

DIRECT: the First Prototype of the PROMISE Evaluation Infrastructure for Information Retrieval Experimental Evaluation

by Nicola Ferro

PROMISE is a network of excellence focused on the experimental evaluation of multilingual and multimedia information access systems. One of its key contributions is to develop and provide an open evaluation infrastructure, which brings automation to the evaluation process, managing, curating, and providing access to the scientific data produced during the evaluation activities.

Experimental evaluation is a key activity for driving and supporting the development of multilingual and multimedia information access systems. It is an essential part of the scientific process since using shared data sets and evaluation scenarios systems can be compared, performances can be better understood, and progress can be pursued and demonstrated.

Large-scale evaluation initiatives, such as Text REtrieval Conference (TREC) in the United States, the Cross-Language Evaluation Forum (CLEF) in Europe, and the NII-NACSIS Test Collection for IR Systems (NTCIR) in Asia, contribute significantly to advancements in research and industrial innovation in the information retrieval sector, and to the building of strong research communities. A study conducted by NIST reports that “for every \$1 that NIST and its partners invested in TREC, at least \$3.35 to \$5.07 in benefits

accrued to IR researchers. The internal rate of return (IRR) was estimated to be over 250% for extrapolated benefits and over 130% for unextrapolated benefits”.

Large-scale evaluation campaigns produce a huge amount of extremely valuable scientific data which provides the foundations for subsequent scientific production and system development and constitutes an essential reference for literature in the field. This data is also economically valuable, due the considerable effort devoted to its production: the NIST study estimates in about 30 million dollars the overall investment in TREC.

Nevertheless, little attention has been paid over the years to modelling, managing, curating and accessing the scientific data produced by evaluation initiatives, despite the fact that the importance of scientific data in general has been highlighted by many institutional organizations, such the European Commission, the US National Scientific Board, and the Australian Working Group on Data for Science.

Objectives

Our goal is to deliver a unified infrastructure and environment for data, knowledge, tools, methodologies, and the user community in order to advance the experimental evaluation of complex multimedia and multilingual information systems. The evaluation infrastructure will:

- manage and provide access to the scientific data produced during evaluation activities;
- support the organization of evaluation campaigns;
- increase the automation of the evaluation process;
- provide component-based evaluation;
- foster the usage and understanding of the scientific data;

A user-centered design approach will be adopted involving the different stakeholders, eg scientists, evaluation campaign organizers, system developers, students, in the development of the infrastructure.

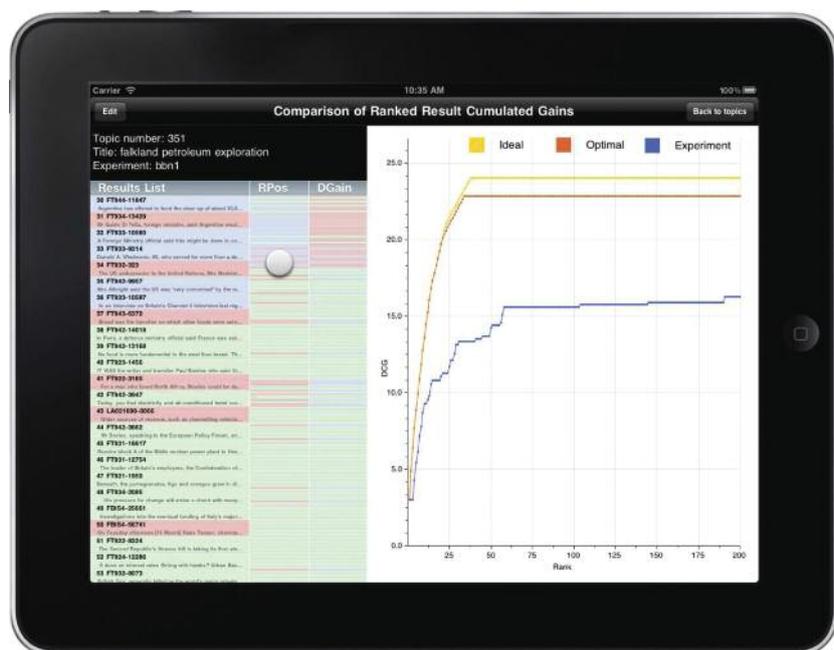


Figure 1: DIRECT provides tools for statistical analyses and reporting of results

DIRECT: the First Prototype

The outcome of this effort is the Distributed Information Retrieval Evaluation Campaign Tool (DIRECT), which

- introduces a conceptual model of the information space of an evaluation campaign;
- provides metadata describing the scientific data managed, to enable sharing and re-use;
- adopts a unique identification mechanism allowing explicit citation of and easy access to the scientific data;
- manages all the aspects of an evaluation campaign, and provides tools for statistical analyses and reporting of results.

