PerAda Magazine
towards pervasive adaptation

10.2417/2201006.003004

# Simplicity in search technology is the ultimate sophistication

**Dag Johansen and Bjørn Olstad**

*Development of next-generation, large-scale information-access systems requires a broad range of computer-science advancements in computing infrastructure and algorithmic user experiences.*

Search technology already impacts a wide range of industries, business models and social patterns. Most users consider it a commodity like electricity. It provides a vital service, is almost always available and reliable enough for frequent use, which in turn requires very little technical knowledge. Despite a simple and familiar interface, complex software and hardware hide below the surface. Major Internet search engines, for instance, link hundreds of thousands of computers to serve thousands (if not millions) of users in parallel. Each user is given the illusion that his or her query is matched, in approximately a quarter of a second, against billions of documents located throughout the entire Internet. Providing such a daunting service draws from a wide range of disciplines within computer science and also from fields such as linguistics. At the same time, we foresee a major shift in how search will impact user experiences in general. The current Web-search pattern, i.e., using a query box and resulting in '10 blue links', will be complemented with a range of more contextual and conversational user experiences. Search enables real-time, algorithmic decisions as foundation for smart dialogues, where the user's intent and context can be effectively leveraged.

The Information Access Disruptions (iAd) project targets core research for next-generation scalable information-access systems. This term indicates an ambition to study access of any kind of data type, not just how search is provided today. Reflecting the technology's complexity and diversity, we are studying numerous research problems, such as how to provide schema-agnostic indexing services by fusing structured, unstructured and multimedia content. We are also developing scalable and fault-tolerant system architectures, including data-processing and mining platforms to capture and extract knowledge from high-speed data streams. Third, we are working towards extreme-precision solutions for access to multimedia content (including social networks) with recommender functions. A fourth problem of interest is to discover, document and understand how advances in information-access technology can lead to disruptions for technology providers and users. The list of potential research problems to tackle is long and ultimately driven by a new breed of contextual user experience across PCs, mobile devices and TVs.

***Figure 1.*** *(© Gunnar3000 | Dreamstime.com)*

We are building and experimenting with concrete prototypes targeting a data-centre environment. We are interested in the full software stack from the operating system to disruptive applications. Rad Labs[1] at Berkeley (California) has a similar focus, although its work does not address the bottom of the stack. Along with several other research groups,[2,3] we are investigating alternative operating-system architectures to better explore multicore systems. We have developed Vortex, a brand-new kernel that provides strong performance-isolation properties, for which implementation of fine-grained accounting and scheduling of system resources is key. On top of this isolation kernel, we are investigating light-weight approaches to virtual machines.

Higher up in the stack, we have developed alternative analysis frameworks to the popular MapReduce paradigm.[4] One such system is Oivos,[5] a high-level declarative-programming model and its underlying runtime, where computations can span multiple heterogeneous and mutually dependent data sets. Second, Cogset[6] represents a next step in the process of re-thinking the original MapReduce architecture. The traditional loose coupling between distributed file systems and the MapReduce processing engine leads to poor data locality for many applications. Hence,

*Continued on next page*

Cogset fuses reliable storage and parallel data processing into a single system that ensures good data locality and performance.

We are also investigating novel application ideas such as Davvi,[7] a video search that composes and streams highly personalized videos in response to traditional queries. Davvi does more than just returning the most relevant results obtained from an inverted index. It extracts video segments embedded in larger videos across huge archives and composes them into a smooth and personalized play-out. We have built concept demonstrators around football and extensions of Microsoft's information-access solutions.

Users of large-scale search systems are supposed to be provided with the illusion of a simple and transparent service. However, complex architectures and their implementations reside behind the top level. The iAd project is an attempt to understand and develop novel solutions that contribute towards the next-generation information-access solutions.

*iAd has been designated and is partially funded by the Research Council of Norway as a Centre for Research-based Innovation. The project draws upon a number of international partners and is directed by FAST Search & Transfer (now a Microsoft subsidiary), in collaboration with research groups at Cornell University, Dublin City University (Ireland), BI (Bedriftsøkonomisk Institut: Institute for Business Economics) Norwegian School of Management and the universities in Tromsø, Trondheim and Oslo (Norway). Accenture is the main industrial collaborator, although several smaller Norwegian companies are also involved in this emerging industry cluster.*

## Author Information

**Dag Johansen**
University of Tromsø
Tromsø, Norway

Dag Johansen is a professor. He is currently on sabbatical leave at Cornell University. His research interests include designing and building large-scale distributed systems.

**Bjørn Olstad**
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway

Bjørn Olstad is a 'distinguished engineer' with Microsoft and adjunct professor at the NTNU. His research interests include information access, search and image analysis.

**References**

1. http://radlab.cs.berkeley.edu/wiki/RAD_Lab Reliable Adaptive Distributed Systems Laboratory. Accessed 19 May 2010.
2. S. Boyd-Wickizer, H. Chen, R. Chen, Y. Mao, F. Kaashoek, R. Morris, A. Pesterev, L. Stein, M. Wu, Y. Dai, Y. Zhang, and Z. Zhang, *Corey: an operating system for many cores*, **Proc. 8th Symp. Operat. Syst. Design Implement.**, pp. 43–57, 2008.
3. A. Baumann, P. Barham, P. E. Dagand, T. Harris, R. Isaacs, S. Peter, T. Roscoe, A. Schüpbach, and A. Singhania, *The multikernel: a new OS architecture for scalable multicore systems*, **Proc. 22nd ACM Symp. Operat. Syst. Princ.**, pp. 29–44, 2009.
4. J. Dean and S. Ghemawat, *MapReduce: simplified data processing on large clusters*, **Proc. 6th Symp. Operat. Syst. Design Implement.**, pp. 137–150, 2004.
5. S. V. Valvåg and D. Johansen, *Oivos: simple and efficient distributed data processing*, **Proc. 10th IEEE Int'l Conf. High Perf. Comput. Commun.**, pp. 113–122, 2008.
6. S. V. Valvåg and D. Johansen, *Cogset: a unified engine for reliable storage and parallel processing*, **Proc. 6th IFIP Int'l Conf. Netw. Parall. Comput.**, pp. 174–181, 2009.
7. D. Johansen, H. Johansen, T. Aarflot, J. Hurley, Å. Kvalnes, C. Gurrin, S. Zav, B. Olstad, E. Aaberg, T. Endestad, H. Riiser, C. Griwodz, and P. Halvorsen, *DAVVI: a prototype for the next generation multimedia entertainment platform*, **Proc. 17th ACM Int'l Conf. Multimedia**, pp. 989–990, 2009. doi:10.1145/1631272.1631482