2015

Assessing Perceived Usability of the Data Curation Profiles Toolkit Using the Technology Acceptance Model

Tao Zhang
Purdue University, zhan1022@purdue.edu

Lisa Zilinski
Carnegie Mellon University, ldz@andrew.cmu.edu

D. Scott Brandt
Purdue University Libraries, techman@purdue.edu

Jacob Carlson
University of Michigan - Ann Arbor, jakecar@umich.edu

Recommended Citation
Zhang, T., Zilinski, L., Brandt, D. S., Carlson, J. (2015). Assessing Perceived Usability of the Data Curation Profile Toolkit Using the Technology Acceptance Model. *International Journal of Digital Curation*, 10(1), 48-67.

IJDC | *Peer-Reviewed Paper*

Assessing Perceived Usability of the Data Curation Profile Toolkit Using the Technology Acceptance Model

Tao Zhang
Purdue University

Lisa Zilinski
Carnegie Mellon University

D. Scott Brandt
Purdue University

Jake Carlson
University of Michigan

Abstract

The Data Curation Profiles Toolkit (DCPT) emerged out of a Purdue University Libraries' 2004 initiative to engage in multidisciplinary research. It is a tool developed to assist librarians and other information professionals to conduct data interviews and identify the needs of researchers when managing, sharing, or curating their data. The DCPT has been widely adopted and applied in various contexts but its usability as a tool has not been formally assessed. To address this need, we have conducted a survey of users of the DCPT. The survey included quantitative measures of potential influencing factors of using the DCPT and its perceived usability (its usefulness as a tool and its ease of use). Open-ended questions about users' experiences with the DCPT were also included to better understand the strengths and weaknesses of the tool, as well as areas that could be improved. Factor analysis of the quantitative results and subsequent regression models revealed several underlying factors that affect the perceived usability of the DCPT. Responses to the open-ended questions revealed several themes of users' concerns: the amount of time required to use the DCPT, the structure and format of the DCPT, alignment of the DCPT with particular contexts, and the use of the DCPT to engage faculty and the library community. By correlating themes identified from the open-ended questions with the analysis of quantitative data, this paper provides the first empirical assessment of the DCPT that could help further improve the toolkit's usability based on user needs and expectations. The methodology used in the study could readily be applied to assess and improve the utility of other tools used by data and information professional.

Received 18 September 2014 | *Accepted* 10 February 2015

Correspondence should be addressed to Tao Zhang, Purdue University Libraries, 504 W. State Street, West Lafayette, IN 47907, USA. Email: zhan1022@purdue.edu

An earlier version of this paper was presented at the 10th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Introduction

The Data Curation Profiles Toolkit (DCPT) emerged out of a Purdue University Libraries' 2004 initiative to engage in multidisciplinary research, and was created in 2010 as a resource for librarians to engage researchers in discussion about their data. Specifically, it is an interview protocol designed to capture information about a particular data set developed or managed by a researcher across its data lifecycle, exploring how the researcher and his or her lab are currently managing and working with the data set, and identifying what the researcher would like to do with the data. The output, a Data Curation Profile (DCP), is a document that represents data management and curation needs from the perspective of data producers using their own language. DCPs can be shared among the researchers, service providers and other stakeholders as a means of informing a plan of action.

Since its creation, we have been informally collecting and analysing data to understand how the DCPT is being employed by its intended users (i.e., librarians and information professionals), and more generally about how DCPT users are engaging with researchers around data. To better explore this issue, a formal and structured usability assessment of the toolkit became necessary. We expected the assessment to reveal significant underlying factors affecting users' perception and intention to use the DCPT, difficulties of using the toolkit, and areas in which the toolkit should be improved. To achieve this, we developed a questionnaire survey approach to the assessment, because the DCPT has a relatively large user base and it is difficult to observe how the toolkit is being used in real time. We applied factor analysis to narrow down the questionnaire items and identify their underlying factors. We also developed regression models to characterize how those factors affect the DCPT's perceived usability and users' intention to use the toolkit.

Background

We first describe how the DCPT was designed in order to explain our approach in assessing user experience with the tool. We then introduce the concept of perceived usability and discuss its utility for assessing the DCPT.

The Use of the Data Curation Profiles Toolkit

Using the DCPT to develop a DCP is a three-stage process (Witt, Carlson, Brandt & Cragin, 2009): preparation, interviews, and constructing the profile. In the preparation stage, the interviewer identifies the specific data set that will serve as the focus of the interview and selects modules of the DCPT to be included in the interview. In the next stage, the interviewer conducts the interview with the researcher, gathering information about their data, their practices with their data, and information about their needs. The final stage of the process involves transforming the information from the interview to sections of the DCP document. The user guide document of the DCPT explains the process for building a DCP.

