

On the utility of HCI practices in improving a commercial information retrieval system

Matt Honeycutt, MS.
matt.honeycutt@inradllc.com

James Kolpack, BS.
james.kolpack@inradllc.com

Nathan Honeycutt, BB.
nathan.honeycutt@inradllc.com
InRAD, LLC, 310 East Broad St., Suite B
Cookeville, TN 38501, United States

Douglas Talbert, PhD.
dtalbert@tntech.edu
Department of Computer Science
Tennessee Technological University
Cookeville, TN 38503, United States

Abstract

Human-computer interaction (HCI) issues are an important but often overlooked part of software application development. InRAD's Automated Knowledge Discovery System (Basu, Merrel, Fellbaum, Talbert, Kolpack, Honeycutt, Alonso, & Bloom, 2005) featured numerous Web 2.0 and Web 3.0 components under the hood, but users of the system frequently reported difficulties in actually leveraging the power and capabilities of the system. By partnering with a team of graduate students studying HCI from a local university, InRAD was able to identify and address a wide array of usability issues and improve user satisfaction. The HCI processes that were employed as well as the result of those efforts, rebranded as the InSpire system, are described in this case study as a guide for others to follow.

Keywords: Information retrieval, prototyping, human-computer interaction, knowledge discovery, InSpire

1. INTRODUCTION

InRAD LLC is a small software company focusing on developing next-generation information retrieval and knowledge management systems utilizing ontologies, semantic web technologies, machine learning, and artificial intelligence. InRAD has always adhered to many industry best-practices (continuous integration, test-driven development, etc.), but usability has been one area of weakness. Prior to our

experiences described in this paper, InRAD had no formal process for assessing usability or any interface design principles to guide development.

In 2005, InRAD attempted to go to market with the Automated Knowledge Discovery System (AKDS) (Basu et al., 2005). While many years were invested in its development and it featured a very sophisticated suite of technologies under the hood, very little time and effort

had been spent on the user interface. InRAD entered the commercial market with what amounted to little more than a testing framework as the entire user experience. Though it was functional, potential customers were turned off by this complicated user interface, and InRAD struggled to gain a market foothold.

Driven by the complaints from users, InRAD sought help from graduate students enrolled in a Human Computer Interaction (HCI) class at Tennessee Technological University. The students adopted InRAD's AKDS system as a class project. Their contributions to the project are discussed in Section 3.

The remainder of this paper is organized as follows. Section 2 describes the original AKDS system, its technical underpinnings, and the AKDS user interface. Section 3 describes the process used by the student HCI team in evaluating the AKDS system. Section 4 explains how InRAD leveraged the recommendations from the student HCI team in developing its next-generation InSpire interface. Section 5 describes user impressions of the new system, presents future work, and concludes the paper.

2. BACKGROUND

The AKDS is a hybrid knowledge management and information retrieval system. It is powered by domain ontologies and crawls both surface and deep web (Bergman, 2000) content to find information of interest to its users. While the underlying components were based on extensive research and testing, the user interface was relegated to an after-thought, and no formal design principles were adhered to during its construction.

The AKDS interface

The original AKDS user interface was called the "End-User Subsystem," (EUSS) a name that indicates the diminished level of concern was given to the user experience while the AKDS was in early development. This is understandable given that the original project did not include funds for a user interface or commercialization efforts. Instead, the user interface was developed piecemeal, each feature being added as it was necessary to demonstrate some backend functionality. InRAD had no design strategy before starting UI development. Each developer would implement new features using an ad hoc design, and then wait to see if any users complained loudly enough to justify a

more extensive redesign. Furthermore, InRAD did not identify target users before beginning development. The target user became the developer's responsible for the feature in some cases. In others, the design specifications for the interface came from corporate-level managers, none of whom were target users for the system itself.

