

Applying Hierarchical Task Analysis Method to Discovery Layer Evaluation

Merlen Prommann
and Tao Zhang

ABSTRACT

While usability tests have been helpful in evaluating the success or failure of implementing discovery layers in the library context, the focus of usability tests has remained on the search interface rather than the discovery process for users. The informal site- and context specific usability tests have offered little to test the rigor of the discovery layers against the user goals, motivations and workflow they have been designed to support. This study proposes hierarchical task analysis (HTA) as an important complementary evaluation method to usability testing of discovery layers. Relevant literature is reviewed for the discovery layers and the HTA method. As no previous application of HTA to the evaluation of discovery layers was found, this paper presents the application of HTA as an expert based and workflow centered (e.g., retrieving a relevant book or a journal article) method to evaluating discovery layers. Purdue University's Primo by Ex Libris was used to map eleven use cases as HTA charts. Nielsen's Goal Composition theory was used as an analytical framework to evaluate the goal charts from two perspectives: a) users' physical interactions (i.e., clicks), and b) user's cognitive steps (i.e., decision points for what to do next). A brief comparison of HTA and usability test findings is offered as a way of conclusion.

INTRODUCTION

Discovery layers are relatively new third party software components that offer Google-like web-scale search interface for library users to find information held in the library catalog and beyond. Libraries are increasingly utilizing these to offer a better user experience to their patrons. While popular in application, the discussion about discovery layer implementation and evaluation remains limited. [1][2]

A majority of reported case studies discussing discovery layer implementations are based on informal usability tests that involve a small sample of users in a specific context. The resulting data sets are often incomplete and the scenarios are hard to generalize.[3] Discovery layers have a number of technical advantages over the traditional federated search and cover a much wider range of library resources. However, they are not without limitations. Questions have remained scarce about the workflow of discovery layers and how well they help users achieve their goals.

Merlen Prommann (mpromann@purdue.edu) is User Experience Researcher and Designer, Purdue University Libraries. **Tao Zhang** (zhan1022@purdue.edu) is User Experience Specialist, Purdue University Libraries.

Beth Thomsett-Scott and Patricia E. Reese¹ offered an extensive overview of the literature discussing the disconnect between what the library websites offer and what their users would like.[1] On the one hand, library directors deal with a great variety of faculty perceptions, in terms of what the role of library is and how they approach research differently. The Ithaca S+R Library Survey of not-for profit four-year academic institutions in the US suggests a real diversity of American academic libraries as they seek to develop services with sustained value.[4] For the common library website user, irrelevant search results and unfamiliar library taxonomy (e.g. call numbers, multiple locations, item formats, etc.) are two most common gaps.[3] Michael Khoo and Catherine Hall demonstrated how users, primarily college students, have become so accustomed to the search functionalities on the Internet that they are reluctant to use library websites for their research.[5] No doubt, the launch of Google Scholar in 2005 was another driver for librarians to move from the traditional federated searching to something faster and more comprehensive.[1] While literature encouraging Google-like search experiences is abundant, Khoo and Hall have warned designers to not take users' preferences towards Google at face value. They studied users' mental models, defining it as "a model that people have of themselves, others, the environment, and the things with which they interact, such as technologies," and concluded that users often do not understand the complexities of how search functions actually work or what is useful about them.[5]

A more systematic examination of the tasks that discovery layers are designed to support is needed. This paper introduces hierarchical task analysis (henceforth HTA) as an expert method to evaluate discovery layers from a task-oriented perspective. It aims to complement usability testing. For more than 40 years, HTA has been the primary methodology to study systems' sub-goal hierarchies for it presents the opportunity to provide insights into key workflow issues. With expertise in applying HTA and being frequent users of the Purdue University Libraries website for personal academic needs, we mapped user tasks into several flow charts based on three task scenarios: (1) finding an article, (2) finding a book, and (3) finding an eBook. Jakob Nielsen's "Goal Composition" heuristics: generalization, integration and user control mechanisms[6] were used as an analytical framework to evaluate the user experience of an Ex Libris Primo® discovery layer implemented at Purdue University Libraries. The Goal Composition heuristics focus on multifunctionality and the idea of servicing many possible user goals at once. For instance, generalization allows users to use one feature on more objects. Integration allows each feature to be used in combination with other facilities. Control mechanisms allow users to inspect and amend how the computer carries out the instructions. We discussed the key issues with other Library colleagues to meet Nielsen's five expert rule and avoid loss in the quality of insights.[7] Nielsen studied the value of participant volume in usability tests and concluded that after the fifth user researchers are wasting their time by observing the same findings and not learning much new. A comparison to usability study findings, as presented by Fagan *et al*, is offered as a way of conclusion.[3]

RELATED WORK

Discovery Layers

The traditional federated search technology offers the overall benefit of searching many databases at once.[8][1] Yet it has been known to frustrate users, as they often do not know which databases to include in their search. Emily Alling and Rachel Naismith aggregated common findings from a number of studies involving the traditional federated search technology.[9] Besides slow response time, other key causes of frustrating inefficiency were: limited information about search results, information overload due to the lack of filters, and the fact that results were not ranked in order of relevance (see also [2][1]).

New tools, termed as “discovery,” “discovery tools,”[2][10] “discovery layers” or “next generation catalogs,”[11] have become increasingly popular and have provided the hope of eliminating some of the issues with traditional federated search. Generally, they are third party interfaces that use pre-indexing to provide speedy discovery of relevant materials across millions of records of local library collections, from books and articles, to databases and digital archives. Furthermore, some systems (e.g., Ex Libris Primo Central Index) aggregate hundreds of millions of scholarly e-resources, including journal articles, e-books, reviews, legal documents and more that are harvested from primary and secondary publishers and aggregators, and from open-access repositories. Discovery layers are projected to help create the next generation of federated search engines that utilize a single search index of metadata to search the rising volume of resources available for libraries.[2][11][10][1] While not systematic yet, results from a number of usability studies on these discovery layers point to the benefits they offer.

The most noteworthy benefit of a discovery layer is its seemingly easy to use unified search interface. Jerry Caswell and John D. Wynstra studied the implementation of *Ex Libris MetaLib* centralized indexes based on the federated search technology at the University of Northern Iowa Library.[8] They confirmed how the easily accessible unified interface helped users to search *multiple* relevant databases simultaneously and more efficiently. Lyle Ford concluded that the *Summon* discovery layer by Serials Solutions fulfilled students’ expectations to be able to search books and articles together.[12] Susan Johns-Smith pointed out another key benefit to users: customizability.[10] The *Summon* discovery layer allowed users to determine how much of the machine-readable cataloging (MARC) record was displayed. The study also confirmed how the unified interface, aligning the look and feel among databases, increased the ease of use for end-users. Michael Gorrell described how one of the key providers, *EBSCO*, gathered input from users and considered design features of popular websites, to implement new technologies to the *EBSCOhost* interface.[13] Some of the features that ease the usability of *EBSCOhost* are a dynamic date slider, an article preview hover, and expandable features for various facets, such as subject and publication.[2]

Another key benefit of discovery systems is the speed of results retrieval. The *Primo* discovery layer by Ex Libris has been complimented for its ability to reduce the time it takes to conclude a search session, while maximizing the volume of relevant results per search session.[14] It was suggested that in so doing the tool helps introduce users to new content types. Yuji Tosaka and Cathy Weng reported how records with richer metadata tend to be found more frequently and lead to more circulation.[15] Similarly, Luther and Kelly reported an increase in overall downloads, while the use of individual databases decreased.[16] These studies point to the trend of an enhanced distribution of discovery and knowledge.

With the additional metadata of item records, however, there is also the increased likelihood of inconsistencies across databases that are brought together in a centralized index. A study by Graham Stone offered a comprehensive report on the implementation process of the *Summon* discovery layer at the University of Huddersfield, highlighting major inconsistencies in cataloging practices and the difficulties it caused in providing consistent journal holdings and titles.[17] This casts shadows on the promise of better findability.

Jeff Wisniewski[18] and Williams and Foster[2] are among the many who espouse discovery layers as a step towards a truly single search function that is flexible while allowing needed customizability. These new tools, however, are not without their limitations. The majority of usability studies reinforce similar results and focus on the user interface. Fagan *et al*, for example, studied the usability of *EBSCO Discovery Service* at James Madison University (JMU). While most tasks were accomplished successfully, the study confirmed previous warnings that users do not understand the complexities of search and identified several interface issues: (1) users desire single search, but willingly use multiple options for search, (2) lack of visibility for the option to sort search results, and (3) the difficulty in finding journal articles.[3]

Yang and Wagner offer one case where the aim was to evaluate discovery layers against a checklist of 12 features that would define a true 'next generation catalogue':

- (1) Single point of entry to all library information,
- (2) State-of-the-art web interface (e.g. Google and Amazon),
- (3) Enriched content (e.g. book cover images, ratings and comments),
- (4) Faceted navigation for search results,
- (5) Simple keyword search on every page,
- (6) More precise relevancy (with circulation statistics a contributing factor),
- (7) Automatic spell check,
- (8) Recommendations to related materials (common in commercial sites, e.g. Amazon),
- (9) Allowing users to add data to records (e.g. reviews),

- (10) RSS feeds to allow users to follow top circulating books or topic related updates in the library catalogue,
- (11) Links to social networking sites to allow users to share their resources,
- (12) Stable URL's that can be easily copied, pasted and shared. [11]

They used this list to evaluate seven open source and ten proprietary discovery layers, revealing how only a few of them can be considered true 'next generation catalogs' supporting the users' needs that are common on the Web. All of the tools included in their study missed precision in retrieving relevant search results, e.g. based on transaction data. The authors were impressed with open source discovery layers *LibraryFind* and *VuFind*, which had 10 of the 12 features, leaving vendors of proprietary discovery layers ranking lower (see figure 1).

