An Investigation of Information Problems

IMT 570 Analytic Methods for Information Professionals

Group 10

Jacob Kovacs | kovjac19@uw.edu

Clint Posey | cposey@uw.edu

Lani Smith | fullmoon@uw.edu

TABLE OF CONTENTS

ABSTRACT

We sought to understand the types, frequency, and impact information problems experienced by graduate students in their workplaces prior to joining the University of Washington's iSchool. We hoped to address the absence of a suitable information problems taxonomy in the literature and strengthen our ability to diagnose and prioritize problems as future information professionals. We synthesized data from six interviews to produce a list of seven information problems, then surveyed a non-random sample of students to identify the frequency and impact of these problems (n=31) . We find that *filtering information overload* and *finding appropriate information* are the most common challenges with commensurate impact, but *understanding context* and *integrating information across systems* have a disproportionately high impact given their lower frequency of occurrence.

INTRODUCTION

Information technologies—and the challenges associated with them—proliferate. A 2010 Forbes Insights report surveying more than 200 global companies found the majority of companies lost millions of dollars each year due to data problems, with one-fifth reporting losses of over $20 million (Forbes, 2010). Ninety-five percent of respondents to this same survey also indicated their view that information management is crucial to their success. With such immediate financial repercussions, it is easy to see why organizations are recruiting information professionals to help them meet information management challenges. However, before an information professional can successfully navigate the growing field of information management and create solutions, she must first know how to identify common problems and understand how to prioritize interventions within an organization. Do all organizations struggle with the best ways to collect marketplace information relevant to their industry? Do most fail to properly disseminate critical information to their employees? Are systems overly complex, creating barriers to information access? As businesses grow, how can they ensure that employees at different levels agree on what is and is not necessary for their information needs?

Graduate students at the University of Washington (UW) iSchool will be expected to answer questions like these and identify key information management issues that businesses are facing across industries; and yet, as the study of information management is new relative to more established disciplines, there is a need for theories, models, and concepts to help students derive these insights. With that in mind, our research aims to contribute to the ongoing project of classifying and identifying information problems in the field of information management.

Literature Review

There are a few distinct bodies of literature concerning the definition and frequency of information problems, namely theorization about information in general; academic and practitioner frameworks for assessing an organization's information management practices; practitioner lists of information challenges; industry surveys of business leaders; and research on individuals' conceptualization of and strategies for information problem-solving. This section offers a brief summary of each and discusses their implications for our research.

Detlor (2009) provides a good synthesis of academic work theorizing information management as a process of managing information throughout its "lifecycle" of "creation, acquisition, organization, storage, distribution, and use". Detlor's lifecycle could perhaps be repurposed as a typology of information problems (used to classify information problems according to the lifecycle stage in which they arise). More promising, though, is Evgeniou and Cartwright's (2005) proposal of "three fundamental types of barriers to information intelligence", specifically behavioral, process, and organizational. Within each fundamental type, they identify two specific problems for a total of six information management problems (Table 1). We imagine additional problems could be classified under their schema, but this speaks to the biggest shortcoming within this literature for our purposes:

potential top-level categories are identified, but concrete problems within categories are not sufficiently enumerated.

Table 1. Three types of barriers to "information intelligence", adapted from Evgeniou and Cartwright (2005).

Behavioral barriers
Confirmatory bias: The strong tendency to ignore information that conflicts with an established personal opinion
Imbalance between creativity and hard data: Allowing data to stifle creativity; allowing consumer tastes to dictate product development, rather than innovating.
Process barriers
Unsuccessful problem definition: Defining a research problem incorrectly, such that research results have diminished utility; a common consequence of delegation.
Research rigidity: Failure to recognize the need for information and research as ongoing processes; fixation on specific research products that may become quickly outdated.
Organizational barriers
Misuse of information asymmetries: When information that should be shared isn't, or information shouldn't be shared is.
The newcomer syndrome: When newcomers innovate purely to make their mark on a product or process, to the long-term detriment of that product or process.

Frameworks for assessing an organization's information management practices are another potential resource for classifying information problems. These frameworks may be developed by governments (e.g., ITRB, 1999; RIMB, 2004; PROV, 2013); academics (e.g., Bailey and Pearson, 1983; DeLone and McLean, 1992; Li and Cheung, 1987; Platiša and Balaban, 2009); or private organizations like ISO, the International Standards Office (NAA, "Phase 1"; NAA, "Phase 2"; NAA, "ISO 16175"). These tools vary widely, from Platiša and Balaban's (2009) synthesis of 46 factors grouped in 8 dimensions to the ISO's measurement of information system functionality across four dimensions (NAA, "Phase 2"). In general, though, these tools are too granular and comprehensive to suit our purposes; they are not taxonomies, but rather checklists that a professional might consult while implementing a system. Moreover, they are phrased as questions to be answered rather than analytic categories that might be helpful in perceiving a system.

Less thorough than frameworks are lists of 'common information challenges' curated by practitioners (e.g., Spiegel Institut, n.d.; Robinson, 2005; Informa, 2011; Iron Mountain, 2012; Alabdan, 2014; Matharu, 2015; Michael, 2016). These are presented without pretense of rigor or generalizability, often in the context of selling information management services. There might be value in them regardless (see Schön, 1983 on the value of practitioner knowledge), but for our purposes they are too arbitrary, lacking any underpinning theory and failing to suggest categories for classification.

There are many industry-specific or cross-industry surveys of business leaders asking about current information challenges, their relative frequency of occurrence, and their impact in terms of labor hours and cost (e.g., Forbes, 2010; McLaughlin, 2014). There are even meta-analyses of these surveys, conducted by academics (e.g., Pearson, 1977; Watson, Kelly, Galliers, and Brancheau, 1997; Palvia, Palvia, and Whitworth, 2002; Palvia & Palvia, 2013). By nature, surveys of this sort are inherently dated and must be conducted periodically to maintain value; they are not meant to serve as a taxonomy. Another possible limitation is their reliance on executive perspectives, trusting that information systems managers will have insight into on-the-ground information challenges.

Finally, the question of how common information problems are depends on another more fundamental question: how likely are people to see problems they encounter as problems of information or information management specifically? This question ventures into the territory of library science, writing composition studies, education, and information science—fields that focus on how individuals satisfy 'information needs' and solve 'information problems' in the course of writing academic papers or consuming written information (e.g., Faibisoff and Ely, 1974; Butler, 1994; Leckie, Pettigrew, and Sylvain, 1996; Brand-Gruwel, Wopereis, and Walraven, 2005; Miranda and Tarapanoff, 2008; Brand-Gruwel, Wopereis, and Vermetten, 2009). While this literature is quite distant from information problems and information management in the workplace, it perhaps points in a direction of interest for our study.

This brief review shows that, for our purposes, a suitable taxonomy of information management problems is absent from the literature. Academic theories of information are too sparse, while information systems assessment frameworks are too detailed. A taxonomy at the right level of abstraction for both information managers and information researchers—one that would help identify information management priorities—is missing. Furthermore, research on the frequency of information management problems fails to ask whether information management problems are being 'seen' in the first place.

Research Questions

We asked UW iSchool graduate students to reflect on their experiences and report on the occurrence and severity of information problems they are familiar with via their recent professional experiences. Our specific research questions were: (1) How do students label the problems they have encountered? (2) What categories of information problems have students encountered? (3) Of these, which information problems are most common and most severe? The scope of our data collection was limited to masters students in the MS Information Management and MS Library Science programs at the UW iSchool. As we are not attempting to prove any causation with our data, nor are we trying to test a theory, the proposed research is exploratory in nature; there is no formal hypothesis.