After the first public release of the AKDS, the company continued to refer to the interface as "End-User Subsystem" even to the customers, reflecting a refusal to acknowledge that the user interface was the first point of judgment of the AKDS. No funding, internal or external, was made available to correct any of the problems caused by the necessarily poor development strategy. New features continued to be added on top of an already failing framework, and customers were given the same user interface that InRAD developed as a testing framework.

For all of the failings, the original EUSS did provide a great deal of functionality by the time InRAD retired it from active development. At the core, the interface provided a document search capability based on a large number of fields. Users built queries by repeatedly selecting these fields and criterion using a complicated interface. This led to a powerful search that allowed users to be very specific about what they were seeking. However, because this search continued to have new options bolted onto it, it quickly became too complex for most users. The document search did not have any real Boolean query support. Users could specify "AND" or "OR" on the whole set of criteria only. This put the focus on telling the system what a user was interested in, instead of making the user spend time fighting query syntax. However, it was difficult to narrow in on a particular set of results. Worse yet, the use of OR combinations led to strange or invalid results more often than not, rendering it unusable.

Topic tree browsing was separated from the search of documents. This functionality was originally intended as a starting point for allowing customers to maintain their own topic trees over time, but was never fully realized. Because this feature was added at a different time than document search, by a different developer, users quickly discovered that lists of documents mapped to nodes on a topic tree differed depending on whether the user browsed to the list or performed a search on

the node. It was difficult to maintain consistency with no master plan driving development.

InRAD implemented a framework to allow for fully-contained pluggable reports. Customers could order new report options at any time which could be deployed into individual systems without having to fracture the code base. This allowed virtually unlimited flexibility when designing reports.

Every possible piece of information in the EUSS was made a clickable link to enable "drill around" navigation. Users could click through any piece of information to see more. They could go from a set of summary results, to a particular document, to the organization that funded the activity described, to a list of other activities funded by the same organization. This allowed the user to truly explore the data set and discover new information. However, with a small subset of users this functionality became objectionable when they were unsure whether the "drill around" view was a more specific or more general view on top of the previous view. Attempts to hack in a set of rules after the fact were inconsistent and proved to be unsuccessful in resolving the issue.

Users were able to save queries of interest to them for the system to run automatically on a set schedule. The user would then be notified of any new results via email. Because of the nature of the repository population, this functionality was seldom used. From a design standpoint, the sole "voice of the customer" had the team implement a complex method of scheduling queries, allowing users to specify any number of days of the week, along with where in the month those days fell. As a result, a user was able to specify that their query was only run on the second Wednesday and last Saturday of the month. While this capability was powerful, it was not a feature that any actual user needed.

Domain ontologies

The shortcomings of the interface had the effect of obscuring the underlying strengths of the system. The majority of the effort that went into designing and developing AKDS occurred in the semantically-based knowledge processing. One of the primary focuses was in creating and integrating a knowledge base, or ontology, to enable the information extraction and retrieval processes.

An ontology provides a formal structure for modeling aspects of knowledge in a domain. The knowledge is primarily organized using concepts and the semantic relations between those concepts. Concepts define both the abstract and concrete ideas and objects in a domain. Relations are defined between two concepts – a "Photoelectric cell is a type of Sensor", for instance, defines a hyponym relationship between Photoelectric cell and Sensor. By defining this domain knowledge into a formal structure, it can be computationally harnessed for information processing.

AKDS leverages the domain ontology to enable semantic-based information retrieval. Conventional keyword-based search approaches will not retrieve a document if it only references synonyms of the given keyword. Keywords can often be used in more than one sense, which leads to the search returning results from an unrelated context. Concept-based searching mitigates these issues by indexing documents using highly relevant concepts rather than plain text.

Document Concept Identification

Before any conceptual-based queries for documents can be executed, the document's concepts must first be identified. The first step is to extract relevant concepts from raw text. The Concept Identification process starts with a classic natural language processing implementation consisting of word tokenization, stemming, part of speech tagging, and named entity recognition. To identify the concepts, a series of text matching and filtering steps are performed.

