

Making measurement of use

Report of .5 Paid Leave June to December 2006

Helen Clarke
Head, Collections Services
University of Calgary Library

Friday, January 26, 2007

1.	Designing Collections with the User in Mind	3
1.1.	Building Sustainable User Models	3
1.2.	Applying use and re-use strategies in the Library	5
2.	Designing the Study	6
2.1.	Using Personas to Capture Variation	6
2.2.	Michael	7
2.3.	Athena.....	8
2.4.	Sandy	9
3.	Tying Behavioural Studies to the Model	10
3.1	Starting.....	11
3.2	Chaining.....	14
3.3	Browsing	17
3.4	Monitoring.....	19
3.5	Accessing.....	22
3.6	Differentiating.....	24
3.7	Extracting	26
3.8	Verifying	28
3.9	Networking.....	29
3.10	Managing.....	32
4.	Design Elements for a Digital Collection	35
5.	Measures for a Digital Collection	37
5.1.	Cataloguing and Metadata	38
5.2.	Direct linking between citations and sources	40
5.3.	Direct Access to Resources	41
5.4.	Communication	44
5.5.	Responsiveness	44
5.6.	Provision of unlicensed content.....	45
5.7.	Creation of Personal Copies	47
5.8.	Collection Stability.....	48
5.9.	Information Management.....	51
6.	Reflection and Conclusions	52
6.1.	What type of library would these measures create?	52
6.2.	Next steps.....	53

1. Designing Collections with the User in Mind

Librarianship shares much with the fields of engineering, ergonomics, industrial design, and human-computer interface design, where practical applied approaches are needed to support human interaction with technology. Insights and approaches from these fields can be useful in enriching library practice. In particular the methods used in human-computer interface design (HCI) provide valuable insights.

In this study I review the use of behavioural models, and HCI practices including task analysis, personas, and claims analysis, to develop a sustainable and renewable model of scholars' information seeking behaviour that can be used to develop performance measures and guide the development of digital libraries.

I chose to focus on scholarly behaviour because it is an area often overlooked in my setting.¹ Recent emphasis on information literacy training and learning leads to decisions predicated on the learner's academic experience, which is bounded in short time lines and constant exploration of new topics, creating a transient, ephemeral experience of the library collection.

Scholarly use which is often deeper and richer, provides a balance in collection development decisions, and perhaps reflects the experience of the life long learner we expect our students to become.

1.1. Building Sustainable User Models

Alistair Sutcliffe¹ examined a problem familiar to libraries when he discussed the challenge of capturing an

understanding of the user that could be used productively and longitudinally in human-computer interface (HCI) design. He recounts the limitations of findings from psychological or cognitive studies in informing design as these fields often produce large scale generalizations that are difficult to apply to practical problems and specific situations.

To help inform the practical problems faced by HCI and other design practices, including librarianship, practitioners seek to understand the user's immediate tasks and motivations and to use this understanding to create approaches to technology based problems. Examples of some approaches include: usability studies, focus groups, and ethnographic observation. However Sutcliffe identified a crucial shortcoming of these approaches, they fail to provide a means for testing and re-using information in order to build re-usable artefacts and models of good practice.

Sutcliffe looked to the work of John Carroll and the idea of the task-artefact cycle for inspiration.^{2,3} Carroll, working in HCI studies, advocates viewing design as a

¹ The setting for this study is the University of Calgary, a mid-size 4 year Medical Doctoral University located in Albe. I work as Head of Collections Services for the Library.

process that flows through cyclical stages. Artefacts are designed based on a current understanding of user's tasks; however users will take any artefact and through use discover unanticipated problems, ignore some features, and find new uses or tasks not previously considered. Designers can, in turn, take the information gleaned from user observation and refine the artefact or create something entirely new, and so the cycle starts again.

The task-artefact cycle has some interesting implications for libraries. We traditionally view ourselves as intermediaries between the vendor and the user, assigning to ourselves the responsibility of understanding and interpreting user tasks. We both evaluate current products and provide feedback to vendors. A crucial weakness in this approach is a tendency to equate library practices and services with user tasks, and to ignore the role of user motivations and work environment in creating the conditions of use. Libraries may have a tendency to guide the development of products that meet the needs of our environment, but which fail to provide tools that are effective for the user's task.

Carroll and Rosson contributed another key concept in their use of claims analysis. In claims analysis, the proposed benefits of specific design elements are articulated, creating an explicit statement that can be tested once the product is in use. As well, the potential negative effects of an element are described as are the steps taken to mitigate these drawbacks. By doing claims analysis the designer creates a verifiable understanding of the product design. ⁴

Sutcliffe integrates these concepts, first that design is a longitudinal, cyclical process, and second, that design decisions can be refined and reused through claims analysis. This supports progressive and accumulative design, rather than being bounded by the single design and usability episode.

Sutcliffe proposes employing user models, task analysis and claims analysis as a way of storing the knowledge gained in one design project and carrying it forward to new projects.

Sutcliffe suggests that an overall structure is needed that can classify claims and tasks and which provides room for the integration of new information. *"To enable effective reuse, claims need to be classified and organized in a library. A framework is needed to organize claims into families so designers can locate the appropriate claims for their current application problem."*

Sutcliffe cautions that the level of abstraction chosen for the any structure or indexing of claims ultimately has an impact on the type of knowledge that can be stored or retrieved

1.2. Applying use and re-use strategies in the Library

The design and evaluation of digital library collections face the challenges articulated by Sutcliffe and Carroll for the field of human computer interface design. Not surprising as

the digital library is a subset of this larger area of inquiry. Like these scholars we need to develop expandable, reusable, robust models of user tasks and conditions of use. To do this, and to participate in the task artefact cycle, we need to find ways to assess how well our products and services support user's needs.

In the area of collection development, this opens the potential for new types of analysis. Traditional measures of collection effectiveness rely on quantitative metrics, for example size and growth, without any direct evaluation of how well the collection supports actual user tasks.

For the digital library use is increasingly employed as a proxy for measures of effectiveness, with increasing emphasis on costing based on use, and evaluation based on cost/use ratios. Recent COUNTER standards in particular seem designed to capture frequency of use and not any other aspects or measures of effectiveness.

However, these traditional measures are not adequate for understanding the effectiveness of the digital library as a tool of scholarship and they blind us to the need to understand the conditions of use, miring us in a view of information that is bounded by the walls of the library and the activities that occur there. The shortcomings of this approach are well defined in the larger dialogue surrounding the growth of the information economy; libraries know they have lost their monopoly on the provision of knowledge and the development of information technology. What is not clear to many is that libraries never had this monopoly, and that the information environment for most individuals has always operated in a much vaster venue than that supplied and designed by the library.

Carroll's conceptualization of the task-artefact cycle provides new insight into the importance of developing measures that move beyond the simply quantitative. Measures not only assess the effectiveness of tools but they also encapsulate two facets of artefact design. First, they articulate and test our theories on how library artefacts support the user's work.

Second, measures declare the direction in which we want the library collection to develop. Understanding that evaluation is part of the process by which collection directions are set helps focus evaluation on the ongoing cycle of improvement. It also opens us to understanding that evaluation measures should evolve as our understanding of the user changes over time.

2. Designing the Study

Sutcliffe suggests that effective use and re-use of user information and design experience requires a framework in which knowledge can be stored. For this study I chose

Ellis's model of scholarly information seeking behaviour combined with amendments suggested by Lokman and Tibbo as a scaffold for behavioural information garnered from a review of studies of scholarly information work.⁵

In his model, based on observations of social science scholars, Ellis detected six categories of activity:

- Starting—activities surrounding the initial search for information
- Chaining—following citations and other types of references
- Browsing—semi directed searching
- Differentiating—filtering material, deciding what should be examined more closely
- Monitoring—monitoring specific sources to maintain awareness of the field
- Extracting—working through sources systematically to locate material of interest

Lokman and Tibbo in their 2003 follow up study examined how well Ellis's model had stood up with the advent of electronic information sources. Their finding was that while the model was still valid, they would add four new elements:

- Accessing—locating and retrieving information
- Networking—connecting with other researchers and individuals as sources of information and feedback
- Verifying—affirming the correctness of information
- Managing—organizing information for later retrieval

In this study the ten behaviours are used to index and organize information extracted from studies on scholarly information behaviour published since 2000. Behavioural studies were chosen in that they tended to observe behaviour within the scholar's environment, rather than focusing on library activities. The 2000 cut-off was set as I felt that older articles would not reflect the impact of the availability of digital information.

2.1. Using Personas to Capture Variation

While the information seeking behaviour of scholars could be captured with Ellis, Lokman and Tibbo's model, the literature review

revealed variations in emphasis and approach that required another dimension to capture.

Scholars' view of information and the relative importance of different activities changes depending on the type of questions they are asking, the structure of their field of study, the stage of their research, and their individual personalities. Documenting variation was key to developing meaningful user models and moving beyond a one size fits all approach to collection evaluation.

I chose to use personas to encapsulate individual variation, developing three characters that represent a range of approaches and activities. I purposely did not develop personas that refer to particular disciplines, even though this is often the first variable we think of when considering why different scholars have varying research habits. My reading showed that there are more reasons than simply discipline for scholars to have different behaviours, particularly the type of research question and the scholar's stage in investigation.

I developed three personas as a way of integrating variations into the user model.

2.2. Michael



Michael is the first persona. I envisioned him as someone who works in a speciality that is focused and self contained, rarely needing to look for information outside his domain.

He has an established reading routine, rarely looking at unfamiliar sources and focusing on journals.

Michael publishes frequently, always in prestigious journals with multiple co-authors. He is almost always cited as the senior researcher.

Michael has a strong reputation; he travels widely and often works on international research teams.

Michael has a very good connection for his computer in his home office. But he doesn't use his computer at work, usually being too busy during the day to do much information seeking.

I pulled the following characteristics from my readings and fit them with Michael's persona.

- Balancing the contained nature of his work is the heavy publishing output by researchers in the field particularly via articles. ⁶ This means Michael has to fit a large amount of reading into his busy schedule.
- Teaches senior classes in his specialty, textbook is primary reading. ⁷ This means that Michael doesn't spend as much time looking for material to use in teaching as many other scholars.
- Likes electronic access, but doesn't think it has affected the quality of his work ^{8,9} Michael values digital access because it saves him time, but he doesn't make use of many of the features that would be important to scholars who are actively looking for new sources or are starting new projects.

- Never asks for help in Library, except for locations, doesn't like changes in Library layout or anything that disrupts his routine ⁹⁻¹¹ Thinks library is overstaffed, he doesn't connect library and access, and believes the Library is mostly for undergraduates ^{9, 12}

Michael is more likely to be annoyed by what he sees as bureaucratic difficulties caused by library procedures, he doesn't see the Library or its staff as partners in his research and realistically the Library's primary service to him is to provide comprehensive and seamless access to the information in his field.

