# An Effective User Interface for a Scientific DLS to Support Information Retrieval Evaluation

Marco Dussin* and Nicola Ferro*

*Department of Information Engineering – University of Padua – Italy
{dussinma, ferro}@dei.unipd.it

## I. Motivations and Goals

The experimental evaluation of information retrieval systems is usually carried out in important international evaluation initiatives, such as the *Text REtrieval Conference (TREC)* [1] or the *Cross-Language Evaluation Forum (CLEF)* [2], which bring research groups together, provide them with the needed tools to measure the performances of their systems, discuss and compare their results.

These evaluation initiatives produce a huge amount of experimental results, which are valuable *scientific data* and need to be properly managed, curated, and enriched [1], [2], [3]. The growing interest in the proper management of scientific data is also witnessed, in the 7th European Community Framework Programme, by the i2010 Digital Library Initiative which clearly states that "digital repositories of scientific information are essential elements building European eInfrastructure for knowledge sharing and transfer, feeding the cycles of scientific research and innovation up-take" [5].

In this context, we have started an effort to design and develop a scientific *Digital Library System (DLS)* able to support the course of an evaluation initiative, and to manage, curate and enrich the scientific data produced during it. We have implemented our system in a prototype, called *Distributed Information Retrieval Evaluation Campaign Tool (DIRECT)* [3], [4]. DIRECT has been adopted and tested in the CLEF 2005 and 2006 campaigns. During CLEF 2005, it was used by nearly 30 participants spread over 15 different countries, who submitted more than 530 experiments; then 15 assessors assessed more than 160,000 documents in seven different languages, including Russian and Bulgarian which do not have a latin alphabet. During CLEF 2006, it was used by nearly 75 participants spread over 25 different countries, who submitted around 570 experiments; 40 assessors assessed more than 198,500 documents in nine different languages.

In this study, we present the *User Interface (UI)* which has been developed as a *communication tool* to support all the stages of an evaluation campaign, among which the submission of the experiments, the creation of the relevance assessments, the provision of computational services, and the visualization of statistical analysis.

Moreover, we discuss the re-design of the UI, which has been undertaken passing from the first prototype, used in CLEF 2005 to the new one used in CLEF 2006.

In conclusion we summarize some of the features planned for the next prototype, which will be used in the CLEF 2007 campaign.

[1] http://trec.nist.gov/
[2] http://www.clef-campaign.org/

## II. DIRECT User Interface

After an initial investigation of user requirements and needs, the UI of DIRECT was designed to meet the following goals:

- to be cross-platform and easily deployable to end users;
- to be as modular as possible, clearly separating the application logic from the presentation logic;
- to be intuitive and capable of providing support for various user tasks, such as experiment submission and consultation of metrics and plots about experiment performances;
- to support different types of users, i.e. participants, assessors and administrators, who need to have access to different kinds of features and capabilities;
- to support the internationalization and localization: the application needs to be able to adapt to the language of the user and his country or culturally-dependent data – such as dates and currencies – without engineering changes or recompilation.

In order to achieve the first goal, we decided to develop a Web UI taking care of addressing cross-browser compatibility issues. With respect to the second goal, we used a *Model-View-Controller (MVC)* approach to obtain a good separation between the application and presentation logics.

Now we discuss how the remaining three goals have been addressed in DIRECT 2005 and 2006 prototypes.

## III. The First Propotype: DIRECT 2005

The main objective of this prototype was to offer all the basic functionalities required for an evaluation initiative and to provide us with some feedback about the actual needs of the target user community.