The DCPT has been widely adopted and used by research librarians all over the world to help them connect with student and faculty researchers to learn more about

 doi:10.2218/ijdc.v10i1.344

their data management and curation needs. Notable uses of DCPs include Cornell's project to re-imagine the services offered through their DataSTaR repository (Wright et al., 2013) and Purdue's work to understand the needs of graduate students developing data in an agriculture field station (Carlson and Stowell-Bracke, 2013). More recently, the DCPT was applied to digital dance preservation to explore its digital outputs and artist expectations (Brandt and Kim, 2014). Completed DCPs can transcend the individual interaction between librarian and researcher and serve as a community resource for librarians seeking to understand researcher needs with data more generally and as a means to inform the development of data services. The Data Curation Profiles Directory[1] was launched in November of 2013 and provides access to a collection of published DCPs. In its first year, more than 4,000 copies of DCPs have been downloaded from the directory.

Previous studies have investigated the effectiveness of the DCPT. Carlson (2013) reported an increase in confidence in discussing data sharing, description and intellectual property among participants of workshops on using the DCPT, although participants noted the time and effort it took to develop a DCP as a barrier to its use. Participants in a symposium discussing the DCPT recognized the utility and impact of DCPT, and strongly suggested that it be enhanced to further facilitate data curation (i.e., archiving and preserving data at the end of a project) as opposed to data management, to further distinguish it from tools such as DMPonline or DMPTool (Brandt and Carlson, 2013).

Perceived Usability

Perceived usability is the core concept of the Technology Acceptance Model (TAM) proposed by Davis (1989), which has been widely used to investigate user adoption and acceptance of new technologies and systems. TAM proposes that two particular beliefs of perceived usability, perceived usefulness and perceived ease of use, are the primary drivers for technology and system acceptance. Perceived usefulness (PU) is "the degree to which a person believes that using a particular system would enhance his or her job performance" and perceived ease of use (PEOU) is "the degree to which a person believes that using a particular system would be free of effort" (Davis, Bagozzi and Warshaw, 1989; Davis, 1989). PU and PEOU jointly affect a person's attitude toward using the system as well as further intentions to use the system, which then translates into actual system use (or not). PU and PEOU are affected by external variables, such as management support, participation in training, tool functionality, task characteristics, prior similar experience and relevant skills (Legris, Ingham and Collerette, 2003).

TAM is an adaptation of the theory of reasoned action (Ajzen and Fishbein, 1980) and has been proven to be effective in explaining user behaviour in information system and tool implementation. The assessment of a software's or tool's perceived usability could be helpful in better understanding the determinants of a quality experience (Hassenzahl and Sandweg, 2004). The amount of time required to use the DCPT makes one-and-one usability tests inappropriate. Furthermore, those usability tests tend to focus on a limited set of performance metrics, such as task success, time on task, errors, efficiency and ease of learning (Albert and Tullis, 2013). It is hard to measure the DCPT in these metrics, as some interview questions of the DCPT may be skipped, while others may be extemporized. On the other hand, studies on PU and PEOU show that they can "predict a user's acceptance and actual usage of a system," which in turn can have

1 Data Curation Profiles Directory: http://docs.lib.purdue.edu/dcp/

"strong implications for IT designers, trainers, and stockholders, enabling them to better strategize their resources and emphases" (Zhang and Li, 2005).

In this study, we extended TAM from technical systems or tools to assessing the DCPT since previous user feedback suggested that the toolkit's perceived usability is critical for its overall user experience and adoption. We used TAM as a basis to survey potential external factors that may affect the perceived usability of the DCPT, identify major categories of external factors, and then characterize the relationship between external factors, perceived usability, and intention to use the DCPT. The results will help understand critical factors affecting the perceived usability and adoption of the DCPT, making targeted training for user groups and deep customization of the toolkit for various scenarios possible. The methodology validated in this study could be useful for evaluating the utility of other tools used in the data management and curation area.

Methodology

We developed an initial pool of 28 possible questions covering potential influencers of the perceived usability of the DCPT (see Appendix 1) based on our experiences with the tool and previous user feedback. The questions covered the following areas: (1) experience in conducting interviews; (2) knowledge of data management issues; (3) current job responsibilities; (4) motivations and perceived time requirements in using the DCPT; (4) considerations of the format, complexity, and adaptability of the DCPT, and requirements placed on interviewees; and (5) user needs for training, support, and help documentation. The items were formatted into statements with five-point Likert scale (e.g., 'strongly disagree' to 'strongly agree', 'very short' to 'very long', or 'too simple' to 'too complex'). The perceived usability was measured by adapting the items in Davis (1989), as shown in Appendix 2.