2.3. Athena



Athena made her reputation on the strength of interdisciplinary work. Her information seeking involves keeping up with core readings and reading widely to stay abreast of developments in many areas.

Much of her work is collaborative, and she publishes widely in a variety of formats, including government reports, book chapters and articles. Much of her work is the result of collaborations with colleagues and is co-authored.

She does most of her information seeking using the good computer connection in her office. She has a computer at home, but mostly uses it for writing as she does not have a very good connection.

Because of her wide ranging reading habits she is a heavy library user. She also buys many books for her personal collection, and is a regular user of interlibrary loan.

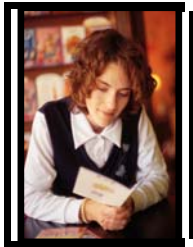
Some of the characteristics found in the literature that can be aligned with this persona are:

- She has a reading routine, but the amount can be overwhelming, electronic access has made much more available ¹³ This means that Athena is open to features that will effectively alert her to new materials and which will help her evaluate materials before reading.
- Has large, loose, network of colleagues, most work is collaborative, supported by large grants, love's working on "big questions"

¹⁴ Athena will want to be able to share information with colleagues both in and outside of the institution. Sharing and exchanging ideas are important parts of how her research community maintains cohesion and conveys status.

- Teaches all levels, adjunct in another department, uses variety of readings, images and maps in classes ⁷ Athena enjoys teaching and she is always looking for information that she can include in the classroom.
- Worried Library is going “too electronic” and will stop buying books ⁶ Athena values the Library for the variety of materials it provides, she is sensitive to any changes in collection or services that she perceives diminishes variety and depth in order to increase the digital collection.
- Uses electronic access, more information means she reads more ¹³ Athena isn’t a Luddite, she uses electronic sources to help her locate and access information. Now that so many of her colleagues use mailing lists, listservs, and blogs to communicate she is finding the rate at which new information emerges overwhelming. At the same time her natural curiosity and need to be knowledgeable means that she isn’t willing to shut out information, or restrict her range of interest.
- Has gotten to know the subject specialist and appreciates ILL staff ⁹ Athena identifies with the individuals in the library as well the services the library offers. This means that changes to either of these affect her adversely.

2.4. Sandy



Sandy is a recent PhD and she has just joined the Department. Her field is characterized by very individualized specialized topics. Each researcher tends to work independently.

Because she is new to the Department she is anxious to prove herself, this includes demonstrating mastery of information related to her topic.

Sandy’s publishing record is limited. So far she has only had two articles published, both based on her PhD. She knows the Department will expect much more output from her.

She still depends on the advisors from her PhD and the faculty in her graduate school as her main contacts and mentors.

Her teaching load includes two freshman classes and a seminar that focuses on the topic of her PhD. Teaching is new to her and she is finding it takes a surprising amount of time to amass the readings and visual sources she wants to use in the classroom.

Sandy has a laptop computer that she takes from work to her home and on business trips. She has good connections both at home and at work.

Some of the characteristics from the readings that can be attributed to Sandy are:

- Discipline is diffuse, reading patterns are idiosyncratic and dependent on research question ¹⁵ Sandy is looking to master her subject area, demonstrating that she is familiar with all the key literature, and to find new information that other scholars haven't used before.
- Her research process is time intensive ¹⁵ Finding information takes enormous amounts of Sandy's time, this creates pressure on her.
- Knows her Department highly values independent work ¹⁶ Sandy doesn't have the option of working as a junior partner in a bigger research project; she is under pressure to prove herself as scholar who can be independently productive over the long term.
- Uses electronic access, also original, archival, and microform sources ¹⁵ Sandy is anxious for information, using a wide variety of sources and even traveling to other libraries and archives.
- Library has only scattered holdings of the books she needs, visits larger libraries with better collections and spends several days scouring whole sections ¹⁵ Sandy needs to be complete, she can't afford to miss information, and missing information can invalidate her research.
- Finds libraries hard to use, likes archives and special collections because staff are helpful and she has gotten to know them ¹⁰ Sandy needs to feel confident in the assistance she gets. She prefers to work with people she thinks are experts on their collections. Subject expertise doesn't mean that much to her, as her topics are wide ranging and idiosyncratic.

3. Tying Behavioural Studies to the Model

This section gathers the observations of user behaviour reported in a variety of studies to the model developed by Ellis and amended by Lokman and Tibbo.

Each section begins with a brief review of the nature of the particular behaviour.

Findings are collated for the studies and organized by behaviour. Each section concludes with reflections on what features in a digital library might support the behaviour, and what collection measures might be used to assess how well the collection is supporting this behaviour.

3.1 Starting

Identifying and accumulating references to serve as starting points

The purpose of starting is to accumulate references and understand the key authors, problems, and resources in the field. Scholars in starting mode are not engaged in deep reading; rather they are scanning and trying to move relatively quickly.

Starting Behaviours

Frequency of starting varies from scholar to scholar. For example, Michael may spend his entire career working in a well defined topic, and starting is something he really only remembers from his early days.¹¹ Athena often finds herself starting as she familiarizes herself with new topics and literature; she may begin new projects and research questions every few years. Finally, Sandy as a new scholar is not only starting, she also lacks an understanding of shortcuts or a well defined approach to beginning research in a new topic.

Some starting behaviours can be aligned with Michael's persona.

Scholars may start by browsing journals, text books, or electronic databases,^{13, 17, 18} this is probably most effective when the discipline is well structured with defined research questions.

Starting often begins from familiar sources that that are known to be productive^{13, 17}, likely this is a good strategy in fields with a well structured and bounded literature. .

Even when starting a scholar working in discrete research areas are unlikely to ask for Library help.^{10, 11}

Athena's persona can be aligned with a less focused approach to starting.

Keyword searching the internet⁸ Google, Amazon¹⁹ the catalogue⁵, and^{13, 17} full text databases.

Amazon is viewed by some scholars as a comprehensive source of recent monographic literature.¹⁹

Athena is experienced and can judge the value of material from catalogue records, abstracts, and other metadata, she is able to make use of students to retrieve material from the Library^{11, 19}

Athena, who frequently starts new topics, can be expected to ask the subject specialist in the library for leads and assistance.¹⁰

Sandy who is interested in archival or primary sources won't tend to consult these sources directly while starting, but will try to identify which sources are important,. She will often use web based finding aids to do this. ^{5, 19}

Other starting activities included

- asking colleagues for leads ^{5, 9, 17}
- reviewing one's personal collection ⁵
- chaining from readings that seemed particularly important. ²⁰
- using historic literature to identify seminal readings ⁹

Problems

Some of the problems reported by scholars included

- Staying oriented. ⁹
Scholars working on a new topic and/or working in a new field may have difficulty keeping track of where they have searched. They often need to return to products and repeat searches as new keywords, authors, or topics come to their attention.
- Delays in access. ¹⁷
Starting is a process where much information needs to be scanned and assessed as the scholar works to build up a picture of the field. This probably works best when there is quick access to material, reducing delays and allowing the scholar to work in a flow, not interrupting the exploration of a new topic with complex or lengthy processes for information retrieval. Examples of interruption to flow include, having to repeat searches in multiple databases or complex retrieval systems.
- Pressure to be complete. ⁸
Clearly the ability to find materials on the internet and the overall expansion of access because of "big deals" means that more information is available. In addition, scholars are engaged in a very active online environment where communication and information sharing increase immediacy and overall awareness. Pressure to be complete arises in an environment where work is open to constant public assessment by peers.
- Poor documentation. ⁵
Some topics may be new or lightly researched. In this case the real lack of documented information may present a challenge. Depending on the experience of the scholar, it may be difficult for them to differentiate lack of information from poor information seeking methodology.

- Access to “hidden information”.
Depending on the discipline, scholars may have difficulty because important information is hidden through poor description or incomplete inventories. Examples include chapters in books, archival sources, and any material that is not fully described in predictable sources such as Abstracting and Indexing tools, the catalogue, or finding aids.

Some disciplines do not map well to natural language searching. ¹⁹ In the humanities, for example, the topic depends on context, and keyword or even subject term searching are often inadequate.

Collections that support Starting

Depth and range of content are clearly central to supporting starting. Scholars who are starting need to follow expansive paths of exploration. Traditional collection measures that emphasize the size of collections and the inclusion of a wide range of materials support of starting.

Starting in the Digital Library

In designing the digital library some tools to look for are:

- Broadcast searching across multiple databases would help limit repetition and aid discovery.
- Search logging, tools that help scholars record where they have searched, capture sources, and add personal annotations. This would help scholars track what work they had done, as well a logging tool may also help with creating a sense of “completeness” as it would give a better sense of the paths followed and items found.

Measuring support for Starting

Starting involves quick scanning and trolling, this means that collections need to properly exposed through the provision of metadata in predictable places. A key digital collection measure is the completeness of metadata provision. To assess this we should track the presence of metadata for distinct items in predictable sources.

- Are e-books and e-journals included in the catalogue?
- Are other formats described appropriately, for example government documents, technical reports, reference tools, or images?
- Is the metadata in various sources open to broadcast searching tools, for example Google Scholar or WorldCat?

3.2 Chaining

Following references gained through reading and personal contacts

Chaining is a favoured activity of scholars. Chaining provides an entry into the literature that is mediated by the prior judgement of relevance of other researchers. It helps with the identification of primary sources and crucial articles.^{10, 11, 19}

Chaining provides a way for scholars to move into unfamiliar subjects, expanding readings by identifying resources outside the familiar discipline and providing clues to identify key readings.¹⁹

Chaining is a path into information that is poorly organized or indexed.⁵ Examples include grey literature, data tables, or government documents. Here scholars are able to leverage the effort of others to save time and to continue a coherent scholarly dialogue.

Scholars assess the importance of following up on references using a variety of criteria including author, novelty, publisher, cost to acquire, effort to locate, colleagues recommendations, instinct and an assessment of the nature of the content being referenced.⁵ Perceptions of the richness of the resource and the level of detail may also be used in assessment.¹⁷

Starting Behaviours

There are many methods of chaining. We are familiar with using the references in a key published source such as a book or article.^{18, 19, 21} But scholars are also using chaining when they follow the links in web sites, follow up on author addresses in databases, or ask colleagues for suggested readings.²⁰

Some scholars also reported making use of forward linking, a relatively new innovation enabled by electronic databases.²¹

Michaels' persona can be linked to certain activities.

Scholars like Michael whose key journals are well defined, and who don't often start researching new topics will use chaining less.¹⁹

Michael, with his tendency to stick to the familiar, will be among those scholars who choose readings by availability.¹⁷

Scholars like Athena, who often works in new areas and use many different types of sources, will have a different approach.

She will use chaining often.

Chaining is useful for identifying methodological and theoretical articles ¹⁹and will appeal to an interdisciplinary researcher like Athena.

Sandy will find chaining productive.

She is working in a new area and needs to identify useful methodological and theory based articles.