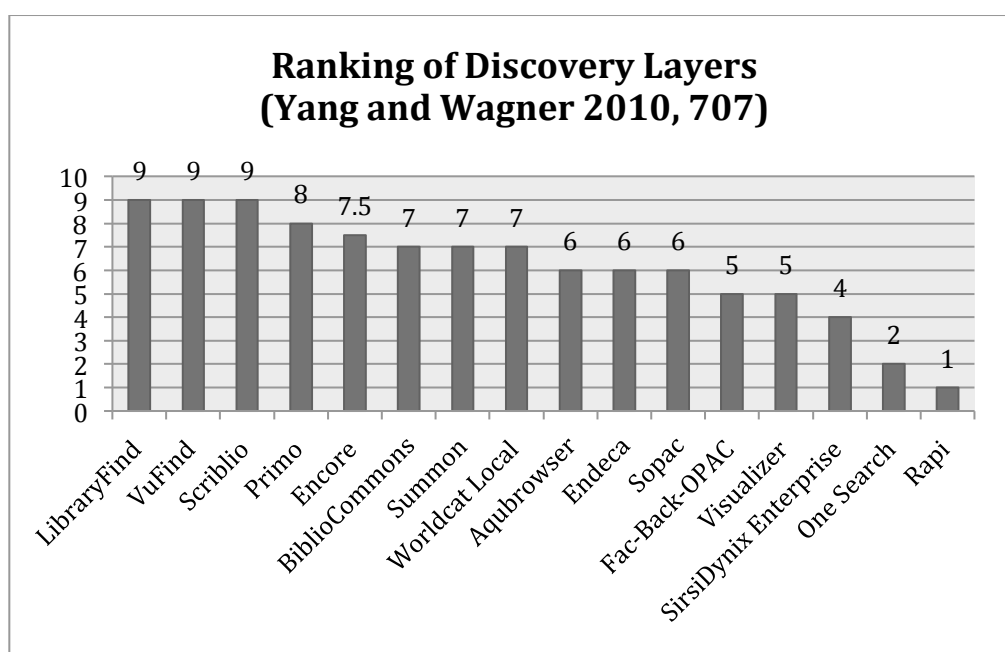


Figure 1. 17 discovery layers (x-axis) were evaluated against a checklist of 12 features expected of the next generation catalogue (y-axis)

Yang and Wagner theorized that the relative lack of innovation among commercial discovery layers is due to practical reasons: vendors create their new discovery layers to run alongside older ones, rather than attempting to alter the proprietary code of the Integrated Library System's (ILS) online public access catalog (OPAC). They pointed to the need for "libraries, vendors and the open source community [...] to cooperate and work together in a spirit of optimism and collegiality to make the true next generation catalogs a reality".[11] At the same time, the University of Michigan Article Discovery Working Group reported on vendors' being more cooperative and allowing coordination among products, increasing the potential of web-scale discovery services.[19] How to evaluate and optimize user workflow across these coordinating products remains a practical

challenge. In this study, we propose HTA as a prospectively helpful method to evaluate user workflow through these increasingly complex products.

Hierarchical Task Analysis

With roots in Tylorism*, industrial psychology and system processes, task analyses continue to offer valuable insights into the balance of efficiency and effectiveness in human-computer interaction scenarios [20][21]. Historically, Frank and Lillian Gilbreth (1911) set forth the principle of hierarchical task analysis (HTA), when they broke down and studied the individual steps involved in laying bricks. They reduced the brick laying process from about 18 movements down to four (in [21]). But, it was John Annett and Keith D. Duncan (1967) who introduced HTA as a method to better evaluate the personnel training needs of an organization. They used it to break apart behavioral aspects of complex tasks such as planning, diagnosis and decision-making (see in[22][21]).

HTA helps break users goals into subtasks and actions, usually in a visual form of a graphic chart. It offers a practical model for goal execution, allowing designers to map user goals to the system's varying task levels and evaluate their feasibility [23]. In so doing, HTA offers the structure with which to learn about tasks and highlight any unnecessary steps and potential errors that might occur during a task performance [24][25], whether cognitive or physical. Its strength lies in its dual approach to evaluation: on the one hand, user interface elements are mapped at an extremely low and detailed level (to individual buttons), while on the other hand, each of these interface elements gets mapped to user's high-level cognitive tasks (the cognitive load). This informs a rigorous design approach, where each detail accounts for the high-level user task it needs to support.

The main limitation of classical HTA is its system-centric focus that does not account for the wider context the tasks under examination exists in. The field of human-computer interaction has shifted our understanding of cognition from an individual information processing model to a networked and contextually defined set of interactions, where the task under analysis is no longer confined to a desktop but "*extends into a complex network of information and computer-mediated interactions*" [26]. The task step focused HTA does not have the ability to account for the rich social and physical contexts that the increasingly mediated and multifaceted activities are embedded in. HTA has been reiterated with additional theories and heuristics, so as to better account for the increasingly more complete understanding of human activity.

Advanced task models and analysis methods have been developed based on the principle of HTA. Stuart K. Card, Thomas P. Moran and Allen Newell [27] proposed an engineering model of human performance – GOMS (Goals, Operators, Methods, and Selection) – to map how task environment features determine what and when users know about the task [20]. GOMS have been expanded to cope with rising complexities (e.g. [28][29][30]), but the models have become largely impractical

* Tylorism is the application of scientific method to the analysis of work, so as to make it more efficient and cost-effective. Modern task

in the process [20]. Instead of simplistically suggesting cognitive errors are due to interface design, Cognitive Task Analysis (CTA) attempts to address the underlying mental processes that most often give rise to errors [24]. Given the lack of our structural understanding about cognitive processes, the analysis of cognitive tasks has remained problematic to implement [20][31]. Activity Theory models people as active decision makers [20]. It explains how users convert goals into a set of motives and how they seek to execute those motives as a set of interactions in a given situational condition. These situational conditions either help or prevent the user from achieving the intended goal. Activity Theory is beginning to offer a coherent foundation to account for the task context [20], but it has yet to offer a disciplined set of methods to execute this theory in the form of a task analysis.

Even though task analyses have seen much improvement, adaptation and usage in its near-40-year-long existence and its core benefit – aiding an understanding of the tasks users need to perform to achieve their desired goals – have remained the same. Until Activity Theory, CLA and other contextual approaches are developed into more readily applicable analysis frameworks, classical HTA with the additional layers of heuristics guiding the analysis remains the practical option [21]. Nielsen's Goal Composition [6] offers one such set of heuristics applicable for the web context. It presents usability concepts such as reuse, multitasking, automated use, recovering and retrieving, to name a few, so as to systematically evaluate the HTA charts representing the interplay between an interface and the user.

Utility of HTA for evaluating discovery layers

Usability testing has become the norm in validating the effectiveness and ease of use of library websites. Yet, thirteen years ago, Brenda Battleson, Austin Booth and Jane Weintrop [32] emphasized the need to support user tasks as the crucial element to user-centered design. In comparison to usability testing, HTA offers a more comprehensive model for the analysis of how well discovery layers support users' tasks in the contemporary library context. Considering the strengths of the HTA method and the current need for vendors to simplify the workflows in the increasingly complex systems, it is surprising that HTA has not yet been applied to the evaluation of discovery layers.

This paper introduces Hierarchical Task Analysis (HTA) as a solution to systematically evaluate the workflow of discovery layers as a technology that helps users accomplish specific tasks, herein, retrieving relevant items from the library catalog and other scholarly collections. Nielsen's [6] Goal Composition heuristics, designed to evaluate usability in the web context, is used to guide the evaluation of the user workflow via the HTA task maps. As a process (vs. context) specific approach, HTA can help achieve a more systematic examination of the tasks discovery layers should support, such as finding an article, a book or an eBook, and help vendors coordinate to achieve the full potential of web-scale discovery services.

METHOD: Applying HTA to *Primo* by Ex Libris

The object of this study was Purdue University's Library website, which was re-launched with *Ex Libris' Primo* in January 2013 (Figure 2) to serve the growing student and faculty community. Its 3.6 million indexed records are visited over 1.1 million times every year. Roughly 34% of these visits are to electronic books. According to Sharon Q. Yang and Kurt Wagner [11], who studied 17 different discovery layers, *Primo* ranked the best among the commercial discovery layer products, coming fourth after the open source tools *Library Find*, *VuFind*, and *Scriblio* in the overall rankings. We will evaluate how efficiently and effectively the *Primo* search interface supports users' of the Purdue Libraries tasks.

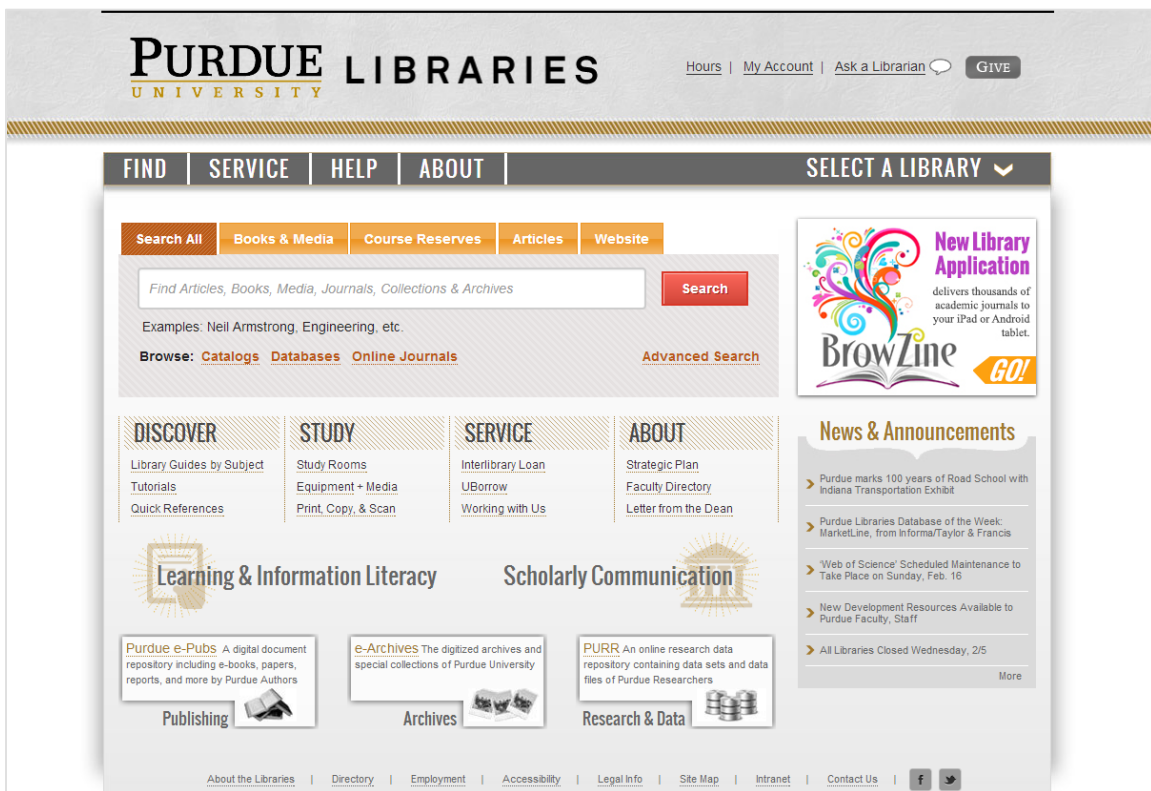


Figure 2. Purdue Library front page and search box

Based on our three year experience of user studies and usability testing of the library website, we identified finding an article, a book and an eBook as the three major representative scenarios of Purdue Library usage. To test how *Primo* helps its users and how many cognitive steps it requires of them, each of the three scenarios were broken into three or four specific case studies. The case studies were designed to account for the different availability categories present in the current *Primo* system, e.g. 'full text available', 'partial availability', 'restricted access' or 'no access'. This is because the different availabilities present users with different possible frustrations and obstacles

to task accomplishment. This system-design perspective could offer a comparable baseline for discovery layer evaluation across libraries. A full list of the eleven case studies can be seen below:

Find an Article:

Case 1. The library has only a full electronic text.

Case 2. The library has the correct issue of the journal in print, which contains the article, as well as, a full electronic text.

Case 3. The library has the correct issue of the journal, which contains the article, only in print.

Case 4. The library does not have the full text, either in print or electronically. A possible option is to use Inter Library Loan (here forth ILL) Request.

Find a book (print copy):

Case 5. The library has the book and the book is on the shelf.

Case 6. The library has the book, but the book is in a restricted place, such as The Hicks Repository. The user has to request the book.

Case 7. The library has the book, but it is either on the shelf or in a repository. The user would like to request the book.

Case 8. The library does not have the book. Possible options are UBorrow[†] or ILL.

Find an eBook:

Case 9. The library has the full text of the eBook.

Case 10. The eBook is shown in search results but the library does not have full text.

Case 11. The book is not shown in search results. Possible option is to use UBorrow or ILL.

It is generally accepted that HTA is not a complex analysis method, but since it offers general guiding principles rather than a rigorous step-by-step guide, it can be tricky to implement [24][20][21][23]. Both authors of this study have expertise in applying HTA and are frequent users of the Purdue Library's website. We are familiar with the library's commonly reported system errors; however, all of our case studies result from a randomized topic search, not from specific reported items. To achieve consistent HTA charts one author carried out the identified use-cases on a part-time basis over a two-month period. Each case was executed on the Purdue Library website, using the *Primo* discovery layer. An on campus Hewlett-Packard (hp) desktop computer with Internet Explorer and a personal MacBook laptop with Safari and Google Chrome were used to identify any possible inconsistencies between user experiences on different

[†] *uBorrow is a federated catalog and direct consortial borrowing service provided by the Committee on Institutional Cooperation (CIC). uBorrow allows users to search for, and request, available books from all CIC libraries, which includes all universities in the Big Ten as well as the University of Chicago, and the Center for Research Libraries.*

operating systems. As per Stanton's [21] statement that "HTA is a living documentation of the sub-goal hierarchy that only exists in the latest state of revision", mapping the HTA charts was an iterative process between the two authors.

According to David Embrey [24] "the analyst needs to develop a measure of skill [in the task] in order to analyze a task effectively" (2). This measure of skill was developed in the process of finding real examples (via a randomized topic search) from the Purdue Library catalog to match the structural cases listed above. For instance 'Case 1. *The library has only the electronic full text*' was turned into a case goal: '0 Find the conference proceeding on *Network-assisted underwater acoustic communication*'. A full list of referenced case studies is below:

Find an Article:

Case 1. Find the article "*Network-assisted underwater acoustic communication*" (Yang and Kevin, 2012).

Case 2. Find the article "*Comparison of Simple Potential Functions for Simulating Liquid Water*" (Jorgensen *et al.*, 1983).

Case 3. Find the journal *Design Annual* "*Graphis Inc*" (2008).

Case 4. Find the article "*A technique for murine irradiation in a controlled gas environment*" (Walb, M. C. *et al.*, 2012).

Find a book (in print):

Case 5. Find the book *Show me the numbers: Designing tables and graphs to enlighten* (Few, 2004).

Case 6. Find the book *The Love of Cats* and place a request for it (Metcalf, 1973).

Case 7. Find the book *The Prince* and place a request for it (Machiavelli).

Case 8. Find the book *The Design History Reader* by Maffei and Houze (2010). (UBorrow or ILL).

Find an eBook:

Case 9. Find the Ebook *Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests* (Rubin and Chisnell, 2008)

Case 10. Find the Ebook *The Science of Awakening Consciousness: Our Ancient wisdom* (Partly available via Hathi Trust)

Case 11. Find the Ebook *Ancient Awakening* by Matthew Bryan Laube (UBorrow).

HTA descriptions are generally diagrammatic or tabular. Since diagrams are easier to assimilate and promise the identification of a larger number of sub-goals [23], diagrammatic description method was preferred (Figure 2). Each analysis started with the establishment of sub-goals, such as 'Browse the Library website' and 'Retrieve the Article', and followed with the identification of individual small steps that make the sub-goal possible, e.g. 'Press Search' and 'Click on 2, to go to page 2' (Figures 3-5). Then, additional iterations were made to include: (1) cognitive steps, where

users need to evaluate the screen in order to take the next step (e.g. identifying the correct URL to open from the initial results set), and (2) capture cognitive decision points between multiple options for users to choose from. For instance, items can be requested either via interlibrary loan (ILL) or UBorrow, presenting the user with an A or B option, requiring cognitive effort to make a choice. Such parallel paths were color coded in yellow (Figure 2). Both physical and cognitive steps were recorded into XMind[‡], a free mind mapping software. They were color-coded black and gray, respectively, helping visualize the volume of cognitive decision points and steps (i.e. cognitive load).

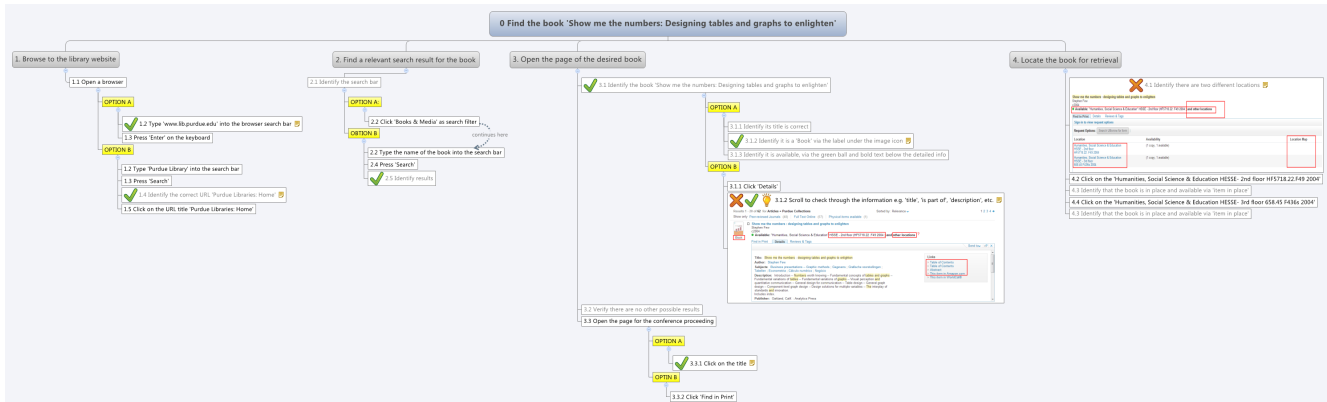


Figure 3. Full HTA chart for 'Find a Book' scenario (CASE 5). Created in Xmind.

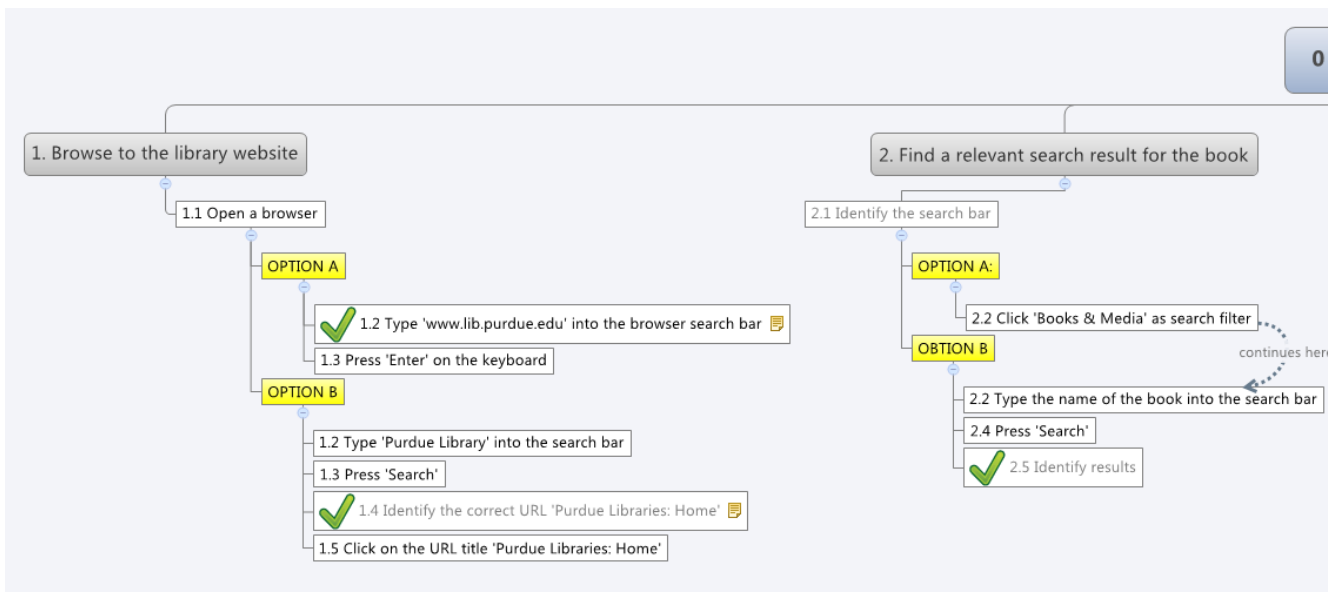


Figure 4. Zoom in to steps 1 and 2 of the HTA map for 'Find a Book' scenario (CASE 5). Created in Xmind.