We included optional open-ended questions in the questionnaire to understand the reasoning behind respondents' answers to the Likert scale questions. The questions covered the difficulties and obstacles of using the DCPT, and areas that user liked and could be improved. After pilot tests of the questionnaire with librarians who have used the DCPT previously, we sent the questionnaire to 895 registered users of the DCPT website[2] in December 2013. Respondents were given a month of time to complete the questionnaire. We collected in total 221 responses (24.7% response rate), of which the majority are professional librarians with data management related responsibilities.

For data analysis, we first calculated descriptive statistics (min, max, mean, and standard deviation) of the ratings of questionnaire items to form an overview of the data. To identify the underlying factors that influence respondents' ratings of perceived usability and intention to use the DCPT, we conducted an exploratory factor analysis of the ratings in SAS 9.2. Exploratory. Factor analysis is a statistical method used to uncover the underlying structure of a relatively large set of measured variables (see Lattin, Carroll and Green, 2002) by extracting a small set of interpretable factors that can adequately explain the correlations among the measured variables. We then used the extracted factors as a basis to develop regression models and examined the relative weights of each factor affecting perceived usability and intention to use. Finally, we analyzed respondents' answers to the open-ended questions and related those responses to the quantitative analysis results.

2 DCP Toolkit website: http://datacurationprofiles.org

 doi:10.2218/ijdc.v10i1.344

Results

Factor Analysis

For the item statements that are in the opposite direction of other items (e.g., "The DCP Toolkit is rigid and inflexible to work with"), we reverted their responses by subtracting the original rating from six to make them consistent with responses to other items (i.e., "strongly disagree" means the lowest/worst rating and "strongly agree" means the highest/best). We conducted statistical tests and showed that the questionnaire items were reliable measures of respondents' perceptions and the ratings data were suitable for factor analysis. Details of those test results are reported in Appendix 3. Seven factors were extracted from the final factor analysis results and they accounted for 84.2% of the total variance. The result of factor pattern with loadings is shown in Appendix 4, with the entries higher than 0.5 in each row highlighted in grey. Factor 1 includes items about current job responsibilities, current data service situation, choosing needed modules, adjusting organization of modules, and adjusting questions of DCPT. These items are related to respondents' understanding of their data management service needs and how to apply DCPT in their own specific contexts. Therefore, we refer this factor as "Applicability". Factor 2 includes items that are related to time commitment of the process of applying the DCPT and creating DCPs. We refer Factor 2 as "Time". Factor 3 includes items that cover time requirement of learning DCPT and its complexity. We thus refer Factor 3 as "Complexity". Factor 4 combines two items that are related to the experience and skills of the interviewer conducting one-and-one interviews, and two items about publishing and sharing DCPs. This combination was probably due to the similar ratings of those questionnaire items from the respondents. We refer Factor 4 as "Experience and Share". Similarly, we refer to Factors 5, 6, and 7 as "Training and Help", "Extensibility", and "Interviewee Requirements" (see Appendix 4).

Regression

To reveal how the identified factors influence perceived usability, we conducted multivariate, stepwise regression analyses, using the means of item responses for each factor as independent variables and PU, PEOU, and intention to use as dependent variables. For the regression analysis with PU as the dependent variable, Factors 1 (Applicability), 4 (Experience and Share), and 5 (Training and Help) were in the final regression model ($R^2 = 0.439$, $F(217, 3) = 17.22$, $p = 0.018$). For the regression analysis with PEOU as the dependent variable, the final regression model ($R^2 = 0.471$, $F(216,4) = 12.93$, $p = 0.011$) included Factors 1, 2 (Time; not significant with $p = 0.064$), 3 (Complexity), and 7 (Interviewee Requirements). Factors 1, 2, 3, 5, and 6 (Extensibility) were in the final regression model of intention to use ($R^2 = 0.515$, $F(215, 5) = 12.10$, $p < 0.0001$). The detailed regression parameter estimates are shown in Table 1. These results show that PU, PEOU, and Intention to Use are all affected by Applicability combined with other different factors. The positive regression coefficients (β) of Applicability, Experience and Share, Training and Help, and Extensibility mean that respondents tended to rate higher PU, PEOU and Intention to Use with higher ratings of these factors. On the contrary, the Complexity and Interviewee Requirements have negative impact on PEOU; and Time and Complexity negatively affect intention to use.

Table 1. Parameter estimates of regression analyses of perceived usability.

