MEETING REPORT

SYMPOSIUM ON HUMAN–COMPUTER INFORMATION RETRIEVAL

Daniel Tunkelang,[1] Robert Capra,[2] Gene Golovchinsky,[3] Bill Kules,[4] Catherine Smith,[5] and Ryen White[6]

Abstract

Human–computer information retrieval (HCIR) is the study of information retrieval techniques that integrate human intelligence and algorithmic search to help people explore, understand, and use information. Since 2007, we have held an annual gathering of researchers and practitioners to advance the state of the art in this field. This meeting report summarizes the history of the HCIR symposium and emphasizes its relevance to the data science community.

Introduction

It's a cliché that we live in a world of Big Data. But the bottleneck in understanding data is not computational. Rather, the biggest challenge is designing technical solutions that effectively leverage human cognitive ability. As big data startups Quid[1] and Opera Solutions[2] have argued, data analysis systems should augment people's capabilities rather than replace them. But this is hardly a recent argument: human–computer information pioneer Doug Engelbart said in 1962 that the goal of technology is "the enhancement of human intellect by increasing the capability of a human to approach a complex problem situation."[3] Algorithms extract signal from raw data, but people fill in the gaps, creating models and evaluating analyses.

In exploring how to best leverage the human analyst in the loop, data science is following in the footsteps of information retrieval. Human–computer information retrieval (HCIR) emerged as a critique of modern information retrieval, recognizing that information retrieval should be "more than a branch of computer science, concerned primarily with issues of algorithms, computers, and computing."[4] We thus believe that the HCIR Symposium holds broad interest for big data researchers and practitioners.

The First HCIR Workshop

In the summer of 2007, Daniel Tunkelang (then chief scientist at Endeca, Cambridge, MA) reached out to Michael Bernstein (then a PhD student at MIT, Cambridge, MA) to create a

[1]LinkedIn, Mountain View, California.
[2]School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.
[3]FXPAL, Palo Alto, California.
[4]School of Library and Information Science, The Catholic University of America, Washington, D.C.
[5]School of Library and Information Science, Kent State University, Kent, Ohio.
[6]Microsoft Research, Redmond, Washington.

workshop that would explore the nexus of information retrieval and human–computer interaction. There was precedent: a pair of workshops in 1996 and 1998 at the University of Glasgow,[5] as well as exploratory search workshops held at the Association of Computing Machinery (ACM)'s international conferences on information retrieval (SIGIR) and human-computer interaction (SIGCHI).[6] But there was no venue or community expressly devoted to this intersection of the two fields.

We held the first HCIR workshop at MIT and Endeca in Cambridge, Massachusetts, on October 23, 2007. We borrowed the name from a lecture by Gary Marchionini, dean of the University of North Carolina's School of Information and Library Science. The lecture, entitled "Toward Human–Computer Information Retrieval," put forth a vision of users being more active participants in the information-seeking process—specifically that systems should increase user responsibility and control, requiring and rewarding human intellectual effort.[7]

"AFTER HCIR 2008, WE KNEW THAT WE HAD MOVED BEYOND A ONE-OFF EVENT INTO AN ANNUAL INSTITUTION."

The HCIR 2007 workshop featured a keynote by Microsoft researcher Ryen White about exploratory search on the web, a talk on visualization by Many Eyes[8] creators Fernanda Viégas and Martin Wattenberg, and a full day of talks on topics that included collaborative search, question answering, and personal information management. The success of the workshop demanded a sequel.

The Workshop Evolves

The following year, the workshop took place at Microsoft Research in Redmond, Washington, and featured Susan Dumais as a keynote shortly before she won the Gerard Salton Award recognizing her for leadership in bridging the fields of information retrieval and human–computer interaction. Again, the workshop brought together leading lights from academia and industry, including Greg Linden and Ashok Chandra. After HCIR 2008, we knew that we had moved beyond a one-off event into an annual institution.

In 2009, the workshop returned to the East Coast, specifically The Catholic University of America in Washington, DC. The keynote speaker was Ben Shneiderman, founding director of the Human–Computer Interaction Lab at the University of Maryland and one of the world's top researchers in human–computer interaction and visualization. The conference location attracted participants from government agencies that supplemented the usual academic and industry mix.