We believe this work is important because students bring valuable academic and nonacademic experiences into graduate programs—experiences that may be overlooked, despite having value to

both managers who decide how to allocate resources and academics who want to research impactful problems. Furthermore, a survey of issues students have already encountered in their workplaces and previous experiences can help take a pulse on where the field of information management may be headed. While current UW iSchool researchers are certainly already aware of many future trends, additional study would add to the ongoing project to classify information problems and identify problems that will most affect students' professional lives. For a graduating professional, knowledge of other students' information problems can prepare her to enter the workforce with a better understanding of what skills and energies will be needed and empower her and her peers.

METHODS

We conducted six semi-structured interviews to generate items for a survey that was administered to a larger sample. Table 2 provides a mapping of our research questions onto our research instruments, and the instruments themselves are included in Appendices B and C.

Table 2. Mapping research questions onto questions posed in study instruments.

How do students label the problems they have encountered?	[interview] *Describe a time when you encountered a difficulty with information at work.* [interview] *How would you personally define an information problem?* [survey] *What's your definition of an information problem?* [survey] *Which of the following would you consider an information problem?*
What categories of information problems have students encountered?	[interview] *How would you personally define an information problem?* [interview] *Can you give me a few more examples of information problems?* [survey] *What's your definition of an information problem?*
Of these, which information management problems are most common and most severe?	[survey] *How often, if ever, did you notice the following information problems at work?* [survey] *How much of an impact, if any, do you think these information problems have on your organization?* [survey] *Thinking of their impact on your organization, sort these information problems from most to least harmful:*

Sample Selection

Our study population consists of graduate students at the iSchool, including full-time and mid-career MSIM, as well as residential and online MLIS students. For interviews, we employed a purposive sampling strategy (Pickard, 2013) and extended study invitations to eight subjects on the basis of their prior work experience, with six subjects consenting to join our study. We thought a sample of more professionally experienced subjects would be a fruitful place to start our search for potential "theoretical generalizations" (Pickard, 2013, p. 60) regarding information problems.

For the survey, we drew a random sample of 50 students from the frame provided by Dr. Matt Saxton. However, after extending an invitation and two reminders, a very low response rate (11 people) and our impending deadline led us to extend the study call to an indeterminately larger pool of students using the MSIM Whatsapp group and personal relationships. In doing this, we were able to obtain 31 survey responses, but we lost our ability to calculate a response rate and cannot claim our results represent the experience of the larger iSchool graduate population. We felt that it was a priority to obtain sufficient data to make the data analysis a worthwhile learning exercise. For a full-scale study to be truly valid, we would need to use a random sample from a sampling frame that itself represented the iSchool population (since the opt-in nature of the sampling frame provided *already* represents a departure from statistical representativity).

Data Collection Procedures

We collected interview data according to the protocol in Appendix A, with our specific interview questions listed in Appendix B. The purpose of in-person interviews was to collect rich, open-ended data on the types of information problems participants have encountered, especially within the contexts of their work experience. Interviews began with simpler general questions both to set the participants at ease and to gather contextual information. After this, we employed critical incident technique (Flanagan, 1954) to gain insight into a particular problem participants encountered, asking probing questions to ascertain root causes, consequences, and potential solutions. We waited until the end of the interview to ask the participants to provide a definition of an "information problem", so the definition would emerge more naturally from their recollection of experiences. We collected survey data using the form provided in Appendix C. The structure of our survey instrument is "flipped" from the interview: we first asked respondents to define information problems before collecting data how they perceive these problems, and ended by asking for demographic data. The first question of the survey is an open-ended field, to collect participants' definitions of information problems without exposing them to our prior data and influencing their definition. For the rest of the survey, we provided respondents with the list of problems generated from interview data, and ask them to rank these information problems in terms of frequency and impact.

Validity and Reliability

As discussed above, our sampling procedures mean that our results cannot be taken as representative of the larger iSchool graduate population (much less a population outside the iSchool). Our data likely suffers from both non-response bias (with only about 20% our original random sample responding) and self-selection bias (with, for example, women comprising 70% of our respondents; it appears men were much less likely to respond). This flaw would be addressed in a full-scale study by proper random sampling. We did take other measures, though, to improve the validity of our research.

In writing questions for our interview and survey, we tried to word them clearly, avoid double-barreled questions and leading questions, and refrain from seeking unnecessary sensitive data

(Cooper and Schindler, 2013). We workshopped questions as a team and, to control order effects, thoroughly discussed the logic behind sequencing questions in a way that would produce the most valid data. We avoided the pitfalls of forced choice questions by allowing participants to opt-out of questions or choose a neutral response option; the one exception is a survey question that asks participants to rank items, which, due to a software limitation in Google Forms, had to be made mandatory. For our survey, an additional threat to validity is that we had no pre-developed scales to draw on. For response options, we tried to choose words with equal semantic distance and limit the total to 5-7 items; but we cannot state with confidence that our scale questions were interpreted accurately by study participants or that they accurately capture the magnitude of people's sentiments.

At a more fundamental level, we were exploring a very complex and abstract concept, and have doubts that our methods were adequate to the challenge. We expect our "content validity" is quite low—that there are many possible dimensions of our core concept not covered by our interview and survey questions (Trochim, 2006). Moreover, our research assumes that participants' recollections and perceptions provide a meaningful picture of the information problems they actually faced at work; this assumption is most questionable when we ask subjects to estimate the *frequency* and *impact* of different information problems. At best, this is very partial data taken from the perspective of entry-level and junior employees. Our method compares favorably with industry surveys that ask only for executive-level perspectives (e.g., Forbes, 2010; McLaughlin, 2014), but does not give us a basis for making strong claims about the actual frequency and impact of specific information problems. A full-scale study would need to harness perspectives throughout an organization and possibly use administrative records and participant-observation as a reality-check for perceptions-based data.

Lastly but not trivially, in our research context we also face the challenge of study participants who were overwhelmed with their finals week workloads. We asked very cognitively-demanding questions to which they may have provided random or close to random answers rather than giving the necessary thought.

Ethical Considerations

With regard to privacy and confidentiality, we did not collect names from the survey, and all recordings of interviews were anonymized and deleted after transcription. In the case of survey respondents, the results were aggregated for analysis, effectively eliminating the possibility of identification. Regarding the potential for harm, we did not collect any data regarding sensitive behaviors or situations. At most, we expected that our questions about information problems in the workplace would provoke mild frustration or anxiety; we tried to ameliorate this risk by assuring all interview and study participants that participation was wholly voluntary and specific questions were optional (with the exception of one survey question that, due to survey software limitations, needed to be mandatory.) In particular, all demographic disclosures were optional in recognition that some

respondents may not wish to provide this personal information. Finally, we tried our best to make the interviewee comfortable by maintaining a friendly, casual demeanor.

RESULTS

Interview Data Analysis

The main purpose of our interviews was to generate items for survey response options. Towards that end, each member of the research team extracted a list of information problems from the surveys they personally conducted (Appendix D), then met to synthesize these lists into a reasonable number of response options by combining and hierarchizing items (Table 3). Developing a coding manual (Woelfer, Duong, and Hendry, 2013) would have been a more robust approach to this analysis, but we could not accomplish this under the time constraints of the study.

Table 3. List of information problems synthesized from interview data.