(*a*) Dependent Variable: Perceived Usefulness

Independent Variable	Parameter Estimate (β)	Standard Error	*F*	*P*
Intercept	-3.60	3.76	0.92	0.34
Factor 1 (Applicability)	3.00	0.67	20.39	< 0.001
Factor 4 (Experience and Share)	2.38	0.61	15.52	0.0002
Factor 5 (Training and Help)	1.17	0.48	5.89	0.018

(*b*) Dependent Variable: Perceived Ease of Use

Independent Variable	Parameter Estimate (β)	Standard Error	*F*	*P*
Intercept	27.22	3.81	51.08	< 0.0001
Factor 1 (Applicability)	1.12	0.52	4.73	0.034
Factor 2 (Time)	-1.11	0.59	3.56	0.064
Factor 3 (Complexity)	-1.59	0.63	6.46	0.014
Factor 7 (Interviewee Requirements)	-1.71	0.66	6.79	0.012

(*c*) Dependent Variable: Intention to Use

Independent Variable	Parameter Estimate (β)	Standard Error	*F*	*P*
Intercept	2.22	1.22	3.29	0.08
Factor 1 (Applicability)	0.60	0.18	11.61	0.001
Factor 2 (Time)	-0.41	0.18	5.00	0.03
Factor 3 (Complexity)	-0.38	0.19	4.04	0.05
Factor 5 (Training and Help)	0.29	0.12	6.18	0.02
Factor 6 (Extensibility)	0.33	0.13	6.62	0.01

We also developed a regression model with PU and PEOU as independent variables and intention to use as the dependent variable. The model R^2 is 0.449, with $F(218, 2) = 28.92$ and $p = 0.03$. The estimated regression parameter (β) for perceived usefulness is 0.156 ($F = 46.43$, $p < 0.0001$). For perceived ease of use the estimated regression parameter (β) is 0.054 ($F = 3.37$, $p = 0.07$). This shows that both PU and PEOU positively affect the intention to use, with PU having the predominant influence.

Open-Ended Questions

We asked four open-ended questions on the survey to augment our understanding of the survey respondents' perceptions and use of the DCPT. The text of the questions and the number of responses received are as follows:

 doi:10.2218/ijdc.v10i1.344

- If you used the DCP Toolkit in the past, did you encounter any difficulties? If yes, please explain. (n=46)
- If you plan to use the DCP Toolkit, what would be the obstacles you may encounter? (n=62)
- What are the things you like about the DCP Toolkit? (n=69)
- What are the things you think should be improved in the DCP Toolkit? (n=54)

We conducted an initial analysis of the content of responses to identify common themes and connections between them. We crafted a list of categorical codes based on the themes identified and assigned relevant codes to each response in a second analysis. Each response was independently reviewed by two of the authors, who then discussed the content of the response and came to consensus on which of the codes to assign. This process required two iterations. The two authors had 66.8% agreement on the initial coding and they reached 100% consensus on the final coding results. Many of the responses included more than one theme and were assigned multiple codes. A few were not substantive and did not receive a code. A complete list of the codes we developed is included in Appendix 5.

There were several themes that emerged in all four of the open-ended questions: the amount of time that is required to use the DCPT; the structure and format of the toolkit; alignment of the DCPT with the particular context, plans, or goals of the library; and the use of the DCPT as a means of engaging faculty and the library community. In particular, there was strong tension in the utility of the DCPT between the thoroughness of the toolkit and amount of time that is required to make use of it. This was illustrated by positive comments on the depth and completeness of the DCPT such as:

- 'I like the depth of information it provides about data practices.'
- '…[T]he questions are all there and you can choose which ones you need.'
- 'When I do use the toolkit, it is very thorough which makes it useful…'
- 'I think the step where you put together a narrative is very important… For me, that was the moment when the data management practices of the lab became a coherent story, rather than disconnected answers to questions. It helped me understand the motivations and key issues for the lab.'

These comments are in stark contrast from those who saw the thoroughness of the DCPT as requiring too much of an investment on their part or on the part of faculty:

- 'Completing the DCP Toolkit to the point that it is truly useful to the data curation process requires a large amount of time on the part of the conductor and the interviewee. …it's a lot to expect.'
- 'I like the structure and the content, but using the full toolkit requires more time than I can usually devote to the average consultation.'

One element that appeared to exacerbate concern over the investment of time was that of the perceived alignment of the DCPT to current practices or intended development. Although the DCPT was recognized as an instrument that could readily be

applied to most situations by some, others felt that it fell outside of their particular interests or situations and therefore was not the right tool for them to apply. Others were not sure how they could apply what they learned from a DCP to their efforts:

- 'The main challenge is to adapt the tool to my work scenarios.'
- 'It didn't fit well into the types of interactions that I have with researchers, whereas a shorter version might.'
- 'I have read some data curation profiles, and I wonder how librarians and researchers make data management decisions based on the information collected by the worksheet. It would be helpful if there was discussion at the end of each profile of what actions were taken…'

The amount of time required and questions surrounding its application made it difficult for some to be able to articulate the value of using the DCPT to their faculty or to their colleagues in the library.

