Engineering researchers' data reuse behaviours: a structural equation modelling approach

Yeon Kyoung Joo
Department of Digital Media, Myongji University, Seoul, Republic of Korea, and
Youngseek Kim
School of Information Science, University of Kentucky, Lexington, Kentucky, USA

Received 16 August 2016
Revised 26 January 2017
Accepted 25 March 2017

Abstract

Purpose – The purpose of this research is to investigate the factors that influence engineering researchers' data reuse behaviours.

Design/methodology/approach – The data reuse behaviour model of engineering researchers was investigated by using a survey method. A national survey was distributed to engineering researchers in the USA, and a total of 193 researchers responded.

Findings – The results showed that perceived usefulness, perceived concerns and norms of data reuse have significant relationships with attitudes toward data reuse. Also, attitudes toward data reuse and the availability of data repositories were found to have significant influences on engineering researchers' intention to reuse data.

Research limitations/implications – This research used a combined theoretical framework by integrating the theory of planned behaviour (TPB) and the technology acceptance model (TAM). The combination of the TPB and the TAM effectively explained engineering researchers' data reuse behaviours by addressing individual motivations, norms and resource factors.

Practical implications – This research has practical implications for promoting more reliable and beneficial data reuse in the engineering community, including encouraging positive motivations toward data reuse, building community norms of data reuse and setting up more data repositories.

Originality value – As prior research on data reuse mainly used interviews, this research used a quantitative approach based on a combined theoretical framework and included diverse research constructs which were not tested in the previous research models. As one of the initial studies investigating data reuse behaviours in the engineering community, the current research provided a better understanding of data reuse behaviours and suggested possible ways to facilitate engineering researchers' data reuse behaviours.

Keywords User studies, Data repositories, Institutional repositories, Data sharing, Metadata standards, Data reuse, Engineering researchers,

Paper type Research paper

1. Introduction

Technological advances in modern research have opened up new opportunities for data sharing and data reuse research. Researchers today have more opportunity to use publicly available shared data or reuse other researchers' data (Borgman, 2012; Davis and Vickery, 2007; Tenopir *et al.*, 2011). Such data sharing and reuse opportunities can be useful for researchers to validate their findings, identify errors and enrich academic discussions by

The authors acknowledge the CoS Pivot (ProQuest) for allowing them to use their scholar database in recruiting survey respondents for this research.

0264-0473
DOI 10.1108/EL-08-2016-0163

exploring new research questions that were not asked in the original research (Borgman, 2012; Tenopir *et al.*, 2011). However, regardless of the importance and value of data sharing and reuse, relatively little scholarly attention has been given to data sharing and data reuse (Faniel *et al.*, 2016; Yoon, 2016; Zimmerman, 2008). Furthermore, most existing studies on data reuse deal with hard science (Tenopir *et al.*, 2015), health science (Yoon, 2016) or social science (Faniel *et al.*, 2016), leaving the field of engineering largely unexplored. As a result, there is not a clear understanding of what factors promote or impede engineers' data sharing or data reuse behaviours. Given this, this research study attempts to examine engineering researchers' data reuse behaviours, an active format of data sharing.

Although earlier studies mainly dealt with the design and development of technical systems, engineering studies in more recent years include a wider range of topics, such as management as well as traditional design and development (Buede and Miller, 2016). In consideration of this trend in engineering, the present study defines engineering research as an academic discipline focusing on diverse aspects of design, development, implementation and use of mechanical and technical systems (Downey, 2005). This definition of engineering research also covers a number of engineering disciplines, including electrical engineering, mechanical engineering, environmental engineering, computer engineering, civil engineering, industrial engineering, chemical engineering and more.

This study investigates possible factors that can influence engineering researchers' data reuse behaviours. Specifically, this research investigates what factors influence engineering researchers' data reuse behaviours, and to what extent those factors influence those behaviours. To this end, a review of the related literature provided a theoretical framework, from which was developed a research model and hypotheses, and an evaluation of the research model was conducted using data analysis of survey responses from 193 engineering researchers in the USA.

2. Literature review

Regardless of the importance of data reuse in the modern scholarly debate, only a limited number of studies have explored data reuse behaviours within specific disciplinary communities, including environmental engineering (Van House *et al.*, 1998), social science (Niu, 2009; Yoon, 2016), ecology (Zimmerman, 2008), archaeology (Faniel *et al.*, 2013a) and astronomy (Carlson and Anderson, 2007). Furthermore, such studies on data reuse often used qualitative research methods (Niu, 2009; Yoon, 2016; Zimmerman, 2008) or descriptive surveys to describe researchers' motivations for data reuse (Faniel *et al.*, 2016; Tenopir *et al.*, 2015). As a result, the findings of the previous studies are often limited to episodic information or a field-restricted understanding of data reuse, which makes it hard to generalize the findings in a broader context, such as the field of engineering. Some of the possible factors that previous studies found to influence data reuse behaviours in other fields will be reviewed to test if these factors can influence data reuse behaviours in the field of engineering as well.

The perceived benefits of data reuse are revealed to be a main motivator of researchers' data reuse behaviour (Faniel and Jacobsen, 2010; Niu, 2009; Yoon, 2015). A recent study (Yoon, 2015), for example, revealed that social scientists reuse data because data reuse is cost-effective. That is, researchers can explore additional research questions in a cost-effective way when they reuse data, as data are already collected. Similarly, Niu (2009) reported that the perceived usefulness of data is a key factor in data reuse. According to Niu, data reusers prioritize the informative value of data when they decide to use secondary data. If researchers believe that existing data are related to their research and, therefore, are worthy of reuse, researchers would actively find a way to reuse data to reveal new findings that are secondary to the purpose of original research.

