A Questionnaire for the Institutional Assessment of Personal Information Management

A.A.J. (Jos) van Helvoort
The Hague University of Applied Sciences, Academy of ICT&Media, The Hague, The Netherlands. a.a.j.vanhelvoort@hhs.nl

Abstract: *The purpose of the research was the development of a questionnaire that can measure the behaviour of groups of students (for instance department cohorts) in Personal Information Management (PIM). Variables for the questionnaire were derived from the international literature on PIM. The questionnaire was tested on 79 students (last year before graduation) from four different departments of the Academy of ICT&Media at The Hague University of Applied Sciences. The students' responses were checked on consistency, item non response, desirability bias and information value of the results. All these criteria indicated that the questionnaire is an adequate tool for the assessment of PIM at an institutional level. The results that have been found for the four departments have not yet been discussed with the managers of the Academy or of the individual departments.*

Keywords*: Information literacy, personal information management, questionnaires*

Introduction

Information literacy is often considered as a set of skills to solve specific information problems (Helvoort, 2010a, p. 63). The Scoring Rubric for Information Literacy that is described in a previous paper (Helvoort, 2010b) is an example of a tool that can be used to assess students' performances for such information problem-solving tasks. However, academics use many other conceptions of the phenomenon information literacy (Bruce, 1998). One of those conceptions is the "information control conception" (Bruce, 1998, pp. 33-34). This conception of information literacy is about "storing information (usually documents), in a fashion which ensures easy retrieval. For example, all the information is selected on the basis of its likely value for future use in research or teaching" (Bruce, 1998, p. 33).

People Still Build Their Personal Information Collections

People appear to prefer the use of information that is easily accessible. McGeachin (2004) reports a 1999 research paper from Hurd, et al. that found that scientists were accustomed to using easily accessible information and often cite "those titles that were in the library closest to them" and David Owen (1997, p. 30) reported that "many scientists and health professionals turn to their personal collection of documents as their primary source of information". But this was before the common use of digital information and electronic journals and nowadays the most convenient ways to obtain information are the World Wide Web and the electronic (journal) collections of the institutions where people work (Hemminger, Lu, Vaughan & Adams, 2007, p. 2211). However, this does not mean that people rely solely on retrieving information from the Web or their institutional electronic journal collections at the moment that they need it. A lot of information users prefer to use information sources which are personally known to them (Singh & Satija, 2007, p. 19) and they still keep personal information collections for future use (Bruce, 2005). While building such collections, they interpret the information they have encountered and evaluate it before deciding to keep it. This process assists them in remembering the information at the moment that it suits them. In other words, it helps them to avoid the situation that the existence of an item is forgotten entirely, at the moment that it could be useful for them (Jones, 2004). The previously mentioned Hemminger, et al. (2007, pp. 2211-2212) found that more than 85% of the scientists they surveyed maintained a personal article collection, in print or electronically.

Coughlan and Johnson (2009, pp. 7-9) discovered a similar practice among creative practitioners (in their case "film makers") who collected interesting materials and ideas which came along more or less accidentally and which were not really related to the projects they were currently working on. The maintenance of collections like these by scientists, scholar authors and more general information workers might also be motivated by the argument that they prefer to have reliable and relevant information at their fingertips at the moment that they compose their creative products rather than having to carry out the labour-intensive work of evaluating a long list of information items after a retrieval session.

The finding that people build personal information collections seems to be true not just for experienced information workers like scientists, scholar authors and creative practitioners. Head and Eisenberg (2011) found that also college students do not rely solely on search engines but use personal collections (notes, books and magazines in print), besides turning to their friends, family and classmates when they need information for daily life. These findings seem to confirm that the "control of information" is really a valuable facet of the information literacy competence.

