

## Cardio Online Reader - An Easy Way to Obtain Valid Scientific Articles in Cardiology

Zvolský, Miroslav 2011

Dostupný z http://www.nusl.cz/ntk/nusl-55979

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 04.05.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz .

# Cardio Online Reader – An Easy Way to Obtain Valid Scientific Articles in Cardiology

Post-Graduate Student:
MUDR. MIROSLAV ZVOLSKÝ

Department of Medical Informatics Instutite of Computer Science of the ASCR, v. v. i. Pod Vodárenskou věží 2 182 07 Prague 8, CZ

zvolsky@euromise.cz

Supervisor: DOC. ING. ARNOŠT VESELÝ, CSC.

Department of Medical Informatics Instutite of Computer Science of the ASCR, v. v. i. Pod Voddrenskou věží 2 182 07 Prague 8, CZ

vesely@pef.czu.cz

## Field of Study: Biomedical Informatics

This work was supported by the project 1M06014 of the Ministry of Education CR and the research plan AV0Z10300504.

#### Abstract

Searching for the clinical valid information in the large bibliographic database can be time consuming and hard work. We designed an easy-to-use web service Cardio Online Reader (COR) specialized on the topic of Cardiology. As a source we use the PubMed database adding simple filter and social functions for sharing the content.

## 1. Introduction

The cumulative total of journal articles exceeded 50 million in 2009 [1]. The most important free accessible resource of biomedical science articles is the PubMed database, which is one of key services provided by the National Center for Biotechnology Information (NCBI). The PubMed database contained 21067999 article citations on the 1 August 2011.

It is extremely complex problem to orientate oneself and to find desired information in this huge amount of papers. Especially it is important when searched information is clearly defined by a clinical domain, demanded time of publishing keywords or authors.

The web interface if the PubMed database [2] accessible on http://www.ncbi.nlm.nih.gov/pubmed/ offers one html form field for searching for key terms in the database. It put the accent on search query syntax, for the definition should be defined precisely. When the search query consists of one or two key terms, the search engine returns often tens of thousand results. The result list is sorted by time in a descendent order, so the most recent articles come at first, but this order says nothing about qualitative parameters of articles.

The NCBI web pages also offers an advanced search tool for the PubMed database http://www.ncbi.nlm.nih.gov/pubmed/advanced.

Users can define the search query in the PubMed Advanced Search in 39 parameters, which stores the PubMed database. Advanced Search saves a history of searched queries for each user. These queries can be repeatedly retrieved.

The web interface of Advanced Search is more complicated to use then one form field in the basic search. The definition of the query is more time consuming and it needs an experience with search query formulation for obtaining high-quality results. The best information sources provide relevant, valid material that can be accessed quickly and with minimal effort [3].

## 2. The Cardio Online Reader Web Application

Clinicians should obtain the information they want easy and quickly. Our purpose was to simplify the process of obtaining searched articles in the stressing and time lacking situation of clinical practice. We wanted to allow clinical workers without an experience with advanced database search tools to utilize the possibilities of large bibliographical databases. In the first phase we decided to limit the area of clinical domains to the Cardiology and developed the Cardio Online Reader application. This application is freely accessible on http://neo.euromise.cz/cor.

## 2.1. The Cardio Online Reader Database

The database of citations and abstracts of biomedical science articles is the main part of the Cardio Online Reader application (COR). This database uses a MySQL database engine. The main data source for our project is the PubMed database, that can be used free of charge.



**Figure 1:** An example of COR application web interface showing the list of articles retrieved for parameters set in the filter form fields above.

Import of the data was realized by the query to the PubMed database defining the domain of Cardiology by using Most important MeSH terms from Cardiology.

We filtered off articles, which do not fulfil our qualitative criteria from the Evidence Based Medicine point of view. We selected only these types of articles:

- Randomized Controled Trials,
- Systematic Reviews,
- Systematic Reviews with Metaanalysis,
- Guidelines,
- Practical Guidelines.

The result of this query was saved in the XML file.