In 2010, we collocated the HCIR workshop with the Information Interaction in Context Symposium (IIiX 2010) at Rutgers University in New Brunswick, New Jersey. Continuing our tradition of distinguished keynote speakers, Dan Russell, Google's uber tech lead for search quality and user happiness, offered insights on what makes search easy or difficult for users. We also added a new element to the program, the HCIR Challenge, which is described in detail in the next section.

In 2011, we held the workshop at Google's headquarters in Mountain View, California. Finally making good on our debt to Gary Marchionini for lending us the workshop name, we invited him to be our keynote. Holding the event in Silicon Valley—and at Google in particular—helped us achieve a record attendance of one hundred people. We also created a special topic issue of the Journal of Information Processing and Management devoted to HCIR as an opportunity for authors to expand on their presentations.

Finally, 2012 brought the workshop back to Cambridge, Massachusetts—now expanded into a two-day symposium. Our growing set of industry sponsors now included FXPAL, IBM Research (which hosted the symposium), LinkedIn, Mendeley, Microsoft Research, MIT CSAIL, and Oracle (which had acquired Endeca in 2011). Our keynote speaker was search user-interface pioneer Marti Hearst, who regaled us with her "Halloween Cauldron of Ideas for Research." And we saw a full third of our 75 attendees come from industry, a welcome mix at a scholarly conference.

Over its six years, HCIR has evolved into a premier venue for exploring ideas at the intersection of information retrieval and human–computer interaction.

The HCIR 2010 Challenge: Exploratory Search

In 2010, we introduced the HCIR Challenge as a new feature of the workshop. The HCIR Challenge invited researchers and practitioners to build and demonstrate systems embodying the spirit of HCIR, effectively increasing the user's participation in the information-seeking process.

We decided to hold our first HCIR Challenge around the topic of exploratory search in a news archive. The Linguistic Data Consortium provided participants free access to the *New York Times* Annotated Corpus; a corpus is a collection of over 1.8 million *New York Times* (NYT) articles published between 1987 and 2007 and annotated with rich metadata.[9] We also offered participants a baseline search system for the NYT corpus built using the open source Apache Solr[10] package.

Each participating team submitted a four-page challenge report describing their work. At the workshop, participants