The first step in extracting concepts is to locate text which explicitly references the ontology concepts by name, synonym or acronym. Terms and phrases extracted from a concept's definition are also compared with the document's text to find potentially relevant concepts. The collection of candidate concepts is analyzed with regard to their interrelationships in the ontology (Magnuitman, Menczer, Roinestad, & Vespignani, 2005). Related concepts which are not explicitly referenced in the document are included for consideration.

After extracting all the potentially relevant matches, a filtering process removes the irrelevant concepts and assigns a relevancy score to the remainder. The scoring mechanism is based on a number of features gathered from

Table 1. Concept scoring features

A conceptual variant of term frequency, inverse document frequency.
The inter-relationship strength between a concept and the other concepts referenced in the document.
An information content measure of the concept derived from hit counts from a large web engine's search results (Etzioni, Cafarella, Downey, Popescu, Shaked, Soderland, Weld, & Yates, 2005).
A rule-based system to override and nudge results according to Subject Matter Expert review.

concepts and how they are used in the document. Several examples of these features are described in Table 1.

Organization disambiguation

From the collected documents, the AKDS gathers information about organizations such as who published the document, which organization sponsored the project as well as any other organizations which are referenced inside the document text. The ways that authors can refer to an organization, or any named entity, can vary drastically from case to case. Derivations of the full organization name by using abbreviations or acronyms are quite common. Many times the shortened name will conflict with organization names, creating an ambiguous scenario. For instance, NSF could refer to either National Science Foundation or NSF International.

To disambiguate, AKDS utilizes a statistical-based pattern matching algorithm using n-grams of the known organization names to populate a Markov chain structure. Shortened organization names can then be probabilistically matched to the most likely full version, or if no match exists, can be tagged for review by a human expert at a later time.

Additional processing

Documents which were written as part of a single project are grouped together to allow the user to evaluate the entire "activity" as a whole. To find these groupings, several properties of the documents are compared. If a group of documents share overlapping dates, the same funding organization, and a similar conceptual model, then it is likely they all come from the same project and can be clustered as a single activity.

Table 2. Usability heuristics (Nielsen, 2005)

U1. System status should always be visible; provide feedback in a timely manner
U2. Conventions, terminology, and concepts should match the users'
U3. Users should be in control and free to explore without risk
U4. Follow conventions, both platform and otherwise
U5. Prevent errors from occurring wherever possible
U6. Rely on <i>recognition</i> rather than <i>recall</i>
U7. Cater to both experienced and novice users
U8. Strive for an aesthetic and minimalist design
U9. Help users recognize and recover from errors
U10. Provide useful, easily-accessible help and documentation

Documents brought into AKDS from external sources are not all worthwhile. In an effort to minimize the amount of noise, a filtering mechanism to remove any advertising or other "spam"-like documents has been built (Roush, 2005). A hand-selected collection of these are then used to train a classifier.

Many times a single document may be retrieved from more than one source. To identify these copies, a hash using (Chowdhury, Friender, Grossman, & McCabe, 2002) is generated and compared with the system's previously added documents. Any near duplicates are grouped together as one.

3. THE HCI PROCESS

The student HCI team followed a simple process in evaluating the AKDS user experience. Their methods included analyzing the existing interface for usability issues, comparing the system to similar systems from other vendors, evaluating current research on the user interfaces in information retrieval systems, low-fidelity mockups, and user interviews. This section describes the process in detail as well as lessons learned through its application.

Analyzing the existing system

Students were given user accounts to a production AKDS system. They were also given a brief training session on how to use the system as well as a summary of the technical capabilities.

ties of the system. The students evaluated the system in terms of ten design heuristics (Nielsen, 1994) listed in Table 2. These heuristics, while not all-encompassing, provide a reasonable starting point for good usability experiences and should be used as guidelines for creating software interfaces.

