextual Query Formulator	Loads the alternative labels for the diagnosis and recommends them to the user as alternative queries. Loads the EHR and presents it as pre-filtering options
Re-finder	Based on the logs, it retrieves the recent documents and queries of the user
Collaborative search UI module	Based on the logs, the ratings and the workflow state, documents are retrieved and recommended to the user on a team base and on users' base. Email functionality and collaborative query recommender are included.
Solr SE	The search engine that parses the queries scores the documents indexed and returns the results. Solr is accessed through Solarium PHP client.
Contextual Query Expansion	Selected concepts from the EHR expand the user query in order to receive more relevant results

Contextual Reranking	The top results of the user's query are reranked based on a Boolean score of containing a concept for the EHR of the patient
Collaborative filtering Reranking	An I-Spy based algorithm reranks the documents based on ratings of the users' base for the same diagnosis and workflow state.
UI	Orchestrates the modules in order to present them in a structured way.

3.2 Data import, contextual expansion and contextual query formulator

The data of the patient following the Common Data Model of the EURECA framework is structured in a way that needs further processing in order to be used by Hippocrates and in order to present the contextual information of the patient to the user and include that information in the searching process. As shown in the section 2.1.2, one patient is linked to multiple acts which are of type Observations, Procedures, and Substance Administration. So further processing of the EHR is required. The NLP processing that takes place before the construction of the EHR is out of the scope of Hippocrates.

Contextual Expansion

The first task of the component Contextual Expansion is to classify the patient according to the date of birth by comparing the current time of the system and the date of birth of the patient. A simple rule is applied in that case, if the difference is greater than 18 years, the patient is an adult, otherwise is a child. As explained in 2.1.2, the *administrativeGenderCode*, which is 248152002 for the female and 248153007 for the male according to the SNOMED CT ontology is translated by the system.

After the retrieval of the Acts, each class is handled in an array internally by the system and expanded depending on the class by the corresponding *Target body structure* or *Observation values*. At this point the system has no knowledge on which is the observation that corresponds to the Diagnosis of the patient e.g. there is no distinction between the Observations *Tumor Size* and the *Infiltrating Duct Carcinoma* which is the actual diagnosis.

The identification of the diagnosis of the patient but also the retrieval of the same type of antineoplastic agents that are administered to a patient, are tasks performed with the provocation of an external service, the Bioportal. The Bioportal contains many of the ontologies used in the Bio-Medical domain. SNOMED CT is one of the ontologies used for encoding concepts used by EURECA and is also one of the ontologies included in Bioportal. The service provides search functionality of codes existing in the ontology, but it is usually time demanding. Hippocrates makes use of another functionality of the Bioportal which is the searching inside a hierarchy. The integration with Hippocrates is achieved through the API of the Bioportal service⁶.

⁶ <http://data.bioontology.org/documentation>

For the identification of the diagnosis of the patient, the Acts of class Observation is being searched in the sub tree Epithelial Neoplasms (SNOMED CT code: 118285006). This is a limitation of the system since there are neoplastic diseases that are outside the scope of the selected sub-tree. The performance though would be significantly decreased, in case the more general term Neoplasm was selected to be used for querying the Biportal. Searches under the Neoplasms sub-tree lasted, during the development phase, up to several minutes and in most of the cases failed to return any results. Besides the identification of the diagnosis, the alternative labels of the diagnosis are retrieved and used as query expansion recommendations to the user.



Figure 15 Use of Diagnosis as Initial Query with alternative labels retrieved from SNOMED CT

For the identification of the antineoplastic agent administered to the patient the system calls the API to search each code of the class *Substance Administration* on the sub-tree of the Antineoplastic Agents (SNOMED CT code: 27867009). If an Antineoplastic agent is identified the parent and the sibling of that agent is retrieved as shown in Figure 16 in order to provide options to the users to search for alternative treatments as well.

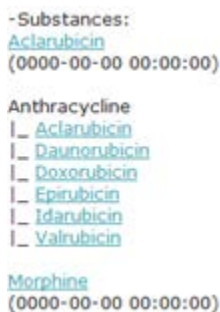


Figure 16 Contextual expansion of the Aclarubicin Antineoplastic Agent

Contextual query formulator

The result of the processing of the EHR data is seen in Figure 17. After this step, the EHR concepts are used for guiding the query formulation of the user by selecting and deselecting concepts to be included in the query. The information included in the EHR is presented in a structured way with the first part including the general information of the patient, the second part includes the Observations with the latest diagnosis on top and the body structure related to that in different color. In the third part, the procedures of the patient and in the last part the substances that are administered to the patient with the sibling of the antineoplastic agents are presented. Where is available the time of registration of each class is presented in order to provide the full picture of the EHR to the user.

In the previous sub section there is a description of an alternative use of the contextual expansion information which is the use of the alternative labels of the diagnosis in the query of the user as seen in Figure 15. This functionality is included in the contextual query formulator module, though it is not migrated to the evaluation server since a collaborative UI module was selected to be evaluated instead.

<p>-livingSubject: <u>administrativeGenderCode</u>: 248152002 <u>birthTime</u>: 1965-03-01 00:00:00</p> <p>-Act: <u>class</u>: OBS, <u>title</u>: Infiltrating duct carcinoma <u>effectiveTime</u>: 2006-02-02 00:00:00</p> <p><u>class</u>: OBS, <u>title</u>: Tumor size <u>effectiveTime</u>: 2006-02-02 00:00:00</p> <p><u>class</u>: PROC, <u>title</u>: Surgical procedure <u>effectiveTime</u>: 2006-02-02 00:00:00</p> <p><u>class</u>: SBADM, <u>title</u>: Aclarubicin <u>effectiveTime</u>: 0000-00-00 00:00:00</p> <p><u>class</u>: SBADM, <u>title</u>: Aclarubicin <u>effectiveTime</u>: 0000-00-00 00:00:00</p> <p>-ActTargetSiteCode (for the Infiltrating duct carcinoma act) <u>code</u>: 76752008 <u>title</u>: Breast structure (body structure)</p> <p><u>code</u>: 7771000 <u>title</u>: left <u>code</u>: 78615007 <u>title</u>: with laterality</p> <p>-ActObservationValue (for the tumor size) <u>valueST</u>: 52 <u>units</u>: mm</p>	<hr/> <p>Filter out:</p> <p>-Patient data: Adult Female</p> <hr/> <p>-Observations: Infiltrating duct carcinoma (2006-02-02 00:00:00) Breast structure left with laterality</p> <p>Tumor (2006-02-02 00:00:00) 52mm</p> <hr/> <p>-Procedures: Surgical procedure (2006-02-02 00:00:00)</p> <hr/> <p>-Substances: Aclarubicin (0000-00-00 00:00:00)</p> <p>Anthracycline _ Aclarubicin _ Daunorubicin _ Doxorubicin _ Epirubicin _ Idarubicin _ Valrubicin</p> <p>Morphine (0000-00-00 00:00:00)</p>
---	---

Figure 17 From the EHR to the Contextual query formulator

Limitations

As mentioned above there are limitations using the Bioportal service. Even in the limited usage that was chosen for Hippocrates, the search and retrieval last several seconds or even a minute before returning results. Currently, for the needs of other parts of the EURECA project an internal development of the SNOMED CT ontology is in progress and expected to be more efficient for the tasks described above. Also a redesign of the NLP processing is taking place, in order to include the information regarding the Act which is the corresponding diagnosis of a patient as indicated in the HL7 RIM standard. These alterations are expected to reduce the time of the contextual expansion of the EHR to tolerable levels.

3.3 Solr configuration

In this section, there is an overview of the Solr configuration details from preprocessing and indexing to the different modules such as the content based recommender. Solr is an open source search platform based on the Lucene search engine libraries. It includes many features some of them configured and included in Hippocrates. For the Solr-Hippocrates integration, the Solarium v 3.2 was used. Solarium is a PHP client for Solr. Many of the Solr components are used in the UI layer, the produced result of which is presented in this section.

3.3.1 Pubmed References Preprocessing and Indexing

The first step in order to build up a search engine is to acquire the document set to be searched. In case of web search engines, crawlers are usually used. In case of Hippocrates, the documents were obtained from the Medline/Pubmed digital library after the corresponding memorandum of understanding was signed. The selected documents from PubMed and in order to scope down the project include the MeSH term *Neoplasm*.

The structure of the documents was given in section 2.2. The fields indexed from each document are a subset of the fields provided for each document and are the following:

Field	Description
PMID	The Unique ID of the Article
Title	Title of the journal that published the Article
ArticleTitle	Title of the Article
AbstractText	The abstract of the Article
PublicationType	The type of Publication. It may be Journal Article or/and Review etc.
ArticleIdPMC	The Pubmed Central ID based on which the link for the FTP retrieval of the full article is created
Affiliation	May include the list of authors
MedlinePgn	The pagination of the reference
ISOAbbreviation	Is used for linking different reference systems
Volume	The Volume of the Journal that the article was published
Issue	The specific issue in which the article was published
Year	The Year the article was published
Month	The Month the article was published
Day	The Day the article was published
Author	The list of authors of the article

The previous fields were stored and indexed. Though the search is taking place on the fields ArticleTitle and AbstractText. The preprocessing includes lowercasing, stemming and stop words removal from the fields, the same as in the query parsing step.

3.3.2 Matching and Scoring

The returned results of medical literature references are a subset of the set of references included in the index of Hippocrates. The references returned must follow a simple rule, to contain more than 80% of the terms included in the user query.