The Web UI has been implemented in plain *HyperText Markup Language (HTML)* in order to guarantee the maximum compatibility among browsers (Figure 1), with a *thin client* pattern approach. This pattern delegates to the server the handling of the data in response to a user action, forcing the interface to a complete reload every time data need to be changed or refreshed, also for tasks that need a lot of

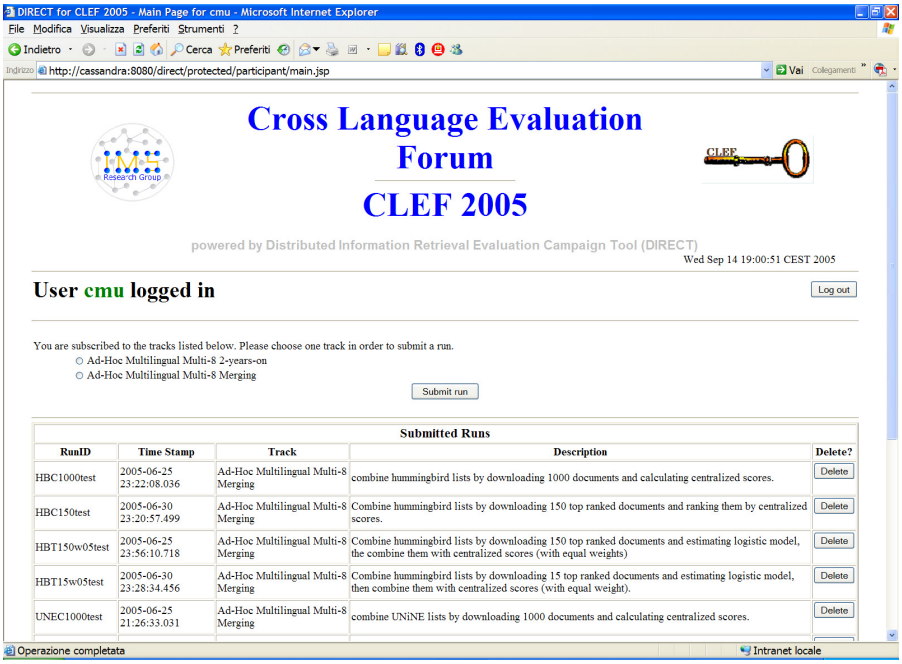


Figure 1. Participant interface of DIRECT 2005

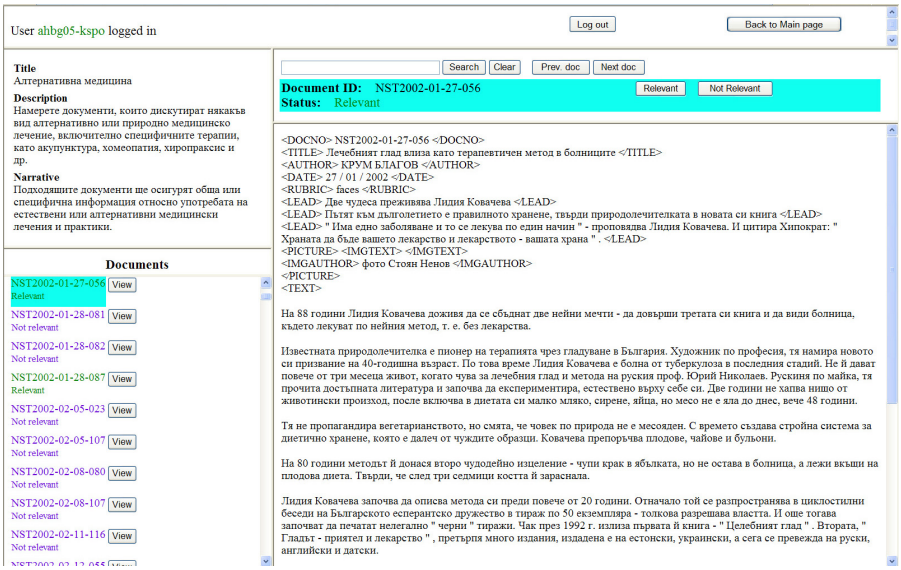


Figure 2. Assessor interface of DIRECT 2005

interaction such the assessor interface (Figure 2). The result was unfortunately an interface lazy to user requests.