- 'It was hard to rationalize the need for these interviews to busy researchers.'
- 'The biggest difficulty is meeting with a professor for two hours. Except during the summer, most faculty don't have that kind of time to spend on something when they don't see the value to themselves.'
- 'Generating interest among librarians… to commit time to becoming familiar with and utilizing the toolkit when data management/curation does not comprise a significant portion of their responsibilities (yet).'

Several respondents mentioned the flexibility of the DCPT and that it could be modified to suit time constraints or a particular interest as one of its strengths. Some mentioned modifying the DCPT as a means of addressing concerns with issues of time commitment. Despite the ability to modify the DCPT, several respondents requested the development of a more compact "lite version". Other respondents would like to see offshoots of the toolkit that are more focused to a particular aspect or type of data or geared to a specific field. In considering how the DCPT could be further improved, several respondents voiced their support of creating an online version of the toolkit that would allow for greater flexibility and utility for interview preparation, interview data collection, and DCP creation.

- 'I really wish that there was a database driven version of the DCP Toolkit. The Word version is a pain. The formatting gets all messed up too easily.'
- '… [I]f it were an online tool, it could probably point to examples of various things that might help interviewer ask better questions or help interviewee offer more fleshed out opinions.'
- 'Adopt an XML format which will allow for changes to sections without need to re-publish the whole page. Subsections within each section can be added, subtracted, based on context and adaptation will be easier, as each institution has different reality.'

 doi:10.2218/ijdc.v10i1.344

Outside the toolkit, a few respondents mentioned a desire to see more emphasis on a community of practice surrounding the creation and use of DCPs.

- 'It would be extremely useful to be able to see a combined view of multiple DCPs… (i.e., look at an individual area or question across multiple DCPs, and allow for slicing and dicing by discipline, data type etc.).'
- '… I also like the idea of linking the DCP to trusted practices, such as the Data Seal of Approval, metadata schemas, curation standards, etc. Ideally there needs to be a suite of tools that helps in developing a collection policy, an archival policy, a curation program … and certification.'

Discussion

The questionnaire survey generated both quantitative and qualitative responses. Results of the factor analysis of the quantitative data showed a number of underlying factors that affect the perceived usability and users' intention to use the DCPT. These factors are the toolkit's applicability to particular contexts, time requirement, complexity, users' interview experience and willingness to share DCPs, training and help, extensibility, and requirements on interviewees.

The regression models showed that applicability, experience and share, as well as training and help are positive determinants of the DCPT's perceived usefulness; and the former two factors have higher weights than training and help. The DCPT's perceived ease of use is positively affected by its applicability to different contexts and negatively by its complexity and interviewee requirements; and the three factors have almost equal weights. Respondents' intention to use the DCPT is affected by its applicability, time requirement, training and help, and extensibility, with time requirement being the negative factor. The intention to use is largely affected by perceived usefulness with a regression coefficient about three times of perceived ease of use, which suggests that respondents consider the utility of DCPT more important than its ease of use. The regression models all have R^2 values around 0.5, which suggests that there could be other factors not covered by the questionnaire items affecting the DCPT's perceived usability and respondents' intention to use (e.g., awareness of alternative tools or methods, and perceived return on investment). The open-ended responses of our survey may provide some insights into additional factors unaccounted for in the regression models, but future studies could provide a more extensive examination of potential factors affecting users' perception and use of the DCPT.

The open-ended responses substantially matched results from the quantitative analysis and provided additional contextual information. The factor analysis showed that the DCPT's applicability to specific contexts is the prevailing determinant of its perceived usability and respondents' intention to use the DCPT. Corresponding to this, the majority of respondents mentioned adapting the toolkit to create questions for their own data service contexts and interview needs. Overall, respondents thought the DCPT is useful, mainly because: (1) the toolkit provides "a vocabulary and questions" for developing a data interview; (2) it provides a good "framework for thinking about how to approach consulting on data management issues"; and (3) the DCPs developed using the toolkit are relevant to ongoing data management work. These reasons summarize

users' evaluation of the DCPT's usefulness and provide directions for further improving the utility of the toolkit.