As presented in the section 2.4, the BM25 scoring function is considered the state of the art algorithm in the probabilistic retrieval framework. An implementation of the BM25 function is available for the Solr search engine. The function is implemented as a class and can be activated at the Schema.xml configuration file. The actual scoring approach selected is based but not limited to the BM25 scoring. The motivation for this approach is that phrases matches should be boosted over simple keywords matching. Another choice made was that scores of the Article Title matching should be scored higher than the score of the Abstract Text. The motivation for this choice is that, if the keyword of a query is matched in the Title it is more probable to cover the given information need than if there was a match inside the Abstract Text.

$$Score = \sum_{t \in q} \max\{BM25(term, k_{Title}), BM25(term, k_{AbstractText})\}$$

Phrases are scored as extra terms in the equation above. Phrases are considered the keywords with distance of less than 7 terms in between.

Sorting by date

Until the October of 2013 the users of Pubmed could only retrieve results based on a Boolean matching sorted by date. It is assumed that the users of such a system are interested in recent developments in their domain of expertise. Therefore a similar functionality might be important to the users of Hippocrates. For that reason, a sort by date functionality was configured in the Solr search engine and can be activated by the user.

This option can be activated through the UI by a dropdown menu with the user choosing between sorting by Relevance and sort by Year. Through the Solarium client for PHP the corresponding Solr query is created, sorting the results by Year of publication. The matching of the Keywords follows the method described in the previous section. There is no scoring taking place in this case.

3.3.3 The More Like This component

The more like this component was configured in Hippocrates in order to provide similar documents to that selected by the user. The component creates two vectors, one for each field ArticleTitle and AbstractText and performs a tf/idf scoring for the documents that have a minimum matching of words. The top 5 results are then presented on the left side of the page where the selected document is presented. The evaluation of this module is left for the general usability evaluation and alternative choices are left as future work.

Similar references

[Evaluation of immunocytochemical staining as a method of improving diagnostic accuracy in a routine cytopathology laboratory.](#)

[\[Use of immunohistochemistry in clinical cytology\].](#)

[\[Cytodiagnosis and marker determination in pleural effusion\].](#)

[A new marker for human cancer cells. 3. Immunocytochemical detection of malignant cells in serous fluids with the Ca1 antibody.](#)

[Primary effusion lymphoma: cytopathologic diagnosis using in situ molecular genetic analysis for human herpesvirus 8.](#)

Figure 18 The MLT component

3.3.4 Autocomplete

Based on the index created from the ArticleTitle and the AbstractText, the Autocomplete module, proposes keywords to be included by the user. There is one limitation for this module. Since the index contains stemmed words, the proposed words are frequently stemmed. To overcome this issue a new index could be created, but due to limited resources this solution is not implemented.

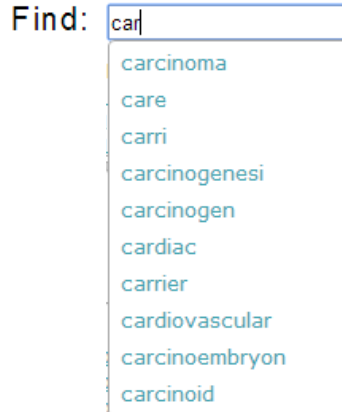


Figure 19 The Autocomplete feature

3.3.5 Text Highlighting

This module returns the result set with a preconfigured html tag in order to highlight the matched words of the query and the returned document. A motivation to include this component to the tool is that the user needs some justification for the results she receives. Matched words is the main cause for the result set she receives and potentially the highlighting of the matched terms might lead to a faster evaluation of the result compared to the information need. A limitation in this component is that the highlighting is case sensitive. In case there is a capital letter in a word of a query then the highlighting is not taking place. This holds for mismatches with the result set as well.

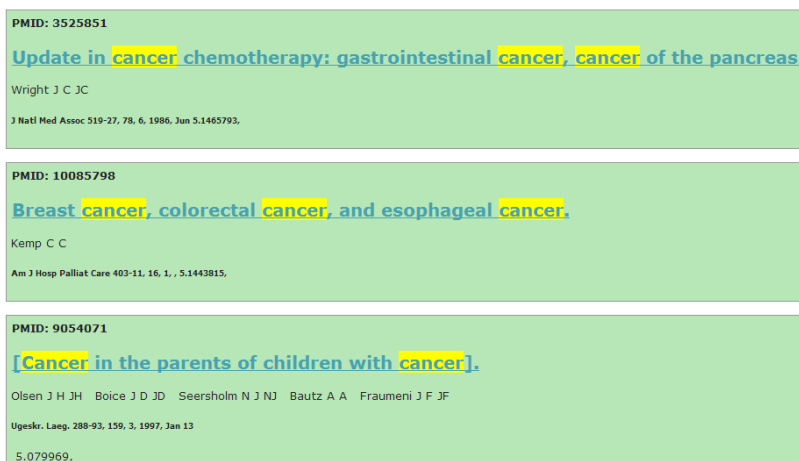


Figure 20 Text highlighting for the query "cancer"

3.3.6 Faceted Search

Faceted search is one of the many features of the Solr. This feature creates a menu of links in the side of the result set of Hippocrates. These links when pressed and activated, narrow down the result set based on a given criteria. In case of Hippocrates, three facet menus were selected to be configured, based on the year of publication, the publication type and the Journal type. The first two were chosen in order to give a familiar environment to the users based on the Pubmed page. The third menu of Journal Title was implemented after the first feedback from the oncologist as presented in section 4.1. The year menu presents in decades interval the number of documents corresponding to each decade. The rest two menus present the top 5 and 10 Publication Types and Journal Title respectively.



Figure 21 Faceted Search menus

3.4 Contextual Search

There are three basic methods to affect the search task on a given document set and document representation; by affecting the query, by changing the matching and scoring algorithm and finally by re-ranking the scored documents. The contextualization of the searching task is described in this section. The graphical representation of contextualized search is give on Figure 22. The left model refers to the context aware search engines which scores the documents taking the context into account. The middle model re-ranks the results based on the context; in our case the concepts extracted by the search engine. In the right model the query is modified in order to include the contextual information from the EHR. In Hippocrates the last two models are implemented into three different components that could be manually selected by the user and are described in the following subsections.

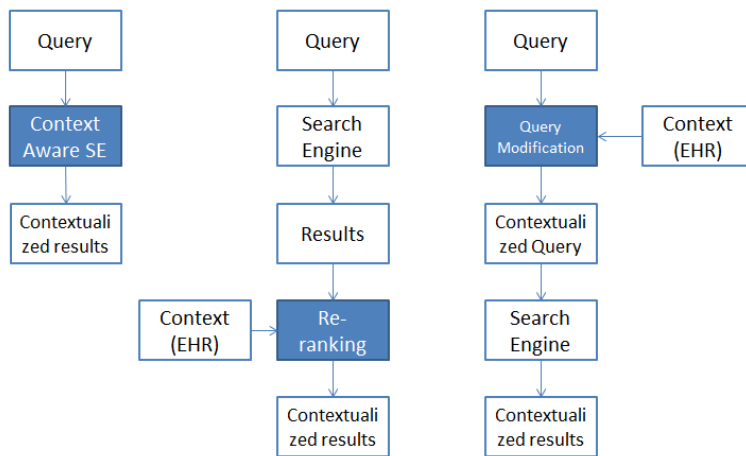


Figure 22 Approaches for contextualizes search

3.4.1 Query Expansion

As presented in 2.6 there are several attempts to capture contextual information from the query history of a user to the contents of the visited pages. These systems expand the user query with additional terms of their previous queries. In our system the same approach to provide contextualized search is applied by including in the query the concepts extracted from the EHR. Initially all the concepts were included in the query with lowering the minimum matching criteria from 80% of the terms to 50%. The experimentation with this approach though suggested that many irrelevant results were retrieved due to very common words such as woman or adult.

Based on a behavior analysis of the users of Pubmed the diagnosis is used by the 20% of the issued queries to Pubmed (Rezarta, Murray , Aurelie, & Zhiyong, 2009). Thus the current setup of this component expands the user query with the observation of the patient and mainly focusing in the diagnosis. A suggested future work is a comparative study with the inclusion of other type of concepts such as the Drugs which are included in a percentage of 11% of the queries issued on Pubmed.

3.4.2 Contextual Pre filtering –Guided Query formulation

This module is presented in 3.2. The aspect not presented is how the filters affect the matching criteria presented in 3.3.2. The selected terms are considered as Boolean queries and the articles returned contain at least one occurrence of the term or concept.

3.4.3 Boolean Based Score for Contextual Re-ranking

In case the corresponded option is selected, the re-ranking component is activated. In that case, the top N=500 documents are assigned to an array internally. Another array is created with the terms extracted from the EHR. Then a Boolean score is calculated for each document which corresponds to the number of matched terms from an EHR to the content of the document. Then the final score is calculated by:

$$Score = Score + k * BooleanScore$$

Where *Score* is the resulted score from the BM25 based score of the search engine and *k* is a normalization weight for the *BooleanScore*. In our case k=10.

In continuation of the current project within Philips Research an alternative scoring algorithm was registered as an invention⁷.

3.5 Collaborative Search

This section includes a presentation of all the components used to enhance the collaboration among a group of users but also to enhance the searching process based on the knowledge that is acquired through the usage of Hippocrates. This is done through a collaborative filtering based re-ranking algorithm and different UI components.

3.5.1 Capturing information

Two plugins were used in order to capture information about the judging the relevance of an article to the given case. The first is a rating system that enables the user to rate an article with 1-5 stars (rating system). The scale selected is:

1. Highly irrelevant
2. Irrelevant

⁷ Registered Invention and assigned to Philips IP&S

Inv. Disc.: 2014ID01348

Title: Contextualized and targeted presentation of clinical knowledge based on available patient data

3. Relevant
4. Highly relevant
5. Perfect match

The ratings captured by this plugin are used in most of the following collaborative search modules as a minimum threshold in order to filter out results or as a part in ranking algorithms.