- Threats to information security or privacy
- Difficulties with exchanging or sharing information
- Difficulties coordinating or collaborating
- Insufficient understanding of context, constraints, or user needs
- Finding information that is complete, relevant and credible
- Filtering a high volume of information (information overload) to keep only what's relevant or credible
- Integrating information across multiple systems and disparate standards

Survey Data Analysis

Our first survey question was open-ended: "What's your definition of an information problem?" A member of the research team conducted in vivo coding on this data, followed by axial coding to produce a short list of codes (Table 4; Woelfer, Duong, and Hendry, 2013). To test inter-rater reliability (IRR), another member of the team applied the resulting codes to the data and Cohen's kappa was calculated for each code (Table 5). We suspect that IRR would have been considerably higher if the second coder were given basic information regarding the initial coding approach, which was very conservative (i.e., much more likely to leave data uncoded if a code was not explicitly mentioned). Ignoring the application of "NO CODE" (kappa of 0.00), our kappa ranged from 0.32 to 0.86 with an average of 0.60—the exact threshold between "moderate" and "substantial" agreement (Landis and Koch, 1977).

Table 4. Codes generated from qualitative survey data.

DIFFICULT TASK (D): ***The respondent describes a difficult task related to information***
OBTAINING (D1), ORGANIZING (D2), ANALYZING (D3), COMMUNICATING (D4), APPLYING (D5)
UNDESIRABLE QUALITY (U): ***The respondent describes an undesirable quality of information***
TOO MUCH (U1), TOO LITTLE (U2), IRRELEVANT (U3), INACCURATE (U4)

NO CODE (XXX): *No other codes apply*

Table 5. Inter-rater reliability of codes as measured by Cohen's kappa.

Code	ID	Cohen's kappa	Percent agreement
Difficult Task	D	0.814	0.909
Obtaining	D1	0.627	0.818
Organizing	D2	0.421	0.818
Analyzing	D3	0.637	0.864
Communicating	D4	0.831	0.955
Applying	D5	0.645	0.955
Undesirable Quality	U	0.581	0.818
Too Much	U1	0.861	0.955
Too Little	U2	0.321	0.773
Irrelevant	U3	0.450	0.909
Inaccurate	U4	0.463	0.909
Not codeable	XXX	0.000	0.773

To capture the full range of language used, we also present responses to this first question as three word clouds (Table 6) showing the most frequently used nouns (*information, data*), verbs (*finding, analyzing, solving*), and adjectives (*too much, not relevant*); note that nouns and adjectives are presented in order of frequency, while verbs are arranged according to Detlor's (2009) information lifecycle (with stages of creation/acquisition, organization, storage, distribution, and use). As basic as this data is, the frequencies align very well with results from quantitative questions in our survey.

Table 6. Word clouds showing language used to define information problems.

Nouns:	**information** (15), data (3), analytics, systems, databases, options, social media, knowledge, insight, fact
Verbs:	managing (2); **finding** (5), collecting, gathering; storing, deleting, removing; categorizing, organizing; accessing (2), retrieving; **analyzing** (4), understanding; sharing (2), portraying, communicating, delivering; **solving** (4) [another problem with], applying
Adjectives:	**too much** (4), **not relevant** (4), not accurate (3), not enough (3), not valuable/useful (2), ambiguous, too complex, addicting

We also asked subjects to perform a categorization task and indicate which items they considered an information problem (Table 7). We used the seven items synthesized from our interview data (Table 3) as well as four additional items that didn't make it into the primary list due to disagreement within the research team: "Too much time spent on information, not enough on action";

"Insufficient resources dedicated to learning and obtaining feedback"; "Denied access to necessary information because of low status in organizational hierarchy"; and "Not enough attention paid to ethics ('should we' versus 'can we')".

In general, for this question, respondents' opinions agree very closely with what might be expected on the basis of common sense. The main result from this survey question is a lack of 100% agreement regarding *any* single item; agreement ranged from 32% to 83%. This startling lack of consensus points points to the need for additional research to understand what—if not an information problem—respondents consider such problems as "Filtering a high volume of information (information overload) to keep only what's relevant or credible".

Table 7. Respondent answers to "Which of the following would you consider an information problem (versus some other kind of problem)?".

Item	#	%
Filtering a high volume of information (information overload) to keep only what's relevant or credible	26	83.87
Finding information that is complete, relevant and credible	25	80.65
Insufficient understanding of context, constraints or user needs	23	74.19
Threats to information security or privacy	22	70.97
Integrating information across multiple systems and disparate standards	21	67.74
Too much time spent on information, not enough on action	18	58.06
Difficulties with exchanging or sharing information	18	58.06
Insufficient resources dedicated to learning and obtaining feedback	16	51.61
Denied access to necessary information because of low status in organizational hierarchy	12	38.71
Difficulties coordinating or collaborating	12	38.71
Not enough attention paid to ethics ("should we" versus "can we")	10	32.26

Respondents were asked to rate both the frequency (Figure 1) and negative impact (Figure 2) of information problems; in addition, we asked respondents to rank seven information problems according to the magnitude of their negative impact (Figure 3).

Figure 1. Per-item distribution of responses to the question "How often, if ever, did you notice the following information problems at work?" using a scale of Never/Rarely/Sometimes/Often/Always. Please note that these are sparkline-style visualizations (Tufte, n.d.), and the absence of numbered axes is deliberate in order to facilitate visual detection and comparison of qualitative trends between images at the same scale.

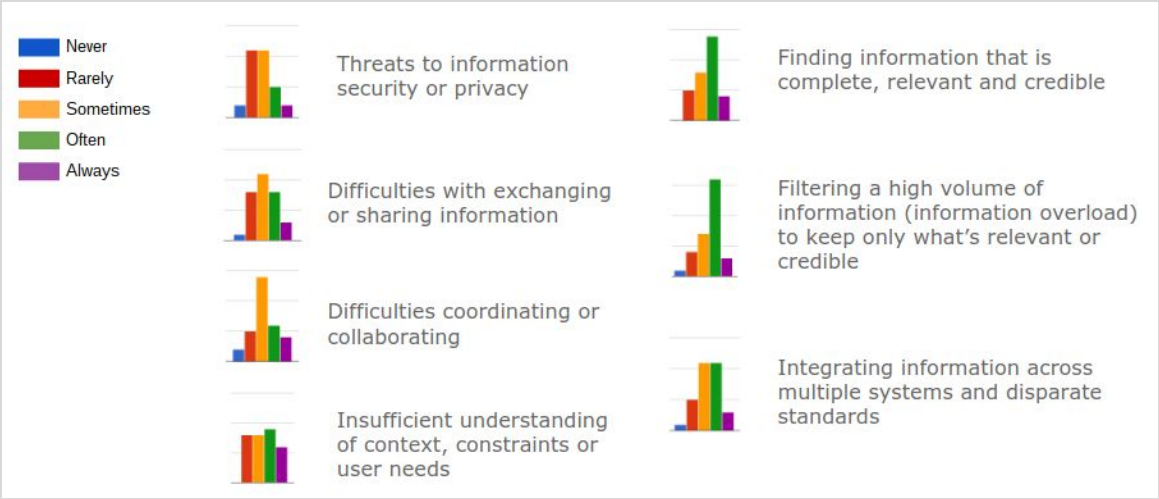


Figure 2. Per-item distribution of responses to the question "How much of an impact, if any, do you think these information problems had on your organization?" using a scale of 1 (No impact) to 5 (Severely negative impact).