[‡] XMind is a free mind mapping software that allows structured presentation of step multiple coding references, the addition of images, links and extensive notes. <http://www.xmind.net/>

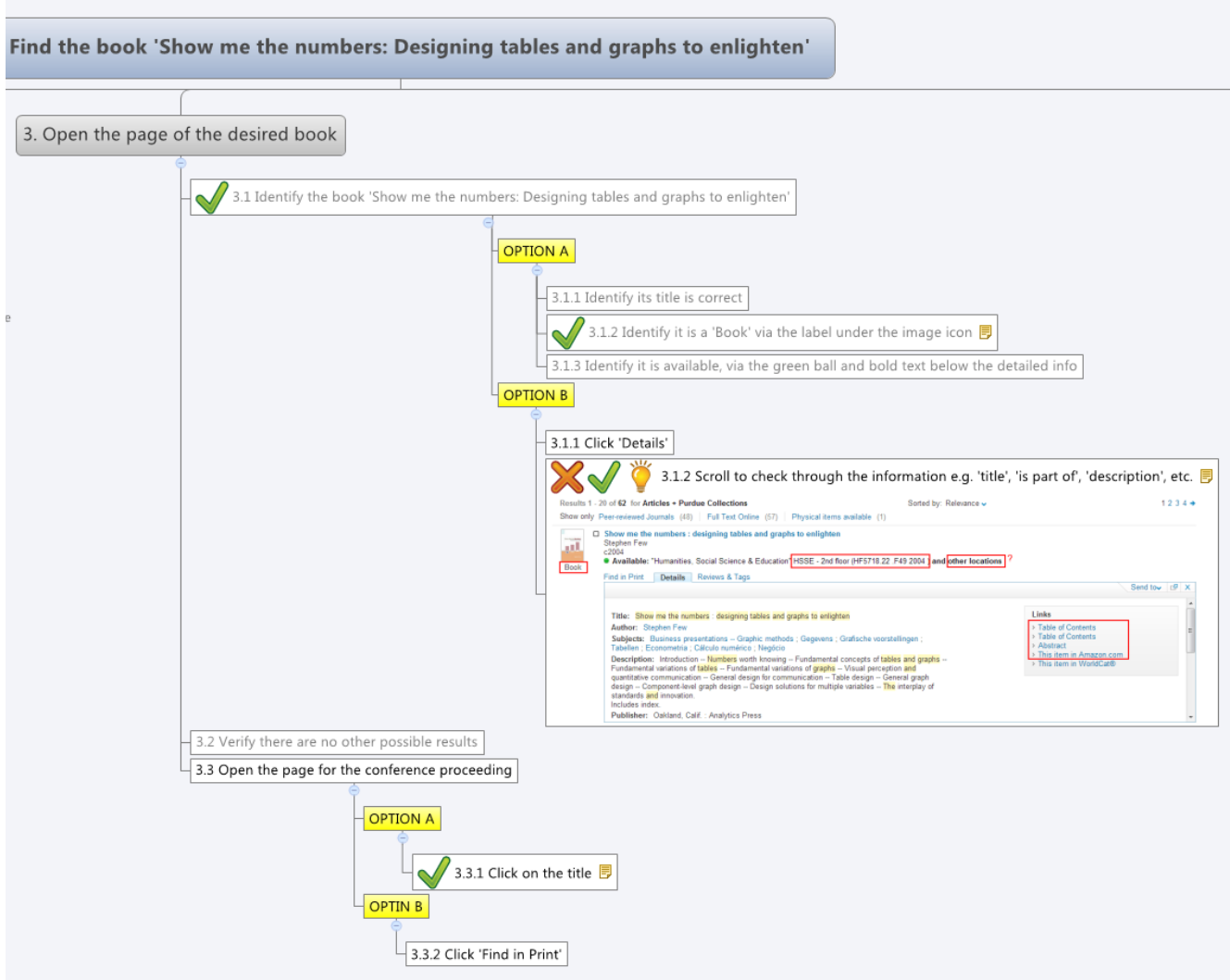


Figure 5. Zoom in to step 3 of the HTA map for the 'Find a Book' scenario (CASE 5). Created in XMind.

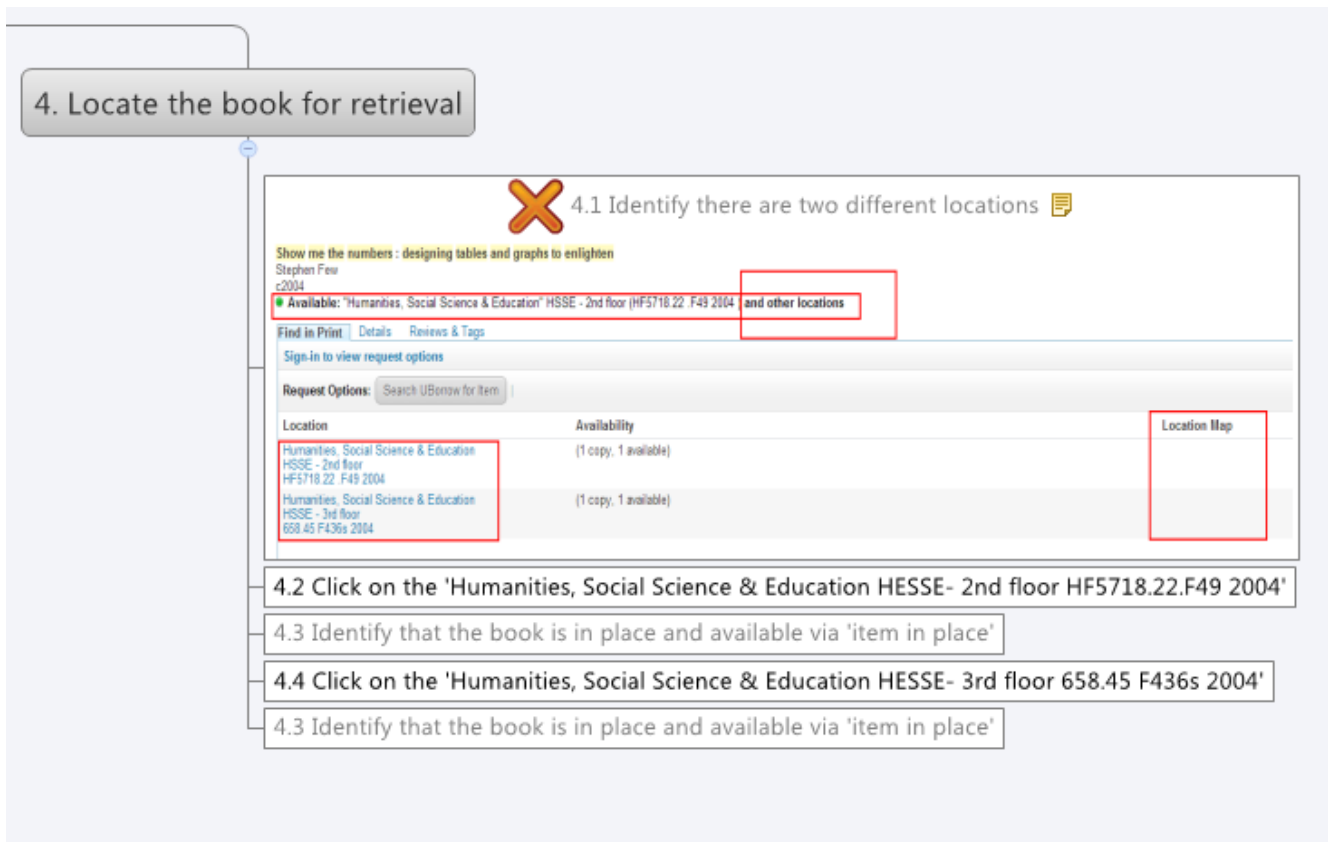


Figure 6. Zoom in to step 4 of the HTA map of the 'Find a Book' scenario (CASE 5). Created in XMind.

To organize the decision flow chart, the original hierarchical number scheme for HTA that requires every sub-goal to be uniquely numbered with an integer in numerical sequence [21], was strictly followed. Visual (screen captures) and verbal notes on efficient and inefficient design factors were taken during the HTA mapping process and linked directly to the tasks they applied to. Steps, where interface design guided the user to the next step, were marked 'fluent' with a green tick (figures 3 and 4). Steps that were likely to mislead users from the optimal path to item retrieval and were a burden to user's workflow were marked with a red 'X' (see figures 4 and 5). One major advantage of the diagram format is its visual and structural representation of sub-goals and their steps in a spatial manner (See figures 2-5). This is useful for gaining a quick overview of the workflow [21].

When exactly to stop the analysis has remained undefined for HTA [21]. It is at the discretion of the analyst to evaluate if there is the need to re-describe every sub-goal down to the most basic level, or whether the failure to perform that sub-goal is, in fact, consequential to the study results. We decided to stop evaluation at the point where the user located (a shelf number or reserve pick up number) or received the sought item via download. Furthermore, steps that were perceived as possible when impossible in actuality were transcribed into the diagrams. Article scenario case 1 offers an example: once the desired search result was identified, its green dot for 'full text available'

was likely to be perceived as clickable, when in actuality it was not. The user is required to click on the title or open the tab 'Find Online' to access the external digital library and download the desired article (See figure 7).