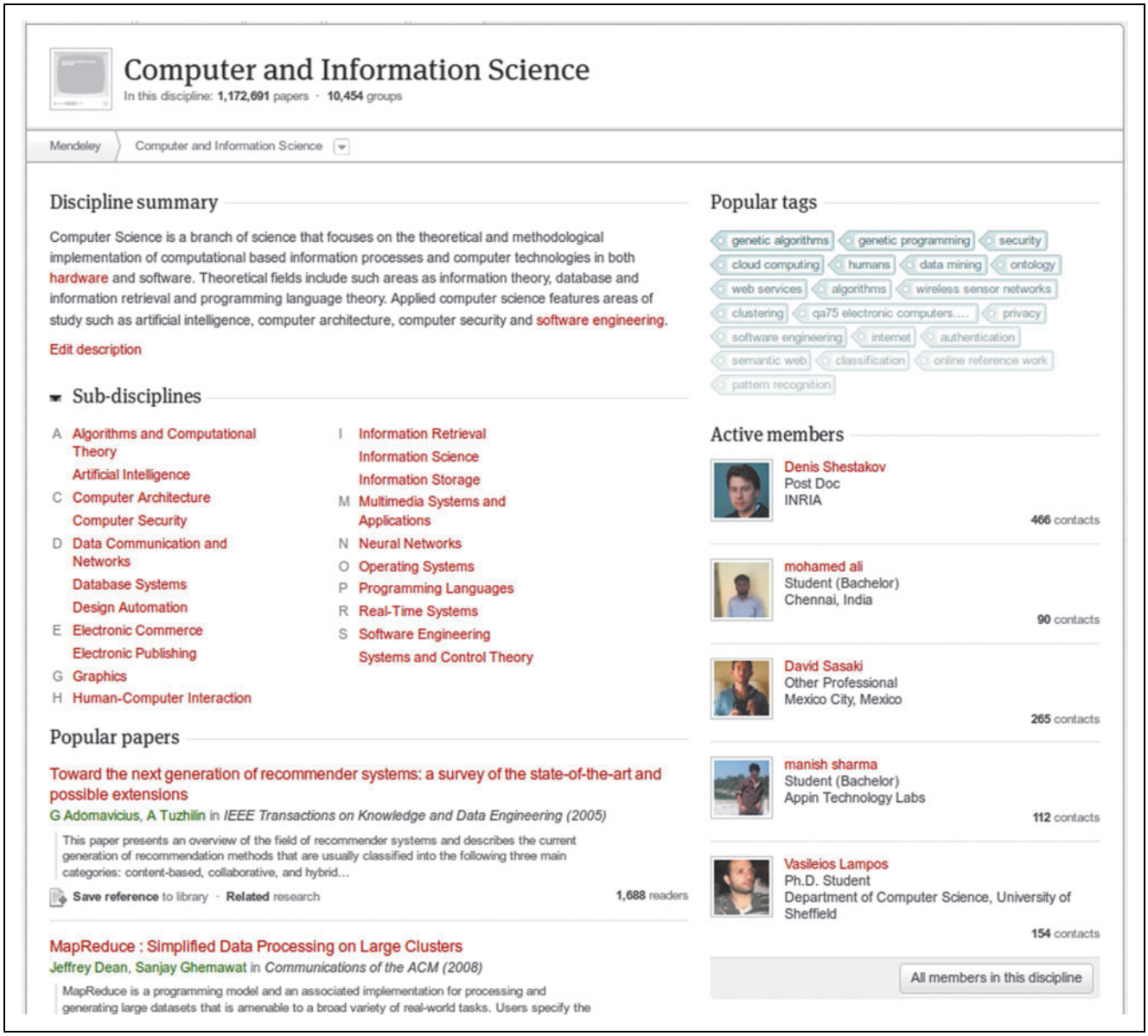


FIG. 1. Mendeley is an open catalog of 180M+ academic documents, associated with a network of 1.6M+ researchers.

presented their systems so that attendees could evaluate them based on the following HCIR evaluation criteria:

- Effectiveness: Are users able to complete tasks?
- Efficiency: How efficiently do users complete tasks?
- Control: Does the system give users control over the information-seeking process?
- Transparency: Do users understand what the system is doing?
- Guidance: Does the system help users refine their search strategy or reach their search goals?
- Fun: Is the system engaging and fun to use?

Participants demonstrated how their systems helped users perform complex tasks, such as tracking the price of a slice of pizza in New York varied over the past two decades, or enumerating the main arguments made for and against rent control.

The quality of the entries was impressive. All of the participants offered interesting ideas: custom facets, visualization of the associations between relevant terms, multi-document summarization to catch up on a topic, and combination topic modeling with sentiment analysis to analyze competing perspectives on a controversial issue. The winning entry, the Time Explorer, [11] came from a team of Yahoo researchers. As its name suggests, it allowed users to see the evolution of a topic over time. It also parsed absolute and relative dates from article tests—in some cases, references to past or future times outside the publication span of the collection. Moreover, the temporal visualization of topics allowed users to discover unexpected relationships between entities at