Moreover, effort is known to play an important role in researchers' data reuse behaviours. Previous studies in data reuse suggest that it is not easy to use secondary data. Data reuse often requires a great deal of effort from researchers, as researchers need to search for appropriate data, as well as spend time and effort to understand the logic and characteristics of the data (Rolland and Lee, 2013; Scaffidi *et al.*, 2006; Zimmerman, 2008). Specifically, finding a source of data is a tedious process in which researchers often struggle with fragmented data (Scaffidi *et al.*, 2006). Limited accessibility to data is also a big obstacle in the process of data reuse (Faniel *et al.*, 2016), as it is sometimes very difficult to get permission from the owner of data for reuse. Even if one finds an appropriate source of data and obtains access to it, there is another important step which makes sense of the data set. McCall and Appelbaum (1991) argued that secondary data analyses involve additional considerations and efforts to understand the nature of original data, including its sample, measures and assessment ages. Thus, if understanding original data requires too much time and effort on researchers' ends, they are not likely to reuse data even if data are relevant to their research. In a similar vein, Faniel *et al.* (2012) conducted interviews with novice data reusers and found that novice researchers' decision to reuse data can be affected by the amount of effort in understanding and reconstructing data.

As the amount of effort is critical in the decision to reuse data, previous studies on data reuse highlighted the importance of contextual information about original data (Berg and Goorman, 1999; Cragin and Shankar, 2006; Jirotka *et al.*, 2005; Zimmerman, 2008). If contextual information is well-documented and available for data reusers, researchers can reduce the effort to understand original data and maximize the usefulness of data. However, in reality, databases often do not have enough contextual information, or, even if information is presented, such information is likely to be difficult for data reusers to decode (Birnholtz and Bietz, 2003; Faniel *et al.*, 2013b; McCall and Appelbaum, 1991; Yoon, 2016). This is because researchers of original data tend to describe only a sufficient amount of contextual information for their original research purpose and, therefore, critical information for data reusers is often missing or hard to identify. As effort for data reuse greatly depends on systematic documentation of contextual information of data, researchers of data reuse often argued that documentation of physical and technical contexts of data is essential to promote data reuse (Baker and Yarmey, 2009; Chin and Lansing, 2004).

Although previous studies suggested diverse factors as possible predictors of researchers' data reuse behaviours (Faniel *et al.*, 2016; Piwowar and Vision, 2013; Tenopir *et al.*, 2015; Yoon, 2016), previous studies have limitations because they used only a limited set of variables when examining data reuse behaviours (in the field of science); their studies focused mainly on phenomenal descriptions of data reuse among scientists but not investigating theoretical mechanisms of data reuse; and there has not been enough empirical supports as their approach was more qualitative rather than quantitative. Taking the current challenges from the previous studies, this research explored and identified a broader set of variables that can possibly affect engineering researchers' data reuse behaviours by using a survey method and investigated the underlying mechanism behind the data reuse behaviours by applying a strong theoretical framework.

3. Theoretical framework

This research used a unique theoretical framework by integrating the theory of planned behaviour (TPB) (Ajzen, 1991; Ajzen and Fishbein, 2005) and the technology adoption model (TAM) (Davis, 1989). TPB is a well-known social psychology theory that explains people's behaviours in terms of their attitudes, subjective norms and resource-facilitating conditions (Ajzen, 1991; Ajzen and Fishbein, 2005). TPB can guide us to understand engineering researchers' data reuse behaviours by providing possible explanatory factors, including:

- their attitudes toward data reuse (the overall evaluation of reusing other researchers' data);
- their norm about data reuse as a member of the research community; and
- resource facilitating conditions, such as availabilities of data repositories and additional technical supports that assist researchers' data sharing and reuse behaviours.

Furthermore, the current research extends its theoretical framework to TAM, as TAM can provide a more precise understanding of engineering researchers' attitudes toward data reuse. TAM explains how an individual's perception can influence their intention to conduct a target behaviour (mainly adopting technologies) with respect to an individual's perception of the target, including utilitarianism (perceived usefulness) and effort expectancy (perceived ease of use) (Davis, 1989). By applying TAM to the present research, engineering researchers' perceptions of data reuse can be better described and it can be more clearly demonstrated how these perceptions are related to researchers' benefit expectancy factors (e.g. usefulness of data reuse) and effort expectancy factors (e.g. effort involved in data reuse). Therefore, the theoretical framework combining TPB and TAM can help us better articulate the underlying mechanism of engineering researchers' data reuse behaviours.

4. Research model and hypotheses development

Based on the theoretical framework presented above, a systematic model was developed, which can be particularly effective to examine engineering researchers' data reuse behaviours. Based on TPB, attitudinal (attitudes toward data reuse), normative (norm of data reuse) and institutional (availabilities of metadata standards and data repositories) factors were selected for the model. Based on TAM, more specific individual factors were also considered, such as perceived usefulness, perceived concerns and perceived effort involved in data reuse behaviours.

4.1 Perceived usefulness

In regard to scientific data reuse, perceived usefulness can be defined as the degree to which the engineering researchers benefit from reusing other researchers' data (Davis, 1989). According to TAM, if the perceived usefulness of performing a target behaviour is high, an individual is likely to conduct the target behaviour (Davis, 1989). That is, if researchers believe that their reuse of other researchers' data is useful for their research purpose, they are likely to formulate a positive attitude regarding data reuse for their own research (Faniel and Jacobsen, 2010; Niu, 2009; Yoon, 2015). Based on this logic, the hypothesis evolved that the perceived usefulness in data reuse would positively influence engineering researchers' attitudes toward data reuse.

H1. Perceived usefulness in data reuse positively influences engineering researchers' attitudes toward data reuse.

4.2 Perceived concerns

Perceived concerns would be another factor for individual researchers' reuse of other researchers' data. Scholars proposed perceived concerns as an important factor influencing data sharing and illustrated its effect on individual perceptions of reusing other researchers' data (Borgman, 2007; Cragin *et al.*, 2010; Vickers, 2006). Previous research indicates that perceived concerns arise through an individual's intrinsic motivation or anxieties encountered in the process of reusing other researchers' data (Benlian and Hess, 2011; Cheng, 2011). According to previous research, researchers' perceived concerns involving data reuse are one of the most important factors influencing scientists' data sharing and reuse behaviours. Therefore, this study assumed that the perceived concerns of data reuse would negatively influence engineering researchers' attitudes toward data reuse.