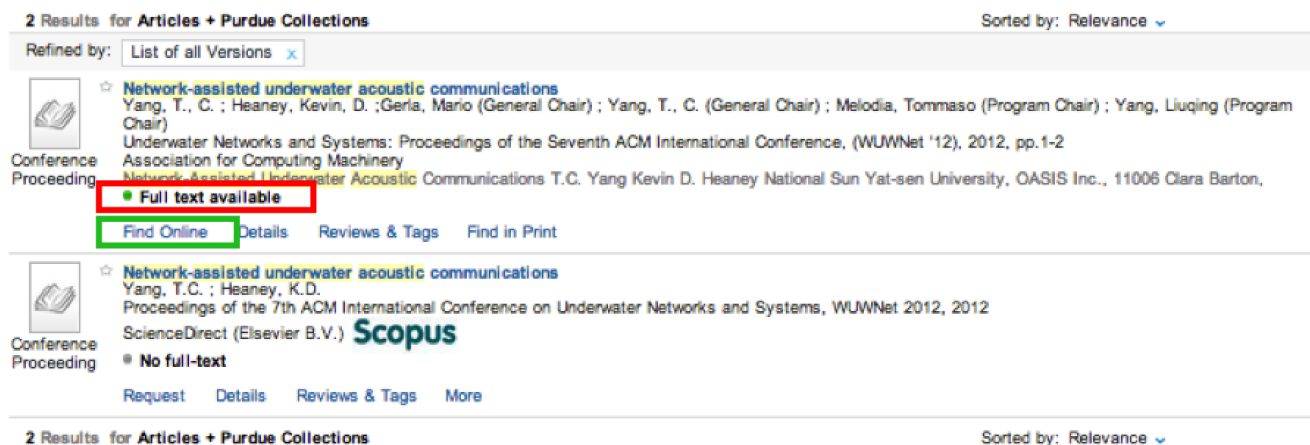


Figure 7. Article scenario (CASE1) two search results, where green 'full text available' may be perceived as clickable.

Task analysis focuses on the properties of the task rather than the user. This requires expert evaluation in place of involving users in the study. As stated above, both of the authors are working experts in the field of user experience in the library context, thoroughly aware of the tasks under analysis and how they are executed on a daily basis. A group of 12 (librarians, reference service staff, system administrators and developers) were asked to review the HTA charts on a monthly basis. Feedback and implications of identified issues were discussed as a group. According to Nielsen [7] it takes five experts (double specialist in Nielsen's terms, is an expert in usability as well as in the particular technology employed by the software.) to not have significant loss of findings (See figure 7). Based on this enumeration, the final versions of the HTA charts offer accurate representations of the *Primo* workflow in the three use scenarios of finding an article, finding a book and finding an eBook at Purdue University Libraries.

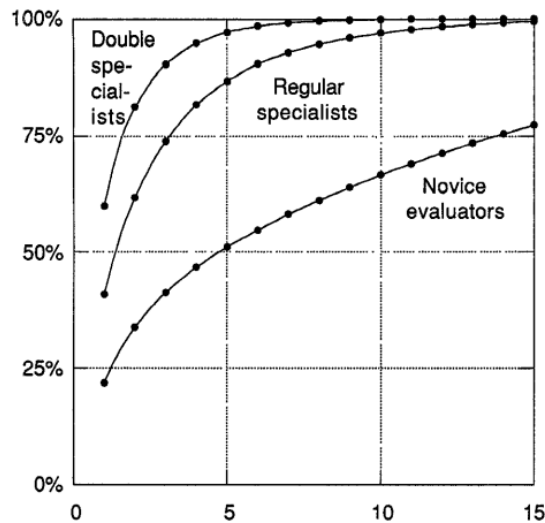


Figure 8. Average proportion of usability problems found as a function of number of evaluators in a group performing heuristic evaluation [7].

RESULTS

The reason for mapping *Primo's* workflows in HTA charts was to identify key workflow and usability issues of a widely used discovery layer in scenarios and contexts it was designed to serve. The resulting HTA diagrams offered insights into fluent steps (green ticks), as well as workflow issues (red 'X') present in *Primo*, as applied at Purdue University Libraries. It is due to space limitations, that only the main findings of the HTA will be discussed. The full results are published on Purdue University Research Repository[§]. Table 1 presents how many parallel routes (A vs. B route), physical steps (clicks), cognitive evaluation steps, likely errors and well guided steps each of the use cases had.

On average it took between 20 to 30 steps to find a relevant item within *Primo*. Even though no ideal step count has been identified for the library context, this is quite high in the general context of the web, where fast task accomplishment is generally expected. Paul Chojecki [33] tested how too many options impact usability on a website. He revealed that the average step count to lead to higher satisfaction levels is 6 (vs. 18,16 average steps at Purdue Libraries). In our study, the majority of the steps were physical pressing of a button or filter selection; however, cognitive steps took up just under a half of the steps in nearly all cases. The majority of cases flow well, as the strengths (fluent well guided steps) of *Primo* outweigh its less guided steps that easily lend themselves to the chance of error.

[§] *Task analysis cases and results for Ex Libris Primo.* <https://purr.purdue.edu/publications/1738>

| CONTENT TYPE <i>CASE NUMBER</i> | ARTICLES | | | | | BOOKS | | | | | EBOOKS | | | |
|---|----------|----|----|----|-----|-------|----|----|----|-----|--------|----|----|-----|
| | 1 | 2 | 3 | 4 | AVG | 5 | 6 | 7 | 8 | AVG | 9 | 10 | 11 | AVG |
| No. of decision points (between A & B), to retrieve an item | 5 | 8 | 4 | 4 | 5 | 4 | 5 | 5 | 2 | 4 | 6 | 3 | 2 | 4 |
| Minimum steps possible to retrieve an item (clicks + cognitive decisions) | 18 | 27 | 16 | 30 | 23 | 18 | 25 | 28 | 24 | 24 | 22 | 19 | 19 | 20 |
| <i>Of these minimum steps, how many were cognitive (information evaluation was needed to proceed)</i> | 4 | 8 | 9 | 13 | 9 | 6 | 9 | 7 | 7 | 7 | 4 | 6 | 4 | 5 |
| Maximum steps it can take to retrieve an item (clicks + cognitive decisions) | 26 | 35 | 23 | 36 | 30 | 22 | 31 | 33 | 28 | 29 | 32 | 23 | 22 | 26 |
| <i>Of these, maximum steps, how many were cognitive</i> | 10 | 17 | 14 | 15 | 14 | 10 | 13 | 16 | 8 | 12 | 9 | 8 | 5 | 7 |
| Errors (steps that mislead from optimal item retrieval) | 3 | 15 | 4 | 8 | 8 | 2 | 2 | 4 | 3 | 3 | 13 | 1 | 2 | 5 |
| Fluent well guided steps to item retrieval | 11 | 11 | 9 | 8 | 10 | 7 | 8 | 7 | 5 | 7 | 6 | 4 | 3 | 5 |

Table 1. Table listing each case’s key task measures, and each scenario’s averages.

Between the three item search scenarios – *Articles*, *Books* and *Ebooks* – the retrieval of articles was least guided and required the highest amount of decisions from the user (5, vs. 4 for books and 4 for eBooks on average). Retrieving an article (between 23-30 steps on average) or a book (24-29 steps on average) took more steps to accomplish than finding a relevant eBook (20-26 steps on average). The high volume of steps (max 30 steps on average) it required to retrieve an article, as well as its high error rate (8), were due to the higher amount of cognitive steps (12 steps on average) required to identify the correct article and to locate a hard copy (instead of the relatively easily retrievable online copy). In the book scenario, the challenge was also two-fold: on the one hand, it was challenging to verify the right book when there were many similar results (this explains the high number of 12 cognitive steps on average); on the other hand, the flow to place a request for a book was also a challenge. The latter was a key contributor to the higher amount of physical steps required for retrieving a book (max 29 on average).

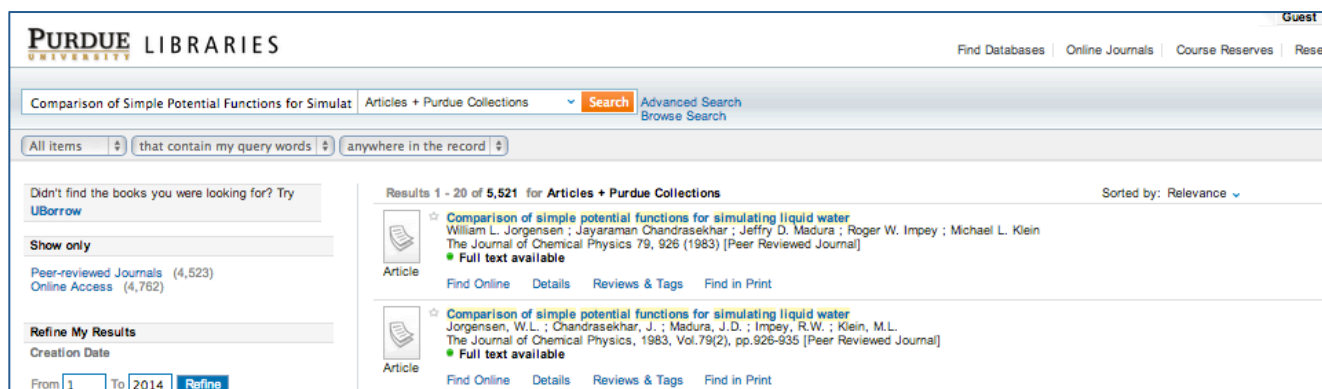
Common to all eleven cases, whether articles or books, was the four sub-goal-process: 1) Browse the Library website, 2) Find results, 3) Open the page of the desired item, and 4) Retrieve, locate or order the item. The first two offered near identical experiences, no matter the search scenario or case. Third and fourth sub-goals, however, presented different workflow issues depending on the product searched and its availability, e.g. 'in print' or 'online'. As such, general results will be presented for the first two themes, while scenario specific overviews will be provided for the latter two themes.

Browsing the Library Website

Browsing the Library website was easy and supported different user tasks. The simple URL (lib.purdue.edu) was memorable and appeared first in the search results. The immediate availability of sub-menus, such as *Databases* and *Catalogs*, offered speedy searching for the frequent users. The choice between: a) general URL, or b) sub-menu, was the first key decision point users of *Primo* at Purdue Libraries were presented with.