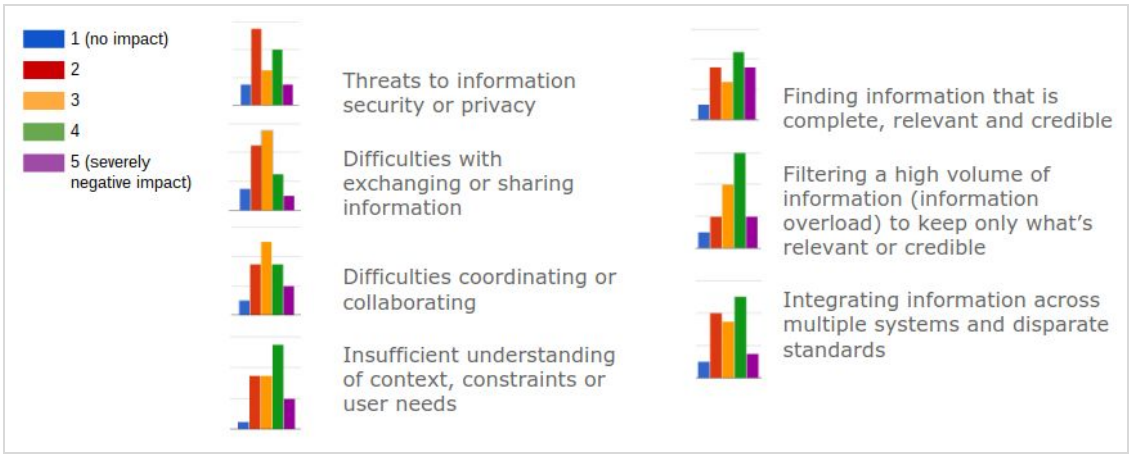
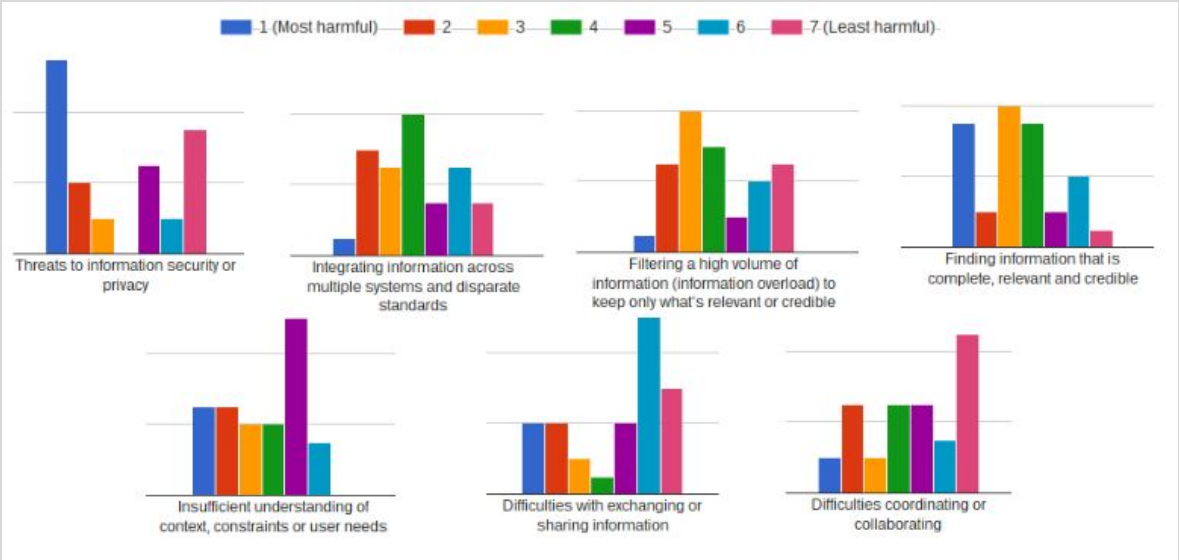


Figure 3. Per-item distribution of responses to the ranking question "Thinking of their impact on your organization, sort these information problems from most to least harmful".



For *frequency*, respondents overwhelmingly chose "Finding information" and "Filtering information" as problems that occur "Often", while "Insufficient understanding of context" varied most and "Threats to security" skewed the most towards "Rarely"/"Sometimes". For *impact*, "Filtering" and "Finding" remained in the lead and were joined by "Insufficient understanding of context" and "Integrating information"; "Finding" rose to the very top concern; "Coordinating" and "Sharing information" were the most popular choice for least impact; and "Threats to security" exhibited a bimodal distribution, perhaps capturing a distinction between workplaces that deal with sensitive data and those that don't. This same bimodal split showed up in our impact rankings data, although there were interesting disagreements that arose here from exploring the same concept through different response formats—mainly the low impact assigned to "Insufficient understanding of context", and the clear ascendence of "Threats to security" to highest ranked.

Finally, we asked for participant demographics (Tables 8-11) in order to perform Chi-square tests of association. For each of our seven information problems, we tested for associations between perceived frequency and length of respondents' work experience; between perceived frequency and the industry in which respondents worked pre-iSchool enrollment; and between perceived frequency and specific iSchool program. We thought that respondents might be more or less likely to observe certain problems depending on their industry or the seniority of their role (proxied by years of experience); we also thought that respondents' choice of iSchool program might reflect the sort of problems they considered most common. In all cases, though, we failed to reject the null hypothesis of non-association (see Appendix E for outputs). In other words, we found no evidence that participants' assessment of the frequency of information problems varied according to their iSchool program affiliation, their industry of work, or the length of their work experience.

Table 8. Program affiliation of survey respondents.

Table 9. Gender identification of survey respondents.

	#	%
Full-time MSIM	24	77.4
Mid-career MSIM	0	0
Residential MLIS	3	9.7
Online MLIS	4	12.9

	#	%
Female	21	67.7
Male	9	29.0
Decline to Respond	1	3.2
Other	0	0

Table 10. Work experience of survey respondents.

	#	%
0-1 years	6	19.4
2-3 years	10	32.3
4-5 years	8	25.8
>5 years	7	22.6

Table 11. Pre-iSchool industry of survey respondents.

	#	%
IT	7	22.6
Finance/Banking	5	16.1
Business/Consulting	5	16.1
Education/Libraries	3	9.7
Science	2	6.5
Other	2	6.5
Government	1	3.2
(No response)	6	19.4

DISCUSSION

How do students label the problems they have encountered? / What categories of information problems have students encountered?

As previously discussed, there appear to be two themes present when students are asked to define an information problem: the problem as task and the problem as a measure of quality. These themes first emerged in our qualitative assessment of interview subjects, and was substantiated by the open ended question in the survey, “What is an information problem?”. However, as the word cloud in Table 6 shows, there is no single label that stands out with a significant level of frequency, outside of the word “information”. This could be an indication that there are other situations applicable to the concept of an “information problem” that did not occur with great enough frequency in our sample to be confidently identified.

As mentioned earlier, when it comes to the idea of information problem as task, responses with the highest frequency of occurrence were the verbs “finding”, “analyzing”, and “solving (another problem)”. These verbs all lend themselves to a more active interaction with information, rather than a more passive task such as storing or collecting (although these were each mentioned by one respondent). This may have been influenced by the context in which we asked about information

problems—through the lens of work experience—in which participants were more likely to have encountered information in a task-oriented setting. Expanding this context to include non-professional situations in a later study may reveal a higher incidence of more passive interactions. Surveying more students in the MLIS program may also yield such results, as the focus of that program is more on theory, preservation, and archiving, rather than the analysis and application focus of the MSIM.

The adjectives describing information problems as a measure can be further broken down into two categories: quantity of information available (too much/not enough) and quality of the information itself (not relevant, not accurate, not useful). Originally we had hoped to be able to construct a taxonomy of information problems by breaking down responses in a similar fashion. Unfortunately, due to a lower response rate and lack of additional identified problems from the survey (no participants opted to respond to the "Other" field), we did not have enough information to work with. A future, more extensive study will hopefully yield more results that can be used to establish a taxonomy.