H2. Perceived concerns of data reuse negatively influence engineering researchers' attitudes toward data reuse.

4.3 Perceived effort

In scientific data sharing and reuse settings, perceived effort represents the degree to which people associate the level of difficulty with the reuse of other researchers' data (Venkatesh *et al.*, 2003). Several recent studies have demonstrated that perceived effort can influence scientists' data sharing and reuse behaviours (Rolland and Lee, 2013; Scaffidi *et al.*, 2006; Zimmerman, 2008). According to the TAM, if researchers believe that reusing other researchers' data sets is difficult and complicated, they may not want to reuse them (Davis, 1989). Therefore, perceived effort would be an important factor to influence engineering researchers' reuse of other researchers' data. A negative impact of perceived effort is expected on engineering researchers' attitudes toward data reuse and intention to reuse other researchers' data. Therefore, this study assumed the following hypotheses:

H3. Perceived effort involved in data reuse negatively influences engineering researchers' attitudes toward data reuse.

H4. Perceived effort involved in data reuse negatively influences engineering researchers' intention to reuse other researchers' data.

4.4 Norm of data reuse

A community norm of data reuse would affect an engineering researcher's attitudes toward reusing other researchers' data. Social norm constructs in the TPB can be applied to researchers' data reuse behaviours, as engineering researchers would be influenced by expectations from their colleague researchers (Ajzen, 1991; Fishbein and Ajzen, 1975). Recent studies in data sharing incorporated social norm construct into their operational models and found some empirical support (Kim and Stanton, 2016; Kim and Zhang, 2015). Hence, this study assumed that the community norm of data reuse would positively influence engineering researchers' attitudes toward data reuse.

H5. Norm of data reuse positively influences engineering researchers' attitudes toward data reuse.

4.5 Attitudes toward data reuse

According to the TPB, attitudes toward data reuse is a summary evaluation about data reuse behaviour, and the attitudes toward data reuse can be developed by their perceptions of data reuse behaviour and norm of data reuse (Ajzen and Fishbein, 2005). Based on the TPB, the researchers' perceptions were identified with three categories, namely, perceived usefulness, perceived concerns and perceived effort. According to the TPB, the attitudes toward data reuse would have a strong positive relationship with engineering researchers' intention to reuse other researchers' data (Ajzen, 1991; Fishbein and Ajzen, 1975). Therefore, this study assumed that the attitudes toward data reuse would positively influence engineering researchers' intention to reuse other researchers' data.

H6. Attitudes toward data reuse positively influence an engineering researcher's intention to reuse other researchers' data.

4.6 Availability of metadata standards

In addition to the aforementioned factors, there are other factors that directly or indirectly affect individual researchers' data reuse behaviours, such as the availability of metadata standards (Bowker and Star, 2000; Michener, 2006; Zimmerman, 2007) and data repositories (Cragin *et al.*, 2010; Fennema-Notestine, 2009; Marcial and Hemminger, 2010). These factors include resource factors. According to the TPB, resource facilitating conditions are important factors influencing people's behaviours. In this research, the availability of metadata standards is included, as it has been expected to influence engineering researchers' data reuse behaviours. Although there is no specific study focusing on the role of metadata standards in data reuse, prior studies in data sharing argued the importance of metadata standards in data sharing (Bowker and Star, 2000; Michener, 2006; Zimmerman, 2007). Thus, this study assumed that the availability of metadata standards would positively influence engineering researchers' intention to reuse other researchers' data.

H7. Availability of metadata standards positively influences an engineering researcher's intention to reuse other researchers' data.

4.7 Availability of data repositories

Similar to the metadata standards, this research also considered the availability of data repositories, as an important resource-facilitating condition, would positively influence engineering researchers' data reuse behaviours. Prior studies in data sharing and reuse pointed out the importance of data repositories as a supporting resource for data sharing and reuse (Cragin *et al.*, 2010; Fennema-Notestine, 2009; Marcial and Hemminger, 2010). Similar to the metadata standards, the availability of data repositories is included in this research, because it has been expected to influence engineering researchers' data reuse behaviours. Therefore, this study assumed that the availability of data repositories would positively influence engineering researchers' intention to reuse other researchers' data.

H8. Availability of data repositories positively influences an engineering researcher's intention to reuse other researchers' data.

Based on the theoretical framework presented above, a systematic model was developed which can account for engineering researchers' data reuse behaviours. Based on the TPB, attitudinal (attitudes toward data reuse), normative (norm of data reuse) and institutional (availabilities of metadata standards and data repositories) factors were considered, which mainly explain engineering researchers' data reuse behaviours. Based on the TAM, specific motivational factors were included, such as perceived usefulness, perceived concerns and perceived effort involved in data reuse behaviours (Figure 1).

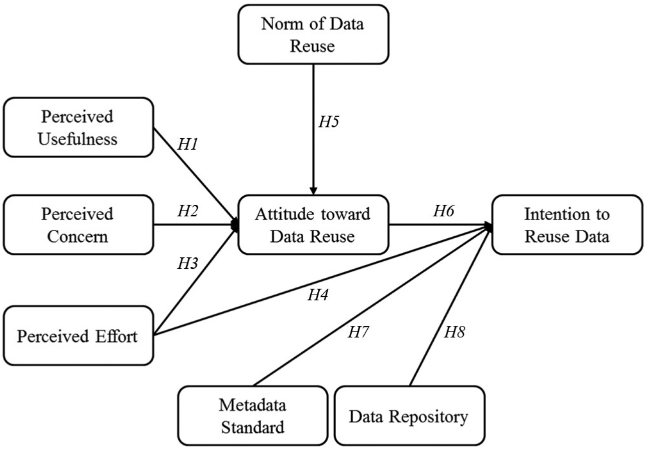

Figure 1. Engineering researchers' data reuse model

5. Research method

This research used a survey method, which has the positivist research perspective. Surveys are one of the most prevalent research paradigms in measuring people's perceptions of certain objects and behaviours, as well as their influences on people's intentions and behaviours (Ajzen and Fishbein, 2005). The survey method has been used to test research hypotheses, extend existing models, validate existing theories and build new theoretical models (Creswell, 2008). Furthermore, it is also most appropriate to investigate the relationships of variables and to predict or identify the level of one variable in comparison to another (Babbie, 1990). Thus, the use of a survey method can be a valuable academic approach in examining engineering researchers' attitudes about data reuse and their intention to reuse others' data.