The student team found that the AKDS interface violated the majority of these heuristics. The major issues found included the following (the violated heuristics are listed in parenthesis):

- The system failed to respond quickly to many queries. Often the system would return no results, but re-running the same query later would return a full result set. (U1, U5, U9)
- The system used terminology that was not intuitive and did not match the user's vocabulary. The search interface differed drastically from the interfaces of other systems. While powerful, the search interface was complex and could be overwhelming to a novice user. (U2, U4, U7, U8)
- The system did not support undo and redo semantics. For example, users could delete a saved search, but there was no mechanism to recover from this action (U3).
- Help was very limited. All documentation was contained in a single PDF file. While this PDF was easily accessible, it still required the user to navigate through its content in order to find information relative to their current task. (U10)
- The system inconsistently used links and buttons to denote actions. The two navigation mechanisms were used seemingly interchangeably and without adhering to any convention. (U4)
- Accessing core system functions required too much navigation. Users must remember the correct series of actions to perform. The path to these functions was not immediately intuitive. (U6, U8)

It was immediately clear from this heuristic assessment that InRAD had to address numerous usability issues in the AKDS interface.

Comparison to similar products

The students also compared the AKDS system to other similar software products. In addition to highlighting violations of heuristics U2 and U4, this comparison identified features that

were becoming common in competitor products but not yet implemented in the AKDS system. These features included hit highlighting, separate basic and advanced search interfaces, copy-and-paste ready URLs for search results, richer subscription options than those supported by AKDS, powerful Boolean query syntax (which has been shown to improve the quality of query results in some cases) (Croft, Turtle, & Lewis, 1991), and social features such as tagging (Morrison, 2007) and collaborative filtering (Wei, Yang, & Hsiao, 2008).

The students were also able to identify common conventions for both user interface design and behavior. These observations were leveraged later in the creation of low-fidelity mockups for a revised interface.

Evaluating existing research

New developments are occurring at a frantic pace in the realm of information retrieval systems. Many of these developments are described in conference papers and research journals. The students surveyed literature and identified recent advances in usability related to information retrieval. The largest contribution from this effort was faceted search and browsing (Hearst, 2006; Käki, 2005). Work in Hearst (2006) showed that facets were well-received by users and enabled them to find information faster than keyword queries alone. Faceting became an important feature of the mockups created by the student team.

Low-fidelity mockups

After identifying usability issues in the current interface, comparing the interface to similar solutions, and conducting a literature survey to find emerging techniques, the student team was ready to create mockups of a revised interface. Though many techniques for creating software interface mockups exist (Arnowitz & Berger, 2006), the students opted to create simple mockups using the Paint application that is bundled with Microsoft Windows. Such primitive mockups force reviewers to focus on the core usability of the system instead of focusing on secondary issues such as fonts and colors. Primitive mockups are also easier to create and as likely to elicit feedback from users as more polished mockups and prototypes (Virzi, Sokolov, & Karis, 1996).

The students completely redesigned the homepage of AKDS, opting to immediately expose the search interface to users. The search in-

AKDS Search:

Intelligent Keyword

Date

Start:

End:

Author

Organization

Funding

Figure 1 – Advanced search mockup

terface itself was also redesigned. Gone was the complex query building UI from the original AKDS system. Instead, users would be greeted with a single text box for queries and a single action button to begin their search. The result is shown in Figure 2.

As stated in heuristic U7, interfaces should cater to both novice and experienced users. The new search interface would certainly meet the needs of novice users, but would likely not provide advanced users with the power and flexibility that they were accustomed to. The students elected to add a separate advanced search interface for these users. This interface is shown in Figure 1. The new advanced search interface was still greatly simplified from AKDS's original search interface, and the students recommended that it include a feature not present in the original: the ability to manually craft advanced queries by embedding operators such as AND, OR, and grouping. To help users discover this functionality on their own, the query interface would automatically display the corresponding advanced query syntax as users entered values for the various criteria available on the advanced search page.