A major obstacle of using the DCPT is its time requirement. Conducting the data interview requires time commitment of the interviewees who are perceived as being very busy people. Preparing for the interview, transcribing the interview content and developing DCPs also require a considerable investment of time. The time requirement on both sides may make it difficult to scale up the effort of developing DCPs. Related to the time requirement, respondents had concerns about the learning needed to fully understand the process of applying the DCPT and the terminology in the toolkit that requires explanations of concepts during interviews. These concerns are also indicated by the complexity and interviewee requirements in the regression model of PEOU. However, the time factor is not significant in the regression model of PEOU ($p = 0.064$), which may be caused by insufficient number of responses for the regression analysis. A post-hoc sensitivity analysis showed that a sample size of at least 250 might be able to show the significance of the Time factor in the regression models.

Regarding how the DCPT could be improved, a lighter and more adjustable version with less time requirements was the priority. As mentioned, a possible approach to reducing time was to create an online application for setting up the DCP structure, recording interview information, and developing the DCP document. The online application may help address other suggestions about improving the structure and making the DCPT more adaptable. However, it is worth noting that the applicability and thus perceived usefulness of the DCPT are the most important determinants of users' intention to use. Future versions of the DCPT should continue improving its comprehensiveness and coverage of typical data management and curation scenarios in scientific research.

The regression models suggest that training and help are important for the DCPT's perceived usefulness and respondents' intention to use. Correspondingly, a number of open-ended responses requested additional help on different adaptations of the DCPT in data service, terminology, and how to transform the collected information into a complete DCP. While the interview experience and willingness to share results emerged in the factor analysis, it did not enter the regression models. This result suggests that most respondents feel they can understand the DCPT well enough to use it but would like support in adapting and customizing it for their own purposes.

Conclusion

The study's goal was to assess the usability of Data Curation Profiles Toolkit and identify factors that affect its perceived usability and users' intention to use the toolkit. We applied the Technology Acceptance Model in the assessment. With open-ended responses, the results have provided comprehensive information about users' perception, difficulties and expectations of the DCPT. This information will feed in to a recently awarded Institute of Museum and Library Services National Leadership Grant, "Enhancing the Data Curation Profiles to help Bridge the Gap between Researcher and Repository." With the grant, we will develop a more extensible and powerful DCPT for librarians and information professionals to connect with stakeholders seeking to deposit their data into a repository. Improving usability is key to the redesign, and perceptions uncovered in this study will be valuable as work progresses on the DCPT 2.0.

This study also provided validation that our assessment approach could be used for the usability assessment of other tools in the data management field by revealing crucial factors that need to be addressed. Because of the exploratory nature of the study, we did not apply strict quantitative analyses, such as structural equation modelling. We also had limited numbers of open-ended responses, as the questions were optional to respondents. Future survey studies could collect more responses and develop different statistical models to offer additional assessment of the DCPT or other similar tools.

Acknowledgements

We thank all the survey respondents for taking their time to complete the questionnaire. We also thank Michael Witt, Courtney Matthews, Judy Nixon, and Amy Van Epps for pilot testing the questionnaire and providing valuable feedback.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Albert, W., & Tullis, T. (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*. Newnes.

Brandt, D.S., & Carlson, J.R. (2013). *Final performance report: Understanding curation through the use of data curation profiles* (Project Report, IMLS Grant No. RG-06-10-0101-10).

Brandt, D.S., & Kim, E. (2014). Data curation profiles as a means to explore managing, sharing, disseminating or preserving digital outcomes. *International Journal of Performance Arts and Digital Media*, *10*(1), 21–34. doi:10.1080/14794713.2014.912498

Carlson, J. (2013). Opportunities and barriers for librarians in exploring data: Observations from the data curation profile workshops. *Journal of eScience Librarianship*, *2*(2). doi:10.7191/jeslib.2013.1042

Carlson, J., & Stowell-Bracke, M. (2013). Data management and sharing from the perspective of graduate students: An examination of the culture and practice at the water quality field station. *Portal: Libraries and the Academy*, *13*(4), 343–361. doi:10.1353/pla.2013.0034

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. doi:10.2307/249008

Davis, F.D., Bagozzi, R.P., & Warshaw, P.R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003. doi:10.1287/mnsc.35.8.982

Hassenzahl, M., & Sandweg, N. (2004). From mental effort to perceived usability. In *Extended abstracts of the 2004 Conference on Human Factors and Computing Systems - CHI '04* (p. 1283). New York, NY: ACM Press. doi:10.1145/985921.986044

Lattin, J., Carroll, D., & Green, P. (2002). *Analyzing multivariate data* (1st ed.). Cengage Learning.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, *40*(3), 191–204. doi:10.1016/S0378-7206(01)00143-4

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Witt, M., Carlson, J., Brandt, D.S., & Cragin, M.H. (2009). Constructing data curation profiles. *International Journal of Digital Curation*, *4*(3), 93–103. doi:10.2218/ijdc.v4i3.117

Wright, S.J., Kozlowski, W.A., Dietrich, D., Khan, H.J., Steinhart, G.S., & McIntosh, L. (2013). Using data curation profiles to design the Datastar dataset registry. *D-Lib Magazine*, *19*(7/8). doi:10.1045/july2013-wright

Zhang, P., & Li, N. (2005). The importance of affective quality. *Communications of the ACM*, *48*(9), 105–108. doi:10.1145/1081992.1081997

 doi:10.2218/ijdc.v10i1.344

Appendix 1

Table 2. Questionnaire items that may influence perceived usability of DCP Toolkit.