A majority of measurement items were adapted from previous studies, including perceived usefulness (Davis, 1989; Davis *et al.*, 1989), perceived concerns (Lee, 2009; Littler and Melanthiou, 2006), perceived effort (Davis *et al.*, 1989; Thompson *et al.*, 1991), attitudes toward data reuse (Ajzen and Fishbein, 2005; Tohidinia and Mosakhani, 2010), norm of data reuse (Ajzen and Fishbein, 2005; Kostova and Roth, 2002), availabilities of metadata standards and data repositories (Thompson *et al.*, 1991; Venkatesh *et al.*, 2003) and data reuse intention (Ajzen and Fishbein, 2005; Tohidinia and Mosakhani, 2010). The measurement items for the research constructs are listed in Appendix 1. Based on the measures of constructs developed from previous studies, items for a survey questionnaire were created by adjusting the previous measures of the constructs for the present study's target group, researchers in engineering. A Likert scale (ranging from 1 for "strongly disagree" to 7 for "strongly agree") was used for all survey items.

Since the objective of this research is exploratory rather than confirmatory, this research used the partial least squares-based structural equation modelling (PLS-SEM) approach rather than covariance matrix-based SEM. The PLS-SEM approach is useful for exploratory studies with a small sample size (Goodhue *et al.*, 2012). The sample size required for the PLS-SEM is calculated as ten times of the largest indicators measured for one construct (Hair *et al.*, 2011; Ringle *et al.*, 2012). Based on the research model developed above, the number of the largest indicators measured for one construct is both attitude toward data reuse and intention to reuse data (three indicators for each construct and four path indicators toward each construct – seven indicators in total). Therefore, a total of 70 valid responses were required as a minimum sample size to appropriately evaluate the research model developed above. To better validate the research model, two to three times more samples (i.e. 140 to 210 responses) were expected with a minimum of a 10 per cent response rate.

The survey instruments included structured and predefined questions about engineering researchers' perceptions of data reuse, availabilities of metadata standards and data repositories and their intention to reuse other researchers' data. The survey was distributed through a self-monitored online survey from 5 October to 30 November 2015. The survey questionnaire is included in Appendix 2. A total of 2,129 potential survey participants were identified from the community of scientists' scholar database (http://pivot.cos.com). From this online survey, a total of 193 valid responses were received from engineering researchers with less than 5 per cent of missing values for the final data analysis. The response rate for this study was 9.07 per cent.

6. Data collection and analyses

The present study investigated to what extent the various types of factors listed above would influence individual researchers' reuse of other researchers' data. To answer the research questions, a total of 193 valid responses were analysed. In the following section, the data preparation procedure and data analysis methods are presented.

6.1 Demographics of the respondents

The respondents' demographics covered gender, age, ethnicity, education, position, status and discipline. Among the 193 survey participants from engineering, there were 160 male participants (82.9 per cent) and 29 female participants (15.0 per cent), while four participants (2.1 per cent) did not report their gender. In terms of age, the survey participants were well-distributed throughout the ranges of age including 25-34 (13, 6.7 per cent), 35-44 (51, 26.4 per cent), 45-54 (44, 22.8 per cent), 55-64 (51, 26.4 per cent) and 65 or more (30, 15.5 per cent). For ethnicity, a majority of participants were Caucasians (131, 67.9 per cent), and the rest of participants included Asians (30, 15.5 per cent), Hispanics (6, 3.1 per cent), African-Americans (3, 1.6 per cent), and other/multiracial participants (8, 4.1 per cent), while 13 participants (6.7 per cent) did not indicate their ethnicity. In terms of education, most of the participants had a PhD (182, 94.3 per cent). Regarding job position, participants listed themselves as full professor (80, 41.5 per cent), associate professor (47, 24.4 per cent), assistant professor (15, 7.8 per cent), professor emeritus (6, 3.1 per cent), professor of practice (1, 0.5 per cent), lecturer/instructor (3, 1.6 per cent), researcher (19, 9.8 per cent), postdoctoral fellow (6, 3.1 per cent) and other positions (9, 4.7 per cent), such as director, associate dean and research professor. The survey participants' demographics are summarized in Table I.

In terms of academic disciplines in the engineering field, the respondents were from diverse engineering disciplines, including aerospace engineering (12, 6.2 per cent), agricultural engineering (18, 9.3 per cent), biomedical engineering (9, 4.7 per cent), chemical engineering (15, 7.8 per cent), civil engineering (28, 14.5 per cent), computer engineering (7, 3.6 per cent), electrical engineering (16, 8.3 per cent), environmental engineering (21, 10.9 per cent), industrial engineering (12, 6.2 per cent), mechanical engineering (23, 11.9 per cent), material engineering (9, 4.7 per cent), nuclear engineering (1, 0.5 per cent) and other engineering disciplines (22, 11.4 per cent). The summary of survey participants' academic disciplines in engineering is shown in Table II.