Another plugin used is a tagging system (tagit-jquery). The functionality of this system is limited to present the query terms as default tags describing the information need that this article covers and by allowing the users to modify this tags by adding or removing terms. The tag sets produced for each article are then used in the collaborative filtering re-ranking 3.5.2. Future extension of this plugin would enable the users to search for their own tags and search for articles tagged with similar tags.

3.5.2 Collaborative filtering re-ranking

The collaborative filtering re-ranking algorithm is based in that of I-SPY presented in 2.5. The result of the algorithm is to boost documents highly rated for similar queries and cases and to penalize documents with lower ratings. The algorithm presented in the I-Spy section is altered in order to make use of the rating component and the contextual information that is available through the EHR and the Hippocrates system. The extension proposed includes the rating of the users instead of assuming as relevant the pages that were clicked and a penalization of the deviation of the ratings. Besides that, a contextual pre-filtering is applied. The set of documents retrieved, are rated in a context of the same diagnosis and workflow state.

As seen in Kwon's paper (Kwon, 2008) by subtracting the standard deviation of the ratings of a movie, it improves the accuracy in predicting an unknown rating. A similar approach is implemented in Hippocrates.

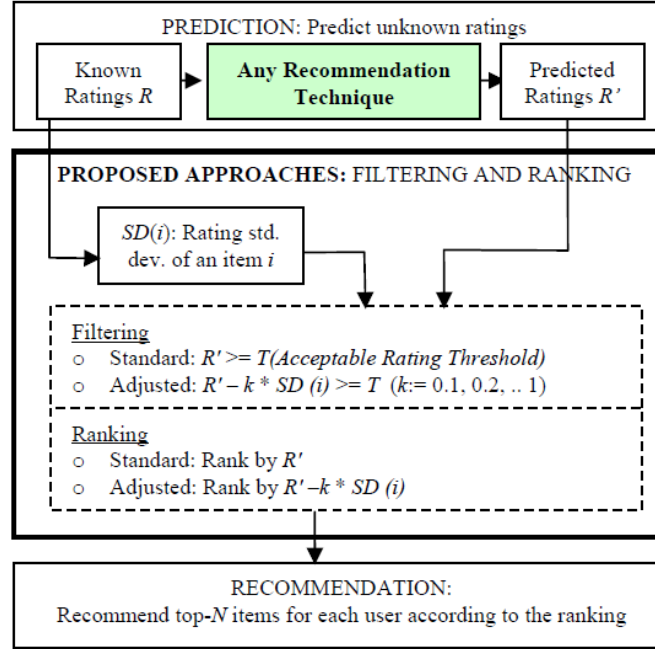


Figure 23 Proposed approach in including SD in the recommendation models (Kwon, 2008)

The I-SPY equations are expanded in the following way:

$$Sim(q, q') = \frac{|q \cap q'|}{|q \cup q'|} \quad (1)$$

$$Relevance(p_j, q_i) = \frac{H_{ij}}{\sum_j H_{ij}} \quad (2)$$

$$RatingScore(p_j, q_i) = \frac{AvgScore_{ij}}{1 + STD_{ij}} \quad (3)$$

$$W Rel(p_j, q_i, q_1, \dots, q_n) = \frac{\sum_{i=1 \dots n} Relevance(p_j, q_i) * Sim(q_T, q_i) * RatingScore(p_j, q_i)}{\sum_{i=1 \dots n} Exist(p_j, q_i) * Sim(q_T, q_i)} \quad (4)$$

Into more detail, the algorithm is described through the following steps:

1. Select the articles, the ratings and the tags where the diagnosis is the same as the diagnosis of the current patient and the workflow state is the same as the workflow state of the current case
2. Calculate the similarity of the tags used for the rating of a document with current query.
3. Filter out the results with lower threshold of 60% (1)
4. Group the documents by tag set
5. Count the frequency that each document is rated on a given tag set (2)
6. Calculate the Rating Score of a document for each tag set (3)
7. Calculate the score (4) of a document by aggregating the scores of the document for the different tag sets.

The scores calculated from the algorithm, are stored in an external file. Then the score of each document is calculated as following:

$$Score = Score + k * MaxScore * WRel$$

The MaxScore is the score of the first document and k is a multiplier in Hippocrates which is set to 0,2.

3.5.3 UI enabled collaborative search modules

Four components are developed in order to promote the collaboration through UI intervention. The first is presented in Figure 24. The module, presents to the users, what the rest team members have rated above a threshold of 3 stars (Rellevant). It also includes other contextual information such as the time of rating the user and the workflow state when the rating was imported to the system. The sorting of the presented recommendations is by Workflow State and then by rating. As a future work is left the task to aggregate comments and ratings in a single entry of an article.

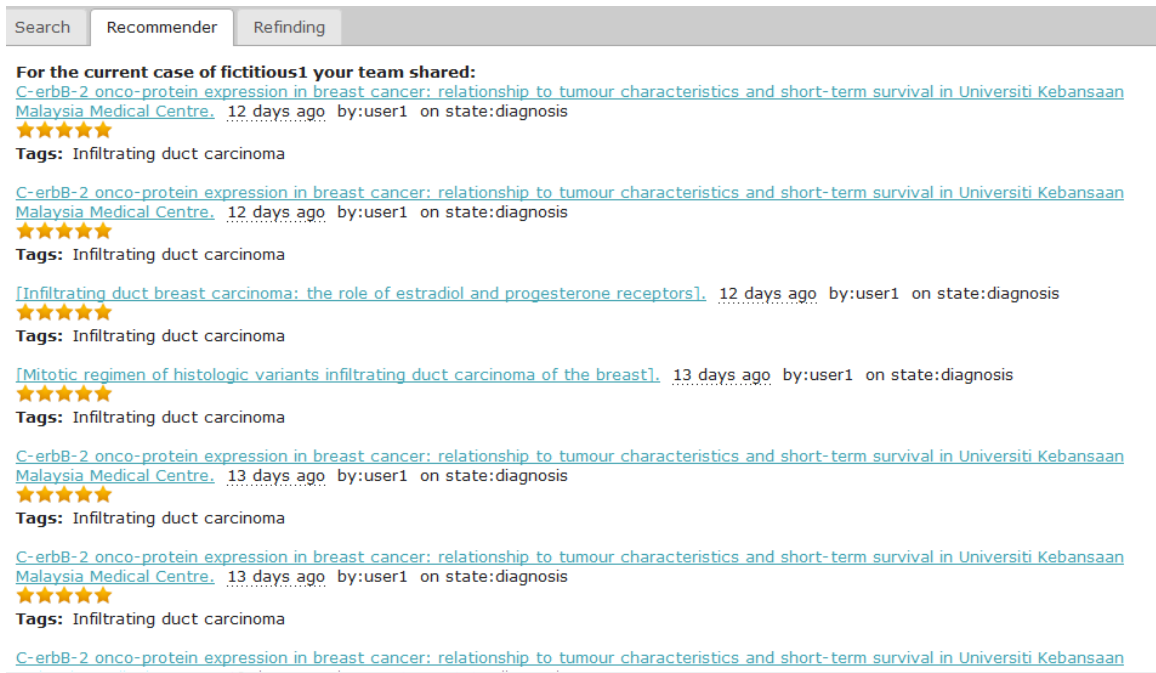


Figure 24 Case based collaboration module

The second module presented in Figure 25 recommends to the user the articles with positive score from the collaborative re-ranking algorithm presented in 3.5.2. This presents to the user which documents other users have rated positively for the same diagnosis, workflow state, and for similar queries. The positive scores are transformed to ratings through the star rating plugin.

For cases with the same diagnosis and similar queries, doctors shared:

- [\[Infiltrating duct breast carcinoma: the role of estradiol and progesterone receptors\].](#) ★★★★★
 - [C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre.](#) ★★★★★
 - [The prognostic value of p53 and c-erbB-2 expression, proliferative activity and angiogenesis in node-negative breast carcinoma.](#) ★★★★★
 - [Selected immunohistochemical prognostic factors in endometrial cancer.](#) ★★★★★
 - [\[Correlation of CerbB-2 to Ki-67 and P53 expressions in hormone-independent breast cancer\].](#) ★★★★★
 - [Expression of JMJD2A in infiltrating duct carcinoma was markedly higher than fibroadenoma, and associated with expression of ARHI, p53 and ER in infiltrating duct carcinoma.](#) ★★★★★
 - [Malignant cystosarcoma phyllodes with simultaneous carcinoma in the ipsilateral breast.](#) ★★★★★
 - [\[Histological classification of malignant breast neoplasms. Recent concepts\].](#) ★★★★★
 - [\[Mitotic regimen of histologic variants infiltrating duct carcinoma of the breast\].](#) ★★★★★
 - [Outcomes of breast-conservation therapy for invasive lobular carcinoma are equivalent to those for invasive ductal carcinoma.](#) ★★★★★
-

Figure 25 Recommender based on the collaborative re-ranking algorithm

Context aware-collaborative Query Recommendations

In order to raise group awareness of the search task but also to enable a learning process of searching strategies among the group members, a query recommender was implemented. Based on the current workflow state and the frequency of issuing a query for a given case, the top 5 queries set is proposed to the user.

Find: Sort by:

Your team has searched for:
[Infiltrating duct carcinoma](#)
[Duct carcinoma](#)
[infiltrating](#)
[ductal breast cancer](#)
[cancer](#)

Figure 26 Context aware-collaborative Query Recommendations

Email for collaboration

As described in the section 2.5, the most common mean used to communicate interesting results of a search process is through email. The use of email was confirmed by the oncologist (section 4.1). There is an implementation of sending email of a selected article in Hippocrates. This component is implemented after the feedback for the necessity of such a feature in the evaluation by Ahmed Ibrahim described in 4.3.