She is actively seeking to understand the breadth and structure of her area of study.

Problems

Some problems scholars may encounter with chaining.

- Lack of access to items cited. ^{17, 18, 21}
As previously noted, delayed or difficult access to material blocks the flow of information exploration. If we believe that chaining is most important to scholars in unstructured or novel areas, then the ability to scan, evaluate, and quickly judge information is an important part of the process.

Reasons that access can be blocked include

- Not publicly available, for example internal reports, researcher owned data sets.
 - Rare, not widely held, for example primary source items.
 - Commercially published information that is not held by the scholar's library.
 - Low use material that is in storage.
 - Format, print materials that require a trip to the Library.
- Evaluating citations. ⁵
Clearly the more unfamiliar the area the more difficult it is for the researcher to employ common measures used for evaluating citations.

Collections that support Chaining

While libraries have long known that chaining is an important technique, there haven't been many tools that directly support it. The best know tool, ISI citation indexes, reflect a view of information seeking that best supports disciplines with a well defined core literature, and where frequency of use is an accepted measure of value. For scholars working in these areas, for example the physical sciences, ISI citation indexes have long been core tools.

However, scholars in many other disciplines may be seeking material that is either novel or which has a small potential readership. The advent of Google, a tool that used the viewing patterns of a very large population of users to rank web pages, revealed the strength of the approach to uncover and give relevancy structure to a wide universe of information sources. Google showed that the reading habits of specialized interest groups could be leveraged to troll for quality.

Chaining in the Digital Library

Many individual products now offer tools that support chaining, what is still needed are products that can offer functionality that spans multiple database and types of sources.

Some important features would include:

- Forward citation. Products should allow scholars to move forward from a work to other works that cite the item at a later date. Many products now offer this feature, but tools that allow this to work outside a single product are still needed.
- Features that allow readers to mark certain articles to know when they are cited at a later date are also beginning to emerge. These should be emphasised in evaluating products.
- Reading clubs. These tools would allow researchers to define areas of interest and then see what was popular among others with the same interests. Elements of this feature can be seen in online tools such as *delicious*. However these are limited by not having direct access to commercially produced sources.
- Live linking from references. This feature should be given more priority in evaluating databases and in requesting new features from publishers. Live linking should be enabled to operate across multiple vendor platforms, to catalogues, and to other potential information sources.

Measuring support for Starting

Chaining is be supported by collections that are well described. So the provision of metadata continues to be important. Measures of depth and range will also indicate if a collection is supporting the type of seamless access that maximises the productivity of chaining.

Another potential measure is to assess how many of our products are Open URL compatible, in particular allowing linking from references to sources outside the database.

We should also measure how open our local collections (digitized collections, local catalogues) are to linking by scholars outside the institution.

3.3 Browsing

Casually looking for information of potential interest

Browsing reflects the scholar's constant interest in finding new information. The readings in this report demonstrate that, although it is defined as casual, browsing is a deliberate, frequent technique employed by scholars who have learned from serendipitous experience to cultivate strategies and a mind set that keeps them constantly open to information discovery.

Browsing is particularly useful in identifying newly published material,¹⁹ and in identifying sources of primary information^{10, 19, 22}

Browsing Behaviours

Browsing methods include the well recognized techniques of scanning library collections of print journals and books, often opening up items to flip through the pages.^{5, 11, 19} Scholars also reported the analogous activity of scanning electronic Tables of Contents.¹⁹

Some less obvious techniques include browsing bookstores, publisher websites and even other colleague's personal collections.^{10, 19, 22}

Entering favourite terms into Amazon, Google, and full text databases are also favoured approaches.^{5, 19}

Perhaps surprisingly scholars reported regularly scanning the library catalogue.^{5, 10, 11, 19, 22}

Colleagues can also be used as browsing agents, with some making a point of asking them to keep an eye open when at conferences or abroad.⁵

While we often hear that electronic resources are not as open to browsing as traditional print formats, scholars can pursue activities akin to browsing using the digital collection. Examples include, browsing key journals online, browsing Amazon for new books, searching in Google using favourite keywords, searching key terms in full text databases, and catalogue browsing.

For Michael, whose subject area is well defined, a routine of browsing specific parts of the physical collections can be quite productive; and he can be comfortable knowing that most relevant information is kept together.^{11, 19}

Michael can be identified with those scholars who report finding almost half of new material through browsing.⁵

Scholars like Michael who rely on electronic access can also be expected to be heavy users of electronic Tables of Contents, and even to have TOC alerts set up.¹⁹

Athena and Sandy work in less structured disciplines, and relevant material may be shelved in many different parts of the library.

These scholars may make a practice of leafing through journal issues or always looking to see which books are shelved near titles they are retrieving.

“Hidden” information such as book chapters, archival sources, and un-indexed journal material is also uncovered through browsing.

Problems

Browsing can be hampered by various issues:

- Changes to access. For example, if the library rearranges or stores physical items then scholars who have a well developed browsing routine will be disadvantaged. Browsing should not be dismissed as a minor strategy simply because it is open, serendipitous and casual. It is a critical activity for scholars. Libraries should be cognizant of the disruption caused by changing web pages, new urls, physical reorganizations etc.
- Multiple platforms. The hallmark of browsing is that it can be casual and spontaneous. Trolling the internet using key terms is a good model of how we would like browsing to work in the digital library. In particular scholars should be able to broadcast a search across multiple platforms.
- Dark documents. Browsing helps uncover hidden information. This is content that is not fully described through metadata. Ebooks are a good example of how a source can be opened to a different kind of browsing as users are able to mine to contents using key terms. So even though the “page flipping” activity of print books may not be present, the digital format offers browsing capacity that different studies show scholars are well aware of.

Collections that support Browsing

Browsing in the Digital Library

Features of the digital library that support browsing are:

- Searchable/viewable contents. While full text journal packages may create the impression that the digital library is open, many products are not open to searching. For example, government documents, standalone journals, standalone e-books, all these can only be viewed by opening them individually; casual scanning for related materials is not well supported. As well, rare materials that are only available in print could be opened to browsing if they were digitized.
- Links from browsing tools into content. We already use software such as SFX as open url linking tools, and this provides a significant support to browsing. Knowing that scholars make use of tools such as Amazon for browsing means that we should also look for applications that take users directly from these external sources into our content. A good example of this is linking from Google Scholar directly to local content.

Measuring support for Browsing

Measures of collection browsability are:

- How much of the collection is completely open and searchable?
- How many key external searching or browsing aids can mine local content?
- What proportion of the rare hard-copy parts of the collection are digitized and open to searching by external sources?

3.4 Monitoring

Maintaining awareness to increase professional knowledge and gain exposure to new ideas. Understand the current state of research, find new methods

Monitoring is the activity that helps scholars keep up to date in their discipline, staying aware of new publications, knowing who is working in the field, and what research questions are being pursued. ^{19, 21, 23}

Routine is an important part of monitoring ^{1, 6, 23, 24} and a key way in which it differs from browsing, which tends to be more spontaneous and opportunistic. Scholars appear to identify what they consider to be productive techniques or sources, and then make a habit of using these to stay up to date.

Monitoring Behaviours

Some of the approaches described included

- use of the literature for leads, for example having personal subscriptions to frequently read journals ⁵

- perusing new book lists and conference proceedings ^{5, 6, 24}
- reviewing and scanning book and journal bibliographies ²⁵
- use of pre-print sources ¹⁹
- electronic alerts¹⁷.

Communication with fellow scholars is also involved in monitoring and includes

- reading listservs ^{5, 6}
- conversations with local researchers ^{5, 10, 25, 26}
- email correspondence ^{5, 6, 26}

Researchers like Michael can probably rely on a regular routine that involves a limited number of sources. ⁶

Consequently scholars working in these areas will feel confident that they are aware of developments in their field.

Such a scholar may also be less likely to use alerting services given that they are not trying to track a wide range of literature. ^{17, 26}

The exchange of pre-prints is a growing trend, ¹⁹ and we can align this characteristic with Michael, who is likely to have good contacts with key scholars.

In contrast a scholar such as Athena reads a much wider range of material and may use tables of contents and alerting services to maintain awareness of materials that are not part of her regular reading routine. ¹⁷

She would also be among the scholars that reports communication with colleagues as a key way of staying up to date. ¹⁰

While Michael is under pressure to read everything relevant to his field, Athena's interests are so wide ranging that she is content to have a good level of awareness and to focus her actual reading on what is directly relevant to her current research project. ¹⁹

Sandy, because she is under pressure to show mastery and because she is new to her subject area, does feel obligated to be both aware of and read new materials in her field. ¹⁹ For a scholar like Sandy, monitoring is a time consuming task.

Problems

Monitoring can be complicated by the same issues as beset browsing, other issues include:

- Expanding literature. ²⁷

In most fields the amount of publication is expanding. Staying aware without becoming overwhelmed will be a challenge for almost all scholars.

- Interdisciplinarity.¹⁷
The trend to interdisciplinary studies may complicate routine, creating a wider range of topics and sources that need to be monitored.

Collections that support Monitoring

Many of the features considered for browsing would support monitoring.

Monitoring in the Digital Library

Additional features of a digital library that would support monitoring include:

- Ability to set alerts.
Many individual databases have this feature, but the ability to broadcast alert searches across sources would increase efficiency. Including sources such as blogs, newsfeeds, and listservs would increase effectiveness, particularly for scholars with wide ranging interests, those who work in rapidly developing fields, or those that incorporate primary sources into their research.

Akin to alerts is the ability to monitor sites for changes, for example if a certain website is critical to studies in an area, then having an alert when the site changed could be an effective feature.

- Ability to share evaluations.
Products such as Faculty of 1000 demonstrate that scholars can share opinions about material to determine which articles are important to read. We should seek more tools that would allow scholars to know what others in the field are reading.

Measuring support for Monitoring

Collection measures that could be used to assess our support for monitoring include:

- What routines do we use to distribute information about new library holdings? For example, new book lists.
- What are the delays in bringing new sources into the library, for example, the extent of embargos and the time between selection and availability?
- What routines do we use to distribute information about new primary source materials to scholars outside the institution, for example new archival holdings or special collections?

- How easy do we make it for scholars to identify our holdings, for example accurate and complete cataloguing?

3.5 Accessing

Obtaining information once sources are identified

Obtaining information is the scholarly information activity that is most closely associated with the Library, even as the range of possible sources expands to include a much wider range of options. The Library may be somewhat too focused on accessing as primarily a prelude to reading. Accessing is also a requirement for activities such as quick review and fact checking.