The search result interface also received a major overhaul by the students. In the mockup, the search bar remained visible at the top of all the search results, enabling users to change or

AKDS Search:

[Advanced Search](#)

Figure 2 - Homepage mockup

refine their query directly from the results. This is in contrast to the AKDS interface that required users to navigate to a new page in order to alter their queries. The revised interface also featured dynamic faceting over the fields and categories computed by AKDS's backend processes. These facets would allow users to "drill down" into their search results and find relevant results more quickly.

User Feedback

After completing the mockups, the students worked with InRAD to conduct reviews of the mockups with project stakeholders and selected users. The feedback was unanimously positive. The stakeholders all seemed to immediately recognize the utility of the team's assessment and were able to evaluate the mockups in terms of usability instead of less relevant concepts such as branding, colors, and fonts.

Users did express concern over the apparent omission of AKDS's reporting capabilities from the mockups. Many stakeholders believed this to be a key distinguisher between AKDS and its competitors. Despite this omission, InRAD gained tremendous value in the assessment. The developers determined that fixing the identified issues in the existing system would not be feasible. Instead, the developers lobbied for a complete rewrite of the interface.

4. THE ROAD TO INSPIRE

The decision to scrap an existing system and start from scratch should never be made lightly. Such rewrites are expensive and may still fail to achieve the desired outcome. In the case of AKDS, all project stakeholders agreed that the existing interface was a major weakness and merited proactive attention. InRAD continued to use the lessons learned from the student HCI team as their feedback was evolved and matured into the final product.

Additional Mockups

The mockups created by the students described a revised interface at a high level, but they only addressed a subset of the capabilities exposed by the underlying AKDS system. InRAD developed new mockups based on feedback from users as well as additional research and development. To better facilitate the creation of mockups, InRAD used Balsamiq Mockups (2010). This tool allowed us to not only create but also easily maintain and refine low-

fidelity models of the interface. The built-in library of common UI controls made creating mockups far faster than would have been possible using a drawing or paint tool.

InRAD revamped and expanded the library of mockups created by the students. The mockups were continually evaluated by all stakeholders, and the process actually led to the creation of new innovative features. Among these came *context-sensitive search* which is a system to allow users to quickly change their searcher roles within the system. Another feature designed was a graphical reporting framework to quickly view trends in the extracted knowledge over any search result set.

The drastic changes and rapid mock-prototyping that the process enabled led to a system that would look and behave quite differently from the previous AKDS system. Because of this, InRAD chose to rebrand the Automated Knowledge Discovery System as the InSpire Search System.

Implementation

Once consensus had been reached that the user experience depicted in the mockups met the goals for usability, simplicity, and innovation, the InRAD developers began the difficult task of translating non-functional mockups into working interfaces. An incremental top-down approach was used wherein the mockups were first converted to non-functional views consisting of HTML and CSS. Once the view matched the corresponding mockup, the underlying logic to bind the view to the backend InSpire services was implemented. This approach allowed testers to evaluate the look-and-feel and usability long before the developers had finished adding all the logic necessary to power the system. Through this early testing, InRAD was able to make minor course corrections as needed and refocus efforts on the features that were most important to InSpire's users.

Some of the weaknesses identified by the HCI team extended beyond the interface and into the underlying components of AKDS. For example, both query execution speed as well as limitations in query syntax were identified as usability problems, yet neither can be addressed solely through interface changes. Instead, InRAD chose to abandon their custom query processing and document ranking APIs and instead adopted the open-source Apache Lucene library (Gospodnetic & Hatcher, 2004). Lucene has been used successfully in a range

of commercial applications and is flexible, fast, and powerful. By switching to Lucene, InRAD was able to reduce average query times from minutes to fractions of a second. Lucene also supports complex queries, which addressed the query syntax limitation issue.