According to Christine Bruce (1998, pp. 35-36) this "information control conception" should be distinguished from the "knowledge construction conception" that is also described by her. Critical analysis and meaning construction are important aspects of knowledge construction which in Bruce's opinion do not have to be a part of the information control conception. However, as has just been argued in this paper, storing information often needs evaluation and that is why it may be a good starting point for the construction of knowledge and may enable a learning process. This is particularly true when a person not only stores formal data or downloaded documents but also interprets the information in it for categorization and when he/she processes it cognitively (Bergman, Beyth-Marom & Nachmias 2003, p. 872; Hardof-Jaffe, Hershkovitz, Abu-Kishk, Bergman & Nachmias, 2009, pp. 5 and 12).

What Kind of Instruments do People Use to Build Their Personal Information Collections?

Personal index cards are of course nearly out of use nowadays and have been replaced by databases or spreadsheet software with search opportunities. Bibliographic citation managers are particularly designed for building personal information collections and for using the references later to generate correct in-text citations and references but people seem to experience a lot of trouble using them. María Pinto and Dora Sales found that the Spanish "translation students" they surveyed were badly motivated to use bibliographic management tools and that these students also believed that their performance was impaired when using tools like these (Pinto & Sales, 2010, pp. 625, 628).

Much more popular are bookmarking tools: either the URLs are stored on a person's local computer or in the cloud (online/social bookmarking with, for instance, Delicious). More recent and still not well known by students are academic social networks like Mendeley (see, for instance, Mead & Berryman, 2010). From the viewpoint of e-science the cloud solutions have the advantage that data can be shared with other people. Some bibliographic citation managers (for instance RefWorks) also provide this opportunity but as mentioned before those tools are not very popular (Pinto & Sales, 2010; Mead & Berryman, 2010).

On the other hand, downloading personal copies of electronic files and storing them locally is a fairly common practice, by students as well as by professionals. Titles in file names and a hierarchical folder system are standard tools of the computer operating system that can be used to relocate those files (McGeachin, 2004). A more advanced solution for the storage and retrieval of downloaded documents is the use of Personal Information Management (PIM) software that has the advantage of the integration of other types of information (email, contact information) in the reference collection (McGeachin, 2004).

My own observations of student groups have taught me that a lot of students keep a kind of log file in MS Word with metadata and URLs of documents that they have used, but also that some of them still rely on a paper note block. The more "net savvy" students send documents to their email account and use this as a kind of archive or they use a digital note tool like Evernote that gives them access to their personal notes at any place with any device.

Development of the Questionnaire

With the insights from the literature and my own observations as described in the introduction section of this paper, an online questionnaire was developed that can evaluate the practice in Personal Information Management of students in Higher Education. It was not the intention to develop a tool for the measurement of students' individual information performance but more an instrument that provides insights into the students' practices at an institutional level and that can provide suggestions for improvement of curricula.

The online questionnaire (in Dutch) is divided into an introduction and three "sections". A translation of the complete text of the questionnaire is aattached as an appendix to this research paper. The first section after the introduction consists of only one question: whether or not they use such a personal information system. If their answer is "No, I don't save data. If I need a document or an information item, I will search for it again" the respondents are directed to section 3 with four questions to categorize the respondents (the department of study, whether a full-time or part-time student, gender and age).

If their answer to the first question is "Yes, I save data about information sources with the aim to use them at a later time" the respondents are guided to section 2. In this section they are asked for 5 variables of the personal collections they use:

- the kinds of tools they use,

- the kinds of data they store,
- whether they share their collections with other people,
- how often they add an item to the collection,
- how often they use their collection(s) to retrieve an information item.

The first three questions of these five are multiple answer check boxes. The suggested choices are all based on the literature that is found on personal information collections as discussed in the introduction section of this paper. Questions number 1 and 2 have an additional free text field for the respondents' own alternative answers.

The final two questions are multiple choice questions for which respondents have to choose one answer. After completing section 2 respondents are guided to the same section 3 as the respondents who answered "No…" in section 1.

All the items in the online survey are "required questions": Respondents cannot leave a section before completing it. There is a "back button" at the end of sections 2 and 3 to allow respondents to change their answers in a previous section.

