PONTE: A Context-Aware Approach for Automated Clinical Trial Protocol Design

George Tsatsaronis[1], Konstantinos Mourtzoukos[2], Vassiliki Andronikou[2], Tassos Tagaris[2], Iraklis Varlamis[3], Michael Schroeder[1], Theodora Varvarigou[2], Dimitris Koutsouris[2], Nikolaos Matskanis[4]

[1]Biotechnology Center, TU Dresden
01307, Dresden
Germany
{george.tsatsaronis,ms}@biotec.tu-dresden.de

[2]National Technical University of Athens
15780, Athens
Greece
{kmour,vandro}@mail.ntua.gr
tassos@biomed.ntua.gr
dora@telecom.ntua.gr
dkoutsou@biomed.ntua.gr

[3]Harokopio University of Athens
17671, Athens
Greece
varlamis@hua.gr

[4]Centre of Excellence in Information and Communication Technologies
CETIC
B-6041 Charleroi
Belgium
nikolaos.matskanis@cetic.be

ABSTRACT

The rapidly increasing volume of published clinical and non-clinical data at a variety of sources and the resulting great effort required for researchers to access them and mine information of interest lead to clinical trials that are based on only a limited set of knowledge in the domain they cover. This restricted view of the clinical trials' context is quite often the reason behind unsuccessful trials and/or successful ones which, however, underestimate drugs' unwanted effects and thus their results are of low external validity in the much more complicated environment of clinical healthcare. In this paper, we present a context-aware approach, which has been developed in the PONTE project, for effectively guiding medical researchers during clinical trial protocol design and allowing for more efficient and effective access to scientific literature. The suggested approach incorporates intelligent services and advanced text mining mechanisms for scientific literature querying and mining during protocol design, taking into account the study context (i.e. active substance, target and disease) and the domain context in literature.

1. INTRODUCTION

Among the key aims of clinical research is the investigation of the therapeutic potential of substances, methodologies and devices and their transformation and development into real-world therapies. Clinical trials comprise a critical step in this process, focusing on the investigation of the efficacy and safety of these candidate treatments initially on animal models and eventually on humans. Given the great impact such research has on the world population and the high investment it requires - with recently reported figures indicating that the average cost may even reach € 9 billion [11] per drug approved - much debate has been held over the years related to the effectiveness and the efficiency of the processes followed in clinical research. In the meanwhile, the therapy development timeline forces patients seeking for therapies to be on the wait for about 11 years after the initial discovery of the potential therapy. And still, figures in drug development demonstrate that of every 5000 molecules which are pre-clinically tested, only 1 will in the end be approved and will enter the market [15]. A great challenge which researchers face within this strongly competitive and complicated environment is access to the continuously rising volume of data and information in life sciences. The informational sources, which are highly important for their research design and implementation are numerous, sharing various formats, structures and levels of granularity. As a result the task of accessing and mining information of interest from them requires, quite often, unmanageable effort and excessive time. The latter results in important pieces of scientific information being hidden and thus not taken into consideration with minor and/or major consequences in the research conducted and its potential for real-world application.

This paper presents the architecture and functionality of the PONTE project platform[1]. PONTE is a knowledge-oriented platform, which provides a set of mechanisms and services facilitating the specification of the "test of hypothesis" on a scientifically-valid basis and the intelligent design of clinical trials. Towards this direction, the platform incorporates a set of advanced data mining and semantic reasoning mechanisms which are applied on a variety of web data sources containing clinical and non-clinical information. The two main components of the PONTE system, which incorporate these mechanisms are the Decision Support (DS) component, and the GoPONTE semantic search engine[2]. Their detailed description is provided in the sections that follow. More specifically, in section 2, the overall architecture as well as the design of the individual components of the proposed approach is presented. Section 3 demonstrates the data flow in the system through a real-world scenario and section 4 presents the related work in the field.

2. A Context-Aware Architecture for Clinical Trial Protocol Design

2.1 Overall Architecture Description

In our approach, *clinical trial design context* has two different, yet related, aspects:

- *study-driven*: the specific trial's main parameters (i.e., study disorder, investigational active substance, target), which are fed by the researcher,
- *domain-driven*: the research environment within which the study is held, i.e. the scientific (non-clinical and

[1] http://www.ponte-project.eu/

[2] Publicly available at: http://www.gopubmed.org/web/goponte/

clinical) findings in the specific domain of the study, which are automatically retrieved by the presented system.