5. RESULTS

The InSpire Search system is currently undergoing extensive internal and limited external beta testing at InRAD. Informal evaluations are conducted frequently to ensure that no usability issues have arisen for the testers. In the future, InRAD plans to conduct formal usability testing in order to quantify the improvements made from the old AKDS interface to the new InSpire interface. The transition from the original AKDS interface to the final InSpire interface is depicted in Figures 3 and 4.

The use of HCI heuristics and mockups enabled us to focus efforts on the functionality that was most important to the users first. By building InSpire incrementally, InRAD was able to identify problems with the user interface earlier and correct them before significant development effort had been wasted.

For software development teams building new user interfaces today, we strongly encourage adherence to a set of HCI design heuristics. We found the heuristics at (Nielsen, 2005) to be very useful, but any set of sound principles will do. Our experience has also demonstrated the utility of rapid, low-fidelity mockups. Through such mockups, user interfaces can be prototyped and refined with minimal effort.

6. ACKNOWLEDGMENTS

We would like to thank the HCI team from the Computer Science Department at Tennessee Technological University.

7. REFERENCES

We would like to thank the HCI team from the Computer Science Department at Tennessee Technological University.

Arnowitz, J., Arent, M., and Berger, N (2006). Effective prototyping for software makers. San Francisco, CA: Morgan Kaufmann Publishers.

Balsamiq Mockups (2010). Balsamiq Mockups Home. URL: <http://www.balsamiq.com/products/mockups>.

- Basu, C., Merrell, M.A., Fellbaum, C., Talbert, D., Kolpack, J., Honeycutt, M., Alonso, R., and Bloom, J. (2005). Automated Knowledge Discovery System: A Business Case Study in Ontology Development and Use. Proceedings of FOMI, Lago di Garda, Italy.
- Bergman, M. (2000). The Deep Web: Surfacing Hidden Value.
- Chowdhury, A., Frieder, O., Grossman, D., and McCabe, M. C. (2002). Collection Statistics for Fast Duplicate Document Detection. ACM Trans. Inf. Syst. 20, 2 (Apr), 171-191.
- Croft, W. B., Turtle, H. R., and Lewis, D. D. (1991). The Use of Phrases and Structured Queries in Information Retrieval. In Proceedings of the 14th Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Chicago, Illinois, United States, October 13 - 16, 1991). SIGIR '91. ACM, New York, NY, 32-45.
- Etzioni, O., Cafarella, M., Downey, D., Popescu, A., Shaked, T., Soderland, S., Weld, D. S., and Yates, A. (2005). Unsupervised Named-Entity Extraction From the Web: An Experimental Study. Artif. Intell. 165, 1 (Jun.), 91-134.
- Gospodnetic, O. and Hatcher, E. (2004). Lucene in action. Greenwich, CT: Manning Publications.
- Hearst, M. A. (2006). Clustering Versus Faceted Categories for Information Exploration. *Commun. ACM* 49, 4 (Apr. 2006), 59-61.
- Käki, M. (2005). Findex: Search Result Categories Help Users When Document Ranking Fails. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 131-140.
- Maguitman, A. G., Menczer, F., Roinestad, H., and Vespignani, A. (2005). Algorithmic Detection of Semantic Similarity. In Proceedings of the 14th international Conference on World Wide Web (Chiba, Japan, May 10 - 14, 2005). WWW '05. ACM, New York, NY, 107-116.
- Morrison, P. (2007). Tagging and searching: search retrieval effectiveness of folksomolies on the web. Masters Thesis, Kent University.
- Nielsen, J. (1994). Enhancing the Explanatory Power of Usability Heuristics. Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158.
- Nielsen, J. (2005). 10 Heuristics for user interface design. URL: http://www.useit.com/papers/heuristic/heuristic_list.html
- Roush, E. (2005). Document quality assessment for automated internet text mining. Masters Thesis, Tennessee Technological University.
- Virzi, R. A., Sokolov, J. L., and Karis, D. (1996). Usability Problem Identification Using Both Low- and High-Fidelity Prototypes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground* (Vancouver, British Columbia, Canada, April 13 - 18, 1996). M. J. Tauber, Ed. CHI '96. ACM, New York, NY, 236-243.
- Wei, C., Yang, C., and Hsiao, H. (2008). A collaborative filtering-based approach to personalized document clustering. *Decis. Support Syst.* 45, 3 (Jun.), 413-428.