Accessing Behaviours

Library focused techniques for accessing includes visits to the physical library and the use of interlibrary loan/document delivery.⁵ The digital library provides access via full text databases.⁸

Access techniques that do not depend on the library include:

- correspondence with colleagues, visiting author web sites, and personal subscriptions.¹⁸
- use of personal and colleagues' collections including subscriptions^{5, 18}

Scholars also reported traveling for the purposes of obtaining information; we can speculate that this is most often done for primary sources or private/proprietary material.⁵

Scholars have a preference for electronic formats for initial access.²⁷

Scholars such as Michael who have a limited range of materials they consult regularly may have cancelled personal subscriptions in favour of the library's electronic access.^{11, 18}

Visits to the library are a nuisance for Michael¹⁷, and he sees electronic access as a time saver that greatly improves his efficiency.¹⁹

Michael will print out significant amounts of material at one time, reading them later when he feels he has time.¹⁷

Athena, who was once a heavy user of ILL and document delivery has seen her use of these services decline as electronic access increases.¹⁸

Sandy as a new scholar has never set up personal subscriptions and relies on electronic access as the first place she goes to when trying to access information.

⁶

Problems

Problems in accessing information are frustrating to scholars, causing delays that are seen as essentially unproductive, interrupting the flow of research. Some issues include:

- Visiting the physical library.
Library related problems include concern that visiting the physical library is time intensive and materials are hard to locate.²⁵ Older materials often require a visit to the library, with an added layer of frustration if the items were in storage.^{11, 18}
- Library doesn't provide access.²⁵
Not only does this interrupt the flow of information work it also has a certain monetary impact. The cost of obtaining information the library doesn't supply is a concern. Buying books is expensive as are document delivery fees.⁵ Concern about this expense makes sense when we remember that access is often a prelude to evaluation, not to sustained reading.²⁷
- Nature of the information.⁵
Other concerns in access are less process oriented. For example some geographic or topic areas are secretive by nature, items such as certain government documents or archives may be difficult to view. Some published items are rare because of age or low publishing runs, obtaining this information can be difficult.

Collections that support Accessing

Libraries provide expanding content through “big deals” and the purchase of other electronic resources. The different studies indicated that this was a greatly appreciated innovation, not only because of the additional content, but also because digital access meant that material was easier to find and saved the time of the scholar.

Digital access to information beyond that supplied by the library is also a critical tool for scholars, and supporting this type of access in addition to more straightforward collection building activities is important to creating a functional digital library.

Accessing in the Digital Library

- Broadcast searching.
To locate known items quickly, searches should be distributed across different databases and external sources.
- Intelligent citations.

Services such as COINS and open urls support linking from citations to sources, even when the citation is embedded in a web page.

Measuring support for Accessing

- The completeness of metadata, both in library provided tools such as the catalogue, but also as in external searching tools such as Google Scholar and Worldcat.
- The extent of the digital collection, including the provision of digital versions of material that the library may already have in print or other formats improves access.
- Licenses that support sharing. Interlibrary loan and individual sharing of information creates networks that not only provide access into our collections but also into the collections of other institutions. Licensing should be monitored to ensure that these rights continue to be supported.
- What processes are in place to provide access to materials outside the collection? For digital materials this may mean that the Library should explore means of providing temporary or individual access for specialized materials that do not fit into the Library's long term collection.

3.6 Differentiating

Evaluating and judging material before reading in depth

Scholars will routinely filter information prior to reading in depth. The variation between amount accessed and amount read will vary by scholar and field. However, it is important to value this filtering process, because it indicates that scholars need access to more information than they actually read or directly use in their research.

Libraries historically offer significant amounts of information that support differentiating; this ranges from commercial abstracting & indexing services, to in-house produced catalogues and finding aids.

Differentiating Behaviours

The literature reviewed for this study revealed a relatively short list of techniques for differentiating.

- Reading article abstracts ²⁰
- Reviewing the elements also used in chaining—author, novelty, publisher, cost, frequency of citation, effort needed to locate, and the critical nature of the information ⁵.

- Prior knowledge of the source (i.e., the journal reputation). ¹⁷
- Flipping pages of print materials or the digital analogue of previewing of online text. ^{9, 28}

Scholars such as Michael who work within familiar fields can be expected to find differentiating relatively easy. For Michael ease of access to allow previewing of sources before printing and reading is likely the most appreciated service the Library can offer.

In contrast, scholars like Athena who read very widely and often in new areas most likely view and filter a great deal more information. ^{10, 19} Athena will also be cautious in her use of information, particularly when she is working in controversial areas, she is aware of bias and always seeking balancing points of view. ⁵

Sandy, like Michael, will appreciate electronic access. In her case it will be because she reviews so much material that the digital format is a great time saver over having to retrieve and evaluate print materials. ²³

Problems

Few real problems were noted in differentiating, with scholars confident of their ability to make decisions as long as metadata and access were provided.

Collections that support Differentiating

Differentiating is a key task for scholars, ensuring that they maximise their time. Mistakes in differentiation can lead scholars astray, at the least causing some embarrassment when key sources are passed over, at worst invalidating or bringing research into question.

The availability of digital information sources has a positive effect on the differentiation process. It saves time by providing easy access to a wide range of material for previewing.

Differentiating in the Digital Library

To support differentiation a digital library should include the following features:

- Background information.
Links to information about authors and sources may aid with differentiation, particularly when working in unfamiliar areas.
- Frequency of reading or citation.
Tools that provide information regarding how often an item has been cited (forward citation) and how often an article has been viewed will be helpful.

Measuring support for Differentiating

Measures of how well a collection supports differentiation include:

- Provision of metadata.
- Licenses that support extensive downloads for later differentiation
- Proportion of the collection that can be previewed electronically...here Google Print may play a future role.

3.7 Extracting

Extracting relevant information. Leads to integration of information into the mental model

Extracting is the process by which scholars assimilate information and create new knowledge or mental models. It allows the critical analysis of ideas and theories put forth in documents.²⁹ Extracting is also a way of acquiring data, evidence, exemplars and sources for theory.²⁹

Extracting Behaviours

Extracting is most often accomplished through intensive reading. Important articles are often reread many times to increase information extraction.¹⁹

The studies showed a strong preference for print copies for intensive reading.^{9, 28}

Scholars often annotate materials as they read, one of the reasons print and/or personal copies are preferred.²⁰ Annotations are how scholars record associations and responses to text, these form an important part of the scholar's personal collection and are often re-read or revisited during the writing phase.³⁰

Scholars tend to keep copies of important items, they prefer to buy their own books and make extensive photocopies and printouts.^{13, 19}

Reading is an activity that seems to vary in amount by type of scholar, but it did not appear to vary in process or activity. The only key difference found was the tendency for some types of scholars to be reading older material.

Scholars such as Michael who are confident of their knowledge of work in the field will mostly read newer material.

Scholars such as Sandy who are building up their knowledge will mine older material more intensively.¹⁸

Problems

Reading can be complicated if certain problems occur.

- Too much information.
There may be too much information to properly absorb through intensive reading.⁵
- Too little information widely scattered.
A topic may be obscure or poorly documented meaning that information is difficult to find, creating a need to read widely and carefully with for limited pay back.⁵
- Difficulties in access.
Reading often takes place for long sustained periods of concentration; difficulties in access may interrupt this flow.¹⁷
- Budgeting time.
The digital environment has increased the available amount of information creating pressure to read more. This is a challenge for faculty trying to balance the time they give to different activities.^{11, 13, 18}

Collections that support Extracting

One of the key challenges facing digital collections is that this is not the preferred format for extracting. Consequently digital collections need to be designed to support printing and re-use.

Extracting in the Digital Library

Some of the features that will help with this are:

- Licenses that allow extensive printing.
This is more important as we move into digital books and other types of extended content. The limited printing options now available for many of these licenses don't support intensive reading and seriously limit the usefulness of these formats.
- Quality of printing.
Scholars who want to make use of highly visual materials—particularly images or historic documents require both high resolution display and printing capacity.

Measuring support for Extracting

Collection measures that indicate how well the library supports extracting include

- Do licenses support realistic amounts of printing?

- What is the quality of reproduction that we support in text, images, or sound files?

3.8 Verifying

Checking the veracity of information

Verification includes checking on factual information, but it also means revisiting sources to verify quotes, citations and interpretations.

Verifying Behaviours

Scholars have different sources they use for verification work. They may use their personal collections for much of this.

They may also seek out older literature prior to citing in publication or to check other scholar's reporting of the work. ¹⁸

Scholars such as Michael that work in well defined, well documented areas probably have less problem with verification. They may also be among the group that is comfortable with their ability to judge online factual information. ²³

Scholars like Athena who use a variety of formal and informal sources will more often encounter problems with previously read works disappearing, particularly online materials. ²³ While she may use the internet to verify quick facts for lectures or correspondence, or use her colleagues to help confirm facts, she will tend to use authored sources for her own publications. Her concern about bias in sources will lead her to seek alternative views. ⁵

Sandy can be imagined to be among the type of scholars who are particularly careful in verifying and using authoritative sources. She will tend distrust web pages or references from colleagues. ⁸

Problems

Some problems that can be encountered during verification include:

- Need for alternative sources.
Scholars who work in controversial areas may feel a higher level of verification is needed than scholars in other fields. ⁵
- Distrust sources.
Again depending on the subject and the type of verification, non-academic documents or websites may be viewed as unreliable. ³¹ The more novel the information is to the researcher the more likely it is that they will have problems evaluating the reliability of sources.

- Stability.
Verification involves citing sources in publication. The stability of these sources is a particular concern. Novel sources such as e-prints or web pages are concerning to scholars because they may not be stable. ¹⁹

Collections that support Verification

Verification is an important, time consuming activity for scholars. The digital library faces the problem of residual distrust of online sources and legitimate concern regarding its stability.

Verification in the Digital Library

To support verification the digital library should.

- Indicate stability.
Collections should indicate the stability of information, for example are these sources that are controlled by the library over the long term, does the vendor have long term rights to the product, or is the material likely to be fluid over time.
- Provide authoritative content.
Scholars still depend on authoritative content, and providing content without regard to this diminishes trust in the digital collection. The library should seek to include scholarly editions and to exclude content without adequate provenance.
- Make factual information easy to find.
Broadcast searching that mines the content of handbooks, reference tools, and dictionaries may significantly enhance use and effectiveness for verification. It would also assist scholars in making it easier to find multiple sources of information.

Measuring support for Verification

Collection measures that assess support for verification include:

- How much of the text in factual sources is open to centralized searching?
- What processes are in place for ensuring long term collection stability?
- What strategies are in place to provide scholarly and authoritative editions?