Exported XML file was parsed by one-purpose PHP import script and selected data fields (title, authors, MeSH terms, abstract, unique identificator PMID, date of the

abstract publication in the MEDLINE database, link to the full text, journal title) were saved to the COR database.

The actualization of the COR database proceeds daily by an automatically started PHP script, which browse through an RSS channel of the PubMed database with the same query parameters as the original import script. The actualization script uses tools from the Entrez Programming Utilities [4] for gathering special data for each article, that are not part of the RSS channel.

## 2.2. The Web Interface of the COR

Contrary to the original PubMed interface we concentrated on the fastest way to reducing the number of search results preserving the focus on results important for the clinical practice.

Users can limit search results by one mouse-click to one category of EBM quality of evidence. Users can also use six form fields of the filter on the home page of the COR application for entering search criteria.



Figure 2: An example of COR application web interface showing the detail of an article record.

Individual form field stands for entering a part of text in the article title or abstract. Another form field specifies requested author or authors. The third large form field stands for entering parts or exact full terms of Medical Subject Headings (MeSH) thesaurus. Users can write down requested MeSH terms or thez can choose them from a generated MeSH Cloud or MeSH List, where terms are displayed in relation to their appearance in articles or sorted alphabetically.

In these form fields it is possible to use logical operators AND and OR. It is also possible to use a dynamically generated autocomplete function in these three fields to simplify entering exact phrases.

Users can limit the list of search results by setting the lowest and the highest date of publishing the article in the MEDLINE database in next two form fields. The dates can be set manually or chosen from the javascript date picker.

The last form field stands for the manual choice of the category of EBM quality of evidence.

For the fast choice of most frequented MeSH terms and their insertion to the filter, there is a "MeSH cloud" in the right part of the application web page, where enlisted terms differs in the text size displaying frequency of each term in the database. Users can use the list of last search queries.

## 2.3. Search Results

The COR application display search results matching entered parameters below the filter. Search results are in descending order sorted by the date of publication in the MEDLINE database. The simple list of results shows article title, names of authors and the date of publication in the MEDLINE database. There can be maximum of 15 results on one page, user can browse through the result pages. The EBM category of article can be differentiate by graphical icon.

By clicking on the article title in the list of results user can navigate to the detail page of the article record. The detail page shows the article title, names of authors, list of assigned MeSH terms, PMID identificator, link to the original record in the PubMed database, link to the full text of the article (if available in the Internet), article abstract, journal title and the date of publication in the MEDLINE database. The User of the COR application can rate the helpfulness of the article in the scale from one to five star symbols. This rating is linked to the IP address, so one user can rate a single article only one time. User can also attach a comment to the article.

#### 2.4. Web 2.0 Social Functions

Second big task for the COR application is to allow easy sharing of clinically important search results with colleagues, friends and professional community. The detail of the article including abstract and bibliographical data can be shared via email, Facebook, Twitter and other social networks contained in the "Share this" web service [5].

Excepting concrete scientific article detail, the COR offers an easy way to share a link to itself via Share this service, Facebook, Twitter, e-mail or one of 21 most common social and bookmarking services like Digg, Delicious, Reddit, Youhoo! od Google Bookmarks.

Users can follow COR own profiles at Facebook and Twitter, Blogger account and Youtube channel. Users can subscribe to RSS channels with last 20 articles or last 20 comments generally or individually for each EBM category of articles.

## 2.5. Future Plans and Improvements

We plan further improvements and simplifications in the web interface of the COR application in the future. One thing which can speed up using the filter and make the work more illustrative is to place a graphical slider and the time plot showing numbers of articles published in the discrete time periods and their selection in the filter.

Long term problem is to optimise the autocomplete function in three form fields in the filter to help users in inserting key terms in the easiest way. This process should be evaluated in the cooperation with common users.

Geotagging can help to make search results more regionally-oriented. Metadata contained in the PubMed database can show in which country the article was published. Geographical information in the field "Affiliation" is even more interesting. It is possible to find out where the article was created and what population is in

the article described. We can draw this information in the map or allow its limitation in the filter.

We assume an individualization of the web interface for registered users in the future development of the COR application. After the registration process and logging in user could browse the history of own search queries, create lists of favourite articles, let the system send him notification on some events in the database or define own RSS channels or add authorized comments and ratings.