Appendix

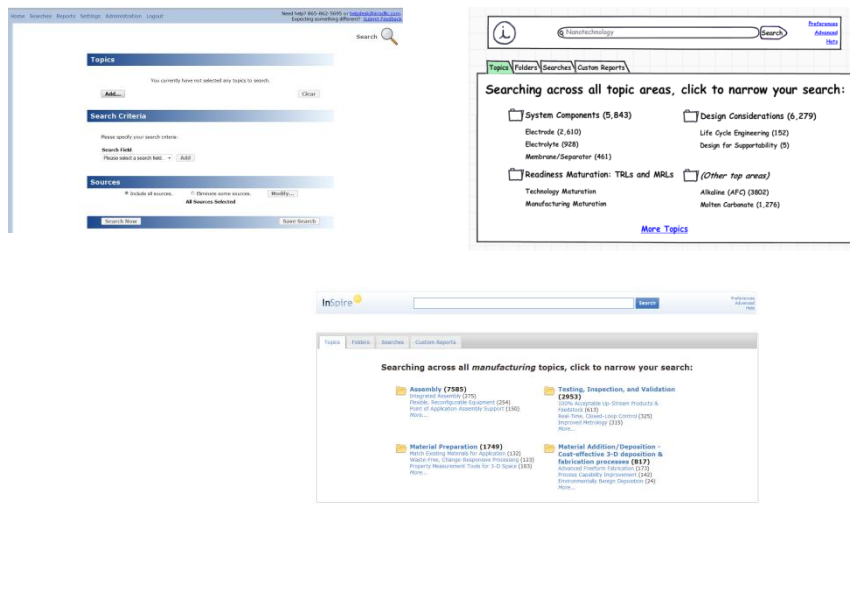


Figure 3 – Search interface before, mockup, and final result

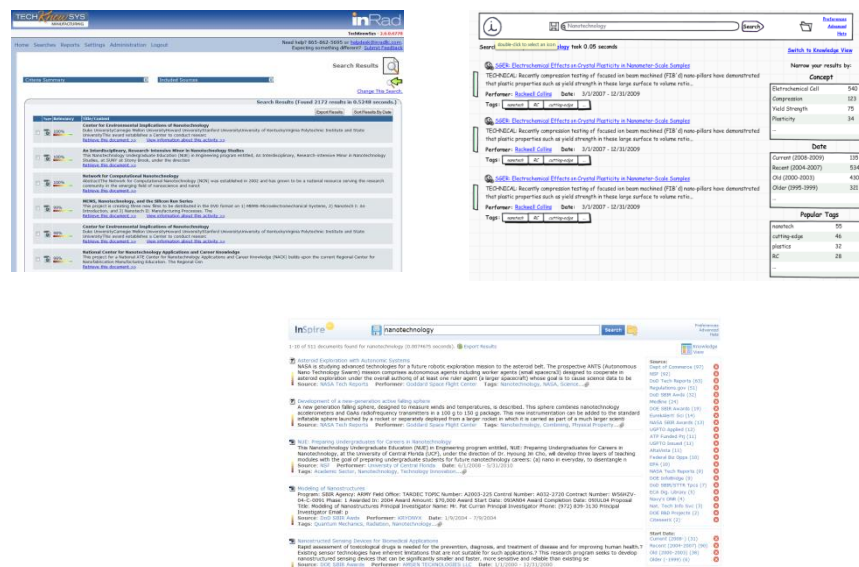


Figure 4 – Search results before, mockup, and final result