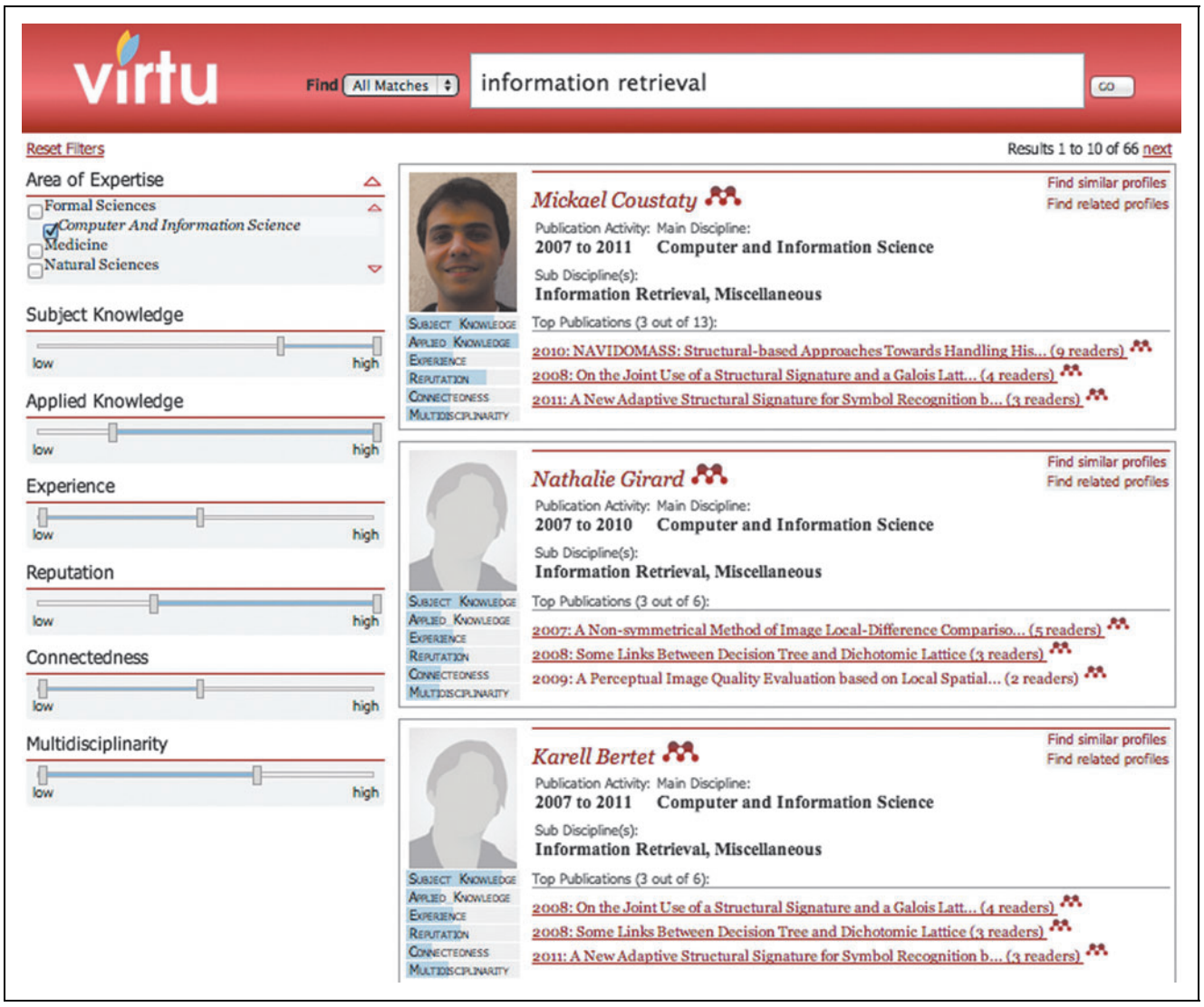


FIG. 2. Virtu, the winner of the HCIR 2012 Challenge. Developed by University of British Columbia researchers Luanne Freund and Kristof Kessler.

particular points in time, e.g., between Slobodan Milosevic and Saddam Hussein.

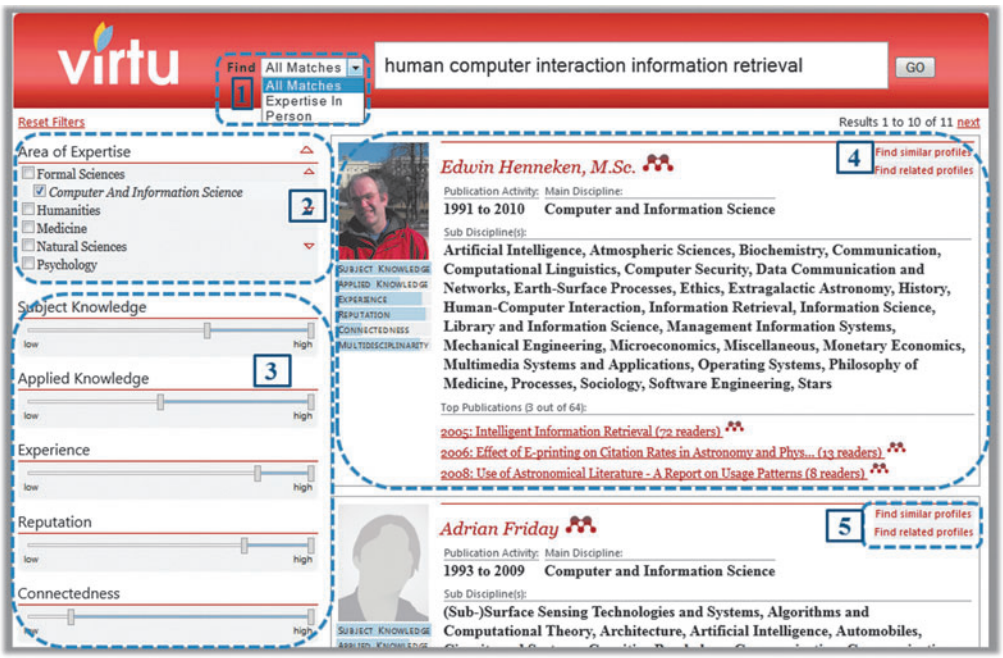
The HCIR 2011 Challenge: Information Availability