The Purdue libraries' home page (revisit figure 1) had a simple design with a clear, central and visible search box. Just above it were search filters for articles, books and the Web. This was the second key decision point users were presented with: a) they could either type into the search bar without selecting any filters, or b) they could select a filter to aid the focus of their results to a specific item type. Browsing the library website offers an efficient and fluent workflow, with eBooks being the only exception. It was hard to know whether they were grouped under *Articles* or *Books & Media* filters. Confusingly (at the time of the study) Purdue Libraries listed eBooks that had no physical copies under *Articles*, while other eBooks that Purdue had physical version of (in addition to the digital ones) under *Books & Media*. This was not explained in the interface, nor was there a readily available tooltip.

Finding Relevant Results



The screenshot shows the Purdue Libraries search interface. At the top, the logo for PURDUE LIBRARIES is visible. Below it, there is a search bar containing the text "Comparison of Simple Potential Functions for Simulat". To the right of the search bar, there are options for "Articles + Purdue Collections" and a "Search" button. Below the search bar, there are filters for "All items", "that contain my query words", and "anywhere in the record". The search results are displayed in a list format, showing two entries for the article "Comparison of simple potential functions for simulating liquid water". The first entry is by William L. Jorgensen, Jayaraman Chandrasekhar, Jeffrey D. Madura, Roger W. Impey, and Michael L. Klein, published in The Journal of Chemical Physics, 79, 926 (1983). The second entry is by Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L., published in The Journal of Chemical Physics, 1983, Vol. 79(2), pp.926-935. Both entries indicate that full text is available. The interface also includes a "Refine My Results" section with a "Creation Date" filter set from 1 to 2014.

Figure 9. Search results for Article (CASE2) 'Comparison of Simple Potential Functions for Simulating Liquid Water'

Primo presented the search results in an algorithmic order of relevance offering additional pages for every 20 items appearing in the search results. The search bar was then minimized at the top of the page, available for easy editing. The page was divided into two key sections, where the first quarter entailed filters (e.g. year of publishing, resource type, author, journal, etc.), and the other three quarters was left for search results (see figure 8). The majority of cognitive decisions across scenarios were made on this results page. This was due to the need to pick up the cues to identify and verify the accurate item being searched. The value of these cognitive steps lies in their leading of the user to the next physical steps. As discussed in the next section, opening the page of the desired item, there were several elements that succeeded and failed at guiding the user to their accurate search result.

Search results were considered relevant when the search presented results in the general topic area of the searched item. Most cases in most scenarios led to relevant results, however, Book CASE 8 and eBook CASE 11, provided only unrelated results. Generally, books and eBooks were easy to identify as available. This was due to their typically short titles, which took less effort to read. Journal articles, on the other hand, have longer titles and required more cognitive effort to be verified.

Article CASE 4, Book CASE 6 and eBook CASE 10 had relevant but restricted results. The color-coding system that indicated the level of availability for the presented search results: green (fully available), orange (partly available) or gray (not available) dots – was followed by an explanatory availability tag, e.g. '*Available online*' or '*Full text available*' etc. Tabs represented additional cues, offering additional information, e.g. '*Find in Print*'. These appeared in a supplementary way where applicable. For example, if an item was not available, its dot was gray and it neither had the '*Find in Print*' nor '*Find online*' tab. Instead, it had a '*Request*' tab, guiding the user towards an available alternative action. Restricted availability items, such as a book in a closed repository, had an orange indicator for partial availability. For these, *Primo* still offered the '*Find in Print*' or '*Find Online*' tab, whichever was appropriate. While the overall presentation of item availability was clear and color-coding consistent, the mechanisms were not without their errors, as discussed below.

Opening the page of the desired item

This sub-goal comprised of two main steps: 1) information driven cognitive steps, which help the user identify the correct item, and 2) user interface guided physical steps that resulted in opening the page of the desired item.

Frequent strengths that helped the identification of relevant items across the scenarios were the clearly identifiable labels underneath the image icons (e.g. '*book*', '*article*', '*conference proceeding*'), hierarchically structured information about the items (title, key details, availability) and perceivably clickable links (blue with an underlined hover effect). The labels and hierarchically presented details (e.g. year, journal, issue, volume, etc.) helped the workflow to remain smooth,

minimizing the need to use side filters. The immediate details reduced the need to open additional pages, cutting down the steps needed to accomplish the task. The hover effect of item titles made the link look and feel clickable, guiding the user closer to retrieving the item. Color-coding all clickable links in the same blue was also an effective design feature, even though bolded availability labels were equally prominent and clickable. This was especially true for articles where the *'full text available'* tags correspond to users goal to immediately download the sought item (figure 8).

The most frequent causes of errors were duplicated search results. Generally, *Primo* displays multiple versions of the same item into one search result and offered a link: *'See all results'*. In line with Graham Stone's [17] study, which highlighted the problem of cataloging inconsistencies, *Primo* struggled to consistently grouping all overlapping search result items. Both book and article scenarios suffered from at least one duplicate search result case due to inconsistent details. Article scenario CASE 2 offers an example, where Jorgensen *et al* "Comparison of Simple Potential Functions for Simulating Liquid Water" (1983) had two separate results for the same journal article of the same year (first two results in figure 8). Problematically, the two results offered different details for the journal issue and page numbers. This may cause likely referencing problems for *Primo* users.

Duplicated search results were also an issue for book scenarios. The most frequent causes for this were instances where authors' first and last names were presented in a reverse order (see also figure 8 for article CASE 2), the books had different print editions, or the editors' name was used in place of the authors'. Book scenario CASE 7: Machiavelli's "*The Prince*" resulted in extremely varied results, requiring 16 cognitive steps and 33 physical steps before a desired item could be verified. This is where search filters were most handy. Problematically, in CASE 7, Machiavelli – the author – did not even appear in the author filter list, while *Ebrary Inc* was listed. Again, this points to the inconsistent metadata and the effects it can have on usability, as discussed by Stone.²

Other workflow issues were presented by design details such as the additional information boxes underneath the item information, e.g. *'find in print'*, *'details'* and *'find online'*. They opened a small scrollable box that maintained the overall page view, were difficult to scroll. The arrow kept slipping outside of the box, scrolling the entire site's page instead of the content inside the box. In addition, the information boxes did not work well with Chrome. This was especially problematic on the MacBook where after a couple of searches the boxes failed to list the details and left the user with an unaccomplished task. Comparably, Safari on a Mac and Internet Explorer on a PC never had such issues.

Retrieving the items (call number or downloading the PDF)

The last sub-goal was to retrieve the item of interest. This often comprised of multiple decision points: whether to retrieve the PDF version from online or identify a call number for the physical

copy or whether to place a request, ordering it via Inter Library Loan (ILL) or UBorrow. Each option is briefly discussed below.

EBooks and *Articles*, if available online, offered efficient online availability. If an article was identified for retrieval, there were two options to access the link to the database, e.g. '*View this record in ACM*': a) via the full view of the item, or b) small '*find online*' preview box discussed above. Where more than one database was available, information about the publication range the Library holds helped identify the right link to download the PDF on the link-resolver page. One of the key benefits of having links from within *Primo* to the full texts was the fact that they opened in new browser windows or tabs, without interference to other ongoing search. While a few of the PDF links to downloadable texts were difficult to find through some external database sites, once found, they all opened in Adobe Reader with easy options to either '*Save*' or '*Print*' the material.

EBooks were available via Ebrary or EBL libraries. While the latter offers some novel uses, such as audio (i.e. read aloud), neither of the two platforms was easy to use. While reading online was possible, downloading an eBook was challenging. The platform seemed to offer good options: a) download by chapter, b) download by page numbers, or c) download the full book for 14 days. In actuality, however, these were all unavailable. EBook CASE 9 had chapters longer than the 60-page limit per day. Page numbers proved difficult to use, as the book's numbers did not match the PDF's page numbers. This made it hard to keep track of what was downloaded and where one left off to continue later (due to imposed time-limits). The 14-day full access option was only available in Adobe Digital Editions software (an ebook reader software by Adobe Systems built with Adobe Flash), which was neither available on most campus computers nor on personal laptops.

The least demanding and most fluent of all retrieval options was the process of identifying the location and call number for physical copies. Inconsistent metadata, however, posed some challenges. Book CASE 5 offered a merged search result of two books, but listed them with different call numbers in the '*Find in Print*' tab. Libraries have many physical copies of the same book, but identifying consistency in call number is a cognitive step that helps verify the similarities or differences between the two results. The different call numbers raised doubts about which item to choose, slowing the workflow for the task and increasing the number of cognitive steps required to accomplish the task.

Compared to books, finding an article in print format was hardly straightforward. The main cause for error when looking up hard copies of journals was the fact that individual journal issues did not have individual call numbers at Purdue Libraries. Instead, they were had one call number per periodical where the entire journal series had only one call number. Article CASE 2, for example, offered the journal code: 530.5 J821 in the '*Find in Print*' tab. In general, the tab suffered from too much information, poor layout and unhelpful information hierarchy, all of which slowed down the cognitive tasks of verifying whether an item was relevant or not. It listed '*Location*' and '*Holdings range*' as the first pieces of information, wherein '*Holdings range*' included not just hard copy related information, but listed digital items as well, even though this tab was for physical version

of the item. To illustrate, Article CASE 2 claimed to have holdings for 1900 – 2013, whereas hard copies were only available for 1900-2000, and digital copies for 2001-2013.

Each scenario had one or two cases where there were neither physical nor digital options available. The sub-goal commonly comprised of a decision between three options: c) Placing a request, d) Ordering an item via Inter Library Loan (ILL), or c) Ordering an item via U BORROW. While the '*Signing in to request*' option and ILL were easy to use with few required steps, there was a lack of guidance on how to choose between the three options. Frequently, ILL and U BORROW appeared as equal options adjacent to one another, leaving the next step unguided. Of all three, placing a request via U BORROW was the hardest to accomplish. It often failed to present any relevant results on the first results page of the U BORROW system, requiring the use of advanced search and filters. For instance, book CASE 6 was '*not requestable*' via U BORROW. When it did list the sought for item in the search results it looped back to Purdue's own closed repository (which remained unavailable).