Table 7 shows the wide range of information problems available for respondents to consider, although as previously discussed, none was indicated with unanimity to be considered a problem. While the potential types of problems presented to survey participants included the very concrete ("denied access") to the more abstract ("not enough attention to ethics"), the frequency of responses suggests that there is no trend toward one end of the spectrum over the other. Further study may illuminate whether or not this is due to a genuine belief that certain possibilities are not problems, or if it is simply a lack of experience or awareness.

Our evidence suggests that information problems are identified primarily by the respondents relationship to information (doing something to or with information), rather than more philosophical concepts (ie, the problem of "what *is* information?", or "what is ethical with regard to information?"). Again, this may be a product of degree program (more MSIM than MLIS) or context, but this is not clear from the results.

Of these, which information problems are most common and most severe?

As discussed in our analysis, the most commonly identified information problems were "Finding information that is complete, relevant and credible" and "Filtering a high volume of information (information overload) to keep only what's relevant or credible". These two problems also had the highest impact, and were joined by "Integrating information across multiple systems and disparate standards" as another high-impact problem. When asked to rank problems according to impact, "Threats to information security or privacy" easily claimed first place as the problem with the highest negative impact.

Conclusions and Future Work

The exploratory nature of this research is clear. Between the time constraints, our inability to obtain a random sample, and the overworked nature of our study population, there is good reason to

avoid generalizing from our results. Still, we think some interesting potential applications and future lines of inquiry emerged from our work. First, there are potential implications for iSchool curricular development. If students and their organizations are mainly impacted by issues of filtering and finding information, are these issues sufficiently covered in the core curriculum? Are students able to discern from course descriptions what existing classes addresses these challenges? In largely rejecting ethics as an information problem, are students perhaps inclined to define ethics as outside their sphere of responsibility as an information professional?

Second, from a more academic perspective, there are implications and questions about the classification of information problems. To develop a robust taxonomy, clearly a larger and perhaps more professionally diverse study population would be required. We regret our inability to produce a finished taxonomy in the scope of a quarter, but our research process and the data in Appendix D represent a first and hopefully illuminating step in this direction. We envision a taxonomy that would assist young information professionals both in diagnosing information problems in the workplace and in helping others (such as employers) to see their work in terms of information problems. Towards this end, a truly interesting line of inquiry would be qualitative research exploring the distinctions that participations are drawing when deciding whether to classify something as an information problem (as in question two of our survey). Finally, with a taxonomy in place, it would be valuable to work backwards by treating the taxonomy as 'symptoms' and enumerating the root and proximate causes underlying them.

REFERENCES

Alabdan, R. I. (2014). Major problems related to information management. *International Journal of Computers & Technology, 12*(3), 3347-3349.

Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science, 29*(5), 530-545.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education, 53*(4), 1207-1217.

Brand-Gruwel, S., Wopereis, I., Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior, 21*(3), 487-508.

Butler, P. (1994). Marketing problems: From analysis to decision. *Marketing Intelligence & Planning, 12*(2), 4-12.

Cooper, D. & Schindler, P. (2013). Measurement scales. In *Business Research Methods*. New York City, NY: McGraw-Hill Education.

De Lone, W. H., & McLean, E. R. (1992). Information systems success: the quest for the dependent variable. *Information Systems Research, 3*(1), 60-95.

Detlor, B. (2009). Information management. *International Journal of Information Management, 30.*

Evgeniou, T. & Cartwright, P. (2005). Barriers to information management. *European Management Journal, 23*(3) 293–299.

Faibisoff, S. G., & Ely, D. P. (1974). *Information and information needs*. New York: Columbia University.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*(4): 327-358.

Forbes. (2010). Managing information in the enterprise: Perspectives for business leaders. Retrieved from http://images.forbes.com/forbesinsights/StudyPDFs/SAP_InformationManagement_04_2010.pdf

Informa. (2011). Information management. Retrieved from http://informa.ie/information_

management

Information Technology Resources Board (ITRB). (1999). *Managing information systems: A practical tool.* Washington, DC: United States Government. Retrieved from http://govinfo.library.unt.edu/npr/howto/feb10ast.pdf

Iron Mountain. (October 23, 2012). Five common problems with information management. *SomersetLive.* Retrieved from http://www.somersetlive.co.uk/common-problems-information-management/story-17156773-detail/story.html

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Leckie, G. J., Pettigrew, K. E., & Sylvain, C. (1996). Modeling the information seeking of professionals: A general model derived from research on engineers, health care professionals, and lawyers. *The Library Quarterly: Information, Community, Policy, 66*(2), 161-193.

Li, H. F., & Cheung, W. K. (1987). An empirical study of software metrics. *IEEE Transactions on Software Engineering, 13*(6), 697-708.

Matharu, K. (October 20, 2015). The problems of information management in life science research. Eagle Genomics. Retrieved from https://www.eaglegenomics.com/problems-information-management-life-science-research/

McLaughlin, E. (June 09, 2014). Top IT management challenges and how to overcome them. *SearchCIO*. Retrieved from http://searchcio.techtarget.com/photostory/2240222213/Top-CIOs-reveal-their-biggest-information-technology-issues/1/Top-IT-management-challenges-and-how-to-overcome-them

Michael, C. J. (July 11, 2016). Six records information management challenges: Pharmaceutical industry. Paragon Solutions. Retrieved from http://www.consultparagon.com/blog/records-information-management-pharmaceutical-industry

Miranda, S. V. & Tarapanoff, K. M. A. (2008). Information needs and information competencies: a case study of the off-site supervision of financial institutions in Brazil. *Information Research, 13*(2), 5.

National Archives of Australia (NAA). (n.d.). Phase 1: Risk Assessment. Retrieved from http://www.naa.gov.au/Images/Phase-1-Risk-Assessment_tcm16-89010.pdf

National Archives of Australia (NAA). (n.d.). Phase 2: Assessment of information management functionality. Retrieved http://www.naa.gov.au/records-management/digital-transition-and-digital-continuity/information-is-interoperable/assessing-information-management-functionality/phase-2-functionality-checklist.aspx

National Archives of Australia (NAA). (n.d.). ISO 16175 Principles and functional requirements for records in electronic office environments. Retrieved http://www.naa.gov.au/records-management/agency/digital/iso-16175/index.aspx#part2

Palvia, P. C. & Palvia, S. C. (2013). Understanding the global information technology environment: representative world issues. In R. D. Galliers & D. E. Leidner (Eds.), *Strategic information management* (2nd ed.). United Kingdom: Routledge.

Palvia, P. C., Palvia, S. C., & Whitworth, J. E. (2002). Global information technology: a meta analysis of key issues. *Information & Management, 39*(5), 403-414.

Pearson, S. W. (1977). *Measurement of computer user satisfaction*. Unpublished dissertation, Arizona State University, Tempe.

Pickard, A. J. (2013). Sampling. In *Research methods in information* (2e). Chicago, IL: ALA \ Neal-Schuman.

Platiša, G. & Balaban, N. (2009). Methodological approaches to evaluation of information system functionality performances and importance of successfulness factors analysis. *Management Information Systems, 4*(2), 11-17.

Public Record Office Victoria (PROV). (2013). Information Management Maturity measurement tool – IM3. North Melbourne, Victoria, Canada: State Government of Victoria. Retrieved from http://prov.vic.gov.au/government/information-management-maturity-measure-tool-im3

Records and Information Management Branch (RIMB). (2004). *Accountability for information management: A model.* Edmonton, Alberta, Canada: Information Services Division, Government of Alberta. Retrieved from http://www.im.gov.ab.ca/3114.cfm

Robinson, J. (November 1, 2005). 10 principles of effective information management. Step Two. Retrieved from http://www.steptwo.com.au/papers/kmc_effectiveim/

Schön, D. A. (1983). *The reflective practitioner: how professionals think in action.* New York: Basic Books.

