

From: Espen Andersen
To: Students interested in writing master thesis with iAD
Date: March 25, 2008
Subject: List of projects iAD subtask 5

The following is meant as an idea generator – a listing of fairly concrete possible projects that could be done as individual pieces of research under the iAD lab, by researchers and students associated with the lab, the project or one of the partners.

In addition to this list, there are possibilities to write under subtask 5.3 (software economics) and 5.4 (search industry in Norway). And, by all means, feel free to suggest your own themes, companies, and projects.

Cheers,

Espen Andersen
(on behalf of the iAD team)

Search as a disruptive innovation

We have identified 6 possible disruptive implications of search technology:

Disruption I: Search as *new standard interface*, obviating or at least reducing the need for categorization and structure as a tool for information access.

Possible theses:

1. *Implications of findability: How findability changes media language.*
A number of writers have pointed out that newspapers and magazines have reduced the use of irony or metaphors in their titles, because a large proportion of their readers arrive via search engines rather than via the front page or other forms of categorical interface. A research activity here would be to use the material available through Schibsted's media archive and FAST analysis tools to investigate and measure this effect – is it real, and to what extent is it effective? Possible researchers could be students from BI's Institute for Language and Communication, perhaps students taking their master's in business journalism. A practical implication could be to compare newspapers with or without an internet presence.
2. *Search technology and public services: Simplification organizations by simplifying information access?* Public services in many countries are charged with becoming simpler to use and more effective in the face of demographic and political changes, but are limited by organizational structures and sector-based traditions. To what extent can search technology be a factor in effectively making information available and simplifying the user's navigation and interaction with the authorities (a question, of course, which can also apply for private companies).
3. *Personalized search.* What are the pros and cons of personalized search: On one hand, the perceived result quality to the individual user increases. On the other hand, the uses of search as storage is limited, in that the results will vary between two people.

4. *Search as an automatic interface/portal builder.* A number of organizations, among them newspapers and financial service companies, are using search technology to create the pages used by their customers and employees. What are the threats and opportunities in doing this? What are best practices?

Disruption II: Search as *enabler of new information access*, where certain kinds of search technologies allow access (as well as monitoring and coordination) into previously unconnected domains, such as audiovisual material

Some possible theses:

5. *Automated rich media monitoring:* Currently, rich media monitoring (TV, audio, video) is to a large extent done with a combination of manual monitoring and metadata analysis. If this can be automated in some form, the costs of media monitoring will be dramatically lowered, and the capability for monitoring increased – but most likely at lower quality. What are key areas for automated media monitoring, and how can new technologies be brought to bear on it?
6. *Retriever as a disruptive company in media monitoring.* Retriever was started as an automated media monitoring company. Initially it sought to sell its technology to traditional media monitoring companies, primarily in the newspaper clipping market. Immature technology and lack of interest forced Retriever to set up its own service tracking web content, an area where human monitoring had less advantage. Over time, the technology improved and more newspaper content was available in digital form, enabling Retriever to gradually out-compete the traditional media monitoring companies in this area. Two new markets – video/audio monitoring and market monitoring – are coming up as candidates for automation. This could be researched as a case study to see to what extent it matches other cases of non-use disruptive innovation.
7. *Metadata vs. intent tracking: Issues of usefulness.* Currently two approaches vie for dominance in analyzing rich media: Metadata generation and analysis (as is done manually with the music service Pandora and partially with the MPEG7 media file standard) and intent tracking or collaborative filtering (as is done by the music service LastFM.com, and, for that matter, Amazon.com.) What are the implications and differences, from a business and user perspective, of these two techniques?
8. *MPEG7 as a disruptive technology: Impacts for artist brand equity.* MPEG7 is a new format for sound and video files which stores metadata (such as music genre) as part of the file format. Genre data and other metadata enables automated (or, at least, automatically configured) music selection. Currently, music consumers choose music based on artist brand name and music categories. What are the implications for artist marketing of these new music formats?
9. *Search and academic publishing:* Academic databases face competition from open content sites such as Google Scholar, which provides a cheap and useful search capability with clearly worse quality than what is available through the database-oriented search available through university libraries. At the same time, academic libraries are loath to pay the stiff fees imposed by the academic databases, when most of the material in them is produced by, yes, unpaid academics. A number of themes come up in this market development

1. Google Scholar vs. academic databases: What jobs to users hire these services to do?
 2. Self-publishing by academics: Many research funding institutions now require research results (including published articles) to be made generally available within a certain time of publication. Many academics self-publish their articles on their own or through their research institutions. How will this affect the academic journals and traditional publishers?
 3. Search technology use in knowledge building (e.g., Zotero).
 4. Defensive strategies from existing vendors, such as [Elsevier](#) Science publishing, which use search technology to develop a search experience more suited to the needs of an academic community.
10. *How can companies use new forms of search technology to find out what their customers think of them?* As many companies have found out, what is being said about them in various Internet forums, from blogs to discussions groups to newsletters, can be very important both as a threat (witness the hit Dell took to its stock price when complaints about their customer service hit the blogosphere in 2004) as well as a possible source of new product ideas or new marketing strategies.
11. *How can companies use search technology to improve customer interaction?* The use of search technology in call center quality advance and control. Companies today are using search technology in combination with speech-to-text applications to monitor conversations between customer service representatives and customers both for quality of the conversations themselves as well as clues to trends in customer issues. How is this done, what are the implications of using this technology, and how useful is it?