3.6 Other UI modules

Until now the majority of the components were presented with the actual result presented to the user in the UI layer. The remaining two components that were not presented until this point are included in this section.

3.6.1 Visual Notification of updated results

The searching process as the Berry-picking model proposes and seen in 2.3 is not a simple IR task but a process that includes several steps until the information need is fulfilled. The user may follow some actions during that course such as generalization of the query or rephrasing the query in order to include more contextual information. In order to improve that process a visual notification of the new results component was implemented. The component stores the results of the previous query and compares them with the current results. The new results are presented with a different background color; green instead of grey.



Figure 27 Visual Notification of updated results module

3.6.2 Re finder

The motivation for implementing this module is presented in section 2.7 and is based on the fact that for 1/3 of the searching is basically re finding previously visited pages. In order to test that hypothesis the *Refinder* module was implemented. It provides information based on the log of each user. The Upper part presents the previously visited documents providing in parallel the rating of the user, her tags and comments. The second part concerns the query history of the user. This section presents to the user her latest queries she issued. Different approaches on incorporating re-finding in the result set exist in the literature and are left out of the scope of the current project.

Your recently accessed documents:

[Biologically based treatment planning.](#) 2 months ago



Tags: complications quality of life risks treatment treatment plans

Comment: t24t3

Your recent searches:

[Infiltrating carcinoma](#) 2 minutes ago

[Infiltrating duct carcinoma](#) 2 minutes ago

Figure 28 Refinder

4 Evaluation

In the previous section, the motivation for the different implementation choices and the description of the various components of Hippocrates were presented. This chapter includes the evaluation in the two different phases in the development of the tool. The first phase includes the feedback received after the presentation of the core components of Hippocrates in which some collaboration and contextualization components were developed. The second phase includes the results of the experiment and evaluation of Hippocrates.

4.1 First evaluation by medical expert

The first evaluation that was the ending of the first implementation phase and the start of the second, took place on January 24th 2014. Through the online demonstration of the tool the writer received valuable feedback, recommendations for the completion of the requirements and suggestion to retain focus on collaboration and contextualization. The system in that phase included the guided query formulation module from the EHR, a team/case history of accessed references, a demo only module of collaborative recommender and the re-finding module.

A first suggestion was to include the impact factor of each reference in the scoring mechanism. The data regarding the impact factor is not accessible through Pubmed, as it can be seen in the schema⁸. Another reason that this requirement was not implemented was because it would affect the novelty of the results. A more elaborate solution could be developed but left as a future work.

Another suggestion was to provide to the user the choice of the results she would like to access by filtering the journals of her choice. This recommendation was implemented as a faceted filter as described in section 3.3.6. Also, it was recommended to develop and provide to the users a rating, tagging and commenting system that could be visible by other users. Such a system was developed and described in section 3.5.1.

Another input was that the doctors use indeed contextual information from the case. The most interesting parts of an EHR concerns a patient's syndrome that could affect the treatment, the diagnosis and those entries that are make rare combinations e.g. child breast cancer.

Email is extensively used for collaboration in information seeking tasks for references of high importance for a specific case or for a recent development on the domain of oncology. This feature is implemented as well (3.5.3).

⁸ The XML files conform to the following DTDs:
http://www.nlm.nih.gov/databases/dtd/nlmmmedlinecitationset_130101.dtd
http://www.ncbi.nlm.nih.gov/corehtml/query/DTD/pubmed_130501.dtd

Finally, the workflow state differentiates heavily the information needs of the oncologists. In the Treatment planning the users would be more interested in available trials, about different treatment options such as the necessity of radiotherapy and the dosage, surgery and drastic options with regard to outcome of the treatment, the quality of life of the patient and the prevention of potential adverse events.

In general and for each module separately the feedback was highly positive.

4.2 Experimental Design/Usability Evaluation

The experimental design was developed to answer whether the objective presented in section 1.2 was achieved. The initial design was based on the participation of at least 5 medical experts with level of expertise higher or equal to that of the oncology resident. This is due to the high degree of difficulty in judging the relevance of a result in a specialized search tool as Hippocrates and in a given patient/case context. Several invitations for participation were sent to Oncology teams in UK, Germany, Italy and Greece, though the participation to the experiment was limited to an oncologist who participated in the first phase as well. Thus the modification of the experiment was necessary in order to enable the inclusion of informaticians in the process mainly for their feedback on the usability of the system.

The experiment/evaluation included three parts namely the Contextual Search evaluation, the Collaborative Search evaluation and the usability evaluation of the tool. The following sub-sections provide further information for each part of the experiment and the motivation for each choice. The directions, the form and the results are attached as Appendices.

Contextual Search evaluation

The goal of this part of the evaluation is to get an estimation of the effect of applying 3 different approaches in contextualized search:

- Query expansion
- Re-ranking
- Pre-filtering

The users were provided with a case namely of fictitious 3 consisting of the following contextual information:

```

-Patient data:
Adult
Female

-Observations:
Infiltrating duct carcinoma
(2007-09-16 00:00:00)
Breast
structure
left
with
laterality

Tumor
(2007-09-16 00:00:00)
6mm

-Procedures:
Surgical procedure
(2007-09-16 00:00:00)

```

The users were asked to judge the relevance of the first 18 results- 2 pages (in order to judge the **precision@18** and the **cumulative gain**) in the 5 star system implemented with a “basic search” and with all three methods for a particular information need with a single query “complete resection side effects”. This information need was formulated in collaboration with an oncologist. For the informaticians’ evaluation, the more general “breast cancer” query was used.

Precision@k (P@k) corresponds to the number of relevant results in the first k results. The Cumulative Gain (CG) and Discounted Cumulative Gain (DCG) of the results are defined as:

$$CG_p = \sum_{i=1}^p rel_i$$

$$DCG_p = rel_1 + \sum_{i=2}^p \frac{rel_i}{\log_2(i + 1)}$$

In the third case of the query formulation support, the users were asked to use the filtering options derived from the EHR until they reached the most relevant result set. Then they were asked to judge the relevance of the top 18 results once more.

Collaborative search evaluation

For the collaborative search evaluation the relevance judgments from the previous step was planned to be used by a user not participated in the first part of the evaluation. After the description of the case and the information need was asked to:

- Judge the relevance of the top 18 results given a default query.
- Judge the relevance of the top 18 results after the collaborative search module is enabled.

This evaluation was designed to provide insight in how **cumulative gain** was affected by the collaborative module. Because of the low participation, an **internal** evaluation took place instead.

Usability evaluation

The usability evaluation includes three parts; the first includes a general usability satisfaction questionnaire⁹, the second refers to the separate usability evaluation of the basic components of Hippocrates and the third part requests the user's input on the most positive and negative aspect of the tool, new features request and any additional comments.

4.3 Results

In the experiment and evaluation participated three persons in total; the Oncologist Dr. Norbert Graf, the Software Engineer ir. Ahmed Ibrahim and the Data Analyst ir. Kosmas Hatzidimitris.

Contextual Search evaluation

The Contextual Search evaluation is resulted from the rating of the searching task with each component by the users. For comparison reasons the users evaluated without the use of any contextualization component, as Basic, the results of a simple query. The clinical expert's results are presented separately for two reasons. The first is that he completed only half the ratings required by the experiment guidelines and the second is that he used a more specialized query as aforementioned in the previous section.

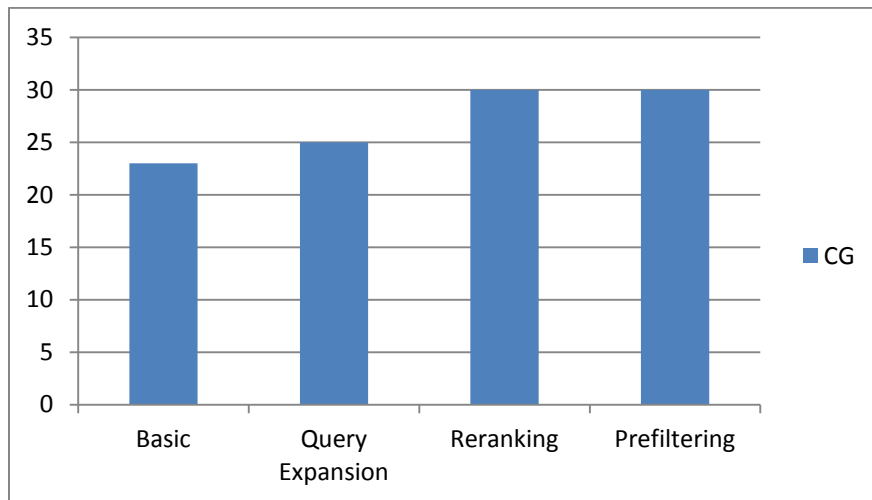


Figure 29 Cumulative Gain Oncologist

As seen in Figure 29 the Cumulative Gain (CG) was increased by the introduction of each of the contextualization components, with Re-ranking and Pre-filtering achieving the same score. The results should be examined under the fact that the maximum CG is 45 derived from the 9 results rated and the 5 as a maximum rating score.

⁹ Adaptation from Lewis, J. R. (1995) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use.