3.9 Networking

Sharing information and maintaining relations with fellow scholars and other associates

Research is a social process; it accrues value and moves forward through its adoption and discussion by other researchers. Scholars seek ways of making their own ideas readily available, and increasing their profile in the community.¹⁹ Novel ideas are prized in the community and networking often involves the exchange of new ideas or information.^{10, 29} This type of social activity also helps scholars identify people working in related areas and provides an entry into new contacts.¹⁹

Networking Behaviour

Some of the ways in which networking is accomplished include

- Distributing print or electronic copies of useful articles, e-mails, or book chapters to colleagues.²⁰
- Sharing annotations and notes when working on common problems or writing collaboratively.³⁰
- Sharing links to useful web pages.³²
- Online communication has been a great boon to sharing allowing participation in online forums and listservs.¹⁰

Networking can even be seen operating across generations, as scholars make use of prior researcher's annotations and notes in creating new knowledge.³⁰

Interestingly, in choosing contacts there is some evidence that scholars deliberately choose to make linkages to researchers that will bring them new skills or knowledge, favouring those with better resources and who are not well known to them. Librarians are rarely seen as part of this network.³³ In choosing contacts scholars may well be making use of the same criteria they use in chaining and differentiating.³⁴

Publication is a type of networking that indicates research performance to colleagues, and helps the scholar gain recognition in their field.¹⁴ Publishing outlets are chosen to maximise the impact of research, features such as peer review and prestige are favoured, although some consideration is given to exposure and speed of publication.¹⁴

A traditional type of networking is to exchange papers with colleagues prior to publication.¹⁴ A scholar such as Michael who works in a competitive field, with many researchers working on the same topic may be expected to have a small circle of trusted colleagues with which he networks.

Athena, in contrast, depends on having a wide circle of colleagues and is more likely to share information freely.²⁹ While Michael may feel articles are his main distribution mechanism for results, for Athena publication is a looser more wide ranging activity including books, book chapters, and government sponsored reports.¹⁴ Her work as an editor and peer reviewer are another way in that she maintains awareness of work in the field.¹⁵ Athena has established routines for developing contacts as she moves into new areas, seeking out scholars who she

feels will be of help to her in learning the vocabulary and discourse of the topic.¹⁵ Because of good reputation and the quality of her scholarship Athena finds that scholars tend to be open to her attempts to initiate contact.¹⁵

Networking is important to Sandy because it helps her confirm the directions her research is taking. She shares drafts of some of her work with the mentors from her graduate school days.^{16, 29} This social dimension is important to her as so much of her everyday research and reading is solitary.²⁹

Problems

Some of the problems encountered in networking include:

- Heavy traffic.²⁹
Maintaining contacts is a commitment and requires ongoing effort. At times when other pressures, such as teaching, reading or writing, require concentration, it can be difficult to sustain networks.
- Quality of discussion.²⁹
Listserves in particular can be problematic for researchers. They are seen to be too open and to have a low quality of discourse. Considering the importance scholars place on novelty, relevance, and authority, these misgivings are understandable.
- Lost recognition.²⁹
Wide communication and dispersal of ideas causes some concern that control over concepts will be lost, and authors will lose recognition for their work.

Collections that support Networking

Digital collections would seem to be ideal for networking, as they potentially allow for easy sharing of content.

Networking in the Digital Library

Features that digital libraries should have to fully support networking include:

- Licenses that allow reasonable sharing.
Scholars should have the ability to share documents with colleagues at other institutions on a limited, non-systematic basis.
- Emailing articles/excerpts.
All databases should have a feature that allows easy distribution via email.
- Collaborative spaces.
There is a need for applications that allow sharing of annotations and even collaborative writing.

Measuring support for Networking

The key measure that indicates how well a collection supports networking is:

- How standard are licensing provisions that support sharing?

3.10 Managing

Filing archiving and organizing information for context and retrieval

While often overlooked in library planning, scholars expend much effort on organizing information to preserve its personal value over time.⁵ Proper organization makes information easier to share with others.³² Interestingly, there appear to be two types of organization, one which preserves access to material for the short term, and another that is performed when a scholarly decides that an item will form part of their personal collection, having long term value.³²

Personal collections support the re-reading that scholars conduct with critical sources.¹⁹ Personal collections are often the first place scholars look for information when preparing for teaching or when beginning a new project.¹² Scholars often take considerable pride in the personal collections they build over time, seeing them as reflections of their careers and the quality of their scholarship.²⁹

Managing Activities

Some of the activities involved in managing include

- The creation of personal libraries through purchasing, photocopying, microfilming, and print outs.^{5, 10, 20, 29} These collections can also include notes, links, and correspondence.¹²
- Use of electronic copies in the personal collection.⁵ It appears likely that storing electronic copies either instead of, or as a supplement to print copies, will grow. Electronic copies having the same advantages of accessibility and transferability that they have in other digital library contexts.
- Manual organization schemes for personal collections.
- Use of commercial software to organize references and text files.³²
- Bookmarking or adding links to toolbars is another way in which electronic information is organized for later retrieval, alternatively scholars sometimes email links to themselves.³²
- If scholars believe information can be easily re-found, they may not make any effort to save personal copies or links.³⁵

- Scholars who keep annotations, journals or other types of written notes consider these an important element in their personal collections. ³⁵

Visits to the library historically result in additions to the personal collection. ¹² The advent of digital sources, with easy finding and printing, has resulted in even more extensive personal collections. ³⁶

Scholars like Michael who read a limited number of journals are likely among those that are cancelling personal subscriptions in favour of using the library's electronic access. ¹⁹

Athena, who reads broadly, will face more challenges in building and organizing a personal collection. She may be among those that often use links and bookmarks rather than print copies. ³² She is also likely to use bibliographic software to organize her readings, particularly if she can use this at the same time as she does her online research. ¹⁰ She has a book collection that is organized by author, and extensive photocopies/printouts that she organizes by date. ³⁶

Sandy represents the young scholar, who is expending much effort both on acquiring material for her personal collection, and in developing strategies for keeping information organized. ^{9, 15} She finds that many items she would like for her personal collection are not available for purchase or are expensive. She has invested in making photocopies of entire texts, even though she knows this is against copyright, she doesn't feel she has any alternative. ¹²

Problems

Some of the problems encountered in managing include:

- Organizing tools are complex.
Software, such as bibliographic management tools, may have a steep learning curve and may not offer the customizability that scholars need. ¹³ Another problem may be that once these products are learned and collections are integrated into them, the product will lapse or no longer be supported. ²⁹
- Inability to retrieve.
Many things can lead to this problem, including instability in online sources, also more prosaic issues such as forgetting titles or how to access information. ²⁹
- Diversity of approaches.
Multiple formats and sources mean that scholars develop different ways of managing different types of information. Unfortunately this increasing diversity can have the result of making management more difficult. ^{12, 35}

Collections that support Managing

Certainly libraries are showing an increased interest in supporting bibliographic management software. However, digital collections increase both the amount scholars may wish to manage and set challenging questions regarding stability and personal storage.

Managing in the Digital Library

Digital collection features include

- Integrated bibliographic software.
What types of software does the library provide or support?
- Personal link management.
Links are key information sources and tools that enable their management, something a bit more sophisticated than bookmarks are needed.
Management tools are beginning to include these features.
- Persistent urls.
Sources may not provide persistence urls, and scholars wishing to have stable links to references for managing, to share in teaching, etc. face a significant failing.

Measuring support for Managing

- Compatibility with bibliographic management software
What proportion of products allows export of formatted citations? This should include non-standard sources, for example image, sound or data resources.
- Persistence of links
How often do resources change link addresses and what mitigation is provided when this happens? Changing addresses can undo significant amounts of work on the part of scholars.
- Downloading/printing rights
Do licenses and products allow for generous downloading/printing rights, at the least reproduction of meaningful units—i.e., chapters, articles, entries. Scholars will want personal control over information in order to use their own managing mechanisms. Restrictive printing or downloading is a significant barrier.

4. Design Elements for a Digital Collection

As expected reviewing studies of user behaviour suggested how different features could enhance the effectiveness of digital collections. Table 1 provides a matrix that ties different elements to specific user behaviours.

Table 1 Design Elements for a Digital Collection

	Starting	Chaining	Browsing	Monitoring	Accessing	Differentiating	Extracting	Verification	Networking	Managing
Broadcast Searching	√				√			√		
Search Logging	√									
Forward Citation		√								
Reading Clubs		√								
Live Linking from References		√			√					
Searchable/Viewable Contents			√					√		
Links from Tools into Content			√							
Sharing Evaluations				√						
Source Information						√				
Frequency of Reading of Citation						√				
Allow Extensive Printing							√			
High Quality Printing							√			
Stability								√		
Authoritative Content								√		

	Managing	Networking	Verification	Extracting	Differentiating	Accessing	Monitoring	Browsing	Chaining	Starting
Allow Sharing		√								
Emailing Contents		√								
Collaborative Spaces		√								
Bibliographic Management Software	√									
Personal Link Management	√									
Persistent URLS	√									

The analysis provides some surprises, linking disparate product features through the support of specific activities. For example, the importance of persistent URLS finds company with personal link management and bibliographic management software. This illustrates how the activities in one library technology, such as the catalogue where we continue to struggle to find a way to easily create persistent URLS, with the long term viability of bibliographic management software.

Another example is how the provision of a technology such as collaborative virtual workspaces is intertwined with negotiating licenses that allow substantial emailing and sharing of contents.

While not the focus of the study, this analysis may prove useful in anticipating how different features will be of use. As well, because it proposes specific ties between activities and features, it is open to ongoing verification, with the potential to improve our understanding of the relationship between features and user behaviour.

5. Measures for a Digital Collection

A key concept in this analysis of user behaviour and measures of effectiveness is that, through the choice of measures, libraries determine how the digital collection will develop. In this way, measurement and evaluation are part of the design process.

In this section, I review the measures suggested by the study and subject them to a claims analysis, a review of the positive and negative potential of each measure. This step is meant to help create an understanding of how application of the measure will impact development of the digital library.

This final section makes specific reference to the University of Calgary Library, pulling the analysis from the theoretical to a more applied specific context.

Table 2 Measures for Digital Collection

Metric		Starting	Chaining	Browsing	Monitoring	Accessing	Differentiating	Extracting	Verification	Networking	Managing
Cataloguing & Metadata	Commercial Content	√		√	√	√	√				
	Local Content										
	Rare or Unique										
	Transparency										
Linking	Direct Linking		√	√		√	√				
	Different Formats										
	Local Content										
Direct Access	Commercial Content	√		√		√	√	√	√		
	Rare or Unique										
	Dual Format										
	Searchable										
Communication Routines					√						
Responsiveness	Time Lags				√						
	Embargos										
	Rush Processes										
Unlicensed Content	Interlibrary Loan					√				√	
	Scholarly Sharing										
	Individual Access										
Personal Copies	Liberal Printing						√	√			√
	Quality of Reproduction										
	Liberal Downloads										
Stability	Control								√		√
	Persistence										

Metric		Starting	Chaining	Browsing	Monitoring	Accessing	Differentiating	Extracting	Verification	Networking	Managing
Management	Exportability										√
	Customization										
	Sharing										

5.1. Cataloguing and Metadata

Cataloguing, which here includes all types of metadata, including provenance, descriptive cataloguing records, subject headings and classification, tables of contents, cover images, and the citation information found in commercial databases, is central to the success of many scholarly information seeking behaviours.

