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

Tao Zhang, D. Scott Brandt
Purdue University
USA

Lisa Zilinski
Carnegie Mellon University
USA

Jake Carlson
University of Michigan
USA
The Data Curation Profiles Toolkit (DCPT)

- Created at Purdue for librarians to engage researchers in discussion about data
  - Interview protocol
  - Capture information about a dataset across lifecycle
  - Explore how data are used and managed
  - Identify data curation needs
  - Document generated from data interview
Using the DCPT

- Three-stage process
  - Preparation
  - Interviews
  - Constructing DCP

- DCPs as community resource
  - Understand researcher needs with data
  - Inform development of data services

- Data Curation Profiles Directory
  http://docs.lib.purdue.edu/dcp/
Assessing the DCPT

- Various data collected on how the DCPT has been used
- Study on effectiveness of the DCPT
  - Carlson (2013): Increased users’ confidence in discussing data sharing, but time and effort for developing DCP as barrier to use the DCPT
  - Brandt & Carlson (2013): Users recognized utility and impact, and strongly suggested further enhancement for data curation needs (as opposed to data management)
Motivation

➢ **Formal** and **structured** assessment of the DCPT to reveal factors affecting:
  ➢ User perception
  ➢ Intention to use
  ➢ Difficulties
  ➢ Areas to be improved

➢ Challenge for usability evaluation
  ➢ Time limit
  ➢ Task performance measures
Technology Acceptance Model (TAM)

- Perceived usability (PU and PEOU)
  - Predict user acceptance and actual usage of technical systems/tools
  - Critical for overall user experience

Source: Davis (1989)
Methodology

- **Survey** of potential determinants of perceived usability (28 questions measured in 5-point Likert scale)
  - “I have experience in conducting one-and-one interviews.”
  - “It takes ___ time to learn the DCP Toolkit.”
  - “I can adjust the questions in the DCP Toolkit for use in different situations.”
- **PU, PEOU, and Intention to Use** the DCPT
  - Questionnaire measures adopted from Davis (1989)
- **Open-ended questions**
  - Difficulties and obstacles
  - Areas that user liked and could be improved
Methodology

- Survey sent to 895 registered users of the DCPT website in Dec. 2013
  - 28 measures of determinants
  - 221 responses (24.7%) in a month
  - Most respondents are professional librarians with data management related responsibilities

- Data analysis
  - Likert ratings of determinants -> Exploratory Factor Analysis -> Regression Analysis
  - Qualitative analysis of open-ended responses
Exploratory Factor Analysis

- Correlations between variables result from sharing of factors
- Uncover underlying structure of a large set of measured variables
Factor Analysis Results

- Seven factors extracted for 84.2% of total variance
  - Applicability
  - Time
  - Complexity
  - Experience and Share
  - Training and Help
  - Extensibility
  - Interviewee Requirements

- Factors in regression models to predict PU, PEOU, and Intention to Use
Regression Results

- Multivariate, stepwise regression models

\[ \text{PU PEOU Intention to Use} = \text{Factors 1-7} + \text{error} \]

<table>
<thead>
<tr>
<th>Perceived Usefulness</th>
<th>Estimated $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicability</td>
<td>3.00</td>
</tr>
<tr>
<td>Experience and Share</td>
<td>2.38</td>
</tr>
<tr>
<td>Training and Help</td>
<td>1.17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perceived Ease of Use</th>
<th>Estimated $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicability</td>
<td>1.12</td>
</tr>
<tr>
<td>Time</td>
<td>-1.11 *</td>
</tr>
<tr>
<td>Complexity</td>
<td>-1.59</td>
</tr>
<tr>
<td>Interviewee Requirements</td>
<td>-1.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intention to Use</th>
<th>Estimated $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicability</td>
<td>0.60</td>
</tr>
<tr>
<td>Time</td>
<td>-0.41</td>
</tr>
<tr>
<td>Complexity</td>
<td>-0.38</td>
</tr>
<tr>
<td>Training and Help</td>
<td>0.29</td>
</tr>
<tr>
<td>Extensibility</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*p-value = 0.064, sample size > 250 needed to show significance
Regression Results

- Multivariate, stepwise regression models
  \[ \text{[Intention to Use]} = \text{PU, PEOU} + \text{error} \]

<table>
<thead>
<tr>
<th>Intention to Use</th>
<th>Estimated $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>0.156</td>
</tr>
<tr>
<td>PEOU</td>
<td>0.054 ($p = 0.07 &gt; 0.05$)</td>
</tr>
</tbody>
</table>

- All regression models $R^2$ around 0.5
Open-ended Questions

- Questions asked in the survey:
  - “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)

- Qualitative analysis method
  - Two independent researchers reviewed and coded responses
  - Two iterations, 66.8% agreement on initial coding and 100% consensus on final coding results
Open-ended Responses

Themes:

- Amount of time required to use the DCPT vs. depth of information from completed DCP
- Structure and format of the toolkit
- Alignment of the DCPT with particular context
- Using the DCPT to engage faculty and library community
Open-ended Responses

- Finding the right balance
  - **Time** required for both researchers and interviewees
  - **Depth of information** in DCP as good utility

- Applicability
  - Adapting structure and format to contexts
  - Making decisions based on results

- Extending the DCPT
  - Compact, “lite version”; online tool
  - Focus on particular data types or fields
  - Community building based on DCPs
Quantitative vs. Qualitative Results

- The balance between time and value
  - Interviewee Requirements affect Perceived Ease of Use
  - Complexity and Time affect Intention to Use
  - Time requirement vs. thoroughness in open-ended responses

- Training and Help
  - Significant in regressions on Perceived Usefulness, Intention to Use
  - Open-ended responses requested additional help on:
    - Adaptations for different purposes
    - Transforming collected information into DCPs and making decisions
Conclusion

- Technology-Acceptance Model useful for assessment
  - Factor analysis & regressions
  - Significant factors identified for Perceived Usability & Intention to Use
- Open-ended responses complement quantitative results
- Usability improvement of the DCPT
  - Reduce time requirement
  - Increase flexibility
  - Training and help
DCP: What We’ve Learned & Going Forward

DCP 1.0
Paper-based

Preparation → Interviews → Constructing DCP

DCP 2.0
Tech-based

Preparation & Training → Interview/Interact → Profiles/Best Practices → Utilizing Outputs

DCP 2.0 Roadmap Workshop, Purdue University, June 2015
Thank you!

Tao Zhang, D. Scott Brandt
Purdue University
USA
zhan1022@purdue.edu
techman@purdue.edu

Lisa Zilinski
Carnegie Mellon University
USA
ldz@cmu.edu

Jake Carlson
University of Michigan
USA
jakecar@umich.edu

IDCC 2015, London