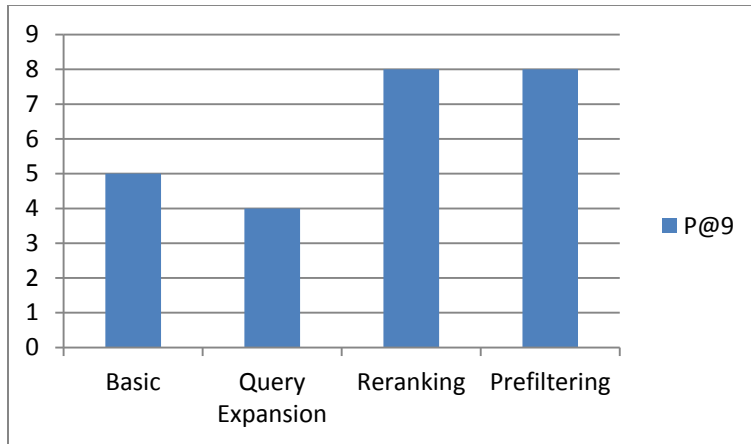


Figure 30 Precision @ 9 Oncologist

The Precision@9 score is presented in Figure 30. It is very positive that with the Re-Ranking and the Pre-Filtering components eight out of nine results were judged as relevant or highly relevant. An early conclusion from the one query tested and the one relevance rating input by a clinician is that the contextualization approaches return results more relevant to a case than the BM25 based scoring function. For the Query expansion this holds partially because it returned one fewer relevant result but the result set in total was more relevant to the case.

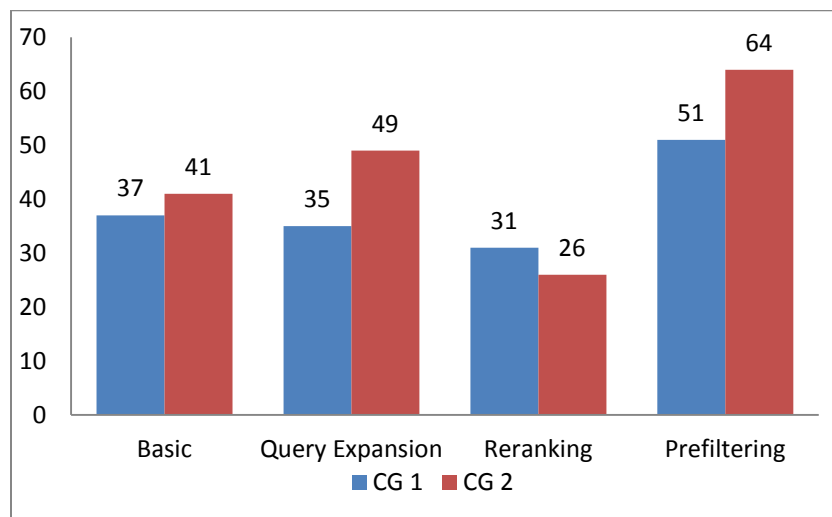


Figure 31 Cumulative Gain Informaticians

The relevance ratings of the informaticians differ. As seen in Figure 31 the Cumulative Gain was definitely lower for the case of re-ranking and for case 1 the query expansion returned results of slightly lower relevance.

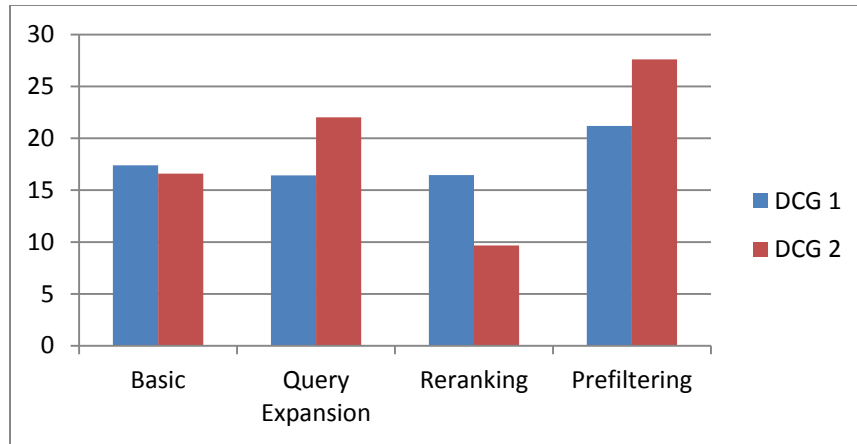


Figure 32 Discounted Cumulative Gain Informaticians

The Discounted Cumulative Gain shown in Figure 32 presents similar results with the CG with the only difference being the Query expansion and the Re-ranking scores for case 1 having closer score to that of the Basic query. This correction especially for Re-ranking is based on ranking the most relevant results in higher rankings.

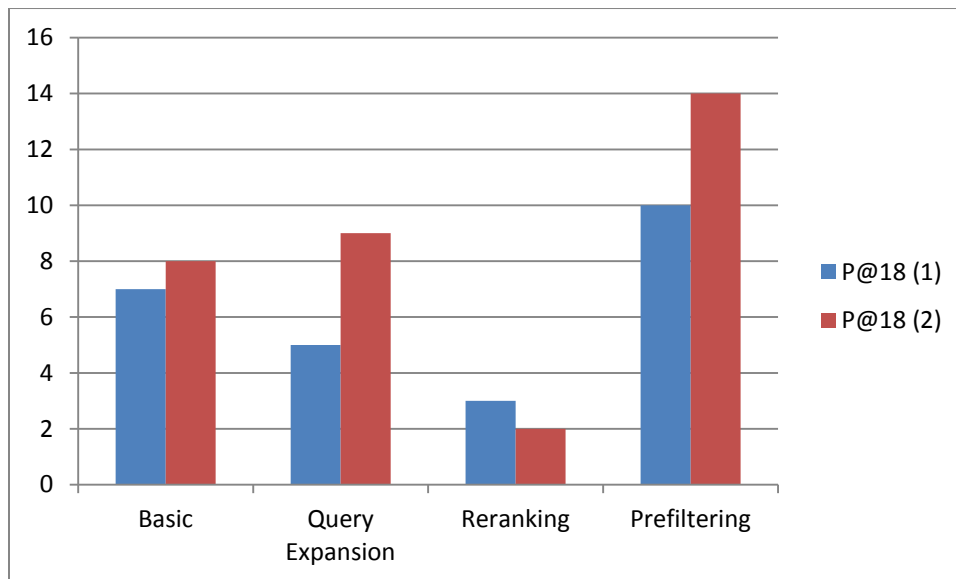


Figure 33 Precision@18 Informaticians

The Precision @18 of Figure 33 resembles to that of the CG metrics. The conclusion of the processing of the relevance input by the informaticians is that re-ranking did not perform as expected, returning results of lower relevance to the users. This applies for the first user for Query Expansion as well. This finding contradicts to the findings from the clinical expert's results. This could be happening for many reasons naming only two; the difference in the background knowledge of the users and the different query used for each case (a more generic for the informaticians). In any case further experimentation for the effectiveness of the algorithms by clinical experts is required in order to reach to safer

conclusions. The component that returned the most relevant results in any case is that of the pre-filtering.

Collaborative search evaluation

The Collaborative search was evaluated internally with the use of the ratings of the users due to low participation in the experiment. This evaluation produced valuable conclusions for the further development of the algorithm. The first finding was that the standard deviation code used in the previous version was not accurate and was replaced by a correct one. A second finding was that as seen in Table 2 the document with PMID 3167232 was ranked higher than before since it received consistently high ratings for the “breast cancer” query (6x4 stars and 1x5 star). For the rest documents the deviation of the ratings was high enough and/or the scoring differences with the next in rank document was so large that did not affect the ranking. Another finding is that the pagination mechanism contained a bug and some of the results were not presented in first place. Lastly, the Sum of scores divided by the sum of similarities in the equation 4 of the section 3.5.2, need to be replaced by a more effective one. The document with PMID 9250077 received a highly negative score. This finding resulted by an entry in the log that recorded the free evaluation of the tool by a user who added an extra term to the query. This was resulted due to the high relative frequency of the document into the slightly altered tag set which was negative-1 star and minimized the influence of 13 ratings with average score of 2.33 stars.

Table 1 Score of document with the basic search with query "Breast cancer"

Ranking	PMID	Score
1	9250077	27.51
2	14574169	26.82
3	16630739	26.81
4	3167232	26.76
5	15711624	26.71
6	22990110	26.39
7	21844186	26.38
8	19996029	26.38
9	15842935	26.34
10	7954390	26.29

Table 2 New ranking after the collaborative ranking algorithm is enabled

New Ranking	Old Ranking	PMID	Old Score	I-spy Score	New Score
1	4	3167232	26.76	0.29	27.05
2	2	14574169	26.82	0.02	26.84
3	3	16630739	26.80	-0.06	26.74
4	5	15711624	26.71	-0.29	26.42
5	6	22990110	26.39	0.02	26.41
6	7	21844186	26.38	-0.05	26.33
7	11	1787782	26.29	N/A	26.29
8	12	16825844	26.27	N/A	26.27
9	13	23464700	26.24	N/A	26.24
10	14	19034644	26.22	N/A	26.22

11	10	7954390	26.29	-0.08	26.21
14	9	15842935	26.34	-0.18	26.16
22	8	19996029	26.38	-0.48	25.90
N/A	1	9250077	27.51	-4.51	23.00

Usability evaluation

The following diagrams were produced from the usability evaluation form that the users were asked to fill after the experiment and the use of Hippocrates.