Cataloguing brings the records of disparate sources together in a uniform and predictable organization. Organized access to metadata not only supports finding, it allows evaluation and accessing. Lack of cataloguing is perhaps the most critical issue that can hinder effective use of information resources.

Cataloguing and Scholarly Behaviour

- **Starting** depends on cataloguing as scholars often search for attributes, such as subject or author, which are not well structured in full text searching. As well, catalogues bring together like materials from disparate sources streamlining the starting process and creating greater confidence for users working in new areas.
- Many scholars use the catalogue, finding aids, and other organized collections of metadata for **Browsing**, entering keywords, authors, or subjects looking for new material. For scholars who depend on hardcopy collections, such as rare books or archives, the metadata found in finding aids may be the primary means of browsing remote collections.
- Cataloguing also supports **Monitoring** by providing centralized, familiar sources that can be mined for new information.
- A key attribute of any information source is its location, and **Accessing** is supported by assuring that all information sources have location information, ideally for digital information this is a direct link from the record to the item.
- **Differentiation** is supported through the provision of metadata, including subject headings, author, source, and abstracts. Even when full text, full images are available, differentiation is greatly simplified by the provision of

appropriate metadata. For scholars using primary information, such as datasets, archives, or images, metadata can be critical in understanding the validity and usefulness sources.

For a scholar such as **Michael**, who works in a well structured field with a limited range of literature, cataloguing may not seem as important. But for scholars such as **Athena**, who consistently start new projects and who monitor a wide range of literature, the information found in metadata and the gathering of like materials together is extremely useful and timesaving. Finally, for **Sandy**, who is working in unfamiliar territory, and who often needs to preview traditional collections prior to visiting other institutions, the catalogue is a key tool for identifying and differentiating resources.

Metrics for Cataloguing

Two metrics are suggested to assess cataloguing progress—Completeness and Transparency.

1. Completeness

- a. What proportion of commercial content in any format is catalogued?
Cataloguing may not be records in the Library catalogue, but for each format a strategy for cataloguing should be established, and the proportion of the collection that is captured through cataloguing should be noted.
- b. What proportion of locally created content in any format is catalogued?
This refers to locally digitized materials, including material in depositories, archival sites, and historic collections.
- c. What proportion of rare/unique local content in any format is catalogued?
Cataloguing of this type of material, whether digital versions are available or not, is a key contribution the individual institution can make to scholarship.

Negatives:

Focusing on cataloguing measures may lead to some unintended consequences. The most critical one identified here is that the current resources applied to cataloguing may be overextended by this emphasis. The current strategy in the Library is to seek new processes and approaches to cataloguing that shift from the individual handling of material to batch processes that acquire cataloguing externally or use vendor supplied information in cataloguing.

Internally, the Library should adopt aggressive measures for the development of new processes for cataloguing, including batch loading, purchased cataloguing, and shelf ready plans.

2. Transparency

What external searching tools is the cataloguing open to, i.e., Google Scholar, WorldCat, LibraryThing etc?

As users move to employ a wider range of information tools, many of which are independent of the Library, it has become clear that Libraries need to place their cataloguing in this open environment. WorldCat is a traditional example of how libraries have carried out this process in the past, deliberately delivering data to cooperative repositories. Google Scholar and LibraryThing are examples of how this process is moving forward, with users employing Library cataloguing without any deliberate decision to use Library based tools, or any intervention on the part of the Library other than the effort of making internal information available.

Negative

We may accept tools that are less effective for certain types of users, because open products support a majority of queries.

We've experienced the overwhelming success of many interdisciplinary tools/sources and this is one of the strengths of the digital library. However, we still need to accept that one type of cataloguing will not suit all products, and format and disciplinary specific cataloguing is still required.

Because this type of cataloguing is really an attribute of individual resources, it is difficult to develop a collection wide measure that will ensure this issue is addressed. It is perhaps best to address this in reviewing the results of the other measures in the cataloguing section, watching for instances where specific types of cataloguing, for example that of images, is not being addressed under either transparency or completeness.

5.2. Direct linking between citations and sources

Direct linking between citations and sources supports the scholarly activities of Chaining, Browsing, Accessing and Differentiating. It has the potential to remove the often frustrating searching step for specific citations or items.

- Direct linking from indexing tools to content clearly supports **Chaining**. If links also work from references embedded in text then this approach is even more effective.
- **Browsing** is best supported when spontaneous interests can be followed, direct links avoids the need to stop and search separately for an item.

- **Accessing** can be a time consuming and frustrating activity, again linking has the potential to greatly simplify this process for scholars.
- Direct linking again allows scholars to drill down to content without using the time more traditional means of retrieval require. This supports more efficient **differentiation**.

All scholars will appreciate linking directly to content from references. In the short term it will likely be of most use to scholars such as **Michael** who work within a fairly standard selection of literature. Scholars such as **Athena** and **Sandy** will need to maintain their locating skills as their searches for grey or rare literature and unusual formats means that it is less likely that linking will meet all their needs. Particularly importance to these scholars will be accuracy and the informative nature of negative results.

Metrics for Linking

1. Completeness

- a. What proportion of products allows direct linking from citations to sources?

In most cases this will mean being open url compliant and having the ability to link from references embedded in full text.

- b. How reliably does our linking software (SFX) find content in specific formats?

One of the critical weaknesses we may find is that linking does not operate well for some disciplines or formats. Assessing the effectiveness of linking is core to understanding how well different user groups are being served.

As well, if the software provides inaccurate information, for instance indicating an item that is available isn't, this could present a significant barrier to users.

2. Transparency

Is locally created content open url compliant allowing the creation of links that point to specific items, and having live embedded references?

Negative

The more prevalent linking is, the more likely traditional ways of finding information will frustrate users. Tools that do not provide linking may not be as popular with users.

5.3. Direct Access to Resources

The strength of the digital library is direct access to information objects, including text, images, numeric/spatial data, and sound files. While paper or originals are still the preferred means of extraction for many types of information, digital presentation supports many key scholarly information activities.

- **Starting** is enabled by searchable full text, particularly where specific key terms are being used, terms that may be too current or which may not translate well into controlled vocabularies.
- **Browsing** is enabled as full text/image supports activities similar to flipping through books, or scanning articles. If the text is searchable, then the electronic equivalent of browsing—searching key terms—is enabled.
- Support for **Differentiating** is similar to that Browsing; again the ability to directly preview items allows more precise evaluation of the suitability of an item.
- As stated earlier, the ability to move from a search directly to a digital item greatly eases **Accessing**, and is highly valued by scholars.
- While we often feel that digital formats do not support **Extraction**, they can play a role. For some formats, such as images, the digital format can provide a surrogate to the original that is appropriate for comparison and reference, even if the original is still preferred for in depth study. For sound or moving image files, the digital format may be preferred to analog.
- **Verification**—rechecking facts, interpretations, citations, or references by other scholars, is greatly supported by the immediacy of digital texts.

For a scholar such as **Michael** digital full text will meet almost all his information needs, and he is more likely to be frustrated when the few items that he does want turn out to not be available digitally. Because Michael tends to focus on journal literature, where articles are easier to print out, he will again want the library to emphasize the development of a digital collection.

For **Athena** who reads a wide variety of material, much of which is not available digitally, there will be more concern that the Library not build a digital collection at the expense of traditional formats for books and grey literature. She is more likely to see the digital library as a move to a homogenous collection without the breadth and depth she requires.

Sandy, will share many of Athena's reservations regarding the digital format, but she will also appreciate the digitization of rare and special collections as they will greatly assist her in her work, allowing her to refer back to materials that she may have already viewed in other collections.

Metrics for Direct Access

Again, two groups of measures are considered to evaluate the provision of digital access.

Completeness

1. What proportion of purchased content is available in digital full text/image format?
This measure tracks the growth of the digital collection and assumes that the long term direction of the collection is the prevalence of digital over traditional formats. Understanding the rate of change will help us respond to scholars concerns regarding the need for traditional formats for some types of work, in particular for extraction.
2. What proportions of rare/unique documents held by the library are available as full text/image?
While such materials may still be viewed by scholars in the original format, and their preservation is critical, the use of digital surrogates has many well documented benefits, both to the survival of the item and the use by scholars.
3. What proportion of the print collection can be previewed online?
There is current trend to dual formats for some type of material; this is in recognition of the preference of print/original formats for some activities, such as extraction, and the preference for digital for other types of activities.

Negatives

The growth of the digital collection is not an unmitigated improvement in library resources, two key issues with digital formats are stability and the unsuitability for close reading. In monitoring the growth of digital collection we need to keep in mind that this is accompanied by some unresolved issues.

A second concern may be the limited computing resources available for the delivery of digital content. While some of this relates to the individual scholar's workstation and is not directly under library control, other issues, such as the notorious lack of bandwidth at the University are problems that need to be addressed centrally.

Transparency

1. What proportions of full text items have searchable full text?
Some of the scholarly information activities that can be supported by full text are only effective if the text is fully searchable. Accompanying this measure of quality is the need for tools that can effectively broadcast searches across multiple text platforms.

Negative

Again there may be a tendency to view full text searching as a substitute for the provision of organized metadata. However, for scholars who use document attributes such as provenance, state, author or sources to differentiate and search, full text is not an appropriate substitute. Care needs to be taken to expand both aspects of the digital collection.

5.4. Communication

The Library engages in many communication activities but in this context we are focusing on communication that directly relates to monitoring, or awareness of new resources or features. Libraries need to be cognizant of the importance of **Monitoring** in scholars' information work. Most scholars develop regular routes and routines for staying aware of developments in their fields.

A scholar such as **Michael** who works in a structured and well defined field may find that a few well chosen activities keep him up to date. Other scholars, such as **Athena** or **Sandy**, who work in a wide range of disciplines or how are starting in their field, may find it less easy to predict where important new information may come from.

Metric for Communication

1. What routines are used to distribute information about new holdings?
The Library's role in this should be to provide easy, transparent, and predictable communication tools for awareness of new resources.

Negatives

Communication is a difficult process, specific problems include users forgetting or misplacing information and library based routines being unsuitable for certain types of researchers.

Routines should be sought that are persistent, for example blogs and downloadable results. As well, routines should be customizable allowing scholars some control in defining the type of information they will receive, for example through customizable alerts and RSS feeds.

5.5. Responsiveness

Critical to **Monitoring** is ability for scholars to access interesting information as soon as they become aware of it. Digital libraries have the potential of increasing support for monitoring as theoretically there shouldn't be a delay between publication and distribution.

However, issues such as embargo, slow internal processes, and gaps between user requests and selection processes can all work against the responsiveness of the collection to individual need.