In addition, to provide a straightforward and functional interface, we designed a set of wizards which guide users to the completion of theirs tasks and we adopted a tabular layout to present all the needed information in a compact and coherent way.

This first UI does not support internationalization of the contents, which were available only in English.

DIRECT 2005 prototype was implemented in Java and the Web UI was based on *JavaServer Pages (JSP)*[3] and the Struts[4] framework, which provides support for the MVC approach.

## IV. The Second Prototype: DIRECT 2006

The experience of DIRECT 2005 allowed us to gain a better knowledge of the target users, of the potential cross-browser compatibility issues, and to have feedback about what improvements can be made to the UI.

In order to make the user interaction more successful and pleasing, we re-designed the UI to provide a completely updated and more usable layout and to be more interactive by using JavaScript client-side code (*fat client* pattern). The user role has been emphasized by different colouring – grey tones for the zone common to all users, green tones for

[3]http://java.sun.com/products/jsp/
[4]http://struts.apache.org/

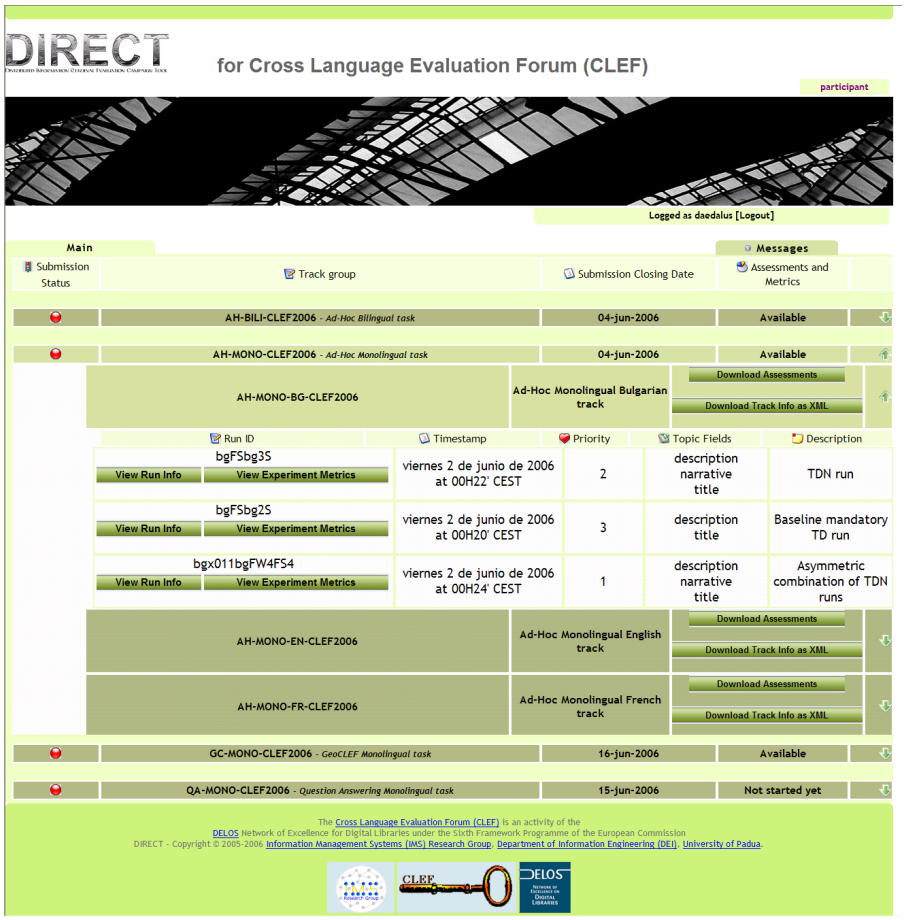


Figure 3. Participant interface of DIRECT 2006

the participants, and violet tones for the assessors – and by a suitable iconography. Furthermore, we have developed a JavaScript library that provides button management, which can be activated, deactivated, and colored in response to user actions, and of expansible tables, so that the user may concentrate on the contents of his own interest. All those improvements are shown in Figures 3 and 4, the new versions for the interfaces described in III.