Spiegel Institut. (n.d.). Problem typology. Retrieved from http://www.spiegel-institut.de/en/method-knowledge/problem-typology

Trochim, W. M. K. (2006). Measurement validity types. Research Methods Knowledge Base. Retrieved from http://www.socialresearchmethods.net/kb/measval.php

Tufte, E. (n.d.). Sparkline theory and practice. Edward Tufte Forum. Retrieved from https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR

Watson, R. T., Kelly, G. G., Galliers, R. D., & Brancheau, J. C. (1997). Key issues in information systems management: An international perspective. *Journal of Management Information Systems, 13*(4), 91-115.

Woelfer, J. P., Duong, T. T., & Hendry, D. G. (2013). Coding manual for barriers and solutions to employment for homeless young people (UW Information School Technical Report IS-TR-2013-01-10). Seattle, WA: University of Washington, The Information School.

APPENDIX A: INTERVIEW PROTOCOL

1. Rooms on campus at the University of Washington will be booked for interviews.
2. Participants who meet criteria from our sample screen will be assigned to the the research team, 2 students for each researcher
3. Participants will be coded to keep data confidential from their interviews
   3.1. Codes will be sequential: e.g, PAR001
4. Interviewers will be coded for tracking data entry and analysis
   4.1. Codes will be sequential: e.g, IVW001
5. Interview Guide will be printed on 8.5" 11" paper prior to each interview
   5.1. Researcher conducting interview will add participant and interviewer codes to interview guide prior to meeting with the participant
6. One blank piece of paper will be paired with the Interview Guide for any additional notes or comments that may emerge in the interview and do not fit on the Interview Guide Form
7. Researchers will bring audio recording device to interview
8. Researchers will dress in business casual and avoid controversial or political symbols and iconography in attire
9. Researcher will meet with participant at location of reserved room
10. Researcher will introduce themselves to participants
11. Researchers will briefly explain the purpose of the interview using the following statement: *Thank you for your availability to meet to help us with our student project. We're trying to learn more about the types of information problems you may have encountered.*
12. Researchers will ask for oral consent for audio recording using the following statement: *Just to remind you, I will be recording this interview for research purposes. Your answers will remain confidential. Are you still willing to participate?*
    12.1. If Participant says yes:
        12.1.1. Circle the "Y" on the Interview Guide and continue with interview, taking thematic notes with recording device active
    12.2. If Participant says no:
        12.2.1. Circle the "N" on the Interview Guide and continue with interview, taking transcriptive notes without a recording device
    12.3. During the interview, Researcher will take written notes using the printed Interview Guide and the additional paper brought to interview
        12.3.1. Notes will be coded to individual questions based on the question numbering: e.g., a written note after question 6 will be coded "(6)"
    12.4. Researcher will note Start Time on the Interview guide
    12.5. Researcher will start recording
    12.6. Researcher will walk through questions listed in the Interview Guide
    Questions 1 - 7 are used to qualify the Participant profile.

12.6.1. No probes are planned for these questions
12.6.2. If a Participant struggles to answer, questions can be skipped
12.6.3. If Participant notes unexpected details in questions 6 and 7, Researcher may probe to get unique details on work experience and interactions with information

Question 8 is designed to instigate discussion of a critical incident related to an information problem.

12.6.4. Probes are planned to help collect data on root causes, consequences, and solutions
12.6.5. If Participant readily provides information on root causes, consequences, and solutions, planned probes can be ignored
12.6.6. If Participant notes unexpected details to question 8, Researcher may probe to get unique details on this scenario

Questions 9 - 10 are used to understand how the Participant defines "Information Problems."

12.6.7. No probes are planned for these questions
12.6.8. If a Participant struggles to answer question 9, Researcher will ask a specific probe from the situation the Participant recounted.
12.6.9. If Participant notes unexpected details to questions 9 and 10, Researcher may probe to get unique details on this scenario
12.6.10. If there is extra time in the interview, Researcher may ask additional probes to question 10

12.7. Researcher will conclude interview
12.7.1. Researcher will stop the recording
12.7.2. Researcher will note the stop time on the Interview Guide
12.7.3. Researcher will thank Participant for their time
12.8. Researcher will copy written notes to exchange with other researchers for analysis
12.9. Researcher will transcribe recording and written notes after interview

APPENDIX B: INTERVIEW GUIDE

Researcher ID:		**Participant ID(s):**	**Consent to recording? Y/N**
Start time:	**End time:**		
Participant profile			
1. What field were you working in prior to coming to the iSchool? 2. About how many people were employed by your company? 3. Did you work in the office, or remotely? 4. Does your organization have multiple locations?			
5. What type of position were you in? 6. Can you briefly tell me what kind of work you do in a typical day? 7. What were your interactions with information like in your job?			
Critical Incident			
8. Describe a time when you encountered a difficulty with information at work. 8.0. [Probe] What do you think were the roots of this difficulty? 8.1. [Probe] What were the consequences of this problem? 8.2. [Probe] What did you do? 8.3. [Probe] Did that work? 8.3.1. Clarification: The [insert subject from previous question]			
Defining Information Problems			
9. How would you personally define an information problem? 9.0. Clarification: That's fine, we can skip it or come back to it. 10. Can you give me a few more examples of information problems? Thanks for your time!			

APPENDIX C: SURVEY

PURPOSE

This is a study on workplace information problems experienced by UW iSchool students. This research is being conducted for a class (IMT 570: Analytic Methods for Information Professionals). We are interested in your thoughts about how to define an information problem as well as your perceptions of the frequency, severity, and impact of information problems.

RISKS AND BENEFITS

This survey should take 5-10 minutes to complete. In appreciation for your time, you will be enrolled in a drawing for a $10 Starbucks gift card. Your participation is voluntary and you may skip questions at your discretion. To safeguard your privacy, responses are anonymous and will be presented in aggregate. There is a small risk that you may experience emotional distress while reflecting on information problems you've encountered.

RESEARCHER CONTACT INFORMATION

Lani Smith (fullmoon@uw.edu)
Clint Posey (cposey@uw.edu)
Jacob Kovacs (kovjac19@uw.edu)