A draft version of the survey was published on the Internet with Google Docs. In a personal blogpost, a message on the HHS Yammer-site and a posting on the Dutch Library 2.0 Ning site, people were invited to test the draft version and to leave their comments. Fifteen professionals completed the questionnaire. Remarks of two of them (both academic librarians) were used to improve the questionnaire.

Results for the Academy of ICT&Media at The Hague University

The questionnaire was administered in regular classes on research methodology in September and November 2011. Seventy nine students (both full-time and part-time) completed it. All students were undergraduate students of the Academy of ICT&Media but from four different departments and each class had participants from the four different departments. The class in which the questionnaire was administered was for most of the students the last one of their major programme before they started work on their bachelor thesis (year 4 of the bachelor programme). To obtain a high response, it was decided to use a non-electronic version of the questionnaire that was completed by the students in the classroom itself. The survey was introduced by their regular teacher who was not engaged in the development of the questionnaire.

Use of PIM

Table 1 gives the results for the first question (whether they use any kind of PIM system or not) for each department and for the complete Academy of ICT&Media.

Table 1. PIM use of students of the Academy of ICT&Media

	Bus. IT&Mgt	Computer Eng.	Libr. & Inf. Sc.	Software Eng.	ICT&Media
Yes	17	11	15	21	64
No	4	0	0	11	15
Total	21	11	15	32	79

A more detailed analysis of the results can be found in Table 2, in which for each department the kind of data stored in the PIM systems is presented, and in Table 3, which shows how those data are related to the different types of PIM systems that the students indicated they used. The numbers in brackets in Tables 2 and 3 refer to the totals for the category concerned. In the group "Other" in Table 3 the students filled in a variety of systems, for instance Dropbox (4 times), Read It Later, Instapaper and iCloud. All the results in these tables will be discussed in more detail in the next section of this paper, in the evaluation of the questionnaire.

Table 2. Data that students of the different departments store in their PIM systems

	Bus. IT&Mgt (17)	Computer Eng. (11)	Libr. & Inf. Sc. (15)	Software Eng. (21)
URL (60)	16	9	15	20
Subjects (33)	9	6	11	7
Bibl. Data (21)	5	1	11	4
Summary (16)	4	4	4	4
Other (4)	1[a]	1[b]	1[c]	1[d]

a: Images, Audio recordings
b: Self created instruction
c: Copy/paste Introduction or Conclusion
d: Picture with IPhone

Table 3. Data that students store related to the types of PIM systems they use

	Bookmarks (48)	Self created dig. doc. (40)	Dig. Copies (37)	Paper Copies (26)
URL (141)	44	38	35	24
Subjects (88)	27	25	20	16
Bibl. Data (56)	16	18	12	10
Summary (46)	10	14	11	11
	Online Bookmarks (16)	Digital notes (16)	Bibl. Mgt. Software (13)	Other (13)
URL (57)	15	16	13	13
Subjects (43)	10	12	10	11
Bibl. Data (26)	9	5	11	1
Summary (21)	5	7	4	5

Subject Related Information

From the viewpoint of learning it is useful to remark that only 37 respondents reported that they store topic-related information (subjects and/or summaries). The scores for these two categories were 33 and 16, respectively (Table 2) but there was a lot of overlap between positively responding students. The other 27 students say that they restrict themselves to the more factual information (URL, title, author). Table 4 gives the distribution for storing subject related information over the four departments and for the complete Academy of ICT&Media.

Table 4. Students who store subject related information

	Bus. IT&Mgt	Computer Eng.	Libr. & Inf. Sc.	Software Eng.	ICT&Media
Yes	9	8	11	9	37
No	8	3	4	12	27
Total	17	11	15	21	64

Information Sharing

In the questionnaire the students were also asked whether they share their references with other students or their friends. Table 5 gives the answers on this question for the four departments. The results for this survey item make clear that LIS students are much more accustomed to sharing their personal collections than students from the other departments and seem to be much more prepared for a collaborative work environment. A chi-square test for the distribution of the answers in Table 5 indicated that those differences for the four departments were extremely significant statistically (*p* value was 0,001). This finding was also true when obvious 'false answers' (see next section on consistency in answer combinations) were deleted. In that case *p* value was 0,003.