As mentioned above, PONTE aims, among others, at providing a SOA (Service Oriented Architecture) -based platform for facilitating clinical trial design through a set of novel decision support services and efficient semantic searching across literature. The following figure presents the part of the PONTE system that provides the clinical trial researchers with context-aware scientific query generation and semantic filtering of the retrieved results:

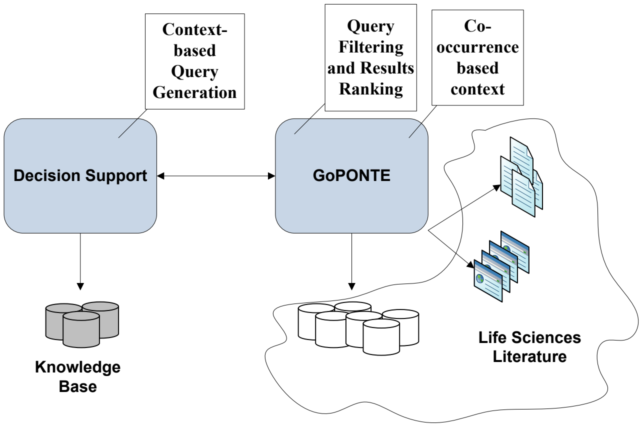

Context-based Query Generation
Query Filtering and Results Ranking
Co-occurrence based context
Decision Support
GoPONTE
Knowledge Base
Life Sciences Literature

Figure 1: Snapshot of the PONTE Architecture

As mentioned above, the two main building blocks of the PONTE system which participate in the provision of the aforementioned functionalities are the *Decision Support (DS) component* and the *GoPONTE* search egnine. The Decision Support component is responsible for generating the queries which are proposed to the researcher designing the clinical trial as highly relevant to the latter and thus their retrieved results could, potentially, be of great value to the design process. The GoPONTE comprises the semantic search engine of the PONTE platform. It acts as the interface to the literature resources on which the queries will be performed and it enhances the querying process through semantic annotation and filtering of the results. The latter is based on the incorporated ontologies which include MeSH [17], GO [8] and UniProt [20]. It applies state of the art annotation based on a set of ontologies it incorporates through the use of Maximum Entropy models for classification purposes [7], trained for all of the underlying ontology concepts. In the following paragraphs, these two key components are presented.

2.2 Decision Support

The main task of the decision support component is to facilitate the researchers' work during clinical trial design by guiding, complementing and enriching their access to literature. More specifically, for each clinical trial under design it generates a series of focused scientific questions which are highly related to this specific trial. These scientific questions are semantically linked to the different sections and parameters of the clinical trial protocol (such as study duration, inclusion and exclusion criteria, treatment schedule and disorder background). As mentioned above, each clinical trial has a different context, which is determined and bounded based on (i) its main parameters (investigational active substance, study disorder, target) and (ii) the scientific findings in the study's domain. The following figure presents the architecture of the decision support component and its interaction with the user and the GoPONTE search engine.

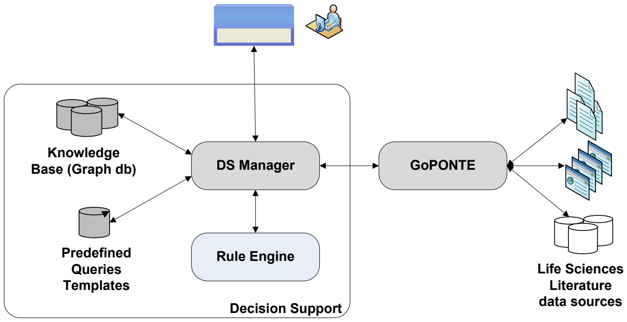

Knowledge Base (Graph db)
DS Manager
GoPONTE
Predefined Queries Templates
Rule Engine
Decision Support
Life Sciences Literature data sources

Figure 2: Decision Support Architecture

The DS Manager comprises the interface of the Decision Support to the user interface and access to literature. It is also the component responsible for initiating and managing the generation of the scientific questions. It receives the main study parameters set by the user and it communicates with the Predefined queries templates database. These templates are abstractions of questions represented both by a human readable description and by a set of concepts from the MeSH, GO and UniProt ontologies. The answers to these questions are potentially of great value for a researcher while specifying the parameters of a clinical trial. For example, while the researcher is working on the study disorder background, linking to literature in terms of the epidemiology of the disease as well as its pathophysiology and its prevalence and incidence in the population is really useful.