Item	Statement	Scale
1	I have experience in conducting one-and-one interviews.	'Strongly disagree' to 'Strongly agree'
2	I have skills necessary for conducting one-and-one interviews.	'Strongly disagree' to 'Strongly agree'
3	I have knowledge of issues in data management and curation necessary for using the DCP Toolkit.	'Strongly disagree' to 'Strongly agree'
4	The DCP Toolkit is relevant to my current job responsibilities.	'Strongly disagree' to 'Strongly agree'
5	The DCP Toolkit is important to my current job responsibilities.	'Strongly disagree' to 'Strongly agree'
6	I think being able to publish data curation profiles with DOIs that can be cited in an open directory is important.	'Strongly disagree' to 'Strongly agree'
7	I think sharing data curation profiles among the data service community is important.	'Strongly disagree' to 'Strongly agree'
8	It takes _____ time to learn the DCP Toolkit.	'Very short' to 'Very long'
9	It takes_____ time to prepare for an interview.	'Very short' to 'Very long'
10	It takes _____ time to conduct an interview.	'Very short' to 'Very long'
11	It takes _____ time to transcribe an interview.	'Very short' to 'Very long'
12	It takes _____ time to analyze an interview record.	'Very short' to 'Very long'
13	It takes _____ time to generate a data curation profile.	'Very short' to 'Very long'
14	The current MS Word document format of the DCP Toolkit works for me.	'Strongly disagree' to 'Strongly agree'
15	I would like to have a XML version of the DCP Toolkit.	'Strongly disagree' to 'Strongly agree'
16	I would like to have a database-driven version of the DCP Toolkit.	'Strongly disagree' to 'Strongly agree'
17	I can apply the DCP Toolkit to the data service situations I am facing.	'Strongly disagree' to 'Strongly agree'

Item	Statement	Scale
18	The DCP Toolkit is _____.	'Too simple' to 'Too complex'
19	The DCP Toolkit is _____.	'Not detailed enough' to 'Too detailed'
20	I can adapt the DCP Toolkit by choosing the modules I need.	'Strongly disagree' to 'Strongly agree'
21	I can adjust the organization of modules of the DCP Toolkit for use in different situations.	'Strongly disagree' to 'Strongly agree'
22	I can adjust the questions in the DCP Toolkit for use in different situations.	'Strongly disagree' to 'Strongly agree'
23	Participating in the interview takes _____ time for the person being interviewed.	'Very short' to 'Very long'
24	Interview questions are generally _____ to understand for the person being interviewed.	'Very easy' to 'Very difficult'
25	Interview questions are generally _____ for the person being interviewed.	'Very irrelevant' to 'Very relevant'
26	How important is training for using the DCP Toolkit?	'Unimportant' to 'Very important'
27	How important is user support for the DCP Toolkit?	'Unimportant' to 'Very important'
28	How important is the help documentation of DCP Toolkit?	'Unimportant' to 'Very important'

 doi:10.2218/ijdc.v10i1.344

Appendix 2

The items for measuring perceived usability and intention to use were as follows.

Perceived usefulness

1. Using the DCP Toolkit improves the quality of the data management consulting I do.
2. The DCP Toolkit enables me to accomplish data management consulting tasks.
3. The DCP Toolkit supports critical aspects of my data management consulting work.
4. Using the DCP Toolkit enhances my effectiveness in my data management consulting work.
5. Using the DCP Toolkit makes it easy to do my data management consulting work.
6. Overall, I find the DCP Toolkit useful in my data management consulting work.

Perceived ease of use

1. Learning to use the DCP Toolkit is easy for me.
2. The DCP Toolkit is rigid and inflexible to work with.
3. Applying the DCP Toolkit is clear and easy.
4. I find it takes a lot of effort to become skillful at using the DCP Toolkit.
5. Overall, I find the DCP Toolkit easy to use.

Intention to use DCP Toolkit

- How likely are you going to use the DCP Toolkit in the future?