A specific library of JSP tags was developed, which deals with the generation of HTML pages to dynamically adapt the UI to the current status of the evaluation campaign and make the proper commands for each stage of the evaluation campaign available to the users. For example, while the experiment submission is open, the user can find buttons to submit new experiments, to delete submitted experiments and to view experiment information; once the submission is closed, the buttons for submitting and deleting experiments are no longer present in the UI while new buttons for accessing evaluation metrics and plots appear, as shown in Figure 3.

Finally the UI has been made more modular by using the Tiles[5] template system which allows for a rapid development and reuse of components, and support for internationalization and localization has been introduced, using dictionaries of couples of labels and translations and using Java internationalization capability[6].

A substantive effort has been made to design a more manageable relevance assessments interface. Assessors have to assess thousands of documents in a limited period of time and thus the responsiveness of the interface is fundamental to their task, so we studied a completely new solution to improve the DIRECT 2005 approach. The new assessor UI, shown in Figure 4, using *Asynchronous JavaScript and XML (AJAX)* can make asynchronous calls and visualize the results without reloading the whole page. For example, a search can now be made and the results displayed with highlighted words without having to reload the structure of the interface, or an assessment can be made by pressing the matching button and iconography and color changes can be seen on the page without having to wait, as the system sends the new data to the server in background.

In this way, we were able to speed up the user interaction by loading only the new document to be assessed without reloading all the other information visualized in the interface, such as the list of documents to be assessed.

## V. Features planned for DIRECT 2007

The prototype of DIRECT for CLEF 2007 aims to improve interaction between users and the system, treasuring the

[5]http://struts.apache.org/1.x/struts-tiles/index.html

[6]http://java.sun.com/javase/technologies/core/basic/intl/

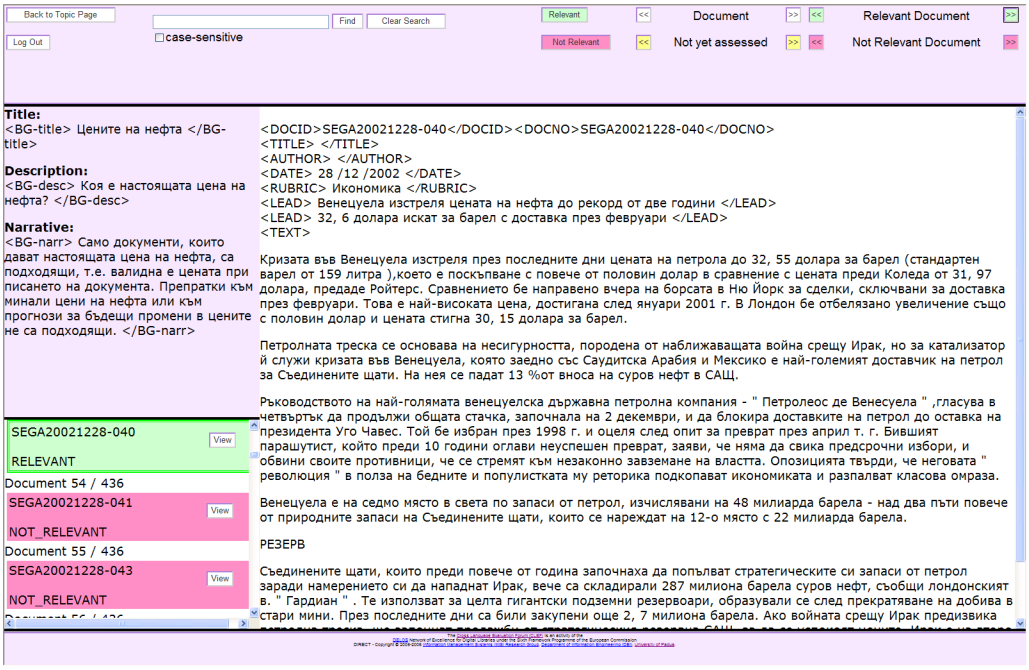


Figure 4. Assessor interface of DIRECT 2006

experience of previous versions and the amount of valuable comments and suggestions on how to improve and make them more usable. Interaction between users is also being studied to create collaborative tools.