An additional set of scientific questions is *dynamically* generated by the DS based on the main study parameters through the application of a set of rules on the knowledge base it stores locally which captures domain knowledge. This knowledge base is a graph database which has been created based on a series of Linked Data sources [16], including DrugBank [4], Diseasome [3], SIDER [19], and KEGG [14]. The rules behind the generation of the questions are linked with the sections and parameters of the clinical trial protocol. Hence, a rule might focus on the interaction between the study drug and any market drug. This rule might be linked with the inclusion/exclusion criteria, the drug background and the adverse events sections in the protocol. When the main study parameters are fed in the DS, for each section the related set of rules is applied on the knowledge base and if triggered the respective questions are generated. An example of such a question generation is presented in section 3.

The set of generated questions – both predefined and customized ones – is presented to the user in their human readable form (such as "What is the epidemiology of myocardial infarction in Italy"). When the user selects this question for retrieving the respective results in literature, the DS sends to the GoPONTE the series of ontology concepts related to this question.

2.3 GoPONTE Semantic Search Engine

The usage of ontologies allows document and query processing at the semantic level [2], which can improve search on the biomedical domain [9]. Since the biomedical literature grows annually by approximately 500000 new citations, the task of annotating them with ontology concepts in real time is very difficult. In PONTE we have developed an Ontology Based

Search Engine (OBSE), namely GoPONTE, for searching efficiently the biomedical literature and the Web. The purpose of GoPONTE is dual: (i) to support the semantic annotation and search of Web documents with the aim to provide useful results for the clinical trial protocol designers, (ii) to provide a Web service to the Decision Support component, in order to extract the most related concepts given an initial input query set of ontology concepts. In Figure 3 the GoPONTE architecture is presented.

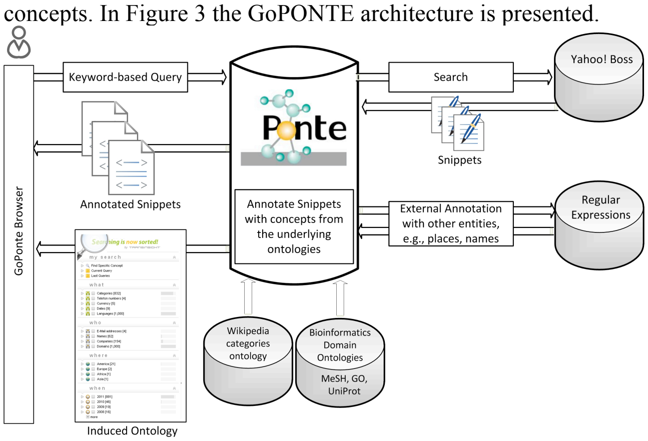


Figure 3: GoPONTE Architecture

GoPONTE offers the abilities of a general search but also classifies and annotates in real time free text with the underlying ontology concepts. Though there are other systems as well that may perform annotation of unrestricted (free) text with biomedical ontology concepts, the Maximum Entropy approach used has been tested against other state of the art techniques in large scale hierarchical classification tasks (PubMed docs with MeSH concepts) and has been show to provide very high accuracy [7]. In Figure 2 we show the overall architecture of GoPONTE. The query may be keywords and/or concepts from the underlying ontologies (auto-filling capabilities are offered). The engine also identifies entities like organizations, and dates, for the purpose of which there exists an in-house repository of regular expressions (rule-based approach). Once a query is submitted, the engine uses the Yahoo! Boss API in order to fetch the most relatedWeb results. The retrieved snippets are annotated in real time with concepts from the aforementioned ontologies, as well as names, places and dates. The snippets' annotation with the concepts uses a Maximum Entropy-based model trained for each of the concepts. However, since the underlying concepts are in the order of magnitude of 150; 000, the text of each snippet is filtered and only a few candidate concepts that are most related with the snippet context are identified. For the filtering, we are using a string matching algorithm based on the GoPubMed engine [2], which utilizes a Radix Tree [10]. As a final step, the maximum entropy models for the respective concepts are applied to each of the snippets. The overall response time is in the order of magnitude of 1 - 2 seconds.