CONSENT

Do you give your informed consent to participate in this study? *

- ◯ Yes
- ◯ No

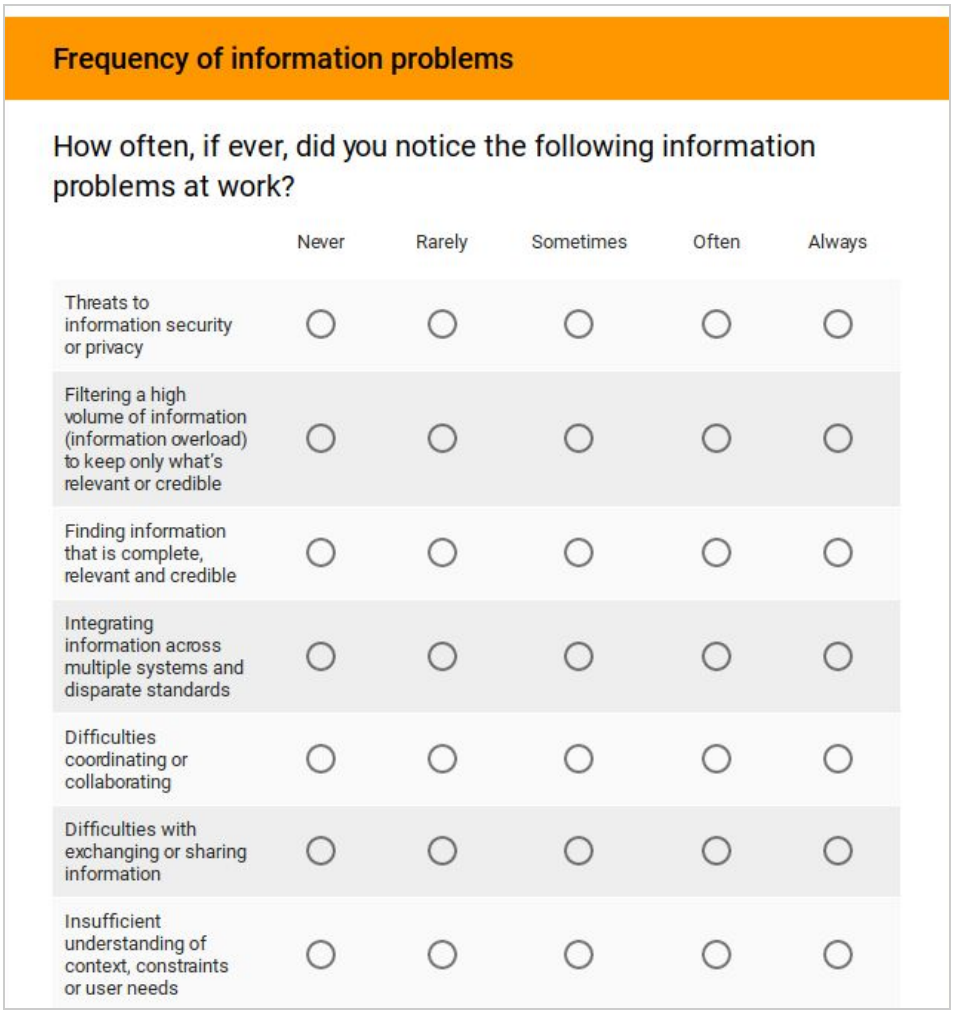
Defining information problems

What's your definition of an information problem?

Your answer

Which of the following would you consider an information problem (versus some other kind of problem)?

- [ ] Insufficient resources dedicated to learning and obtaining feedback
- [ ] Threats to information security or privacy
- [ ] Insufficient understanding of context, constraints or user needs
- [ ] Finding information that is complete, relevant and credible
- [ ] Not enough attention paid to ethics ('should we' versus 'can we')
- [ ] Integrating information across multiple systems and disparate standards
- [ ] Too much time spent on information, not enough on action
- [ ] Filtering a high volume of information (information overload) to keep only what's relevant or credible
- [ ] Denied access to necessary information because of low status in organizational hierarchy
- [ ] Difficulties coordinating or collaborating
- [ ] Difficulties with exchanging or sharing information
- [ ] Other:



Frequency of information problems

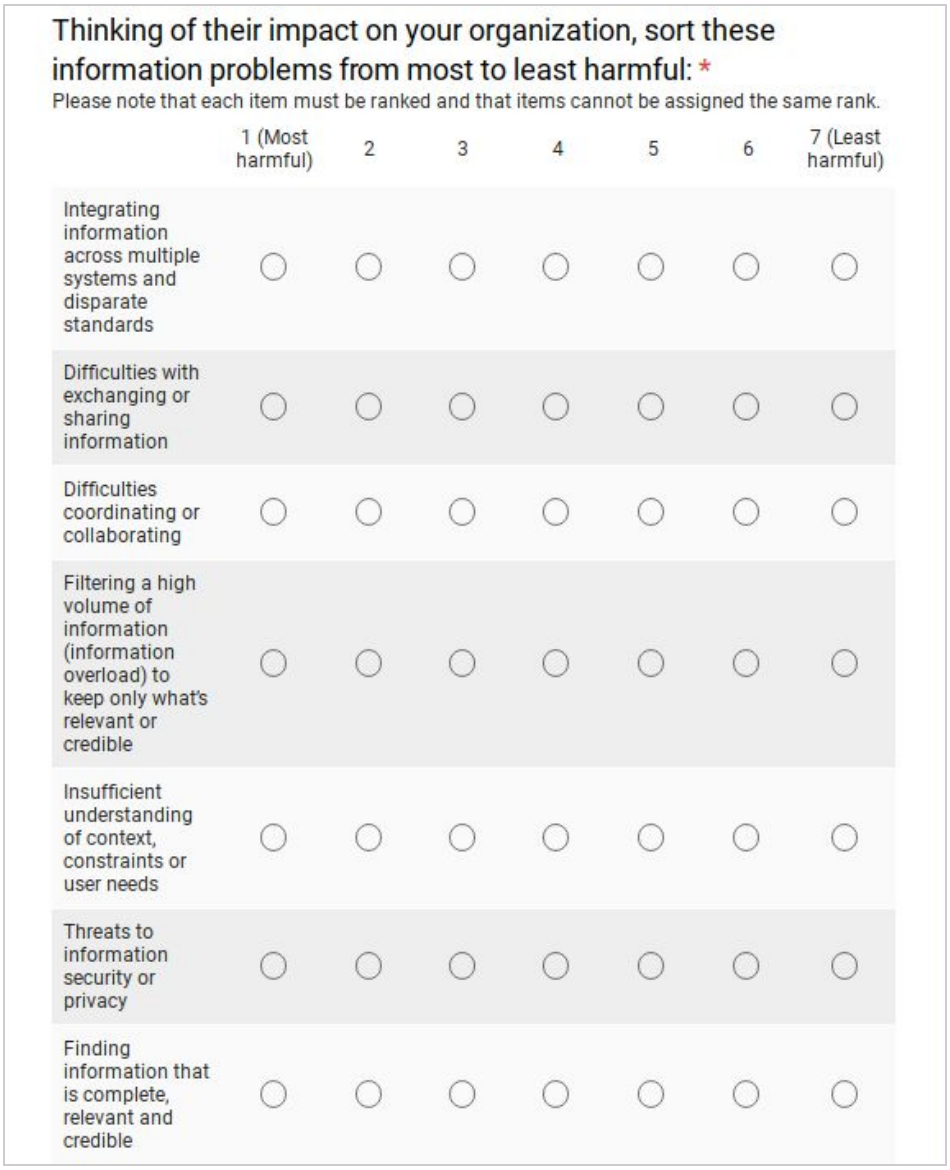
How often, if ever, did you notice the following information problems at work?

	Never	Rarely	Sometimes	Often	Always
Threats to information security or privacy	○	○	○	○	○
Filtering a high volume of information (information overload) to keep only what's relevant or credible	○	○	○	○	○
Finding information that is complete, relevant and credible	○	○	○	○	○
Integrating information across multiple systems and disparate standards	○	○	○	○	○
Difficulties coordinating or collaborating	○	○	○	○	○
Difficulties with exchanging or sharing information	○	○	○	○	○
Insufficient understanding of context, constraints or user needs	○	○	○	○	○

Impact of information problems

How much of an impact, if any, do you think these information problems had on your organization?

	1 (No impact)	2	3	4	5 (Severely negative impact)
Integrating information across multiple systems and disparate standards	○	○	○	○	○
Finding information that is complete, relevant and credible	○	○	○	○	○
Filtering a high volume of information (information overload) to keep only what's relevant or credible	○	○	○	○	○
Threats to information security or privacy	○	○	○	○	○
Insufficient understanding of context, constraints or user needs	○	○	○	○	○
Difficulties with exchanging or sharing information	○	○	○	○	○
Difficulties coordinating or collaborating	○	○	○	○	○



Thinking of their impact on your organization, sort these information problems from most to least harmful: *

Please note that each item must be ranked and that items cannot be assigned the same rank.