Again a scholar such as **Michael**, who tends not to have unusual or unique demands may find that they rarely request specific items for purchase, he is more likely to be bothered by issues such as journal embargos. **Athena** will be more likely to make special requests, and she may well judge the library's effectiveness by how quickly it responds. Finally, a scholar such as **Sandy** who is working in a new area and has a limited personal collection may find that an effective and rapid ILL process is critical to her wide ranging review of the literature.

Metrics for Responsiveness

Collection responsiveness can be measured by examining the ability of the Library to deliver information on demand.

1. What is the time between selection and availability?
Monitoring performance in this area will help discern trends, identifying when and why time lags vary from the norm. As well, they will support more accurate feedback to selectors that can in turn be used to help scholars plan their work and whether items should be acquired for their personal collections.
2. What proportion of the content is subject to embargo?
Issues of embargo are most recognized in journal collections; however the problems can also arise with other formats, particularly ebooks. Again, this measure will help us understand if the collection is trending towards delayed access or toward more immediate delivery.
3. What processes are in place to allow users to request rush materials?
Processes should be transparent providing easily identified places or sites for making requests. As well the process should be open, allowing scholars many avenues for place a request.

Negatives

Without stated delivery goals and directed feedback, scholars may easily lose confidence in the Library's ability to provide information on request. To mitigate this process should incorporate a track back process that can be called up by the scholar on demand.

5.6. Provision of unlicensed content

Digital resources are generally purchased under licenses that restrict access to defined user populations. Remote access is particularly subject to restriction. The

Library's practice is to license walk in access for the general public and assure remote access for students, faculty and staff.

For scholars this approach can have some negative consequences.

- **Accessing** may be prevented when the Library does not wish to license products for small groups or individuals. Licensing models for digital products favour the purchase of materials that will have broad appeal.
- **Networking**, particularly the sharing of information among scholars, may be blocked by license restrictions that prevent distributing information outside the licensed community, this may include direct scholar to scholar sharing, or more formal arrangements including inter library loan.

Scholars such as **Michael** and **Athena** who work collaboratively will have a strong desire to share information. For them, being the first to share a key piece of information or a critical text enhances their reputation within the scholarly community. For the Library to negotiate licenses that make this difficult or risky is in conflict with the natural way scholars wish to work.

For scholars such as **Sandy** who work in relatively obscure areas with small readership, the Library's practice of negotiating site licenses for the entire campus may prevent the acquisition of materials that meet short term or very specific needs.

Metrics for the provision of Unlicensed Content

1. What proportion of licensed resources is open to Inter Library Loan?

Current practice is to try to include these provisions in licenses; however, we do not track how successful we are in doing this. As well, including these provisions for formats such as images, e-books and data files is problematic as many vendors are not open to this practice.

2. What proportion of licensed resources allows direct scholar to scholar sharing with unauthorized users?

Our current practice is not to include this as a factor in license negotiation. However, this study has highlighted how important this feature is to effective scholarly communication. A first step is to begin monitoring licenses and discussing this activity with vendors.

Negatives

Any non standard element in a license may slow negotiations, blocking core user access. However, in the long term, scholars will benefit from the inclusion of these provisions.

Another negative of non standard license elements is the confusion they can cause for both users and staff. For this reason a central, easy to use source is needed that clarifies permitted activities.

3. What processes are in place for temporary or individualized access to specialized materials?
Concentrating on overall collection strength and “big deals” has been very successful in rapidly building a substantial digital collection. However, scholar with individualized needs can be left out of this process.

Negative

Some scholars and groups within the campus may view “special” purchases for individuals or small groups as unfair. There is a constant tension in collection development between the needs of large active user groups and smaller more specialized disciplines. Monitoring processes and types of licenses are one way of ensuring that a balance is achieved and demonstrated.

5.7. Creation of Personal Copies

Creating a hard copy or personal download of digital information is a basic scholarly activity.

- **Differentiating** may require the creation of temporary personal copies. This allows items to be viewed at different times and places, and allows the comparison of multiple items.
- **Extracting** through close study often requires personal copies and copies may become permanent additions to the scholar’s personal library. As well, while the most obvious version of extracting is close reading, extraction may also involve manipulation of the digital version—for example text mining or statistical analysis. In these cases the creation of digital versions that can be manipulated is a key feature.
- Personal collections are a critical element in how scholars **Manage** information. These collections can contain temporary items such as bookmarks, or in the case of items considered to be of long term value and requiring frequent rereading or viewing, hard copies may become part of the scholar’s personal library.

All scholars seem to build personal collections and to value the importance of copies they can control. Likely scholars such as **Athena** and **Sandy**, who work with more obscure, fugitive, or grey literature, will value this ability more, as they will be more likely to predict a need for control.

Metrics for Copying Provisions

Measures of copying and its support for scholarly activity include:

1. What proportion of licensed products allow liberal printing of entire units (book chapters, articles, images)

The Library does not have established standards for printing permission with electronic resources and language provided by vendors is often vague. Particularly irksome for many users are the restrictions placed on printing from monographs.

2. What proportion of products has known problems with quality of reproduction, images, full text, and sound?

Not having a record of which products present problems prevents us from understanding where improvements need to be found, and also from being able to communicate known problems to users.

Quality issues should become a central concern when negotiating ownership type agreements.

3. What proportion of licensed products allows unrestricted downloads of entire units (book chapters, articles, images)?

Downloading is growing in importance for scholars, digital versions of full texts eases retrieval, sharing, transportation, and searching within previously viewed materials. Again, having measures will indicate to the Library trends and areas for improvement.

Negatives

Vendor resistance to negotiating generous printing and downloading provisions is a known issue. However, if we have internal standards established and can provide ready examples of more cooperative vendors, negotiations will be eased.

5.8. Collection Stability

In building digital collections, libraries have experienced many positives. However the enduring question of stability dogs the future of digital libraries. Scholars depend on libraries to be stable sources of information. Two critical activities are affected most strongly by lack of stability.

- **Verification** is not a trivial problem for scholars. Verification applies not only to the individual's ability to return to previously viewed work, it also is the basis for a research dialogue that involves all members of the scholarly community,

and which allows the dialogue to span time. In this way libraries are the generational ships of scholarship, carrying information in a stable, predictable form between generations of scholars.

- Akin to verification is the **Management** of information. Scholars depend on the library to hold information for them over at least the span of their careers. This means that scholars are able to concentrate their personal collections on working tools, using the library as a common warehouse of other information sources.

Libraries and universities appear to clearly understand their role in the preservation of rare or unique items. This particularly supports scholars such as **Sandy** and **Athena** to whom these items are important.

However, a scholar such as **Michael** who focuses on information that is commercially available will be just as concerned about the long term stability of information. It is in our approach to stabilizing commercial information sources that the digital library faces challenges and uncertain solutions.

Measures that assess collection stability can be sorted into two categories—Control and Persistence.

Control

These measures assess what level of control the Library is attempting to exercise over the long term stability of the digital collection.

1. What proportions of products of each format are purchased using an ownership model?

A first step in gaining control over information is to purchase information with licenses that convey some type of long term ownership. For example, e-journals purchased from publishers generally include some form of ownership for the annual subscription fee.

Negatives

Ownership models are generally more expensive than other types of licensing. Consequently, favouring an ownership model negatively affects the depth and range of materials that can be offered. However, once basic scholarly needs in a discipline are met, stability is a critical measure of quality. Understanding the proportion of materials in any subject that are “owned” will help us understand if we are creating a balance between content and stability.

In reality vendors are unstable over time, our experience shows that the information marketplace is subject to constant change, and any ownership model that depends on vendors for stability is flawed. In addition to establishing local control, we need to create more stability in relationships

with vendors, including strong standards for license elements that make ownership effective. This includes extension of the license if the vendor or content should be sold, download or extraction permissions at any time, and the existence of a reliable 3rd party holder of the archive.

2. What activities are in place for establishing control over products purchased using an ownership model?
Even with good licenses, long term collection stability cannot be assured by reliance on vendors. The Library currently lacks any type of plan for exercising control over purchased content, and activities are limited and uncoordinated. Introducing a measure for this question is meant to provide direction for the Library to begin planning in this area.
3. What activities are in place for establishing control over locally created content?
Local digitization projects represent intense investment of resources; locally created metadata is included in this measure as well as digital images, texts, or other formats. Such investment usually reflects a belief that these projects provide information of long term scholarly value.

Negatives

Local solutions may be redundant and unsupportable for commercial and locally developed content. There is increasing understanding that standalone, local solutions may not be effective ways to deal with long term stability of commercial products. Activities that seek cooperative solutions (LOCKSS, Synergies) may become the focus of any emerging strategy.

Persistence

These measures are concerned with the importance of preserving the “refindability” of information over time.

1. What proportion of licensed content provides persistent links to information units?

In order for scholars to depend on links rather than downloading items into their personal collections, products must be able to provide persistent links. For example, scholars who expend time in creating links for documents in bibliographic management software will want to be able to include links to the object.
2. How frequently do individual products change links?
Changes in persistent links by vendors can cause difficulty in re-finding documents. Each time such a change is made; consideration should be given to the upheaval this will cause.

Negative

Changes in URLs may be required to provide improved authentication or service. In this case vendors should provide mitigation tools, including URL redirects.

5.9. Information Management

All scholars engage in **Management** routines that allow them to keep track of information viewed, annotations, and future readings. Management in turn supports activities such as **Networking** and **Verification**. This is an activity that has traditionally been neglected by libraries, except for providing assistance in correct citation. We do not have a well developed body of practice for supporting information management. An additional complication is that scholars have individualized information management routines, creating systems that respond to their specific needs.

For scholars such as **Michael** managing may be more straightforward, as the number of unique or obscure items they need to include is limited. **Athena** will need tools that allow her to manage a wide range of products including web pages, images, government documents, even personal correspondence. **Sandy** will need all of this, plus an ability to cite sources and most likely include extended annotations as she closely reads text or questions issues of provenance.

Metrics for Information Management

1. What proportion of licensed products is compatible with standard bibliographic management (BM) software?

Widespread use of BM software depends on ease of use. Understanding trends in how well this is supported will help us know which scholars will benefit most from awareness of the software.

Negative

Scholars will wish to be able to customize management; we can expect a wide range of approaches and software to be used. To mitigate this product should also support the export of citations in a generic, easily imported format.

2. Does software allow individualized management including annotation and tagging?

Libraries are not known for developing environments that accommodate or support the idiosyncratic needs of individuals. However, support for scholarly information management requires that we look to products for exactly these features.

Negative

Our inexperience in this area means that we do not have an understanding of how different features will be accepted or applied. Seeking feedback and working closely with scholars is an important part evaluating and guiding the design of these applications.

3. Does software allows scholars to share citations and commentary?

Personal collections are resources that scholars regularly draw from when collaborating with colleagues. Tools that primarily operate as citation tracking and bibliography generating applications fall short of supporting networking in the digital environment.

Negative

Just as with tagging and annotation, we are not yet certain how/if scholars will use software to support collaboration.