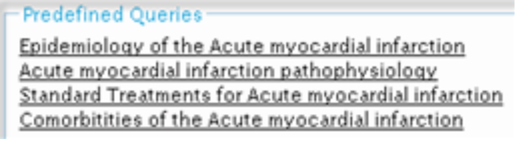
3. Data Flow through a Real-World Scenario

In this section the data flow in the presented system is demonstrated through a real-world scenario. For this purpose, the

Figures 6 and 7 demonstrate the next step of the process. In the presented case, the researcher has chosen to search biomedical literature for retrieving published work (such as articles, reports, surveys and so on) for the Epidemiology of Myocardial Infarction.

main study parameters used include *amiodarone* as the investigational active substance and *acute myocardial infarction* as the study disorder. These two parameters set the study-driven context.

When the DS component receives these parameters during trial design, it instantiates the predefined questions templates it retrieves from the respective database for this particular context. At this point, a first set of *study-driven context-based scientific questions* is generated for each clinical trial protocol. In Figure 4 a list of the questions built for the protocol dedicated to the Background of the study disorder is presented. These questions are highly related to the information expected to be provided in this section of the clinical trial protocol, such as the disorder epidemiology, its pathophysiology, any standard treatments used as well as its comorbidities. Each one of these questions encapsulates the relevant ontology concepts, which in turn are sent to the GoPONTE search engine for retrieving literature findings as soon as the researcher clicks on one of them.


Predefined Queries

Epidemiology of the Acute myocardial infarction
Acute myocardial infarction pathophysiology
Standard Treatments for Acute myocardial infarction
Comorbitities of the Acute myocardial infarction

Figure 4: Example of Predefined Questions for Acute Myocardial Infarction related to the Background of the study disorder section

In the meantime, the DS applies these main study parameters and through the rule engine exploits the graph-based knowledge base for extracting additional scientific questions. This resulting set of *domain-driven context-based scientific questions* is also presented to the research for each protocol section based on their relevance to the section's parameters. In Figure 5, an example of such questions is presented. In this case, the researcher is working for specifying the background of the investigational active substance. Information expected in such a section includes the substance's pharmacokinetics (such as metabolizing gene, metabolite, etc) and pharmacodynamics (such as effective, lethal and toxic dose). Hence, the DS crawls through the knowledge base aiming at identifying, for example, genes and proteins which are highly related to the investigational active substance. Each one of these DS-generated couples is presented to the researcher, who can use the underlying links to GoPONTE (which are annotated by DS with the respective ontology concepts) for instantly accessing literature.

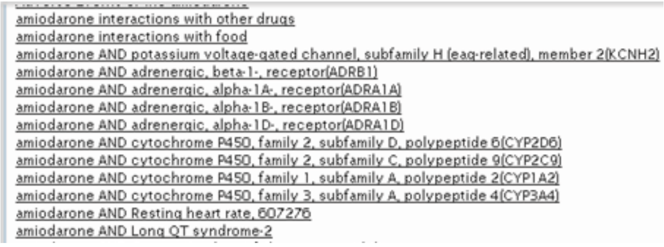
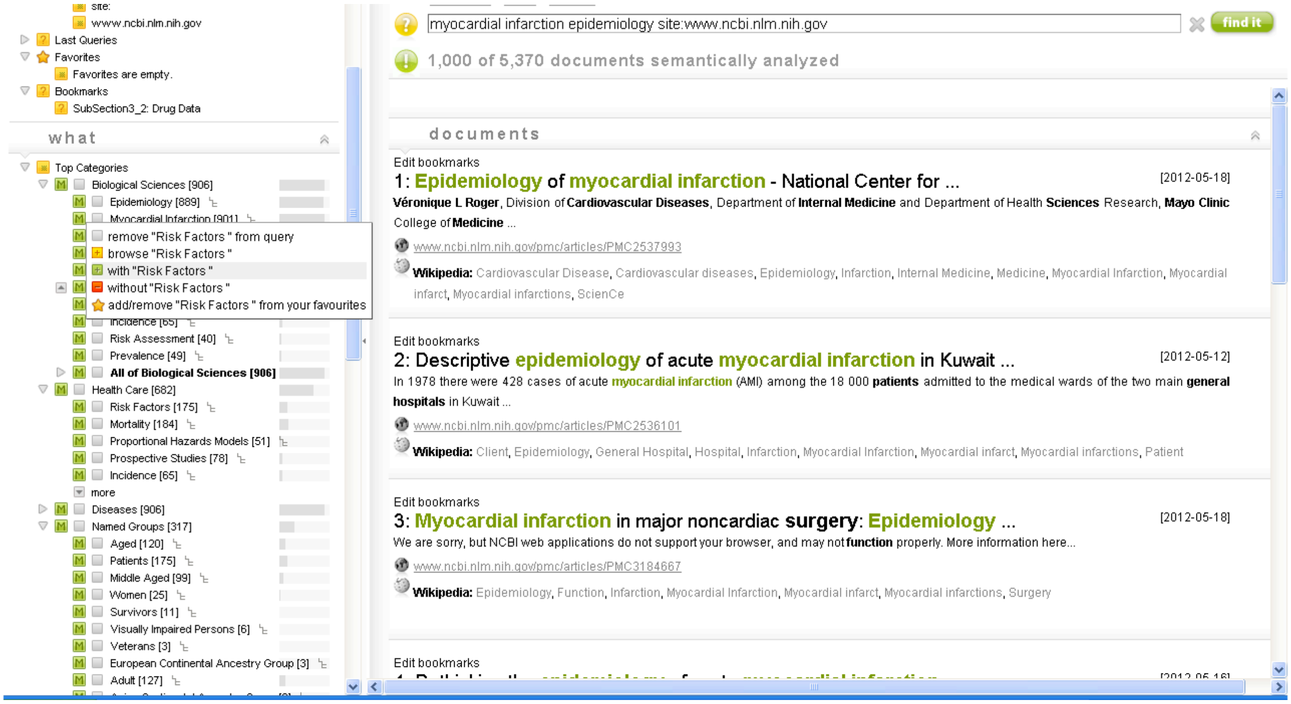