DIRECT has been developed and tested in the course of the annual CLEF evaluation campaigns since 2005. It now manages and provides online access to much of the data produced over ten years of CLEF. It also aims at improving

interaction with the experimental results by researchers and system developers. We are now investigating the adoption of innovative devices, such as the iPad, which can allow for a natural and easy interaction with the experimental results and scientific data in real time.

Next Steps

PROMISE is a three year project beginning in September 2010. It will issue releases of the evaluation infrastructure with new functionalities, such data annotation and visual analytic techniques, annually. In order to achieve a better representation, interaction, and understanding of experimental results, we are investigating how best to exploit human-computer interaction and the principles of visual analytics. This will be the topic of the PROMISE Winter School Information Retrieval meets Information Visualization, which will be held in January 2012, Zinal, Switzerland.

The information retrieval area is now beginning to explore and exploit the scientific data produced during the evaluation studies by making use of methods typical of the database and knowledge management areas. The aim of the Data infrastructureEs for Supporting Information Retrieval Evaluation (DESIRE 2011) workshop, co-located with CIKM 2011, the 20th ACM Conference on Information and Knowledge Management, in October 2011, Glasgow, UK is to bring together experts from the three communities in order to discuss the challenges involved. The intention of the organizers is to produce a roadmap and a set of initial best practices guiding the development of evaluation infrastructures to manage experimental data.

Links:

PROMISE: <http://www.promise-noe.eu/>

DIRECT: <http://direct.dei.unipd.it/>

DEMO: <http://www.youtube.com/watch?v=fDsXDCUPkiM>

CLEF 2011: <http://www.clef2011.org/>

CLEF: <http://www.clef-campaign.org/>

DESIRE 2011 Workshop:
<http://www.promise-noe.eu/events/desire-2011/>

PROMISE Winter School 2012:
<http://www.promise-noe.eu/events/winter-school-2012/>

TREC Economic Impact Study:
<http://trec.nist.gov/pubs/2010.economic.impact.pdf>

Please contact:

Nicola Ferro, University of Padua, Italy
E-mail: ferro@dei.unipd.it

Potential Cell-phone Dangers: Latest Studies

by Harry Rudin

The results of two studies — actually collections of studies — have been published just recently. These are the Europe-wide Interphone study and the Swiss National Research Program NRP 57, Entitled “Non-ionizing Radiation – Health and Environment”. In few words, neither has identified measurable health dangers arising from cell-phone use — but both recommend further study and caution.

Swiss NRP 57 Study

The Swiss National Science Foundation financed eleven research projects in four categories: dosimetry and exposure assessment, laboratory studies and epidemiology, cell biology, and risk perception. The projects were carried out over a three-year period.

In the first category, a study was made to discover the major source of non-ionizing radiation. There are many wireless devices that radiate but, based on an extensive set of week-long measurements made in the city of Basel, the major source of radiation remains the cell phone. How much radiation depends on the transmission system used (UMTS radiates less than the older GSM) and using a headset reduces the exposure by a factor of ten. A surprising and upsetting result of the study is that the modern induction stove causes very strong electromagnetic fields and could well be a substantial danger for pregnant women and their unborn children. This is particularly true for professional cooks.

In the second category, it was clearly shown that the pulse-modulated radio-frequency fields emitted by cell phones do affect the human brain. First, these induce short-term effects observable in the brain’s electro-encephalogram. Some effects continued after the exposed person went to sleep but these effects did not have a sleep-disturbing result. Changes in the cerebral blood circulation and heart rate were also observed but these were small and no negative effect on a person’s well being was observed.

In the cell-biology category, DNA fragmentation in mammalian cells was shown. The frequency of the radiation was not cell-phone frequencies but normal household 50 Hz. The observed effect on cellular processes was very weak and presumably of no medical consequence.

As to results of the risk-perception study, the Swiss public accepts wide-spread cell-phone use but at the same time is somewhat apprehensive of the radiation coming from base stations. Some 5 % believe that they are hypersensitive to electromagnetic fields and suffer health problems as a result.

While the NRP 57 studies found no direct evidence for health problems caused by cell-phone radiation, the recommendation is that research be continued, particularly as cell-phone usage increases and technology changes so rapidly. The first link below is a summary of the findings.