Table 5. Students who share the data in their PIM system with other people

	Bus. IT&Mgt	Computer Eng.	Libr. & Inf. Sc.	Software Eng.	ICT&Media
Yes	7	7	15	8	37
No	10	4	0	13	27
Total	17	11	15	21	64

Evaluation of the Questionnaire

The students' responses reported in the previous section have been analysed to evaluate the questionnaire. Five indications were used to estimate the quality of the responses.

Consistency in Answer Combinations ("Reliability")

At first, an attempt was made to reveal any unexpected and therefore inconsistent answer combinations**.** A high occurrence of inconsistent answers from individual respondents may be an indication that the survey is not reliable.

Reasons for such inconsistent answers can be that questionnaire items are badly formulated or that respondents did not answer the questions seriously, because of survey fatigue, for instance (Chen, 2011, p. 660). This last phenomenon can be affected by the survey's length and the relevance of the survey in the eyes of the respondents (Porter, 2004, pp. 11-14).

Three combinations of answers were found that could be indications of low-quality responses:

- Question2Option1-Question3Option2 (Bookmarks but not URLs) was found for 4 respondents of the 48 who answered that they used Bookmarks,
- Question2Option2-Question3Option2 (Online bookmarks but not URLs) was found for 1 of the 16 respondents who answered that they used Online Bookmarks,
- Question2Option6-Question3Option1 (Bibl. Mgt. Software but no bibliographic data) was found for 2 of the13 respondents.

A second inspection of these questions and answer options did not lead to the conclusion of bad formulation although the concepts of Bibliographic Management Systems and Bibliographic Data may indeed lead to misunderstandings for non LIS students. For the two other inconsistencies the occurrence was rather small and they did not influence the main outcomes of the survey (see also the last paragraph of this section on the information value of the results). Hence, the occurrence of some inconsistent answers for these survey items was not a reason to doubt the reliability of the survey.

Inconsistency was also found within the answer options for question 6 and in this case it really was due to bad survey construction: 27 students answered that they do not share their references with other people in any way but four of them indicated that they do email them to classmates or their friends or share them in a closed community. In a later version of the survey which can be retrieved from Google Docs[1], this question has therefore been changed to a Yes/No question (Do you share your data with other people?) and only respondents who answer 'Yes' are directed to the other three options.

Item Nonresponse

Item nonresponse or "missing data" is another indication of low quality responses (Chen, 2011, p. 664). In the online version of the questionnaire it cannot appear because all the items are "required" questions. In the non-electronic version of the questionnaire that was used at The Hague University it was of course possible for respondents to miss an item but this did not happen for any of the 79 completed forms. In other words, there was no indication that the students had not completed the questionnaire seriously and willingly.

Desirability Bias

Extreme numbers of positive answers can be considered an indication of desirable answering of questionnaire items. Particularly in an educational context there is the danger that students give answers that they believe to be desired by the researcher or by their teachers.

Although the questionnaire forms were completed anonymously, there was indeed a danger of desirability bias with at least item number 1. The results however did not indicate that students were reluctant to answer that they do not use personal information management tools at all. The number of respondents who chose option "No" was not high but neither was it negligible. Also, for the other items, there was no indication that the students chose the most desirable answers.

Consistency Between Different Groups (Full Time/Part-Time Students; September/November)

Students from the Department of Software Engineering score more negatively for Question 1 than students from other departments. This was most clear for the full time students of the September group and was confirmed by the results for the full time group in November. The number of SE students in the part time groups was too small to draw any conclusions. Respondents for the Department of Business IT & Management differed in their answers for the different groups but these differences were so small that they can be put down to coincidence.

The answers for question 1 for the Departments of Computer Engineering and Library & Information Science were 100% positive.

In the following paragraph the statistical significance of the answers to question 1 will be discussed in more detail.