DISCUSSION

The goal of this study was to utilize HTA to examine the workflow of the *Primo* discovery layer at Purdue University Libraries. Nielsen's [6] Goal Composition heuristics were used to extend the task-based analysis and understand the tasks in the context of discovery layers in libraries. Three key usability domains: generalization, integration and user control mechanisms were used as an analytical framework to draw usability conclusions about how *Primo* was supporting, if at all, successful completion of the three scenarios. The next three sub-sections evaluate and offer design solutions on the three usability domains mentioned above. Overall, this study confirmed *Primo's* ability to reduce the workload for users to find their materials. *Primo* is flexible and intuitive, permitting efficient search and successful retrieval of library materials, while offering the possibility of many search sessions at once [14]. A comparison to a usability test results is offered as a way of conclusion.

Generalization Mechanisms

Primo can be considered a flexible discovery layer as it helps users achieve many goals with minimum amount of steps. It makes use of several generalization mechanisms that allow users to utilize their tasks towards many goals at once. For instance, the library website result in Google offers not only the main URL but also seven sub-links to specialist Library site locations, such as opening hours and databases. This makes *Primo* accessible and relevant for a broader array of people who are likely to have different goals. For instance, some may seek to enter a specific Database, instead of having to open *Primo's* landing page and entering the search terms. Another may wish to utilize 'Find', which guides the user, one step at a time, via a process of definition elimination, closer to the item they are looking for know the opening times.

Similarly, the *Primo* search function saves already typed information, both on its landing page and its results page. This facilitates search by requiring query entry only once, while allowing end

users to click on different filters to narrow the results in different ways. As a part of the work done towards one search can be used towards another, e.g. by content, journal, or topic type, the system can ease the work effort required of users. This is further supported by the system saving already typed keywords when returning to the main search page from research results and allows for a fluid search experience where the user adjusts a keyword thread with minimal typing, until they find what they are looking for.

A key problem for *Primo* is its inability to manage inconsistent meta-data. The tendency to group different versions of the same search results together is helpful as it clarifies information noise. In an effort to enhance the speed it takes to evaluate the relevancy of search results, the system seeks to shighlight any differences in the meta-data. If inconsistencies in meta-data cause same search results to appear as separate items, it is likely to affect the cognitive steps and therefore the workload and efficiency with which the user is able to accomplish identification.

It is clear from previous studies that if discovery layers were to become the next generation catalogs [11], and were to enhance the speed of knowledge distribution as has been hoped by Tosaka and Weng [15] and Luther and Kelly [16], then mutual agreement is needed on how meta-data from disparate sources [17]. Understanding that users' cognitive workload should be minimized (by offering fewer options and more directive guidance) for more efficient decision-making, library items should have accurate details in their meta-data, e.g. consistent and thorough volume, issue and page numbers for journal articles, correct print and reprint years for books, and item type (conference proceeding vs. journal article).

Integration Mechanisms

The discovery layer's ability to increase the number of search sessions [14] at any one time is possible due to its flexibility to support multitasking. *Primo* achieves this with its own individual features used in combination with other system facilities and external sources. For instance, *Primo's* design allows users to review and compare several search results at once via the '*Find in Print*' or '*Details*' tabs. Although not perfect, since the small boxes are hard to scroll within, the information can save the user the need and additional steps of opening many new windows and having to click between them just for reviewing search results. Instead, many '*detail*' boxes of similar results may be opened and viewed at once, allowing for effective visual comparison. This integration mechanism allows a fluent transition from skimming the search results to another temporary action of gaining insight about the relevance of an item. Most importantly, this is accomplished without requiring the user to open a new browser page or tab, where they would have to break from their overall search flow and remember the details (instead of visually comparing them), making it hard to resume from where they left off.

A contrary integration mechanic that *Primo* makes use of is its smooth automated connectivity to external sites, such as databases, Ebrary, ILL, etc. New browser pages are used to allow the continuation of a task outside of *Primo* itself without forcing the user out of the system to the

library service or full text. *Primo* users can skim search results, identify relevant resources and open them in new browser pages for later reviewing.

What is missing, however, is the opportunity to easily save and resume a search. Retrieving the search result or saving it under ones' login details would benefit users who recall items of interest from previous searches and would like to repeat the results without having to remember the keywords or search process they used. It is not obvious how to locate the save search session option in *Primo's* interface.

User Control Mechanisms

Yang and Wagner [11] ranked *Primo* highest among the vendors, primarily for its good user control mechanisms, which allow users to inspect and change the search functions on an ongoing basis. *Primo* does a good job at presenting search results in a quick and organized manner. It allows for the needed 'undo' functionality and continued attachment and removal of filters, while saving the last keywords when clicking the back button from search results. The continuously available small search box also offers the flexibility for the user to change search parameters easily. In summary, *Primo* offers agile searching, while accounting for a few different discovery mental models.

However, if *Primo* wants to preserve its current effectiveness and make the jump towards a single search function that is truly flexible and allows for much needed customizability [18][2], it needs to allow for several similar user goals to be easily executable without confusion about the likely outcome. The most prominent current system error for *Primo*, as it has been applied in the Purdue Libraries, is its inability to differentiate eBooks from journal articles or books. It would support users goals to be able to start and finish an eBook related tasks at the home page's search box. Currently, users have the cognitive burden to consider whether eBooks are more likely to be found under '*Books & Media*' or '*Journals*'. Currently, *Primo*, as applied to its implementation at Purdue Libraries at the time of this study, does not support goals to search for content type, e.g. an eBook. This however, is increasingly popular among the student population who want eBooks on their tablets and phones instead of carrying heavy books in their backpacks.

Another key pain-point for current users is the identification of specific journals in physical form, say for archival research. Currently, each journal issue is listed individually in the '*find in print*' section, even though the journals only have one call number. Listing all volumes and issues of each periodical overwhelms the user with too much information and prevents the effective accomplishment of the task of locating a specific journal issue. Since there is only one call number available for the entire journal sequence, it may lead to better clarity and usability if the information was reduced. Instead of listing all possible journal issues, a range or ranges (if incomplete set of issues) that the library has physically present should be listed. In Article CASE 2, for instance, there are five items for the year 1983. Why lead the user to look at a range where there is no possible option?

Comparing HTA to a Usability Test

Usability tests benefit from the invaluable direct input from the end user. At the same time usability studies, as constructed conditions, offer limited opportunities to learn about users' real motivations and goals and how the discovery layers support or fail to support their tasks. Fagan *et al* [3] conducted a usability test with eight students and two faculty members to learn about usability issues and user satisfaction with discovery layers. They measured time, accuracy and completion rate for nine specific tasks, and obtained insights from task observations and post-test surveys. They reported on issues with users not following directions (93), the prevalence of time outs, users skipping tasks, and variable task times. These results all point to a mismatch between the user goals and the study tasks and offer an incomplete picture about the system's ability to support user goals that are accomplished via specific tasks.

Expert evaluation based HTA method does not require users' direct input. HTA offers a method to achieve a relatively complete evaluation of how low-level interface facets support users' high-level cognitive tasks. HTA measures the system designs quality in supporting a specific task needed to accomplish a user goal. Instead of measuring time, physical and cognitive tasks are measured in number of steps. Instead of accuracy and completion rate, fluent workflow steps and mistaken steps are counted. The two methods offer opposite strengths, making them a good complements. Given HTA's system-centric approach, it can better inform which tasks would be useful in usability testing.

To compare the our research findings with usability tests, Fagan *et al* [3] confirmed some of the previously established findings that journal titles are difficult to locate via the library home page (vs. databases), that filters are handy when they are needed and that users' mental models have a preference for a Google-like single search-box. For instance, students and even librarians, struggle to understand what is being searched in each system and how results are ranked (See also [5]). The HTA method applied in this study was also able to confirm that journal titles are more difficult to identify than books and eBooks, the flexibility benefit offered by filters and identify the single search box as a fluent system design. Since, HTA does not rely on the user to tell why these results are true, HTA, as applied in this study, helped expert evaluators understand the reasons for these findings via self-directed execution and discussion with colleagues later. Depending on the task design, either usability testing or HTA offer the capabilities to identify cases such as confusion about how to start an eBook search in *Primo*. Taking a system design approach to task design offers a path to a systematic understanding of discovery layer usability, which lends itself to easier comparison and external validity.

In terms of specific interface features, usability tests are good for evaluating the visibility of specific features. For example, Fagan *et al* [3] asked their participants to (1) search on speech pathology, (2) find a way to limit search results to audiology, and then (3) limit their search results to peer-reviewed (task 3 in [3], p. 95). By measuring completion rate, they were able to identify the relative failure of 'peer-reviewed' over 'audiology' filters, but they were left "*unclear*

[about] why the remaining participants did not attempt to alter the search results to ‘peer reviewed,’ failing to accomplish the task [3]. In comparison, HTA as an analytical rather than observational methodology, leads to more synthesized results. In addition to insights into possible gaps between system design and mental models, HTA as a goal-oriented approach, concerns itself with issues of workflow (how well the system guides the user to accomplishing their task) and efficiency (minimizing the number of steps required to finish a task). These are less obvious to identify with usability tests, where participants are not impacted by their routine goals, time pressures and consequently their patience may be more tolerant as a result.

The application of HTA helped identify key workflow issues and map them to specific design elements. For instance, the lack of eBooks as a search filter meant that the current system did not support content form based searching well for two main forms: articles and books. Compared to usability tests that focus on specific fabricated search processes, HTA aims to map all possible routes the system’s design offers to accomplish a goal, allowing for their parallel existence during the analysis. This system-centered approach to task evaluation, we argue, is the key benefit HTA can offer towards a more systematic evaluation of discovery layers, where different user groups would have varying levels of assistance needs. HTA task-analysis allows for the nuanced understanding that results can differ as the context of use differs. That applies even to the contextual difference between user test participants and routine library users.

CONCLUSION

Discovery layers are advancing the search experiences libraries can offer. With increasing efficiency, increased ease of use and more relevant results, scholarly search has become a far less frustrating experience. While Google is still perceived as the holy grail of discovery experiences, in reality it may not be quite what scholarly users are after [5]. The application of discovery layers has focused on eliminating the limitations that plagued the traditional federated search and improving the search index coverage and performance. Usability studies have been effective in verifying these benefits and key interface issues. Moving forward, studies on discovery layers should focus more on the significance of discovery layers on user experience.