6.2 Scale assessment

Before the actual data analysis was conducted, scale assessment had been performed to ensure the reliability of the measurement scales. Cronbach's alpha (Cronbach's α) and

Demographic category	No.	(%)
Gender		
Male	160	82.9
Female	29	15.0
Missing	4	2.1
Age		
25-34	13	6.7
35-44	51	26.4
45-54	44	22.8
55-64	51	26.4
65+	30	15.5
Missing	4	2.1
Ethnic		
Asian/Pacific Islander	30	15.5
Black/African-American	3	1.6
Caucasian	131	67.9
Hispanic	6	3.1
Native American/Alaska Native	2	1.0
Other/multiracial	8	4.1
Missing	13	6.7
Education		
Bachelor's degree	1	0.5
Master's degree	10	5.2
PhD/Doctoral degree	182	94.3
Status		
Tenured	128	66.3
On tenure track	13	6.7
Not on tenure track	38	19.7
Retired	10	5.2
Missing	4	2.1
Position		
Assistant professor	15	7.8
Associate professor	47	24.4
Full professor	80	41.5
Professor emeritus	6	3.1
Professor of practice	1	0.5
Lecturer/instructor	3	1.6
Postdoctoral fellow	6	3.1
Researcher	19	9.8
Graduate student	6	3.1
Other	9	4.7
Missing	1	0.5
Total	193	100

Table I. Demographics of respondents

composite reliability (CR) were used to confirm that each research construct was reliable to conduct further data analyses. The scale assessment results revealed that the values of Cronbach's α, ranging from 0.71 (perceived effort and attitudes toward data reuse) to 0.94 (intention to reuse other's data), were under the acceptable value of 0.70 (Nunnally and Bernstein, 1994), and composited reliability values, ranging from 0.69 (perceived effort) to

Table II. Discipline information of respondents

Disciplines	Frequency	(%)
Aerospace engineering	12	6.2
Agricultural engineering	18	9.3
Biomedical engineering	9	4.7
Chemical engineering	15	7.8
Civil engineering	28	14.5
Computer engineering	7	3.6
Electrical engineering	16	8.3
Environmental engineering	21	10.9
Industrial/manufacturing engineering	12	6.2
Mechanical engineering	23	11.9
Metallurgical and materials engineering	9	4.7
Nuclear engineering	1	0.5
Engineering, other	22	11.4
Total	193	100

0.96 (availability of data repositories and intention to reuse other's data), were also under the acceptable value of 0.60 (Chin, 1998). The reliability and validity values of this study are presented in Table III.

6.3 Non-response analysis

As the response rate is low for this small sample study, non-response analysis was conducted as suggested by Babbie (1990) by comparing early and late responses (considering the late responses as a proxy for non-responses). The first 20 per cent of responses (i.e. those who took the survey immediately when the first e-mail was sent out) were compared to the last 20 per cent of responses (i.e. those who took the survey at the very end after the last e-mail was sent out), which was 39 cases for each group. Then, an ANOVA test was conducted to compare the mean differences between those two groups.

The ANOVA test indicates no statistically significant differences between the first and last groups of survey respondents for explanatory and outcome variables including perceived usefulness ($F = 0.69$, $p = 0.41$), perceived concerns ($F = 0.06$, $p = 0.81$), perceived effort ($F = 0.00$, $p = 0.99$), norm of data reuse ($F = 0.02$, $p = 0.89$), attitudes toward data reuse ($F = 2.59$, $p = 0.11$), availability of metadata ($F = 0.62$, $p = 0.43$), availability of data repositories ($F = 0.37$, $p = 0.54$) and intention to reuse data ($F = 0.81$, $p = 0.37$). Therefore, the non-response bias was considered to be marginal in this research; so this research did not use any weighting method.

Table III. Reliability and validity values

Variables	Cronbach's α	Composite reliability	Average variance extracted
Perceived usefulness	0.82	0.89	0.74
Perceived concerns	0.78	0.87	0.70
Perceived effort	0.71	0.69	0.46
Norm of data reuse	0.91	0.94	0.85
Attitudes toward data reuse	0.71	0.84	0.64
Availability of metadata	0.92	0.96	0.93
Availability of data repositories	0.92	0.96	0.93
Intention to reuse data	0.94	0.96	0.89

7. Data analysis results

Based on the research model, the eight hypotheses were tested regarding the factors influencing engineering researchers' data reuse behaviours. For the data analyses, the partial least squares-based structural equation modelling (PLS-SEM) technique was used. The SmartPLS software was used for the PLS-SEM analysis (Ringle *et al.*, 2005).

7.1 Measurement model

The measurement model was evaluated first to check the convergent and discriminant validity of the constructs. The square root of each construct's average variance extracted (AVE) (shown in bold font in Table IV) was larger than the inter-construct correlations, presenting reliable convergent and discriminant validity. The results of the measurement model showed that the survey measurements were reliable and valid for structural model evaluation. Table IV shows the square roots of AVEs and correlation matrix.

7.2 Structural model

As a final step of the data analyses, the structural model was evaluated by using a structural equation modelling statistical technique. The data analysis showed that individual motivations, norm and institutional resource factors significantly influence engineering researchers' intention to reuse other researchers' data either directly or indirectly. In terms of individual motivations, perceived usefulness in data reuse ($\beta = 0.662, p < 0.001$) was found to have a significant positive relationship with attitudes toward data reuse, and perceived concerns about data reuse were found to have a significant negative relationship with engineering researchers' attitudes toward data reuse ($\beta = -0.116, p < 0.05$). However, perceived effort involved in data reuse was not found to have any significant relationships with engineering researchers' attitudes toward data reuse ($\beta = 0.101, p > 0.05$) or intention to reuse other's data ($\beta = -0.107, p > 0.05$).

Research construct	Perceived usefulness	Perceived concerns	Perceived effort	Norm of data reuse	Attitudes toward data reuse	Availability of data repositories	Availability of metadata	Intention to reuse data
Perceived usefulness	*0.86*							
Perceived concerns	−0.34	*0.83*						
Perceived effort	−0.11	0.39	*0.68*					
Norm of data reuse	0.23	−0.21	−0.20	*0.92*				
Attitudes toward data reuse	0.74	−0.35	−0.06	0.37	*0.80*			
Availability of metadata	0.13	−0.13	−0.22	0.41	0.13	*0.96*		
Availability of data repositories	0.18	−0.10	−0.20	0.47	0.23	0.65	*0.96*	
Intention to reuse data	0.67	−0.27	−0.18	0.39	0.65	0.20	0.32	*0.94*

Table IV. Square roots of AVEs and correlation matrix

In addition, the community norm of data reuse was found to have a significant positive influence on the attitudes toward data reuse ($\beta = 0.216$, $p < 0.001$). The three variables including perceived usefulness, perceived concerns and norm of data reuse accounted for 60.4 per cent of total variance in attitudes toward data reuse ($R^2 = 0.604$), which is a high rate of explanation power. Consequently, engineering researchers' attitudes toward data reuse were found to have a significantly positive influence on researchers' intention to reuse others' data ($\beta = 0.605$, $p < 0.001$).