[1] https://docs.google.com/spreadsheet/ccc?key=0Ao6yiKblkJ5udGpCenNrRWRLY2FzT1E5d2xiM3ZicUE#gid=0

Information Value of the Results

A survey provides added value when the results are different for the various groups of respondents. In the case of the Academy of ICT&Media at The Hague University, students from the Department of Software Engineering scored much more negatively on question number 1 (whether one uses such a "Personal Information System") than those from the other departments. The results of a chi-square test for the distribution of the answers in Table 1 indicated that those differences for the four departments were very statistically significant (*p* value was 0,012). When obviously inconsistently answering respondents were deleted the chi-square test with the data from Table 1 resulted in an even more statistically significant *p* value (0,009 with N=69). It can be deduced from these results that Software Engineering students pay less attention to the keeping of personal information than students from the three other departments. The rather low appreciation of Software Engineering students for PIM is confirmed by their preference for the keeping of more shallow data (URLs) and the small number of students who store subjects and/or summaries when they reported that they do use some kind of a PIM (Table 4).

Conclusions and Discussion

The questionnaire that was developed has proven to be an adequate tool for the assessment of PIM at an institutional level. A test with 79 students from the Academy of ICT&Media of The Hague University resulted in statistically significant outcomes for the four departments that participated in this Academy. There were no indications that the survey was not completed seriously and willingly by the students. The limited length of the questionnaire (it was easy to complete in about 5 minutes) and the context of the course in which it was administered (the course itself was on research skills) may have contributed to this.

At the moment of writing this paper the outcomes have not yet been evaluated with the managers of the Academy and those of the four departments. An interesting question would be whether the differences in results for the various departments can be explained by the contents of their curricula. The value of the outcomes for the departments also depends on the extent to which the managers, curriculum developers and teaching staff value the personal information management behaviour of their students.

The results for the complete group of respondents from the Academy of ICT&Media at The Hague University confirm the suggestion in the literature that people still build personal information collections and that they do not rely solely on retrieval of information via the World Wide Web.

In the context of this conference on information management in a changing world, however, it was quite disappointing that only 37 students of the 64 who answered that they used any PIM system, reported that they share their references with other people. This is an indication that extra attention needs to be given in the curricula to collaborative learning and collaborative work.

It is worth mentioning that in The Hague University test, more than 50% of the students reported that they do not restrict themselves to factual data when they keep references but that they also store subject related information.

References

Bergman, O., Beyth-Marom, R. & Nachmias, R. (2003). The user-subjective approach to personal information management systems. *Journal of the American Society for Information Science and Technology, 54* (9), 872-878.

Bruce, C. (1998). The phenomenon of information literacy. *Higher Education Research & Development, 17* (1), 25-43.

Bruce, H. (2005). Personal, anticipated information need. *Information Research, 10* (3). Retrieved December 14, 2010 from http://informationr.net/ir/10-3/paper232

Chen, P. (2011). Finding quality responses: the problem of low-quality survey responses and its impact on accountability measures. *Research in Higher Education, 52* (7), 659-674.

Coughlan, T. & Johnson, P. (2009). Designing personal information management systems for creative practitioners. *Journal of Digital Information, 10* (5). Retrieved December 4, 2011 from http://journals.tdl.org/jodi/article/view/440

Hardof-Jaffe, S., Hershkovitz, A., Abu-Kishk, H., Bergman, O., & Nachmias, R. (2009). Students' organization strategies of Personal Information Space. *Journal of Digital Information, 10* (5), 1-17. Retrieved December 4, 2011 from https://journals.tdl.org/jodi/article/view/438

Head, A. & Eisenberg, M. (2011). How college students use the Web to conduct everyday life research. *First Monday, 16* (4). Retrieved April 6, 2011 from http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3484/2857

Helvoort, J. v. (2010a). Impact of recent trends in information and communication technology on the validity of the construct Information Literacy in Higher Education. *Communications in Computer and Information Science, 96* (2), 61-73.