Disruption III: Search as the *basis for (disruptive?) business models.*

Possible themes:

12. *The electronic advertising marketplace: Matching search intent to ad supply.* Google's Adwords (auctioning off search terms) and AdSense (placement of context-specific ads on webpages) has garnered the company a dominant market share in a very rapidly growing market. How does this model work, what are its weaknesses and strengths?
 1. One example: AdSense and content: Forcing specialization? Context-specific ads work best in focused publications, meaning that publications with more numerous or more general themes fare worse, since advertising precision can be reduced. This is not a new phenomenon: The excellent computer magazine BYTE shut down in the mid-90s because advertisers were more interested in placing their ads in specialized magazines such as PC Magazine or MacWorld. Does this effect happen in online advertising as well, or does the ability to tailor ads not only to the content, but to the individual reader change this dynamic?
13. *Use of search and text analysis on customer service interfaces.* A number of companies are beginning to use search technology to analyze their customer interface as well as their reputation. Speech-to-text translation coupled with text analysis tools allows companies to determine how well their customer service interface works.

14. *e-WOM – (e-Word of Mouth) uses of search technology for company and product reputation analysis.* Analysis of Internet communities (blogs, online communities, multi-user games) allows companies both to discover what is being said about their products and services as well as who says it. By monitoring (through search technology) and selectively responding to discussions in these communities, companies can influence perceptions about their products and services, allowing more precision and more rapid reactions than with traditional techniques such as surveys and focus groups.
15. *Who determines what is important?* Search technology and national interest: Where you search determines what you see: There is currently little overlap between search engines, in the sense that they will display relatively disparate results based on the same search terms. A number of regulators and cultural critics are skeptical to search technology, seeing a danger of outsourcing much of the determination of what is important and relevant. Global search engines can become local much faster than local search engines and other forms of specialized knowledge can influence the global. How to deal with this?

Disruption IV: Search as *alternative interface to structured data*, where search technology can decompose the functionality of database software, reducing their role from overall information management to mainly maintaining transactional integrity

16. *Is search really faster?* A central point in the marketing of search engines has been that a search engine is much faster and easier to use for the end user (compare Google Scholar vs. any number of academic databases, for instance.) But how true is this claim, and what downsides are there in an enterprise setting? (Bjørn Olstad)
17. *How can we speed up configuration of search engines?* Configuration of search: Cost and time estimation? How perfect? Cost implications of refining search results to approach database functionality.
18. *How does search change investigative work, such as journalism, insurance investigations and police investigations?* New versions of search technology interfaces allow the end user to sift through data using a graphical interface, a technology that may be at its most useful when doing work that is investigative in nature and involves a great deal of information – such as police investigations, or searches for insurance fraud. How do these new technologies help investigative work?
19. *What drives adoption of search technology?* Even with excellent interfaces and access to enormous amounts of data, search technology adoption will vary between users and organizations – some will use it a lot, others not? Which factors account for differences in search technology adoption?

Disruption V: Search as a *disruptive element in information and enterprise architecture*, where search technology increasingly will replace keys and codes as linkages between information elements, eating into the role now filled by various kinds of middleware

20. *Architecture implications of search implementation* – is it really true that you can just put search in without changing your architecture and your organization? What happens as the use of search increases and applications that make use of search technology automatically are implemented?

21. What is the best way of organizing search implementation and operations refinement? In particular, how do you structure 1st, 2nd and 3rd line support and development? How does this organization differ from other applications?

Disruption VI: *Search as a disruptor of power relations in social networks.* Power in social networks is determined by measures such as centrality, number of connections and to what extent a node in a network has connections that encompass *structural holes* (Burt 1992). Currently, social networks are being established on the Internet in record numbers, most of them “walled gardens” in the sense that each network only provides access to its own participants. Over time, it is to be expected that various standards¹ for network and connection descriptions and communication will emerge outside these standards, and that search technology will play a large role in establishing them.

22. Searching and communicating in Nettby. Nettby is one of Norway’s largest online communities, used by all kinds of groups for interaction and discussion. A possible thesis is a study of communication patterns and search use in Nettby.no – how do people use an online community, and what implications do these communities have for media companies? (Collaborator: Schibsted, Accenture)
23. LinkedIn – an online community for professionals. LinkedIn is a well established online community that focuses on professional relations, shaped to support loosely coupled professional network relations. It prides itself on leaving its users relatively alone. How do people use this service, and how has use changed over time? What are the implications of services like LinkedIn for professional job and consulting markets?
24. The sociologist Ronald Burt created a number of mathematically oriented measures for network attributes. How does the availability of search technology change these measures?

¹ In fact, Google has proposed one such standard already, the Social Network API, which proposes search-related descriptors using the “rel” parameter in http links (rel=’me’; rel=’friend’, much as the company previously established the “rel_nofollow” for authors wanting to link to sites without increasing their GoogleRank.)