Building on this success, the HCIR 2011 Challenge focused on the problem of information availability. This problem arises when the seeker faces uncertainty as to whether the information of interest is available at all. Instances of this problem include some of the highest-value information tasks, such as those facing national security and legal/patent professionals, who might spend hours or days searching to determine whether the desired information exists. We used the CiteSeer digital library of scientific literature as a corpus. The CiteSeer corpus contains over 750,000 documents and provides rich metadata about documents, authors, and citations.[12]

We offered the following example task to give participants an idea of what we expect users to be able to do with their systems:

> *Latent Semantic Indexing (LSI) is an indexing and retrieval method that uses a mathematical technique called singular value decomposition (SVD) to identify patterns in the relationships between the terms and concepts contained in an unstructured collection of texts. Deerwester et al. published seminal papers on LSI in 1988. Is there earlier work that anticipates some or part of this approach?*

Two weeks before the workshop, we told participants which tasks would be used to judge their systems. The topics ranged from validating David J. DeWitt and Michael Stonebraker's claim that MapReduce was not novel[13] to finding research articles applying collaborative filtering to people search.



1. **Query input:** A drop down box offers the option of searching for a person, for expertise in an area, or for all matches (default).
2. **Discipline Facets:** The main disciplines listed in profiles are presented as hierarchical facets, grouped into broad categories at the top level. Checkboxes allow the searcher to select one or more disciplines as limits on a keyword query, or to browse through profiles within a given discipline. The discipline facets are dynamically updated for each query and are combined as a Boolean OR search if multiple disciplines are selected.
3. **Expertise Sliders:** The six expertise dimensions are manipulated directly using dual-control sliders that set lower and upper bounds on each measure. The underlying values for each dimension are calculated as a linear combination of basic measures and mapped to the range 0 to 100. The median value of each dimension is set to the mid-point of the slider. Search results are updated as soon as a slider handle is released.
4. **Profile:** Search results are displayed as index cards that include basic profile information such as name, publication activity, main discipline, sub disciplines derived from publications, the user's top three publications, and an image and title if available in the dataset. Links to the user's Mendeley page and the pages of the publications are also provided. Each profile shows the user's expertise score components as six horizontal bars representing the score for each dimension.
5. **Find Similar and Related:** Two links provided in each index card trigger prototype queries based on the selected profile of interest. *Find similar* retrieves profiles in the same discipline that have similar values on the expertise measures (+/-10 %), by adjusting the values of the Expertise Sliders. *Find related* uses data on co-authors, connections and groups to retrieve profiles that are most closely linked to the profile of interest.

FIG. 3. Virtu takes a task-based approach to expertise, exposing and allowing the user control over dimensions of expertise that are more or less desirable depending on the type of expert-finding task.

Systems were expected to optimize the user experience for three criteria:

- Correctness of the outcome. Do users correctly conclude whether the information of interest is available?
- Efficiency. How much time or effort do users invest, regardless of outcome?
- User confidence in the outcome. Do users believe the results, particularly if they are negative?

Again, the competition was fierce—all of the systems were industry grade, and the participants represented startups, research labs, and established publishers.

The winning entry was a system called Querium, developed by Gene Golovchinsky and Abdigani Diriye. Querium used relevance feedback, faceted search, and query fusion to deliver a compelling HCIR solution to the information availability problem.

The HCIR 2012 Challenge: People Search

> "AGAIN, THE COMPETITION WAS FIERCE—ALL OF THE SYSTEMS WERE INDUSTRY GRADE."

The HCIR 2012 Challenge focused on the problem of people and expertise finding. Mendeley[14] generously provided the corpus: a database of over one million researcher profiles with associated metadata—including published articles, academic status, disciplines, awards, and more—taken from Mendeley's network of 1.6M+ researchers and 180M+ academic documents (Fig. 1). Participants built systems to enable efficient discovery of experts or expertise for applications such as collaborative research, team building, and competitive analysis.