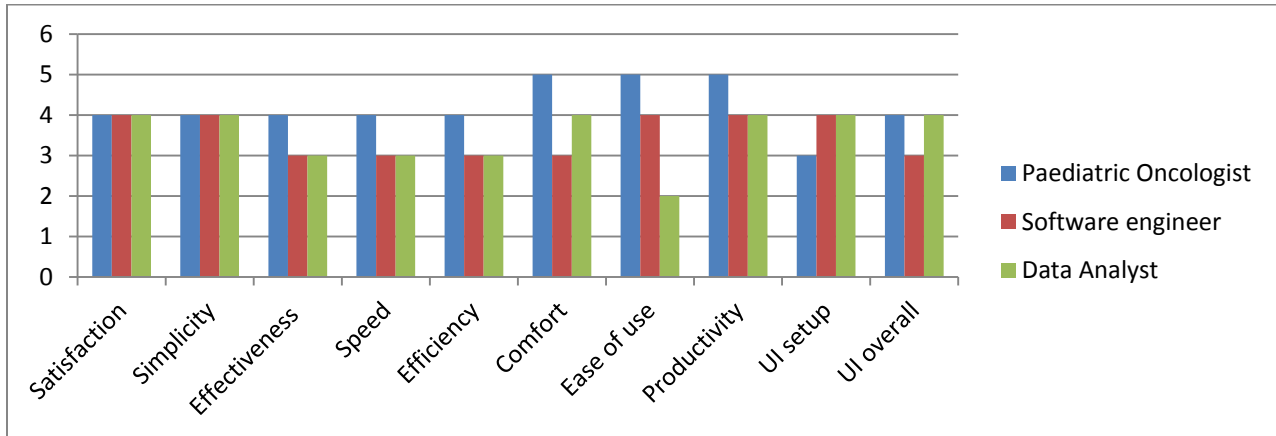


Figure 34 Overall usability evaluation of Hippocrates

In Figure 34 is shown that the tool received an above than average (1-5 rating) evaluation for most of the aspects measured and scored even higher by the domain expert. Notably, comfort, ease of use and productivity gain were perfectly scored by the Oncologist.

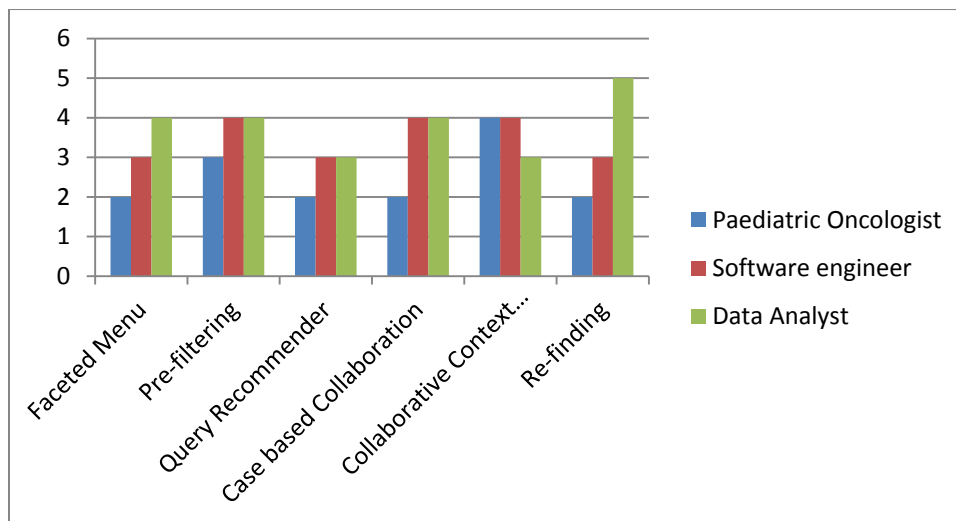


Figure 35 Modules evaluation

The evaluation of some basic components for the improvement of the searching process is presented in Figure 35. The evaluation results are average or above average for all the components. The numbers though should be seen along with the comments of the participants. On the Oncologists' evaluation the comments concerned limitations of the implementation and not the concepts and logic of the components per se e.g. for the Faceted Menu the comment was "you cannot select going back, or search all again to do a new refinement". For the query recommender and case based collaboration he suggested a larger scale validation in order to judge the effectiveness of those components.

The Faceted Menu according to the informaticians needs refocusing on components that would add more value to the end user such as making use of the tagging system and adding a tag-cloud component was a proposed alternative.

On the Collaborative Context aware recommender appeared a bug in the data analyst though he commented positively on the concept.

On the Positive side of the tool the Oncologist mentioned the ease of use. The Software Engineer marked as positive the two out of three contextualization approaches namely that of the query expansion mentioning that he liked the fact that he immediately noticed the results corresponded to a more refined information need by the inclusion of the diagnosis and that of the pre-filtering option. Fast response was another positive aspect of the tool. The re-finding component concluded the positive remarks on the tool. Re-finding was also mentioned by the Data Analyst as a positive aspect of Hippocrates. He also mentioned the importance of the different types of recommendations and the usefulness of the different searching approaches.

On the other side for negative remarks on the tool, the Oncologist included the fact that the search engine returned results from different type of cancer. Another recommendation was to permit full control to the user on which Journals will appear on the faceted menu. He missed the use of the MeSH vocabulary as an option to filter on the results and the lack of a searching component for the tagged words. The Software Engineer suggested the use of the tagging system to be fully functional as previously mentioned and to permit the user to select from a tag cloud and get the most positively rated documents from her rating or from the users' basis rating.

As general comments the overall improvement was suggested by the Data Analyst along with some UI improvements and more specifically on pointing more effectively the selected filters and the changes on the result set after the application of each filter. The Software Engineer suggested Emailing (implemented), saving (only when pdf is available) and exporting the search results. The Oncologist concluded to the final remark that Hippocrates is an interesting and easy and useful tool.

5 Conclusions and Future Work

In the previous chapter, the evaluation of Hippocrates for different aspects was presented. The conclusions of the thesis are presented in this chapter among with recommendations for future work.

5.1 Conclusions

There are strong indications that the concepts and the algorithms used, improve the searching process of the users. The contextualization concepts as shown in 4.3 improved the Cumulative Gain for the Oncologist and especially the Re-ranking and Pre-filtering approaches improved the precision as well. The Query expansion and the Re-ranking did not perform that well for the Informaticians who evaluated the tool but there are multiple reasons as aforementioned that this could happen (different background, query etc.) and in any case further evaluation is required.

The Pre-filtering component succeeded in improving Cumulative Gain and Precision @ the first results for all the participants. The component is differentiated to the related work presented in 2.6 by pre-filtering the documents not based on the MeSH vocabulary matching terms but in the total set of terms. Also it is differentiated by presenting the results, structured in fields including time of registration and extra information by the use of an external service.

The evaluation of the Collaborative Ranking algorithm based on the i-spy resulted into valuable conclusions. The algorithm could succeed in improving the searching process by boosting in the ranking, highly rated documents and placing in lower ranking documents rated low, by summing the score of the document and the score based on the ratings. A weakness of the equations proposed by the i-spy and used in Hippocrates is that unique ratings could affect the score more than multiple ratings. This issue should be addressed in future versions of the tool.

The overall usability evaluation of the tool was rated higher than average and the feedback from the Oncologist indicates that certain aspects of the tool are perfect such as; comfort, ease of use and in productivity gain.

For the rest of the components average ratings were received combined with valuable feedback for future improvement and some of them are mentioned in the next section. Overall Hippocrates could be considered as a framework on which different algorithms for contextualization, collaboration and other concepts could be built, tested and evaluated by the users, already including algorithms and concepts with indications of contributing to that direction. The small number of participants in the evaluation of the tool though, does not lead in statistically confident results about the effectiveness of the algorithms and concepts implemented.

5.2 Future work

There are a lot of directions that researchers could follow from this point. Some of them are already mentioned in the previous chapters. As a research project, Hippocrates was oriented into examining the effectiveness of different concepts and algorithms. If in the future will it be included in a Development project there are issues need to be addressed such as minor bugs fixes, a more sophisticated Graphical User Interface, and additional functionality as proposed in the feedback such as enabling the user to select multiple filters simultaneously. Other aspects include the security of the tool as in the present state only SQL injection protection is implemented. Integration with other tools is considered important and need to be addressed especially with the EURECA framework with which there is a certain level of integration in the ability to process CDM based EHRs.

On the present state, Hippocrates could be used in order to evaluate different IR functions and alternative algorithms for contextualization or concepts and algorithms to enable collaboration. Some examples for classical IR are:

- Adjusting k and b on the BM25 scoring in order to optimize the recall and relevance of the Search engine.
- Comparing the results with the BM25F scoring function.
- Implementing other state of the art scoring functions and comparing the results.
- Adjusting the minimum matching criteria to optimize recall and relevance.

For the context aware part of the system, further evaluation of the currently implemented algorithms should take place in order to statistically prove their effectiveness. Besides that, research could be done in optimizing different aspects with some examples being:

- Applying different weight of importance for each type of information (Patient's Information, Observation, Substance Administration, and Procedure).
- Applying a time based weighting function such as a logarithmically decreasing function.
- Test the approaches with real life Electronic Health Records that include patients' history and design and implement algorithms that would filter out irrelevant information and include in the searching process chronic conditions.
- Applying Personalization techniques through the inclusion of Machine learning and Data Mining algorithms on the log of a person or of a subspecialty. Decision trees and association rules techniques could be applied on the user logs and make use of the query history the click through rates and observe which terms are used for which diagnosis on which workflow state.
- Model the information needs of the doctor per workflow state and include it in the searching process making use of the input received described in 2.1.1.
- Instead of trying to meet information needs in sequence, a UI intervention could be used in order to parallelize the process by retrieving results for the most common aspects for one diagnosis in a Google now way.

- The use of the MeSH vocabulary was included in the evaluation by the Oncologist. A use of such a system should be considered.
- The diagnosis of the patient is critical and should be a pre-filter the results of the search engine for all the modules.

For the collaboration part the options for future work are unlimited as well. To start with, the Tagging system should be fully implemented and as suggested in the evaluation, should give the option to the user to search for documents tagged in an individual or in a set of users' basis. The i-spy based algorithm could be further developed or adjusted and combined with re-finding approaches. Different types of collaboration modules for synchronous or asynchronous collaboration could be tested for the present use case. Exporting the results of the search engine for sharing or for citation purposes besides the pdf download link is necessary.