New features are planned, such as supporting assessors in the whole topic creation cycle (searching, browsing, viewing, and assessing the documents that may be retrieved in response to a topic), and offering the functionalities needed for easing and improving the collaboration among the assessors, like the sharing of notes and comments about the topic being created. In this way, the system will keep track of all the context which led to the creation of a given topic [6].

Another feature worth developing is the opportunity of keeping track and making available information about the results of the current and previous campaigns, using the *Digital Object Identifier (DOI)* System [7]. A DOI is a name that persistently identifies an entity on digital networks. We plan to provide DOIs support for information produced and stored in DIRECT, to actual user types and also to a new class of users called *visitors*, which after registration will browse and/or download papers, graphs, *eXtensible Markup Language (XML)* and other information, according to the privileges assigned by the system.

In addiction, DIRECT offers the functionalities needed for easing and improving the collaboration among the assessors, who can add and share notes and comments about the topic being created. In this way, the discussion, which is usually conducted while creating a topic, does not need to be done by exchanging e-mails among the assessors but can be carried out directly in the system. In this way, the system keeps track of all the context which led to the creation of a given topic.

In view of all this, we plan to develop a more usable, richly interactive and responsive DIRECT UI, by studying how it is used, increasing the iconography expressivity, devising efficient layout issues and developing a new set of tools and libraries built on Javascript and AJAX technologies.

## Acknowledgements

The work reported in this paper has been partially supported by the DELOS Network of Excellence on Digital Libraries, as part of the Information Society Technologies (IST) Program of the European Commission (Contract G038-507618).

## References

[1] M. Agosti, G. M. Di Nunzio, and N. Ferro, “A Data Curation Approach to Support In-depth Evaluation Studies,” in *Proc. International Workshop on New Directions in Multilingual Information Access (MLIA 2006)*, http://ucdata.berkeley.edu/sigir2006-mlia.htm, 2006, pp. 65–68.

[2] M. Agosti, G. M. Di Nunzio, and N. Ferro, “Scientific Data of an Evaluation Campaign: Do We Properly Deal With Them?” in *Working Notes for the CLEF 2006 Workshop*, http://www.clef-campaign.org/2006/working_notes/workingnotes2006/agostiCLEF2006.pdf, 2006.

[3] G. M. Di Nunzio and N. Ferro, “Scientific Evaluation of a DLMS: a service for evaluating information access components,” in *Proc. 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006)*, LNCS 4172, Springer, Heidelberg, Germany, 2006, pp. 536–539.

[4] G. M. Di Nunzio and N. Ferro, “DIRECT: a System for Evaluating Information Access Components of Digital Libraries,” in *Proc. 9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005)*, LNCS 3652, Springer, Heidelberg, Germany, 2005, pp. 483–484.

[5] European Commission Information Society and Media, “i2010: Digital Libraries,” http://europa.eu.int/information_society/activities/digital_libraries/doc/brochures/dl_brochure_2006.pdf, October 2006.

[6] G. M. Di Nunzio and N. Ferro. Queries and Relevance Assessments: The Right Context for the Right Topic. In *Proc. First International Workshop on Adaptive Information Retrieval (AIR)*. http://www.dcs.gla.ac.uk/workshops/air/ [last visited 2007, January 4], October 2006.

[7] N. Paskin, editor. *The DOI Handbook – Edition 4.4.0*. International DOI Foundation (IDF). http://dx.doi.org/10.1000/186 [last visited 2006, October 2], September 2006.