	1 (Most harmful)	2	3	4	5	6	7 (Least harmful)
Integrating information across multiple systems and disparate standards	○	○	○	○	○	○	○
Difficulties with exchanging or sharing information	○	○	○	○	○	○	○
Difficulties coordinating or collaborating	○	○	○	○	○	○	○
Filtering a high volume of information (information overload) to keep only what's relevant or credible	○	○	○	○	○	○	○
Insufficient understanding of context, constraints or user needs	○	○	○	○	○	○	○
Threats to information security or privacy	○	○	○	○	○	○	○
Finding information that is complete, relevant and credible	○	○	○	○	○	○	○

Demographic information

What field were you working in, prior to coming to the iSchool?

Your answer

How many years of work experience do you have, including internship experiences?

- ◯ 0 - 1
- ◯ 2 - 3
- ◯ 4 - 5
- ◯ > 5

Gender:

- ◯ Male
- ◯ Female
- ◯ Decline to respond
- ◯ Other :

Age (in years):

Your answer

Which program are you in?

- ◯ Full-time MSIM
- ◯ Mid-career MSIM
- ◯ Residential MLIS
- ◯ Online MLIS

APPENDIX D: INFORMATION PROBLEMS FROM INTERVIEWS

Researcher #1:

- Coordination across time zones
- Coordination across roles
- Lack of scope in design/software production process
- Lack of consensus on which tools to use
- Lack of trust in data
- Difficulty finding information you need
- Having too much information/needing to wade through or analyze to find meaningful information
- Ensuring completeness of information in decision-making
- Too much information, not enough focus
- Always trying to identify what information was valuable
- Scrapping useless metrics, having to develop new ones
- Collecting and storing large quantities of data for individual patients
- Learning how to be concise and informative in presentations
- Large quantities of data being too complex to comprehend/find patterns in
- Deriving simple output form complex input
- Ethics in user testing/research - just because you can, does that mean you should?
- Harmonizing compatibility of information across updated systems and technologies
- Losing out on information in order to maintain compatibility with older, lower-quality information for comparison

Researcher #2:

- for documentation: getting feedback on work from subject matter experts and development team
    - due to others' overwork
    - due to others' difficulty switching mindsets, stepping outside their own workflow
    - due to attitude that underestimates the value of documentation
    - due to under-commitment of resources to documentation
    - due to inefficiency of email
- for documentation: low quality (readability, accuracy)
    - due to rushing work because of delayed input
    - due to rushing work because of last-minute product changes
- being able to filter out what's relevant from a high volume of information
- being able to frame the problem sufficiently, to guide information search and facilitate effective information filtering
- resource costs of storing such irrelevant data
- being able to evaluate whether information (or the outputs of predictive models) is reliable, accurate, true, representative of the larger population
    - doubts whether researchers have made due effort, attained rigor
    - being able to evaluate whether researchers have made due effort
    - exacerbated by social media
    - perhaps a generational gap, with elders more prey to misinformation
- for database design: gaps in communication user requirements
    - between user and boss
    - between boss and database designers
    - leading to legal threats
    - leading to threat of client loss
- for database design: delays in projects due to changing user requirements
- for database design: junior-level technical errors that evaded normal safeguard mechanism (superior on sick leave)
- tension between what information is easiest for humans to comprehend (pictures, multimedia) and what information is easiest to analyze (textual, numeric)
    - textual information is more amenable to classification
    - we're making more multimedia data but it's harder to filter and classify

- security risks
    - due to IoT-based DDoS attacks
    - due to increased volume of data collected
- tension between consumer desire to share data and consumer fears about privacy
    - "hypocritical"
    - places too much responsibility on companies

Researcher #3:

- Collaboration
- Authority
- Empowerment
- Getting consensus
- Getting the right information within time constraints
- Forced to make decisions without understanding contexts
- Too much information, information overload
- Trade-offs between quality vs getting something out on time
- Getting information at the right point of time, in the right form
- Bad project management
- Intellectual property
- Not enough time spent in learning, lack of training
- Focusing on short-term objectives
- Losing sight of customer [end-user] experiences
- Don't have the right information, incomplete picture
- Don't have access to the right information
- Don't know whether information is useful for you or accurate, misinformation
- Lacking means to extract information
- Authenticity of information, no authority for credibility of sources
- Downward flows of information (from top - bottom in hierarchies)
- No information feedback
- Goals without consensus
- Turnover—knowledge gaps
- Lack of standardization
- Not enough information for management
- Not enough troubleshooting for information
- No point person for information
- Unknown unknowns—unaware of what we don't know
- Unaware of what we do know, of what exists
- Losing focus on important things
- Inaccessible information
- Collaboration across different teams--coordinating with information
- Authority and access to information
- Lack of information system to facilitate consensus
- Overload of information in communication (e.g., too many emails)
- Lack of information on competing goals between stakeholders
- Product usability information gaps
- Billing information gaps
- Lack of information from end-users
- Retrieval of information
- Information gaps in collaborative systems
- Last minute changes to a system without communication
- Lack of information strategies
- Information addiction (spending too much time absorbing information, not enough time acting on information

APPENDIX E: STATISTICAL OUTPUTS

Chi-square test of association between frequency (*How often, if ever, did you notice the following information problems at work?*) and experience (*How many years of work experience do you have, including internship experiences? 0-1; 2-3; 4-5; >5*):

Information problems	χ^2	df	p-value
Threats to information security or privacy	10.420	12	0.5792
Difficulties with exchanging or sharing information	9.368	12	0.6712
Difficulties coordinating or collaborating	7.175	12	0.8458
Insufficient understanding of context, constraints, or user needs	14.312	9	0.1117
Finding information that is complete, relevant and credible	8.532	9	0.4815
Filtering a high volume of information (information overload) to keep only what's relevant or credible	9.769	12	0.6362
Integrating information across multiple systems and disparate standards	11.302	12	0.5032

Chi-square test of association between frequency (*How often, if ever, did you notice the following information problems at work?*) and industry (*What field were you working in, prior to coming to the iSchool? IT; Finance/banking; Business/consulting; Education/libraries; Science; Government; Other*):

Information problems	χ^2	df	p-value
Threats to information security or privacy	19.762	24	0.7102
Difficulties with exchanging or sharing information	22.746	24	0.5348
Difficulties coordinating or collaborating	16.084	24	0.8850
Insufficient understanding of context, constraints, or user needs	23.322	18	0.1785
Finding information that is complete, relevant and credible	23.250	18	0.1812
Filtering a high volume of information (information overload) to keep only what's relevant or credible	24.188	24	0.4509
Integrating information across multiple systems and disparate standards	19.777	24	0.7094

Chi-square test of association between frequency (*How often, if ever, did you notice the following information problems at work?*) and iSchool program (*Which program are you in? Full-time MSIM; Mid-career MSIM; Residential MLIS; Online MLIS*):

Information problems	χ^2	df	p-value
Threats to information security or privacy	3.803	4	0.4333
Difficulties with exchanging or sharing information	9.246	4	0.0552
Difficulties coordinating or collaborating	9.277	4	0.0545
Insufficient understanding of context, constraints, or user needs	3.749	3	0.2898
Finding information that is complete, relevant and credible	2.072	3	0.5576
Filtering a high volume of information (information overload) to keep only what's relevant or credible	2.620	4	0.6233
Integrating information across multiple systems and disparate standards	0.769	4	0.9425