6 Works Cited

- Adomavicius, G., & Tuzhilin, A. (2011). Context-Aware Recommender Systems. *Recommender Systems Handbook*, 217-253.
- Balfe, E., & Smyth, B. (2007). Case-Based Collaborative Web Search. *Case-Based Reasoning Research and Development*, Volume 4626 pp 329-343.
- Bates, M. J. (1989). THE DESIGN OF BROWSING AND BERRYPICKING TECHNIQUES FOR THE ONLINE SEARCH INTERFACE. *Online Information Review*, 407-424.
- Böhm, T., Klas, C.-P., & Hemmje, M. (2014). Collaborative Information Seeking in Professional Work-Settings:A Study of Equipment Utilization. *Springer*.
- Brown, B., & Jones, J. (2001). Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering. *Personal and Ubiquitous Computing*, 5:253–263.
- Cimino, J. J., Johnson, S. B., Aguirre, A., Roderer, N., & Clayton, P. (1993). The Medline Button. *AMIA*, 0195-4210/92.
- Fonseca, B. M., Golgher, P., Possas, B., Ribiera-Neto, B., & Ziviani, N. (2005). ConceptBased. *ACM*.
- Golovchinsky, G., Adcock, J., & Pickens, J. (2008). Cerchiamo: a collaborative exploratory search tool. *In Proceedings of Computer Supported. CSCW*.
- Hong, J.-y., Suh, E.-h., & Kim, S.-J. (2009). Context-aware systems: A literature review and classification. *Expert Systems with Applications*, 8509–8522.
- Huang Xiaoli, Lin Jimmy, & Demner-Fushman, D. (2006). Evaluation of PICO as a Knowledge Representation for Clinical Questions. *AMIA Symposium Proceedings*, 359.
- Jiang, D., Kenneth, L., & Wilfred, N. (2011). Context-Aware Search Personalization with Concept. *ACM*, 978-1-4503-0717-8/11/10.
- Jimenez-Castellanos, A., Fernandez, I., Perez-Rey, D., & Viejo, E. (2013). PATIENT-BASED LITERATURE RETRIEVAL AND INTEGRATION: A USE CASE FOR DIABETES AND ARTERIAL HYPERTENSION. *Communications in Computer and Information Science*, 312-323.
- Jones, K. S. (1972). A STATISTICAL INTERPRETATION OF TERM SPECIFICITY AND ITS APPLICATION IN RETRIEVAL. *Journal of Documentation*, Vol. 28 Iss: 1, pp.11 - 21.
- Kwon, Y. (2008). Improving top-n recommendation techniques using rating variance. *In Proceedings of the 2008 ACM conference on Recommender systems* (pp. 307-310). ACM.
- Manning, D. C., Raghavan, P., & Schütze, H. (2008). 6 - Scoring, term weighting, and the vector space model. *In Introduction to Information Retrieval* (pp. 100-123). Cambridge University Press.

- Mendonça, E., Cimino, J., & Johnson, S. (2001). Using narrative reports to support a digital library. *Proceedings AMIA symposium*, 458-62.
- Mendonça, E., Cimino, J., & Johnson, S. (2002). Using Patient Data to Rank Records of Literature Retrieval. *Proceeding of the AMIA symposium*, 1103.
- Morri, M. R. (2013). Collaborative Search Revisited. *ACM*, 978-1-4503-1331-5/13/02.
- Morris, M. R. (2008). A Survey of Collaborative Web Search Practices. *Proceedings of CHI*, 1657-1660.
- Noh, Y. (2013). A study on next-generation digital library using context-awareness technology. *LHT*, 31-2.
- Perez-Rey, D., Jimenez-Castellanos, A., Garcia-Remesa, M., Crespo, J., & Maojo, V. (2012). CDAPubMed: a browser extension to retrieve. *BMC Medical Informatics and Decision Making*.
- Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., & Back, M. (2008). Algorithmic mediation for collaborative exploratory search. *Proceedings of the Annual ACM*.
- Rezarta, I. D., Murray, C. G., Aurelie, N., & Zhiyong, L. (2009). Understanding PubMed user behavior through log analysis. *The Journal of Biological Databases & Curation*.
- Richarson, S. W., Wilson, M., Nishikawa, J., & Hayward, R. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP J Club*.
- Robertson, S. E. (1977). Theories and Models in Information Retrieval. *Journal of Documentation*, 126-148.
- Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed. *BMC Medical Informatics and*.
- Shah, C. (2012). *Collaborative Information Seeking*. Springer.
- Sieg, A., Mobasher, B., & Burke, R. (2007). Web Search Personalization with Ontological User Profiles. *ACM*, 978-1-59593-803-9/07/0011.
- SMYTH, B., BALFE, E., FREYNE, J., BRIGGS, P., COYLE, M., & BOYDELL, O. (2004). Exploiting Query Repetition and Regularity in an Adaptive Community-Based Web Search Engine. *User Modeling and User-Adapted Interaction*, 14: 383–423.
- Smyth, B., Coyle, M., & Briggs, P. (2011). On Communities, Collaboration, and Recommender Systems in Personalized Web Search. *Recommender Systems Handbook*, 579-614.
- Teevan, J. (2007). The Re:Search Engine: Simultaneous Support for Finding and Re-Finding. *Proceedings of the 20th annual ACM symposium on User interface software and technology*. *ACM* (pp. 23-32). *ACM*.

- Teevan, J., Adar, E., Jones R, & Potts, M. A. (2007). Information re-retrieval: repeat queries in Yahoo's logs. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM*, 151-158.
- Walker, S., Jones, K. S., & Robertson, E. S. (2000). A probabilistic model of information retrieval: Development and comparative experiments. *Information Processing and Management*, 36, 809-840.
- Xiang, B., Jiang, D., Pei, J., Sun, X., Chen, E., & Li, H. (2010). Context-Aware Ranking in Web Search. *ACM*, 978-1-60558-896-4/10/07.
- Yadav, N., & Poellabauer, C. (2012). An Architecture for Personalized Health Information. *ACM*.

Appendices

Appendix A Experiment Directions

Overview

Hippocrates was developed in order to test the effect of different algorithms in the searching process of an oncologist. The main focus is the use of the contextual information available from the Electronic Health Record of a patient and the collaboration among the doctors. It was developed in the context of a master's project within Philips Research.

The core of the system is a search engine that searches over 1.6 million documents from the Pubmed/Medline medical literature digital library.

Experiment

In order to test our hypothesis, relevance feedback from experts (Oncologists, residents or students) is necessary.

Hippocrates is in an early development phase. If you get warnings or notifications please refresh your current page. If the problem is persistent please skip your current step and proceed or contact me: georgios.aravanis@philips.com. Please use Chrome or Opera web browsers

Step 1: Main page

- Please visit the following address: <http://hippocrates.ehv.campus.philips.com/> and fill in the required fields:

entityId (patient):

fictitious3

entityId (user):

(Your name e.g. Georgios Aravanis)

Search method:

Basic Search

- And press Submit.

-Patient data:
[Adult](#)
[Female](#)

-Observations:
[Infiltrating duct carcinoma](#)
(2007-09-16 00:00:00)
[Breast](#)
[structure](#)
[left](#)
[with](#)
[laterality](#)

[Tumor](#)
(2007-09-16 00:00:00)
6mm

-Procedures:
[Surgical procedure](#)
(2007-09-16 00:00:00)

The selected data from the Electronic Health Record of the patient is presented in the figure above.

Step 2: Basic Search

Explanation: With the Basic Search, a reference base is established in order to compare the efficiency of the developed algorithms.

- Please fill in the Find field the following query:
[complete resection side effects](#)
and press Search.
- Rate the Relevance of the 2 first pages of results (20 results), by opening each reference clicking on the stars you think corresponds to the relevance of the query and the patient profile and pressing share.
- If you didn't find a perfect match (5stars) continue searching by changing the query or going to next pages for at most 2 more minutes. You do not need to rate any reference, only the most relevant one.

Step 3: Query Expansion

Explanation: This module expands the query as a background process with selected data from the patients' file.

- In the Search tab on the bottom right corner press the link "New case".
- Fill in the same data as in Step 1 except in the Search Method, insert:
[Query expansion](#)
- Repeat the same procedure as in Step 2.

Step 4: Re-ranking

Explanation: This module re-ranks the result with a score based on the matching of the terms of the query with each reference AND the terms of the patient data with each reference.

- In the Search tab on the bottom right corner press the link “New case”.
- Fill in the same data as in Step 1 except in the Search Method, insert:
Reranking
- Repeat the same procedure as in Step 2.

Step 5: Pre-filtering

Explanation: This module provides an extra filtering menu on the right side of the search tab based on the Electronic Health Record of the patient.

- In the Search tab on the bottom right corner press the link “New case”.
- Fill in the same data as in Step 1 except in the Search Method, insert:
Prefiltering
- Try different settings of the filters in the right side menu until you reach the optimal result set.

Rate the Relevance of the 2 first pages of results (20 results), by opening each reference clicking on the stars you think corresponds to the relevance of the query and the patient profile and pressing share.

Step 6: Collaborative search (For the last participant)

Explanation: This module re-ranks the result based on the frequency that a document was accessed and on the average rating received.

Repeat the Step 5 but in the first page click on the choice Collaborative reranking.

Step 6: Evaluation (Step 7 for the last participant)

Please try without inserting a new case, all the features of the system i.e. the rest 2 tabs of Recommender and Re-finding. Try other queries as well until you have a clear picture of the tool. Next you are kindly requested to fill the following evaluation form:

https://docs.google.com/forms/d/1La0p0U3LrtuWYS2uKAbXQ06eHxwX0LW9_gUNXWp2BW4/viewform

Thank you for your participation!