In terms of resource factors, this research examined how the availability of metadata standards and data repositories influences engineering researchers' intentions to reuse other researchers' data. The availability of data repositories was found to have a significant positive influence on the intention to reuse data ($\beta = 0.155$, $p < 0.05$); however, the availability of metadata standards was not found to have a significant relationship with the intention to reuse data ($\beta = 0.001$, $p > 0.05$).

Finally, both attitudes toward data reuse and the availability of data repositories accounted for 46.0 per cent of the variance of the intention to reuse data ($R^2 = 0.460$). This shows that attitudes toward data reuse and the availability of data repositories can provide a powerful explanation on intentions to reuse data. Figure 2 presents the results of hypotheses testing based on the engineering researchers' data reuse model developed above.

8. Discussion

The objective of this research was to explore factors that can influence data reuse behaviours of engineering researchers and identify factors that facilitate or hinder data reuse behaviours in engineering. The findings of this research demonstrated that engineering researchers' data reuse behaviours are influenced by researchers' motivational factors, norm of data reuse and an institutional resource factor. Specifically, this study revealed two motivational factors influencing engineering researchers' data reuse, namely perceived usefulness and perceived concerns involved in data reuse. The current study also revealed that a norm of data reuse based on their peers' expectations (norm) and the availability of data repositories (an institutional resource factor) affect data reuse among engineering researchers.

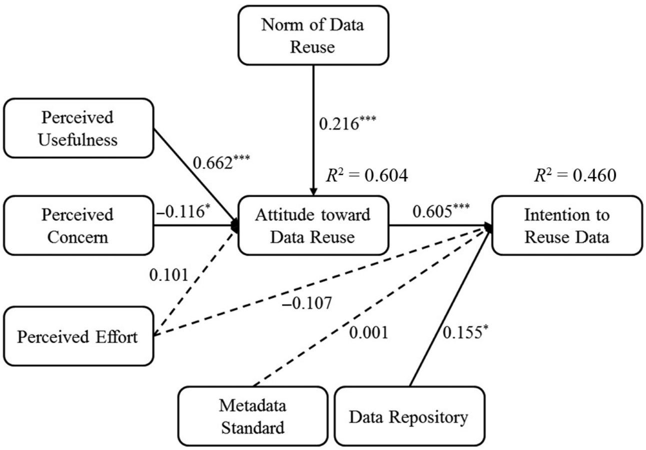

Figure 2.
Results of hypotheses testing based on engineering researchers' data reuse model

8.1 Individual motivations

As mentioned previously, the current research showed that individual motivations, perceived usefulness and perceived concerns of data reuse influence engineering researchers' data reuse by shaping their attitudes toward data reuse. The findings suggest that data reusers in engineering research are motivated by their belief that secondary data analyses would increase efficiency in their research. Furthermore, the current study indicates that if data reuse provokes concerns owing to potential copyright infringements or decreased publication opportunities, engineering researchers would formulate negative attitudes toward data reuse and would not bother to reuse data. Given the findings of the current study, promoting data reuse in the field of engineering is dependent on how the academic community highlights the benefits of data reuse and resolves issues and problems related to data reuse.

It is also noteworthy that the norm of data reuse had a significant influence on engineering researchers' attitudes toward data reuse. The results of this research demonstrated that engineering researchers' norm of data reuse can increase their intention to reuse data by improving their attitudes toward data reuse. That is, engineering researchers care about their colleagues' expectations about data reuse, and as their colleagues have a more positive norm of data reuse, they are more likely to conduct research with a pre-existing data set. This result suggests that it is important to develop positive norms of data reuse to better facilitate engineering researchers' data reuse behaviours. Providing workshops about data reuse in major conferences or in-class lectures about data use can be an effective intervention to educate community members about the norm of data reuse.

It should also be mentioned that the perceived effort involved in data reuse was not a predictor of researchers' attitudes toward data reuse or intention to reuse data. This suggests that the nature of engineering data can be different from social science research, where secondary data analyses require a great deal of effort (Faniel *et al.*, 2016; Niu, 2009; Yoon, 2016). As engineering data tend to be more systematic and structured, compared to data in social science, the expected effort in data reuse might not have affected engineering researchers' attitudes and intention to a significant extent.

8.2 Institutional resources

With regard to institutional resources, this research found that only the availability of data repositories had a significant and positive influence on engineering researchers' intention to use secondary data. This means that engineering researchers would actively consider reusing other researchers' data if appropriate data repositories were available for them. This result suggests that it is important to establish data repositories that are easily accessible to promote data reuse among engineering researchers.

It should be mentioned that, in the current research, the availability of metadata standards was not a significant predictor of researchers' intention to reuse data. This may be explained by the characteristics of the engineering community. Unlike other fields in which metadata are frequently used and, therefore, a widely accepted standard of metadata exists, the engineering research community has just started to use metadata and does not have a strong standard of metadata (Michener, 2006). However, to confirm this inference, future studies should investigate in more detail how metadata standards would influence intention to reuse data among engineering researchers.

9. Implications and limitations

9.1 Theoretical implications

The current study has important theoretical contributions. First, the current study elaborated a model of data reuse by applying both the TPB and the technology acceptance model (TAM). As a theoretical framework of the TPB basically focused on general explanations of engineering researchers' data reuse behaviours and, therefore, cannot fully describe those behaviours, this research extended the theoretical framework by using the TAM, which provides a more concrete understanding of engineering researchers' perceptions involved in data reuse. The combination of the TPB and the TAM more precisely explained engineering researchers' data reuse behaviours by addressing individual motivations, norms and institutional factors.