amiodarone interactions with other drugs
amiodarone interactions with food
amiodarone AND potassium voltage-gated channel, subfamily H (eag-related), member 2(KCNH2)
amiodarone AND adrenergic, beta-1-, receptor(ADRB1)
amiodarone AND adrenergic, alpha-1A-, receptor(ADRA1A)
amiodarone AND adrenergic, alpha-1B-, receptor(ADRA1B)
amiodarone AND adrenergic, alpha-1D-, receptor(ADRA1D)
amiodarone AND cytochrome P450, family 2, subfamily D, polypeptide 6(CYP2D6)
amiodarone AND cytochrome P450, family 2, subfamily C, polypeptide 9(CYP2C9)
amiodarone AND cytochrome P450, family 1, subfamily A, polypeptide 2(CYP1A2)
amiodarone AND cytochrome P450, family 3, subfamily A, polypeptide 4(CYP3A4)
amiodarone AND Resting heart rate, 607276
amiodarone AND Long QT syndrome-2

Figure 5: Example of Dynamically-generated Domain-based Questions for Amiodarone related to the Background of the drug

Hence, the user is directed to the GoPONTE interface in which the ontology concepts related to this question are fed automatically by the DS (Figure 6). In the left part of the figure the ontology concepts with which the retrieved results have been annotated are presented. For example, among the 1000 retrieved results, 184 include the *Mortality* concept and 65 the *Incidence* one. Next to each such ontology concept the number of results within which it appears is included. The right part of the interface includes the question (on the top) and the annotated retrieved results.

The semantic filtering functionality of the GoPONTE is demonstrated in Figure 7. The user has clicked on the "Risk Factors" concept at the ontology menu and selected to filter the retrieved results with this one. The initially retrieved 1000 results have now been reduced to 164, including the ones that include epidemiology data on myocardial infarction which include information about the risk factors of the disorder.



myocardial infarction epidemiology site:www.ncbi.nlm.nih.gov

1,000 of 5,370 documents semantically analyzed

Edit bookmarks

1: **Epidemiology** of **myocardial infarction** - National Center for ... [2012-05-18]

Véronique L Roger, Division of **Cardiovascular Diseases**, Department of **Internal Medicine** and Department of Health **Sciences** Research, **Mayo Clinic** College of **Medicine** ...

www.ncbi.nlm.nih.gov/pmc/articles/PMC2537993

Wikipedia: Cardiovascular Disease, Cardiovascular diseases, Epidemiology, Infarction, Internal Medicine, Medicine, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, ScienCe

Edit bookmarks

2: Descriptive **epidemiology** of acute **myocardial infarction** in Kuwait ... [2012-05-12]

In 1978 there were 428 cases of acute **myocardial infarction** (AMI) among the 18 000 **patients** admitted to the medical wards of the two main **general hospitals** in Kuwait ...