The COR application not only can serve users of the web interface or RSS readers. By creating a XML data interface we can connect another information systems and send them search results or record details on their demands. Possible service for hospital information systems could be to offer relevant document for the concrete clinical situation defined by MeSH and geographical terms.

#### 2.6. Discussion

Widely accepted PubMed database of biomedical citations has a free accessible web interface with a basic or advanced version of the search. We can use other web services for searching for scientific articles by clinical terms or other parameters. These services are more general (Google, Google Scholar) or focused on natural sciences (Scopus). Why create another search tool?

The amount of scientific articles indexed in electronic databases increase steeply. Recent question "where to find" will surrender to questions "how to search" and "how to search the easiest way". The COR offers simple and fast way how to search the PubMed database for articles in the field of Cardiology and with the focus on highest evidence. It copes only one thousandth of the PubMed database and provides easy-to-use tools for setting the search query, that can acquire small amount of articles appropriate to the clinical need.

There are another web services specialized on searching in large databases of scientific bibliography (http://demos.vivisimo.com, http://www.tripdatabase.com, http://www.pubmeddy.com - discontinued). The COR is unique in its focus on one domain (Cardiology), on few defined EBM categories most important for clinical practice and in the simplicity of use.

The key question for the progress of the COR application will be the interest of expert medical community. The COR contains tools for sharing scientific information between experts, tools for subjective rating of their quality and tools for expert discussion. Experts could be motivated by functions for registered users, in-

dividualized search functions, the ease of use and the fact, that this application is free of charge.

The COR application is designed specially for Cardiology. Filters used for extraction from the PubMed data-

base firmly defined. Same technology could be used for another one purpose (one expert domain) web portals. Similar filters could be also individually set for registered users, so the scope could be widened to another domains

Name of service	Number of results
Google.com	approx. 66,000,000
Google.com last year	approx 288,000,000
Google Scholar	approx 1,700,000
Scopus	243,866
Scopus - only Health Sciences	179,511
PubMed	144,097
COR	1,695
COR - Practice Guidelines	79
COR - Pracice Guidelines in last 5 years	33

Table 1: Comparison of the number of search results in different web services - searching for the MeSH term "heart failure".

#### 3. Conclusion

We created the Cardio Online Reader application for an easy search for clinical relevant scientific articles in the field of Cardiology. This application is freely accessible on **http://neo.euromise.cz/cor**. The PubMed database is the main data source for our application.

The application contains a filter consisted from six form fields. Search results are in descending order sorted by the date of publication in the MEDLINE database. The detail page shows the article title, names of authors, list of assigned MeSH terms, PMID identificator, link to the original record in the PubMed database, link to the full text of the article (if available in the Internet), article abstract, journal title and the date of publication in the MEDLINE database.

The COR application offers an easy access to services for content sharing as the "Share this" service, social and bookmarking services, comments and ratings and sharing via email.

#### References

- [1] A. Jinha, "Article 50 million: an estimate of the number of scholarly articles in existence", *Learned Publishing*, Vol. 23, No. 3, pp. 258-263, 2010.
- [2] "PubMed: MEDLINE Retrieval on the World Wide Web" United States National Library of Medicine 2002-06-07 http://www.nlm.nih.gov/pubs/factsheets/pubmed. html.
- [3] R. Smith, "What clinical information do doctors need?" BMJ, Vol. 313, No. 7064, pp. 1062-1068, 1996.
- [4] D. Wheeler, T. Barrett, D. Benson, S. Bryant, K. Canese, V. Chetvernin, D. Church, M. DiCuccio, R. Edgar, S. Federhen, L. Geer, Y. Kapustin, O. Khovayko, D. Landsman, D. Lipman, T. Madden, D. Maglott, J. Ostell, V. Miller, K. Pruitt, G. Schuler, E. Sequeira, S. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. Tatusov, T. Tatusova, L. Wagner, and E. Yaschenko, "Database resources of the National Center for Biotechnology Information", Oxford University Press, 2006.
- [5] ShareThis, http://www.sharethis.com.