6. Reflection and Conclusions

This project marks the first step in developing measures of collection effectiveness. At this stage it is appropriate to reflect on how well this approach provided insight into

user needs and inspired metrics that will help guide development of the digital library.

6.1. What type of library would these measures create?

Part of evaluating these measures is trying to understand what kind of library we would create if we relied on these metrics.

Current valuing of content and use lead to collections that emphasize quantitative increase as positive, the greater the number of items and the more a collection is used, the more it is viewed as successful and healthy. Certainly, the University of Calgary collection performs well when assessed in this way. Big deals and increased accessibility have made this success possible. However, the review of scholarly behaviour indicates that the collection could develop in additional ways to become a more effective research tool.

First, the proposed measures promote increased attention to description, through cataloguing and other types of metadata. The library houses a very large, “silent” backlog including digital resources such as e-books, rare materials such as special collections, and alternative formats such as images and sound files. This study revealed scholars who continue to rely on libraries providing access to this information through descriptions collated into central repositories. We may not be certain if these descriptions will reside in a single catalogue, or a series of internal and external collections linked through searching mechanisms, but it is clear that a complete lack of description is a barrier to use.

While we have made steps in the direction of transparency, for example our catalogue records are available on LibraryThing and we are currently involved in a

reclamation project that will make our records more completely available in WorldCat, the Library lacks an overall strategy for this important collection element.

Second, linking between sources and from descriptions directly to sources was highlighted as a key benefit of new technology. Time spent navigating to locations, either virtual or actual, is time that scholars could spend more productively on other activities. While we have implemented SFX linking technology, we still lack effective integration with inter library loan processes and to the full range of information formats. Linking from the Google Scholar directly to the University of Calgary collection is extremely useful, illustrating the value of pushing information about our collections outside Library spaces into the user environment.

Third, these measures would push the Library towards more defined communication routines and individualized user services. Perhaps this is the area where Library 2.0 concepts are most clearly highlighted. Developing a collection that allows users to establish a personalized environment of alerts, tags, and collaboration, and which leverages this information to provide other users with better service (for example, sharing information about tags or views) is a challenge we have neither discussed nor yet developed a strategy to explore.

Fourth, our current licensing practice emphasizes maintaining a print model for scholars' use of information. This arose from the early days of digital sources, when both librarians and publishers were unsure of the impact new formats would have. Some of the key limitations of our practice include restrictions on downloading, printing and sharing, and tying access to scholarly affiliation. This latter issue will become more relevant over time, as scholars moving from institution to institution lose rights to information sources. Pushing for more liberal licensing, and accommodating the scholar's need to have a career "information portfolio" are issues that are rarely discussed in a licensing environment that places more emphasis on effectively limiting use to stay within current contract parameters.

Finally, this study re-emphasized the importance of stability and preservation. The library role as repository is core to scholarship. It allows scholars to continue dialogs over generations by assuring "refindability". The passive model of preservation libraries have exercised in the past, a model based on refraining from discarding is no longer sufficient. While the University of Calgary library is involved in many preservation projects, e.g., Synergies, Portico, and has migration plans for local content, an overall plan that assures stability over generations is lacking. By establishing measures we will be able to measure real progress and define the extent and nature of the challenge.

6.2. Next steps

Positively, the behavioural models did inspire metrics that could be directly tied to scholarly actions, promising a better link between evaluation and effectiveness.

However, the approach supports a high level view of collections. While this is useful for collection development with measures that are stable and widely applicable, individual product evaluation and design requires different types of user studies, ones that focus on the interaction with tools rather than overall information seeking strategy.

Some of the measures suggested by the study are relatively novel (persistence of links), other measures reaffirmed the importance of traditional activities (preservation) in the digital environment. This combination indicates that the approach is creative and realistic, informed by past practice and responsive to the new environment.

Project COUNTER has established quantitative and usage based metrics as standard for evaluating digital collections, the measures suggested in this study lie outside standard metrics, and do not come with ready made tools for gathering and storing results. If we wish to employ the measures in this study we will need to develop tools for internal management. The Library is currently planning to acquire an Electronic Resource Management System and the implementation of this application could include planning for gathering and archiving many of the measures suggested here.

Validation of the measures and design elements is also needed. The approach that may seem most obvious—usability or focus groups—are not likely to be effective. They may enrich the user model, but they would be unlikely to invalidate it. Metrics are needed that operate at a higher level, for example log analysis which can aggregate the behaviour of entire groups.

More work is certainly needed to assess the measures. The completeness of these measures is not established, and this is only the beginning of a process that is cyclical and must remain open to change over time. Presentation and discussion of the results is the first stage of this process, and this report marks the initial action in this phase. Testing the measures in application is also anticipated as is continuing to up date the user model and assess the applicability and completeness of these measures.

Finally, this project only examined the behaviour of scholars, depending on interest and effectiveness of the work done here, a study that applied similar techniques to learners may prove desirable.

References

1. Sutcliffe, A. On the effective use and reuse of HCI knowledge. *ACM Trans. Comput. -Hum. Interact.* **7**, 197-221 (2000).
2. Carroll, J. M. Making use is more than a matter of task analysis. *Interacting with Computers* **14**, 619-627 (2002/10).
3. Carroll, J. M., Rosson, M. B., Chin, G., Jr. & Koenemann, J. Requirements development in scenario-based design. *Software Engineering, IEEE Transactions on* **24**, 1156 (1998).
4. Carroll, J. M. & Rosson, M. B. Getting around the task-artifact cycle: how to make claims and design by scenario. *ACM Trans. Inf. Syst.* **10**, 181-212 (1992).
5. Lokman I. Meho, Helen R. Tibbo. Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology* **54**, 570-587 (2003).
6. Core Scholarly Informaton Tasks and E-Journal Features: Expanded Discussion. (2002).
7. Borgman, C. L. et al. Comparing faculty information seeking in teaching and research: Implications for the design of digital libraries. *J. Am. Soc. Inf. Sci. Technol.* **56**, 636-657 (2005).
8. Jankowska, M. A. Identifying University Professors' Information Needs in the Challenging Environment of Information and Communication Technologies. *The Journal of Academic Librarianship* **30**, 51-66 (2004/1).
9. Researcher's Use of Libraries and other Information Sources: current patterns and future trends. (2002).
10. East, J. W. Information Literacy for the Humanities Researcher: A Syllabus Based on Information Habits Research. *The Journal of Academic Librarianship* **31**, 134-142 (2005/3).
11. Tenopir, C., King, D. W., Boyce, P., Grayson, M. & Paulson, K. Relying on electronic journals: Reading patterns of astronomers. *Journal of the American Society for Information Science and Technology* **56**, 786-802 (2005).

12. Lee, H. L. Information spaces and collections: Implications for organization. *Library and Information Science Research* **25**, 419-436 (2003).
13. Rapp, D. N., Taylor, H. A. & Crane, G. R. The impact of digital libraries on cognitive processes: psychological issues of hypermedia. *Computers in Human Behavior* **19**, 609-628 (2003/9).
14. Houghton, J. W., Steele C. & Henty, M. Research Practices and Scholarly Communication in the Digital Environment. *Learned Publishing* **17**, 231-249 (2004).
15. Palmer, C. L. & Neumann, L. J. THE INFORMATION WORK OF INTERDISCIPLINARY HUMANITIES AND SCHOLARS: EXPLORATION AND TRANSLATION. *Library Quarterly* **72**, 85 (2002).
16. Wiberley, S. & Jones, W. G. Time and technology: a decade-long look at humanists' use of electronic information technology. *College & Research Libraries* **61**, 421-431 (2000).
17. Fidel, R. & Green, M. The many faces of accessibility: engineers' perception of information sources. *Information Processing & Management* **40**, 563-581 (2004/5).
18. Tenopir, C. et al. Patterns of Journal Use Through Three Evolutionary Phases. *D-Lib Magazine* **9**.
19. Talja, S. & Maula, H. Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. *Journal of Documentation* **59**, 673 (2003).
20. Final Synthesis Report of the e-Journal User Study December 2002
Prepared for
the Stanford University Libraries e-Journal User Study by Institute for the Future. **SR-786**.
21. Su, S. Desirable search features of web-based scholarly e-book systems. *The Electronic Library* **23**, 71 (2005).
22. Wisneski, R. Investigating the Research Practices and Library Needs of Contingent, Tenure-Track, and Tenured English Faculty. *The Journal of Academic Librarianship* **31**, 119-133 (2005/3).
23. Tenopir, C. Use and Users of of Electronic Library Resources: an Overview and Analysis of Recent Research Studies. **pub120** (2003).
24. Sandstrom, P. E. Scholars as subsistence foragers. *American Society for Information Science. Bulletin of the American Society for Information Science* **25**, 17 (1999).

25. de Tiratel, S. R. Accessing information use by humanists and social scientists: A study at the Universidad de Buenos Aires, Argentina. *The Journal of Academic Librarianship* **26**, 346-354 (2000/9).
26. Murphy, J. Information-Seeking Habits of Environmental Scientists: a study of Interdisciplinary Scientists at the Environmental Protection Agency in Research Triangle Park, North Carolina. *Issues in Science and Technology Librarianship* (2003).
27. Hallmark, J. Information-Seeking Behavior of Meteorologists and other Atmospheric Scientists: Access and Retrieval of Cited References. *Issues in Science and Technology Librarianship* (2003).
28. Porter, D. C. *MEDIEVALISTS'USE OF ELECTRONIC RESOURCES: THE RESULTS OF A NATIONAL SURVEY OF FACULTY MEMBERS IN MEDIEVAL STUDIES* (2002).
29. Brockman, W. S., Neumann, L., Palmer, C. L. & Tidline, T. J. Scholarly work in the humanities and the evolving information environment. (2001).
30. Ruvane, M. B. *Annotation as Process: A Vital Information Seeking Activity in Historical Geographic Research* Ser. 42, (2005).
31. Nicholas, D., Huntington, P. & Watkinson, A. Digital journals, Big Deals and online searching behaviour: A pilot study. *Aslib Proceedings* **55**, 84 (2003).
32. Bruce, H., Jones, W. & Dumais, S. Information behaviour that keeps found things found. *Information Research* **10**, May 2005 (2004).
33. Johnson, C. A. Choosing people: the role of social capital in information seeking behaviour. *Information Research* **10**, 10-11 (2004).
34. Nicholas, D., Huntington, P., Monopoli, M. & Watkinson, A. Engaging with scholarly digital libraries (publisher platforms): The extent to which 'added-value' functions are used. *Information Processing & Management* **42**, 826-842 (2006/5).
35. Jones, W. Finders, keepers? The present and future perfect in support of personal information management. *First Monday* **9**, August 9, 2006 (2004).
36. McGeachin, R. B. The Impact of Electronic Bibliographic Databases and Electronic Journal Articles on the Scholar's Information Seeking Behavior and Personal Collection of "Reprints" *Science & Technology Libraries* **25**, 127-137.