www.ncbi.nlm.nih.gov/pmc/articles/PMC2536101

Wikipedia: Client, Epidemiology, General Hospital, Hospital, Infarction, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, Patient

Edit bookmarks

3: **Myocardial infarction** in major noncardiac **surgery**: **Epidemiology** ... [2012-05-18]

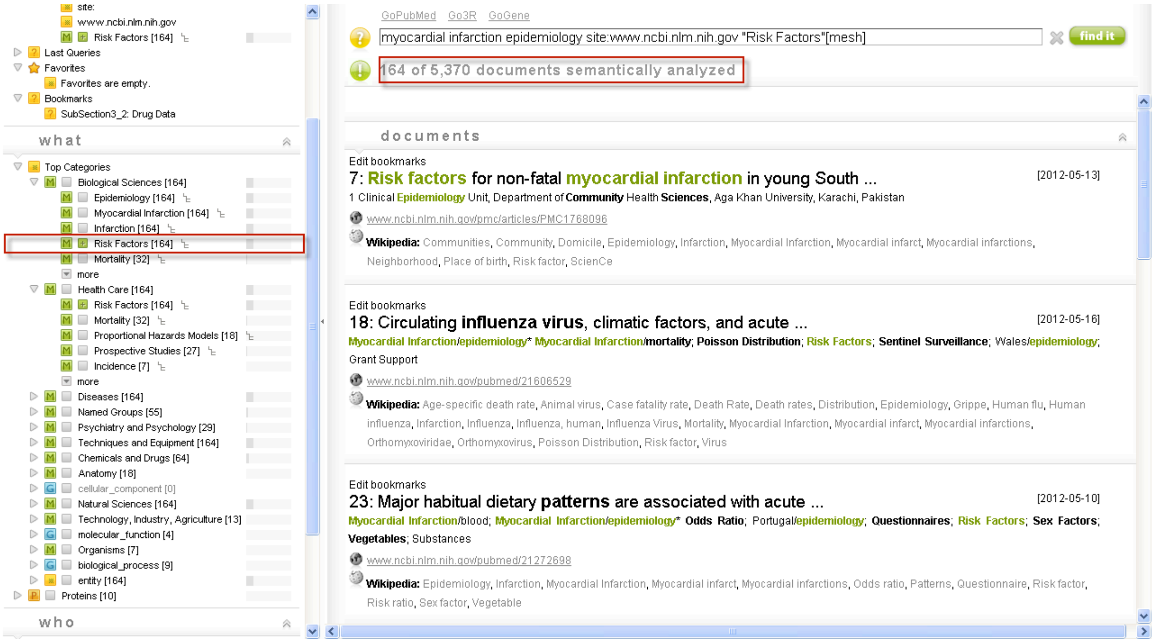
We are sorry, but NCBI web applications do not support your browser, and may not **function** properly. More information here...

www.ncbi.nlm.nih.gov/pmc/articles/PMC3184667

Wikipedia: Epidemiology, Function, Infarction, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, Surgery

Edit bookmarks

Figure 6: Submission of the first query in the GoPONTE



myocardial infarction epidemiology site:www.ncbi.nlm.nih.gov "Risk Factors"[mesh]

164 of 5,370 documents semantically analyzed

Edit bookmarks

7: **Risk factors** for non-fatal **myocardial infarction** in young South ... [2012-05-13]

1 Clinical **Epidemiology** Unit, Department of **Community** Health **Sciences**, Aga Khan University, Karachi, Pakistan

www.ncbi.nlm.nih.gov/pmc/articles/PMC1768096

Wikipedia: Communities, Community, Domicile, Epidemiology, Infarction, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, Neighborhood, Place of birth, Risk factor, ScienCe

Edit bookmarks

18: Circulating **influenza virus**, climatic factors, and acute ... [2012-05-16]

Myocardial Infarction/**epidemiology*** **Myocardial Infarction**/**mortality**; **Poisson Distribution**; **Risk Factors**; **Sentinel Surveillance**; Wales/**epidemiology**; Grant Support

www.ncbi.nlm.nih.gov/pubmed/21606529

Wikipedia: Age-specific death rate, Animal virus, Case fatality rate, Death Rate, Death rates, Distribution, Epidemiology, Grippe, Human flu, Human influenza, Infarction, Influenza, Influenza, human, Influenza Virus, Mortality, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, Orthomyxoviridae, Orthomyxovirus, Poisson Distribution, Risk factor, Virus