This study presents the expert evaluation based HTA methods as a complementary way to systematically evaluate popular discovery layers. It is the system design and goal-oriented evaluation approach that offers the prospects of a more thorough body of research on discovery layers than usability alone. Using HTA as a systematic preliminary study guiding formal usability testing offers one way to achieve more comparable study results on applications of discovery layers. It is through comparisons that the discussion of discovery and user experience can gain a more focused research attention. As such, HTA can help vendors to achieve the full potential of web-scale discovery services.

To better understand and ultimately design to their full potential, systematic studies are needed on discovery layers. This study is the first attempt to apply HTA towards systematically analyzing user workflow and interaction issues on discovery layers. The authors hope to see more work in

this area, with the hope of achieving true next generation catalogs that can enhance knowledge distribution.

REFERENCES

- [1] Beth Thomsett-Scott and Patricia E. Reese, "Academic Libraries and Discovery Tools: A Survey of the Literature," *College & Undergraduate Libraries* 19, no. 2–4 (April 2012): 123–143. <http://dx.doi.org/10.1080/10691316.2012.697009>.
- [2] Sarah C. Williams and Anita K. Foster, "Promise Fulfilled? An EBSCO Discovery Service Usability Study," *Journal of Web Librarianship* 5, no. 3 (Jul. 2011): 179–198. <http://dx.doi.org/10.1080/19322909.2011.597590>.
- [3] Jody Condit Fagan, Meris A. Mandernach, Carl S. Nelson, Jonathan R. Paulo, and Grover Saunders, "Usability Test Results for a Discovery Tool in an Academic Library," *Information Technology and Libraries* 31, no. 1 (Mar. 2012): 83–112, Mar. 2012. <http://dx.doi.org/10.6017/ital.v31i1.1855>.
- [4] Roger C. Schonfeld and Matthew P. Long, "Ithaka S+R US Library Survey 2013," Ithaka S+R, survey 2, Mar. 2014. <http://sr.ithaka.org/research-publications/ithaka-sr-us-library-survey-2013>.
- [5] Michael Khoo and Catherin Hall, "What Would 'Google' Do? Users' Mental Models of a Digital Library Search Engine," in *Theory and Practice of Digital Libraries*, ed. Panayiotis Zaphiris, George Buchanan, Edie Rasmussen, and Fernando Loizides, 1-12 (Berlin Heidelberg, Springer: 2012). http://dx.doi.org/10.1007/978-3-642-33290-6_1.
- [6] Jakob Nielsen, "Goal Composition: Extending Task Analysis to Predict Things People May Want to Do," *Goal Composition: Extending Task Analysis to Predict Things People May Want to Do*, 01-Jan-1994. <http://www.nngroup.com/articles/goal-composition/>.
- [7] Jakob Nielsen, "Finding Usability Problems Through Heuristic Evaluation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 373-380 (New York, NY, ACM: 1992). <http://dx.doi.org/10.1145/142750.142834>.
- [8] Jerry V. Caswell and John D. Wynstra, "Improving the search experience: federated search and the library gateway," *Library Hi Tech* 28, no. 3 (Sep. 2010): 391–401. <http://dx.doi.org/10.1108/07378831011076648>.
- [9] Emily R. Alling and Rachael Naismith, "Protocol Analysis of a Federated Search Tool: Designing for Users," *Internet Reference Services Quarterly* 12, no. 1/2, (2007): 195–210. http://dx.doi.org/10.1300/J136v12n01_10.
- [10] Susan Johns-Smith, "Evaluation and Implementation of a Discovery Tool," *Kansas Library Association College and University Libraries Section Proceedings* 2, no. 1 (Jan. 2012): 17–23.

-
- [11] Sharon Q. Yang and Kurt Wagner, "Evaluating and comparing discovery tools: how close are we towards next generation catalog?," *Library Hi Tech* 28, no. 4 (Nov. 2010): 690–709.
<http://dx.doi.org/10.1108/07378831011096312>.
- [12] Lyle Ford, "Better than Google Scholar?," presentation, Advance Program for Internet Librarian 2010, Monterey, California, 25-Oct-2010.
- [13] Michael Gorrell, "The 21st Century Searcher: How the Growth of Search Engines Affected the Redesign of EBSCOhost," *Against the Grain* 20, no. 3 (2008): 22, 24.
- [14] Sian Harris, "Discovery services sift through expert resources," *Research Information*, no. 53, (Apr. 2011): 18–20.
http://www.researchinformation.info/features/feature.php?feature_id=315.
- [15] Yuji Tosaka and Cathy Weng, "Reexamining Content-Enriched Access: Its Effect on Usage and Discovery," *College & Research Libraries* 72, no. 5 (Sep. 2011): pp. 412–427.
<http://dx.doi.org/10.5860/>.
- [16] Judy Luther and Maureen C. Kelly, "The Next Generation of Discovery," *Library Journal* 136, no. 5 (March 15, 2011): 66-71.
- [17] Graham Stone, "Searching Life, the Universe and Everything? The Implementation of Summon at the University of Huddersfield," *LIBER Quarterly* 20, no. 1 (2010): 25–51.
<http://liber.library.uu.nl/index.php/lq/article/view/7974>.
- [18] Jeff Wisniewski, "Web Scale Discovery: The Future's So Bright, I Gotta Wear Shades," *Online* 34, no. 4 (Aug. 2010): 55–57.
- [19] Gaurav Bhatnagar, Scott Dennis, Gabriel Duque, Sara Henry, Mark MacEachern, Stephanie Teasley, and Ken Varnum, "University of Michigan Library Article Discovery Working Group Final Report," University of Michigan Library, Jan. 2010,
<http://www.lib.umich.edu/files/adwg/final-report.pdf>
- [20] Abe Crystal and Beth Ellington, "Task analysis and human-computer interaction: approaches, techniques, and levels of analysis" in *AMCIS 2004 Proceedings*, Paper 391,
<http://aisel.aisnet.org/amcis2004/391>.
- [21] Neville A. Stanton, "Hierarchical task analysis: Developments, applications, and extensions," *Applied Ergonomics* 37, no. 1 (2006): 55–79.
- [22] John Annett and Neville A. Stanton, eds. *Task Analysis*, 1 edition. London ; New York: CRC Press, 2000.
- [23] Sarah K. Felipe, Anne E. Adams, Wendy A. Rogers, and Arthur D. Fisk, "Training Novices on Hierarchical Task Analysis," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54, no. 23, (Sep. 2010): 2005–2009,
<http://dx.doi.org/10.1177/154193121005402321>.

-
- [24] D. Embrey, "Task analysis techniques," *Human Reliability Associates Ltd*, vol. 1, 2000.
- [25] J. Reason, "Combating omission errors through task analysis and good reminders," *Quality & Safety Health Care* 11, no. 1 (Mar. 2002): 40–44, <http://dx.doi.org/10.1136/qhc.11.1.40>.
- [26] James Hollan, Edwin Hutchins, and David Kirsh, "Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research," *ACM Trans. Comput.-Hum. Interact* 7, no. 2 (Jun. 2000): 174–196, <http://dx.doi.org/10.1145/353485.353487>.
- [27] Stuart K. Card, Allen Newell, and Thomas P. Moran, *The Psychology of Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1983.
- [28] Stephen J. Payne and T. R. G. Green, "The structure of command languages: an experiment on task-action grammar," *International Journal of Man-Machine Studies* 30, no. 2 (Feb. 1989): 213–234.
- [29] Bonnie E. John and David E. Kieras, "Using GOMS for User Interface Design and Evaluation: Which Technique?," *ACM Transactions on Computer-Human Interactions* 3, no. 4 (Dec. 1996): 287–319, <http://dx.doi.org/10.1145/235833.236050>.
- [30] David E. Kieras and David E. Meyer, "An Overview of the EPIC Architecture for Cognition and Performance with Application to Human-computer Interaction," *Human-Computer Interaction* 12, no. 4 (Dec. 1997): 391–438, http://dx.doi.org/10.1207/s15327051hci1204_4.
- [31] Laura G. Militello and Robert J. Hutton, "Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands," *Ergonomics* 41, no. 11 (Nov. 1998): 1618–1641, <http://dx.doi.org/10.1080/001401398186108>.
- [32] Brenda Battleson, Austin Booth, and Jane Weintrop, "Usability testing of an academic library Web site: a case study," *The Journal of Academic Librarianship* 27, no. 3 (May 2001): 188–198.
- [33] Paul Chojecki, "How to increase website usability with link annotations," in *20th International Symposium on Human Factors in Telecommunication. 6th European Colloquium for User-Friendly Product Information. Proceedings*, 2006, p. 8.

Case Study References:

Find an Article:

Case 1. Yang, T. C., and Kevin D. Heaney. "Network-Assisted Underwater Acoustic Communications." In *Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems*, p. 37. ACM, 2012.

Case 2. Jorgensen, William L., Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. "Comparison of Simple Potential Functions for Simulating Liquid Water." *The Journal of Chemical Physics* 79 (1983): 926.

Case 3. "Design Annual." Graphis Inc., 2008

Case 4. Walb, M. C., J. E. Moore, A. Attia, K. T. Wheeler, M. S. Miller, and M. T. Munley. "A Technique for Murine Irradiation in a Controlled Gas Environment." *Biomedical Sciences Instrumentation* 48 (2012): 470.

Find a Book (physical):

Case 5. Few, Stephen. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Vol. 1, no. 1. Oakland, CA: Analytics Press, 2004.

Case 6. Metcalf, Christine. *The Love of Cats*. Crescent Books, 1973.

Case 7. Machiavelli, Niccolò, and Leo Paul S. De Alvarez. 1989. *The Prince*. Prospect Heights, Ill: Waveland Press.

Case 8. Lees-Maffei, Grace, and Rebecca Houze, eds. *The Design History Reader*. Berg, 2010.

Find an eBook:

Case 9. Rubin, Jeffrey, and Dana Chisnell. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. Wiley Technical Communication Library, 2008.

Case 10. Rubin, Jeffrey, and Dana Chisnell. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. Wiley Technical Communication Library, 2008.

Case 11. Laube, Matthew Bryan. *Ancient Awakening*. 2010.

Copyright of Information Technology & Libraries is the property of American Library Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.