Based on the integrated theoretical framework, this research demonstrated that perceived usefulness, perceived concerns, norm of data reuse and the availability of data repositories are important factors influencing engineering researchers' data reuse behaviours. Among these factors, the current research found that perceived usefulness is the most significant factor influencing engineering researchers' attitudes toward reuse behaviours. Moreover, the current research revealed that engineering researchers' intention to reuse data is significantly influenced by engineering researchers' overall attitudes toward data reuse and the availability of data repositories.

9.2 Practical implications

This research also has significant practical implications for promoting more reliable and beneficial data reuse in the engineering community. As the field of engineering is becoming more global and more transdisciplinary, engineering researchers in the current era are facing more needs to understand societal issues, which often requires them to conduct research based on data sharing or data reuse. The findings of this research indicate that to facilitate engineering researchers' data reuse behaviours, it is important to provide assistance to researchers who are planning to use secondary data. Specifically, it would be beneficial if such assistance could increase benefits of data reuse, reduce researchers' concerns involving data reuse and contribute to establishing a positive norm about data reuse. In addition, it is also important to set up more data repositories available to engineering researchers.

9.3 Limitations

A few limitations of the current research should be mentioned. First, this research only focused on the engineering discipline, so it is limited from being generalized to other scientific disciplines. Future research needs to extend the scope of disciplines to other academic disciplines including biological sciences, health sciences, social sciences and more. Furthermore, future research needs to investigate how engineering researchers' data reuse behaviours differ across different subfields of engineering and what causes those differences across different engineering sub-fields. Second, this research measured engineering researchers' data reuse intentions rather than their actual data reuse behaviours. The data reuse intention may not exactly reflect researchers' actual data reuse behaviours. Therefore, future research should not only include intention to reuse data but also actual data reuse behaviours in the model. Third, a survey was implemented to investigate researchers' data reuse behaviours. Although the self-administered survey method is one of the most frequently used research techniques, results from a survey can be limited as it may not be able to present a very exhaustive description of engineering researchers' data reuse behaviours. Therefore, future studies should consider using both quantitative and qualitative methods to better triangulate engineering researchers' data reuse behaviours. Furthermore, the response rate to the survey was relatively low. Although non-response analysis showed no significant differences between the first group of respondents and the last group of respondents as the proxy of non-respondents, the low response rate can raise issues of research validity or generalizability of research findings. Therefore, future research will need to find a way to improve the response rate.

10. Conclusion

This research demonstrated that engineering researchers' data reuse behaviours are affected by their individual motivations, norms of data reuse and institutional resources (i.e. the availability of data repositories). With respect to the individual motivations, this research showed that perceived usefulness of data reuse and concerns about data reuse are significant predictors of attitudes toward data reuse among engineering researchers. With respect to institutional resources, only the availability of data repositories has a significant and positive influence on researchers' intention to use secondary data. The results of this research would be useful to facilitate data reuse among engineering researchers as the study guides where and how resources for data reuse should be planned and spent. Such endeavours can improve the practice of engineering researchers' data reuse behaviours and, eventually, it will enhance the quality of engineering research.

References

Ajzen, I. (1991), "The theory of planned behavior", *Organizational Behavior and Human Decision Process*, Vol. 50 No. 2, pp. 179-211.

Ajzen, I. and Fishbein, M. (2005), "The influence of attitudes on behavior", in Albarracin, D., Johnson, B. T. and Zanna, M.P. (Eds), *Handbook of Attitudes and Attitude Change*, Lawrence Erlbaum Associates, Mahwah, NJ.

Babbie, E.R. (1990), *Survey Research Methods*, Wadsworth Publishing, Belmont, CA.

Baker, K.S. and Yarmey, L. (2009), "Data stewardship: environmental data curation and a web-of-repositories", *International Journal of Digital Curation*, Vol. 4 No. 4, pp. 12-27.

Benlian, A. and Hess, T. (2011), "Opportunities and risks of software-as-a-service: findings from a survey of IT executives", *Decision Support Systems*, Vol. 52 No. 1, pp. 232-246.

Berg, M. and Goorman, E. (1999), "The contextual nature of medical information", *International Journal of Medical Informatics*, Vol. 56 Nos 1/3, pp. 51-60.

Birnholtz, J.P. and Bietz, M.J. (2003), "Data at work: supporting sharing in science and engineering" *ACM Conference on Supporting Group Work, Sanibel Island, FL*, pp. 339-348.

Borgman, C.L. (2012), "The conundrum of sharing research data", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 6, pp. 1059-1078.

Borgman, C.L. (2007), *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*, MIT Press, Cambridge, MA.

Bowker, G. and Star, S.L. (2000), *Sorting Things out*, MIT Press, Cambridge, MA.

Buede, D.M. and Miller, W.D. (2016), *The Engineering Design of Systems: Models and Methods*, John Wiley & Sons, Hoboken, NJ.

Carlson, S. and Anderson, B. (2007), "What are data? the many kinds of data and their implications for data re-use", *Journal of Computer-Mediated Communication*, Vol. 12 No. 2, pp. 635-651.

Cheng, J.H. (2011), "Inter-organizational relationships and knowledge sharing in green supply chains: moderating by relational benefits and Guanxi", *Transportation Research Part E-Logistics and Transportation Review*, Vol. 47 No. 6, pp. 837-849.

Chin, G., Jr and Lansing, C.S. (2004), "Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory" *ACM Conference on Computer Supported Cooperative Work*, *Chicago, IL*, pp. 409-418.

Chin, W.W. (1998), "The partial least squares approach to structural equation modeling", *Modern Methods for Business Research*, Vol. 295 No. 2, pp. 295-336.

Cragin, M.H. and Shankar, K. (2006), "Scientific data collections and distributed collective practice", *Computer Supported Cooperative Work (Cscw)*, Vol. 15 Nos 2/3, pp. 185-204.

Cragin, M.H., Palmer, C.L., Carlson, J.R. and Witt, M. (2010), "Data sharing, small science and institutional repositories", *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, Vol. 368 No. 1926, pp. 4023-4038.

Creswell, J.W. (2008), *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*, Sage Publications, Thousand Oaks, CA.