Edit bookmarks

23: Major habitual dietary **patterns** are associated with acute ... [2012-05-10]

Myocardial Infarction/blood; **Myocardial Infarction**/**epidemiology*** **Odds Ratio**; Portugal/**epidemiology**; **Questionnaires**; **Risk Factors**; **Sex Factors**; **Vegetables**; Substances

www.ncbi.nlm.nih.gov/pubmed/21272698

Wikipedia: Epidemiology, Infarction, Myocardial Infarction, Myocardial infarct, Myocardial infarctions, Odds ratio, Patterns, Questionnaire, Risk factor, Risk ratio, Sex factor, Vegetable

Figure 7: Example of the Semantic Filtering of the query results in the GoPONTE

4. Related Work

Bioinformatics and medical informatics are interesting and rich domains for context awareness research and applications. In the majority of the cases, context awareness in these two domains is defined by parameters such as location, time and role. In the general case, context-awareness in the framework of healthcare can be classified according to the aspects considered as context [18]. From this perspective, three main classes can split the items of context (C) into: (Ci) people, (Cii) environment, and, (Ciii) activities. From another perspective, also analyzed by Dey et al.[1], context-aware methods and applications may also be classified according to the purpose of usage of the considered context (U) into: (Ui) presentation of information and services to a user, (Uii) execution of a service, and, (Uiii) tagging of context to information for later retrieval. The majority of the context-aware methods in the examined domain fall under Ci and Cii following the former categorization, and Ui according the latter. In the following we give some examples of previous works regarding context-aware methods in bioinformatics and medical informatics that cover some or all of the aforementioned categorizations, and we explain how the current work differs and how is categorized with regards to classifications C and U.

In [21], Stanford describes the Vocera communication system, which is wearable badge developed as a hospital mobile device, that allows users to answer calls in a hand-free fashion. The device supports speaker verification, and delivers the information directly to the users who wear it, saving them from the effort to visit the nearest device (phone or PC). From the usage point of view, Vocera belongs to Ui, while the context used is the professionals' ids and roles, as well as their location, thus Ci and Cii.

Bardram [12] describes a context-aware prototype for the hospitals of the future, in which context-aware beds are considered. The hospital beds of the prototype have a built-in display for patients' entertainment, which can also be used by clinicians to access the patients' medical data. The bed is aware of who is using it, as well as who is near it, and, thus, displays relevant information according to this context, classifying the system into Ui, and Ci and Cii respectively.

In another work, Kjeldskov and Skoy [13] present MobileWARD, a prototype that supports the tasks of the hospital wards, and can display patient lists and patient information. The prototype considers the location of the nurse and the time of the day, simulating events linked to the location of the staff member, thus, classifying the approach, as before, in the categories Ui, Ci, and Cii.

Another work which combines more categories, i.e., Ci-iii, and Ui, is the prototype scenario presented by Swanson et al. [5], which describes a set-up in a medical setting that takes advantage of the context by combining the hierarchy of the environment (hospital room), persons (patients) an activities (doctor-patient status discussions).

Finally, towards the direction of supporting clinical decision making, the Health Level Seven International (HL7) Context-Aware Knowledge Retrieval Standard has been developed [11][6], which specifies how computerized information retrieval tools known as "infobuttons" can deliver contextually-relevant knowledge resources into clinical information systems. An example of how an infobutton works is the following: In the context of a patient's problem the infobutton displays directly educational material regarding the evaluation and treatment of a specific disease. The main advantage of infobuttons is that they allow for the integration of several different knowledge sources. In this work, we take a step further towards expanding the notion of the HL7 standard to the direction of supporting automatically the design of a clinical trial protocol. In our case the expansion is two-fold: Primarily, the context is not only limited to a disease, or symptoms, but rather to the triangle "drug-target-disease" which define the basis for the creation of hypothesis testing in the framework of a clinical trial; from this perspective, the current work expands the C categorization into a new axis, namely: domain of the study, with the aforementioned three items of the triangle. Secondarily, besides combining a plethora of domain knowledge bases, it combines additionally intelligent services, such as the decision support and the semantic search engine components described in Section 2. Overall, it is for the first time to the best of our knowledge that a context-aware approach for design of clinical trial protocols is suggested, following the trends and the adoption tendency of the HL7 standard.