Appendix 3

Tests Results for Factor Analysis of the Questionnaire Ratings Data

To check the reliability of responses across the items (i.e., whether those items measured the same set of underlying factors), the overall standardized Cronbach's coefficient was 0.79 and the lowest Cronbach's coefficient of the item responses is 0.76. These two numbers are greater than the suggested value of 0.70 given by Nunnally and Bernstein (1994). This suggests that the questionnaire items are reliable measures of respondents' perception of factors affecting perceived usability and intention to use the DCPT. We performed a Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) test to check whether the raw data were suitable for factor analysis. The KMO test result is 0.518, which is higher than the acceptable exploratory research norm of 0.5 established by Nunnally and Bernstein (1994).

 doi:10.2218/ijdc.v10i1.344

Appendix 4

Table 3. Factor analysis result of items influencing perceived usability of DCP Toolkit.

Item	Factor 1: Applicability	Factor 2: Time	Factor 3: Complexity	Factor 4: Experience and share	Factor 5: Training and help	Factor 6: Extensibility	Factor 7: Interviewee requirements
1	0.000	0.219	0.087	0.794	0.054	0.033	0.065
2	0.041	-0.025	0.208	0.809	-0.050	0.016	-0.098
3	0.423	0.291	0.251	0.410	-0.147	-0.160	0.077
4	0.785	0.133	-0.069	0.240	-0.039	-0.143	0.031
5	0.690	-0.012	-0.200	0.275	0.022	0.031	-0.109
6	0.147	-0.249	-0.231	0.592	0.168	0.286	0.007
7	0.209	-0.194	-0.254	0.678	0.146	0.195	-0.068
8	0.061	0.294	0.513	0.014	0.348	0.190	0.363
9	-0.201	0.539	0.452	-0.221	0.261	0.080	0.138
10	-0.123	0.648	0.339	-0.073	-0.066	0.059	-0.021
11	-0.203	0.725	0.009	-0.013	-0.096	0.216	-0.141
12	-0.039	0.745	0.034	0.121	0.328	-0.020	0.031
13	0.171	0.703	0.387	0.000	0.070	-0.036	0.050
14	0.078	-0.042	0.003	0.012	0.072	-0.498	-0.193
15	0.225	0.101	-0.109	0.142	0.162	0.619	0.107
16	0.186	0.070	0.164	0.200	-0.140	0.695	-0.239
17	0.715	0.074	0.012	0.061	0.018	0.010	0.042
18	-0.158	0.156	0.834	0.032	0.078	-0.008	0.006
19	-0.189	0.070	0.796	-0.110	0.027	-0.057	0.033
20	0.810	-0.188	-0.114	-0.004	-0.084	0.218	-0.103
21	0.810	-0.253	-0.141	-0.054	-0.081	0.095	-0.020
22	0.858	-0.216	-0.098	-0.098	-0.058	0.174	-0.127
23	-0.042	0.235	0.524	0.177	-0.021	0.017	-0.019
24	0.041	0.025	0.099	0.012	0.015	0.252	0.655
25	0.047	0.158	-0.049	-0.056	0.826	-0.149	0.136
26	-0.105	0.038	0.018	0.084	0.917	0.021	-0.125
27	-0.165	-0.035	0.273	0.155	0.708	0.065	-0.224
28	-0.250	-0.104	-0.027	-0.068	-0.188	-0.119	0.756

Appendix 5:
Lists of codes of open-ended question responses

Table 4. Combined responses to the "Difficulties" and "Obstacles" questions (154 codes applied to 108 responses).

Code	Frequency
Time - Faculty	20
Alignment	18
Time (General)	15
Articulating Value	14
Time - Interviewer	11
Learning Curve	10
Structure	8
Modification	7
Time - Transcription	6
Engagement - Faculty	5
Format	5
Language	5
Scope	5
Advance Knowledge	4
Result	4
Engagement - Community	3
Lite Version	2
Resource	2
Time - Develop DCP	2
Time - Interview	2
Transcription	2
Translation	2
Engagement - Library	1
Thorough	1

 doi:10.2218/ijdc.v10i1.344

Table 5. Responses to the "Likes" question (121 codes applied to 69 responses).

Code	Frequency
Structure	27
Thorough	21
Modification	16
Engagement - Faculty	12
Engagement - Library	11
Ease of Use	8
Alignment	4
Result	4
Neutral	4
Scope	2
Advance Knowledge	2
Engagement - Community	2
Lifecycle	2
Standardization	2
Articulating Value	1
Format	1
Language	1
Documentation	1

Table 6. Responses to the "Improve" Question (68 codes applied to 54 responses)

Code	Frequency
Format	17
Support	8
Modification	8
Engagement - Community	5
Lite Version	5
Structure	5
Articulating Value	4
Language	4
Engagement – Faculty	3
Time	2
Time – Develop DCP	2
Alignment	1
Ease of Use	1
Engagement – Library	1
Resource	1
Translation	1