Helvoort, J. v. (2010b). A scoring rubric for performance assessment of information literacy in Dutch Higher Education. *Journal of Information Literacy, 4* (1), 22-39. Retrieved June 30, 2010 from http://ojs.lboro.ac.uk/ojs/index.php/JIL/article/view/PRA-V4-I1-2010-2/1444

Hemminger, B., Lu, D., Vaughan, K. & Adams, S. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, (14), 2205-2225. Retrieved May 3, 2011 from http://www.ils.unc.edu/bmh/pubs/Information_Seeking_Behavior_of_Academic_Scientists-JASIST-2007.pdf

Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday, 9* (3). Retrieved November 28, 2010 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1123/1043

McGeachin, R.B. (2004). 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* (1/2), 127-137.

Mead, T. & Berryman, D. (2010). Reference and PDF-manager software: complexities, support and workflow. *Medical Reference Services Quarterly, 29* (4), 388-393.

Owen, D. J. (1997). Using personal reprint management software to teach information management skills for the electronic library. *Medical Reference Services Quarterly, 16* (4), 29-41.

Pinto, M. & Sales, D. (2010). Insights into translation students' information literacy using the IL-HUMASS survey. *Journal of Information Science, 36* (5), 618-630.

Porter, S. R. (2004). Raising response rates: what works? *New Directions for Institutional Research*, (121), 5-21.

Singh, K. P. & Satija, M.P. (2007). Information seeking behaviour of agricultural scientists with particular reference to their information seeking strategies. *Annals of Library & Information Studies, 54* (4), 7-220. Retrieved May 4, 2011 from http://nopr.niscair.res.in/bitstream/123456789/3240/1/ALIS%2054(4)%20213-220.pdf

Appendix: Translation of the Questionnaire

1.
This survey (10 questions) is about the use of Personal Information Management strategies for maintaining personal collections of information sources. We would like to know whether you knowingly use one or more 'systems' to keep track of the information that you have found or discovered and that you may want to use again in the future. Examples of such systems are: a collection of bookmarks for websites, a digital system with references to journal articles, study books and URLs or a simple paper memo book to write down such data. Will you please indicate in this first survey question whether you use such a 'Personal Information System'.

0 Yes, I save data about information sources with the aim to use them at a later time
0 No, I don't save data. If I need a document or an information item, I will search for it again.

If you answered 'Yes…', continue to question number 2
If you answered 'No…', continue to question number 7

2.
Which of the 'tools' from the list below do you use to relocate information (sources) that you once discovered? (Multiple answers possible.)

0 Bookmarks or Favourites on my own computer
0 Online bookmarks (for instance Delicious)
0 My own memo book (paper)
0 A digital notebook app (for instance Evernote)
0 A self-created digital document (for instance with MS Word, Excel or Google Docs) with URLs and hyperlinks
0 Bibliographic Management Software (for instance RefWorks, Endnote or Mendeley)
0 I keep copies (downloads) of digital documents on my own computer
0 Other ……

3.
What information regarding the source do you keep? (Multiple answers possible.)

0 Bibliographic data like author, title and publication year
0 URL (address on the WWW)
0 Subjects (subject headings, labels or tags)
0 Summaries
0 Other ……

4.
How often do you add an information source or an information item to one of the systems that you mentioned in survey item 2?

0 Daily
0 Weekly
0 Monthly
0 Less than once a month.

5.
How often do you retrieve an information item from one of the systems that you mentioned in survey item 2?

0 Daily
0 Weekly
0 Monthly
0 Less than once a month.

6.
Do you share the data from one of the systems that you mentioned in survey item 2 with other people, for instance by publishing it on the internet? (Multiple answers possible.)

0 Yes, I publish them on the public internet; everyone can access them.
0 Yes I share them in a ‘closed community’, for instance in a shared folder; people have to log in to access the data.
0 Yes, I share them with my classmates or my friends by sending them by e mail
0 No, I keep them for myself and nobody else has access to my data.

Finally, there are some questions about your personal position. Your answers will only be used for this research and will not be shared with your department’s staff or your teachers.

7.
What is your department?

……

8.
How are you enrolled?

0 Full time
0 Part time

9.
What is your gender?

0 Male
0 Female

10
What is your age?

… years old.

//