Appendix B Evaluation form

The evaluation form is accessible through the following link:

https://docs.google.com/forms/d/1La0p0U3LrtuWYS2uKAbXQ06eHxwX0LW9_gUNXWp2BW4/viewform

Usability evaluation of Hippocrates

* Required

Name/Surname *

Expertise *

1) Overall, I am satisfied with how easy it is to use this system *

1 2 3 4 5

Strongly disagree Strongly agree

2) It was simple to use this system *

1 2 3 4 5

Strongly disagree Strongly agree

3) I can effectively complete my work using this system *

1 2 3 4 5

Strongly disagree Strongly agree

4) I am able to complete my work quickly using this system *

1 2 3 4 5

Strongly disagree Strongly agree

5) I am able to efficiently complete my work using this system *

1 2 3 4 5

Strongly disagree Strongly agree

6) I feel comfortable using this system *

1 2 3 4 5

7) It was easy to learn to use this system *

1 2 3 4 5

Strongly disagree Strongly agree

8) I believe I became productive quickly using this system *

1 2 3 4 5

Strongly disagree Strongly agree

9) The organization of information on the system screens is clear *

1 2 3 4 5


Strongly disagree Strongly agree

10) I like using the interface of this system *

1 2 3 4 5

Strongly disagree Strongly agree

Continue »

 33% completed

Faceted search

Year:

[1950 to 1960 \[0\]](#)
[1960 to 1970 \[0\]](#)
[1970 to 1980 \[26\]](#)
[1980 to 1990 \[140\]](#)
[1990 to 2000 \[300\]](#)
[2000 to 2010 \[703\]](#)
[2010 to 2020 \[228\]](#)

Publication Type:

[Journal Article \[1612\]](#)
[English Abstract \[400\]](#)
[Case Reports \[228\]](#)
[Review \[188\]](#)
[Comparative Study \[169\]](#)

Journal Title:

[Hepato-gastroenterology \[54\]](#)

[Cancer \[47\]](#)

[Annals of surgery \[37\]](#)

[The Journal of urology \[33\]](#)

[World journal of surgery \[29\]](#)

[Surgical endoscopy \[26\]](#)

[Annals of surgical oncology \[21\]](#)

[Archivos españoles de urología \[20\]](#)

[Journal of hepato-biliary-pancreatic surgery \[20\]](#)

[Journal of surgical oncology \[20\]](#)

The faceted search contributed significantly in the searching process *

1 2 3 4 5

Strongly disagree Strongly agree

Comment

EHR based filtering

Filter out:

-Patient data:
Adult
Female

-Observations:
Infiltrating duct carcinoma
(2006-02-02 00:00:00)
Breast
structure
left
with
laterality

Tumor
(2006-02-02 00:00:00)
52mm

-Procedures:
Surgical procedure
(2006-02-02 00:00:00)

-Substances:
Aclarubicin
(0000-00-00 00:00:00)

Anthracycline
I_ Aclarubicin
I_ Daunorubicin
I_ Doxorubicin
I_ Epirubicin
I_ Idarubicin
I_ Valrubicin

Morphine
(0000-00-00 00:00:00)

The Electronic Health Record based filtering contributed significantly in the searching process *

1 2 3 4 5

Strongly disagree Strongly agree

Comment

Collaborative query recommendation

Find: Sort by: relevance ▾

Your team has searched for:

[infiltrating duct carcinoma](#)
[Duct carcinoma](#)
[infiltrating](#)
[ductal breast cancer](#)
[cancer](#)

The query recommender improves the awareness of the search task among the members of the team *

1 2 3 4 5

Strongly disagree Strongly agree

Comment

Case-based collaboration module

Search Recommender Refining

For the current case of fictitious1 your team shared:

[C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 12 days ago by: user1 on state: diagnosis

★★★★★

Tags: Infiltrating duct carcinoma

[C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 12 days ago by: user1 on state: diagnosis

★★★★★

Tags: Infiltrating duct carcinoma

[\[Infiltrating duct breast carcinoma: the role of estradiol and progesterone receptors\]](#), 12 days ago by: user1 on state: diagnosis

★★★★★

Tags: Infiltrating duct carcinoma

[\[Mitotic regimen of histologic variants infiltrating duct carcinoma of the breast\]](#), 13 days ago by: user1 on state: diagnosis

★★★★★

Tags: Infiltrating duct carcinoma

[C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 13 days ago by: user1 on state: diagnosis

★★★★★

Tags: Infiltrating duct carcinoma

The case-based group search history increased awareness among the members of the team *

1 2 3 4 5

Strongly disagree Strongly agree

Comment

Refinding

Search Recommended **Refinding**

Your recently accessed documents:

- [C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 21 days ago
★★★★★
Tags: Infiltrating duct carcinoma
- [C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 21 days ago
★★★★★
Tags: Infiltrating duct carcinoma
- [\[Infiltrating duct breast carcinoma: the role of estradiol and progesterone receptors\]](#), 21 days ago
★★★★★
Tags: Infiltrating duct carcinoma
- [\[Molecular regimen of histologic variants infiltrating duct carcinoma of the breast\]](#), 22 days ago
★★★★★
Tags: Infiltrating duct carcinoma
- [C-erbB-2 onco-protein expression in breast cancer: relationship to tumour characteristics and short-term survival in Universiti Kebangsaan Malaysia Medical Centre](#), 22 days ago
★★★★★
Tags: Infiltrating duct carcinoma

Your recent searches:

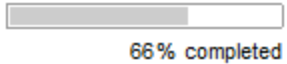
- [Infiltrating duct carcinoma](#) 22 minutes ago
- [Infiltrating duct carcinoma](#) 6 days ago
- [Infiltrating](#) 9 days ago
- [Infiltrating duct carcinoma](#) 9 days ago
- [Infiltrating](#) 10 days ago
- [Infiltrating](#) 10 days ago
- [Infiltrating duct carcinoma](#) 10 days ago
- [Infiltrating duct carcinoma](#) 10 days ago
- [Infiltrating duct carcinoma](#) 10 days ago
- [duct](#) 12 days ago

The refinding modules contributed significantly into the search process *

1 2 3 4 5

Strongly disagree Strongly agree

Comment



Usability evaluation of Hippocrates

* Required

List the most positive aspect(s) of the tool *

List the most negative aspect(s) of the tool *

Request for new features-improvements *

What did we miss? What would you like to see in a next version of the tool?

Comments

« Back

Submit

Never submit passwords through Google Forms.

100%: You made it.

Appendix C Evaluation results

Timestamp	4/12/2014 10:47:30	5/27/2014 17:51:54	6/1/2014 14:50:57
Name/Surname	Norbert Graf	Ahmed Ibrahim	Kosmas Hatzidimitris
Expertise	Pediatric Oncologist	Software engineering	Data Analyst
1) Overall, I am satisfied with how easy it is to use this system	4	4	4
2) It was simple to use this system	4	4	4
3) I can effectively complete my work using this system	4	3	3
4) I am able to complete my work quickly using this system	4	3	3
5) I am able to efficiently complete my work using this system	4	3	3
6) I feel comfortable using this system	5	3	4
7) It was easy to learn to use this system	5	4	2
8) I believe I became productive quickly using this system	5	4	4
9) The organization of information on the system screens is clear	3	4	4
10) I like using the interface of this system	4	3	4
The faceted search contributed significantly in the searching process	2	3	4
Comment	you cannot select going back, or search all again to do a new refinement	I would use it only when I have too many results. I would not use it to access publications regarding year or publication type	Journal title helps when you recognize the shown categories
The Electronic Health Record based filtering contributed significantly in the searching process	3	4	4
Comment	you cannot go back to select all	I would definitely use this feature since it allows me to filter on	

		patient characteristics, which I am interested in	
The query recommender improves the awareness of the search task among the members of the team	2	3	3
Comment	This needs to be validated before one can answer this question		I had expected that the queries would change if I change my search query.
The case-based group search history increased awareness among the members of the team	2	4	4
Comment	This needs to be validated before one can answer this question		
That module would contribute significantly in the search process	4	4	3
Comment	this is correct		Nothing appeared but as an idea yes it is good!
The refinding modules contributed significantly into the search process	2	3	5
Comment	one should be able to go back to all		
List the most positive aspect(s) of the tool	easy to use	<ul style="list-style-type: none"> - query expansion based on the basic search - filter out possibility - fast response of the search engine - recently accessed documents 	Recommendations, different ways to search etc. are needed and provided. Refindings is a great options as a lot of times you forget how you find something in the past..
List the most negative aspect(s) of the tool	despite searching for breast cancer other cancer types are displayed	<p>The "To which cases is this article relevant?" does not work properly or the meaning is not clear to me. E.g., I added the keyword "detection" to the following article "Breast cancer detection by daughters of women with breast cancer."</p> <p>So, what does this mean? Can I search for "detection"? Does the system add a new label</p>	It is difficult to understand what the difference of the search types is. Is it better to search for breast cancer with query expansion or do a basic search on Infiltrating duct carcinoma?! I think an oncologist needs a lot more explanation to understand and to use the system on the right way so he can find what he wants!

		"Detection" which contains e.g. the article above?	
Request for new features-improvements	<p>It is not clear how the journals are selected. This should contain all journals. I did miss e.g. the Journal of clinical oncology.</p> <p>To search with Mesh terms</p> <p>To store search terms to find new papers</p> <p>To select more than 1 item</p>	<ul style="list-style-type: none"> - email results - save results - export functionality of the results 	<p>The most important would be to get as good as possible the existing issues. For example when you are filtering out results it would be nice to be more clear what happens when I click on one of the filtering options and to be more clear which ones are selected. From the other side it is difficult to understand which of the 4 options Basic search, Reranking etc. should be used when?!</p>
Comments	<p>Interesting tool, easy to handle, seems quite useful</p>		