5. Conclusions and Future Work

This paper presented a context-aware approach for the automatic generation of scientific questions and the semantic filtering of the retrieved results, which can facilitate researchers during clinical trial design. This approach, which has been developed within the PONTE project, involves the communication between two main components, i.e. the Decision Support and the GoPONTE. As indicated by the presented architecture and data flow demonstration, the first component is involved in the generation of the scientific questions based on the study- and domain-driven context. The second one applies these questions to literature and provides semantic filtering capabilities for the user to have an instant glance over research taking place in the field of their interest as well as digging out important information for their trial design. Our future work will involve further enrichment of the context through literature which will involve communication of filtered out information from GoPONTE to the DS. The latter will allow for further scientific questions generation driven by current trends in research rather than only established knowledge in the study domain. Moreover, the knowledge graph-based database will incorporate more data sources and relationships among their data. And what is more, further analysis of the clinical trial protocol and its informational requirements will be held. This analysis will be performed towards producing the respective mappings of these requirements into rules which will in turn result in enrichment of the rules at the DS side.

ACKNOWLEDGMENTS

This work is being supported by the PONTE project and has been partially funded by the European Commission's IST activity of the 7th Framework Programme under contract number 247945. This paper expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this paper.

6. REFERENCES

[1] A. Dey, G. Abowd, D. Salber. A conceptual framework and toolkit for supporting the rapid prototyping of context-aware applications. Hum. Comput. Interact. J. 16(2-4):97-166, 2001.

[2] A. Doms and M. Schroeder. Semantic search with GoPubMed. In REWERSE, pages 309–342, 2009.

[3] Diseasome, available at: http://diseasome.eu/

[4] DrugBank, available at: http://www.drugbank.ca/

[5] E. Swanson, A. Calvao, K. Sato. A framework for understanding contexts in interactive systems development. In Proceedings of the 7th World Multi-Conference on Systemics, Cybernetics and Informatics, 2003.

[6] G. Del Fiol, V. Huser, H.R. Strasberg, S.M. Maviglia, C. Curtis, J.J. Cimino. Implementations of the HL7 Context-aware Knowledge Retrieval ("Infobutton") Standard: Challenges, strengths, limitations, and uptake. Journal of Biomedical Informatics 2012.

[7] G. Tsatsaronis, N. Macari, S. Torge, H. Dietze, and M. Schroeder. A Maximum Entropy Approach for Accurate Document Annotation. Journal of Biomedical Semantics, 3(Suppl1), 2012.

[8] GO (Gene Ontology), available at: http://www.geneontology.org/

[9] H. Dietze and M. Schroeder. Goweb: a semantic search engine for the life science web. BMC Bioinformatics, 10(S-10):7, 2009.

[10] H. Dietze. GoWeb: Semantic Search and Browsing for the Life Sciences. PhD thesis, Technical University of Dresden, Germany, 2009.

[11] Herper, M. (2012) "The Truly Staggering Cost Of Inventing New Drugs", Forbes, 10th Feb 2012, available at: http://www.forbes.com/sites/matthewherper/2012/02/10/the-truly-staggering-cost-of-inventing-new-drugs/

[12] J. Bardram. Applications of context-aware computing in hospital work – examples and design principles. In Proceedings of the ACM Symposium on Applied Computing, pp. 1574 – 1579, 2004.

[13] J. Kjeldskov, M. Skov. Supporting work activities in healthcare by mobile electronic patient records. In Proceedings of the 6th Asia-Pacific Conference on Human-Computer Interaction, 2004.

[14] KEGG: Kyoto Encyclopedia of Genes and Genomes, available at: http://www.genome.jp/kegg/

[15] Kraljevic, S., Stambrook, P. J. and Pavelic, K. (2004). Accelerating drug discovery, EMBO reports (2004) 5, 837 - 842 doi:10.1038/sj.embor.7400236

[16] Linked Data, available at: http://linkeddata.org/

[17] MeSH, available at: http://bioportal.bioontology.org/ontologies/3019

[18] N. Bricon-Souf, C. R. Newman. Context awareness in health care: A review. International Journal of Medical Informatics, 76(1):2 – 12, 2007.

[19] SIDER, available at: http://sideeffects.embl.de/

[20] UniProt, available at: http://www.uniprot.org/

[21] V. Stanford. Beam me Up, doctor McCoy. IEEE Pervasive Computing Magazine 2(3), 13-18, 2003.