We asked participants to build systems that could perform three kinds of tasks:

- Hiring. Given a job description, produce a set of candidates for the position.
- Assembling a conference program. Given a conference's past history, produce a set of candidates for keynotes, program committee members, etc.
- Finding people to deliver patent research or expert testimony. Given a patent, produce a set of candidates who could deliver relevant research or expert testimony for use in a trial.

The entries represented our most sophisticated systems to date. One of them was a fully functional iPad app supporting swipe and multi-touch gestures. Another cross-referenced the Mendeley user profiles with data from Academia.edu and used Microsoft Academic Search to categorize publication and journals.

The winning entry was Virtu,[15] a system developed by University of British Columbia researchers Luanne Freund and Kristof Kessler (Fig. 2). Virtu took a task-based approach to expertise, exposing and allowing the user control over dimensions of expertise that are more or less desirable depending on the type of expert-finding task (Fig. 3).

Conclusion

The HCIR Symposium has not only become a leading venue for new research and development but also promoted a research and development program for information-seeking systems that places users first. While we expect to see continued advances in ranking, information extraction, and other algorithmic techniques, information access is fundamentally empowering human users.

While HCIR has focused on revolutionizing the field of information retrieval, its agenda applies to data science generally. Let us never forget as technologists that our job is to help people help themselves. While continuing to invest in novel techniques for managing and analyzing large data stores, data scientists must pay particular attention to interfaces that provide control and interpretability, ensuring that technology augments rather than replaces the role of human intellect. And we look forward to an increasing participation of data scientists in the HCIR Symposium!

Author Disclosure Statement

No competing financial interests exist.

References

1. Gourley S. A little story about Quid. 2011. Published online at http://seangourley.com/2011/10/a-little-story-about-quid/ (Last accessed on January 19, 2013).
2. Woods D. The Man-Machine Framework: How to Build Machine-Learning Applications the Right Way. 2012. Published online at www.forbes.com/sites/danwoods/2012/10/18/the-man-machine-framework-how-to-build-machine-learning-applications-the-right-way/ (Last accessed on January 19, 2013).
3. Engelbart D. Augmenting Human Intellect: A Conceptual Framework. Summary Report AFOSR-3233. Menlo Park, CA: Stanford Research Institute, 1962.
4. Saracevic T. (1997). Users lost: reflections of the past, future and limits of information science. *SIGIR Forum* 1997; 31.
5. Information Retrieval and Human Computer Interaction Workshop. Published online at www.dcs.gla.ac.uk/irhci/. (Last accessed on January 19, 2013).
6. First Workshop on Exploratory Search and HCI. Published online at http://research.microsoft.com/en-us/um/people/ryenw/esi/. (Last accessed on January 19, 2013).
7. Marchionini, G (2006). Toward human-computer information retrieval. *ASIST Bulletin* 2006; 32.
8. Many Eyes visualization platform. Available online at www-958.ibm.com/software/data/cognos/manyeyes/ (Last accessed on January 19, 2013)
9. The NYT Annotated corpus. Available online at www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2008T19. (Last accessed on January 19, 2013).
10. Apache Solr. Available online at http://lucene.apache.org/solr/ (Last accessed on January 19, 2013).
11. Simonite T. A search service that can peer into the future. *MIT Technology Review* 2010, available online at www.technologyreview.com/news/420424/a-search-service-that-can-peer-into-the-future/ (Last accessed on January 19, 2013).
12. Giles C., et al. CiteSeer: An automatic citation indexing system. In *Proceedings of 3rd ACM Conf on Digital Libraries* 1998; 89–98.
13. DeWitt D. and Stonebraker M. MapReduce: A major step backwards. The Database Column 2011. Available online at http://homes.cs.washington.edu/~billhowe/mapreduce_a_major_step_backwards.html (Last accessed on January 19, 2013).
14. Mendeley: Free reference manager and PDF organizer. 2013. Available online at www.mendeley.com/ (Last accessed on January 19, 2013).
15. Freund L, et al. Virtu – HCIR Challenge 2012 Expertise Finder. Available online at www.diigubc.ca/virtu (Last accessed on January 19, 2013).

Address correspondence to:

Dr. Daniel Tunkelang
Director of Data Science, LinkedIn
2029 Stierlin Court
Mountain View, CA 94043

E-mail: dtunkelang@linkedin.com