Davis, F.D. (1989), "Perceived usefulness, perceived ease of use and user acceptance in information technology", *MIS Quarterly*, Vol. 13 No. 3, pp. 319-340.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989), "User acceptance of computer technology: a comparison of two theoretical models", *Management Science*, Vol. 35 No. 8, pp. 982-1003.

Davis, H.M. and Vickery, J.N. (2007), "Datasets, a shift in the currency of scholarly communication: implications for library collections and acquisitions", *Serials Review*, Vol. 33 No. 1, pp. 26-32.

Downey, G. (2005), "Are engineers losing control of technology?: from 'problem solving' to 'problem definition and solution' in engineering education", *Chemical Engineering Research and Design*, Vol. 83 No. 6, pp. 583-595.

Faniel, I.M. and Jacobsen, T.E. (2010), "Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data", *Computer Supported Cooperative Work (Cscw)*, Vol. 19 No. Nos 3/4, pp. 355-375.

Faniel, I.M., Kriesberg, A. and Yakel, E. (2012), "Data reuse and sensemaking among novice social scientists", *Proceedings of the American Society for Information Science and Technology*, Vol. 49 No. 1, pp. 1-10.

Faniel, I.M., Kriesberg, A. and Yakel, E. (2016), "Social scientists' satisfaction with data reuse", *Journal of the Association for Information Science and Technology*, Vol. 67 No. 6, pp. 1404-1416.

Faniel, I.M., Barrera-Gomez, J., Kriesberg, A. and Yakel, E. (2013a), "A comparative study of data reuse among quantitative social scientists and archaeologists", *iConference*, *Fort Worth, TX, iSchools*, pp. 797-800.

Faniel, I.M., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J. and Yakel, E. (2013b), "The challenges of digging data: a study of context in archaeological data reuse", *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, ACM, *New York, NY*, pp. 295-304.

Fennema-Notestine, C. (2009), "Enabling public data sharing: Encouraging scientific discovery and education", in Astakhov, V. (Ed.), *Biomedical Informatics*, Humana Press, Totowa, NJ, pp. 25-32.

Fishbein, M. and Ajzen, I. (1975), *Belief, Attitude, Intention and Behavior*, Addison-Wesley, Reading, MA.

Goodhue, D.L., Lewis, W. and Thompson, R. (2012), "Does PLS have advantages for small sample size or non-normal data?", *MIS Quarterly*, Vol. 36 No. 3, pp. 891-1001.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), "PLS-SEM: indeed a silver bullet", *Journal of Marketing Theory and Practice*, Vol. 19 No. 2, pp. 139-152.

Jirotka, M., Procter, R., Hartswood, M., Slack, R., Simpson, A., Coopmans, C., Hinds, C. and Voss, A. (2005), "Collaboration and trust in healthcare innovation: the eDiaMoND case study", *Computer Supported Cooperative Work (Cscw)*, Vol. 14 No. 4, pp. 369-398.

Kim, Y. and Stanton, J.M. (2016), "Institutional and individual factors affecting scientists' data-sharing behaviors: a multilevel analysis", *Journal of the Association for Information Science and Technology*, Vol. 67 No. 4, pp. 776-799.

Kim, Y. and Zhang, P. (2015), "Understanding data sharing behaviors of STEM researchers: the roles of attitudes, norms and data repositories", *Library and Information Science Research*, Vol. 37 No. 3, pp. 189-200.

Kostova, T. and Roth, K. (2002), "Adoption of an organizational practice by subsidiaries of multinational corporations: institutional and relational effects", *Academy of Management Journal*, Vol. 45 No. 1, pp. 215-233.

Lee, M.C. (2009), "Predicting and explaining the adoption of online trading: an empirical study in Taiwan", *Decision Support Systems*, Vol. 47 No. 2, pp. 133-142.

Littler, D. and Melanthiou, D. (2006), "Consumer perceptions of risk and uncertainty and the implications for behaviour towards innovative retail services: the case of internet banking", *Journal of Retailing and Consumer Services*, Vol. 13 No. 6, pp. 431-443.

McCall, R.B. and Appelbaum, M.I. (1991), "Some issues of conducting secondary analyses", *Developmental Psychology*, Vol. 27 No. 6, pp. 911-917.

Marcial, L.H. and Hemminger, B.M. (2010), "Scientific data repositories on the web: an initial survey", *Journal of the American Society for Information Science and Technology*, Vol. 61 No. 10, pp. 2029-2048.

Michener, W.K. (2006), "Meta-information concepts for ecological data management", *Ecological Informatics*, Vol. 1 No. 1, pp. 3-7.

Niu, J. (2009), *Overcoming Inadequate Documentation*, American Society for Information Science and Technology, Vancouver, British Columbia, pp. 1-14.

Nunnally, J.C. and Bernstein, I.H. (1994), *Psychometric Theory*, 3rd ed., McGraw-Hill, New York, NY.

Piwowar, H.A. and Vision, T.J. (2013), "Data reuse and the open data citation advantage", *PeerJ*, Vol. 1, pp. 1-25.

Ringle, C.M., Sarstedt, M. and Straub, D. (2012), "A critical look at the use of PLS-SEM in MIS Quarterly", *MIS Quarterly*, Vol. 36 No. 1, pp. 3-14.

Ringle, C.M. Wende, S. and Will, S. (2005), "SmartPLS 2.0 (M3) Beta", available at: www.smartpls.de (accessed 11 January 2017).

Rolland, B. and Lee, C.P. (2013), "Beyond trust and reliability: Reusing data in collaborative cancer epidemiology research" *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW), San Antonio, TX, ACM, New York, NY*, pp. 435-444.

Scaffidi, C., Shaw, M. and Myers, B. (2006), "Games programs play: obstacles to data reuse", *Paper presented at the 2nd Workshop on End User Software Engineering (WEUSE) at the Conference on Human Factors in Computing Systems (CHI), Montreal, Canada*, 22-27 April.

Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. and Dorsett, K. (2015), "Changes in data sharing and data reuse practices and perceptions among scientists worldwide", *PloS One*, Vol. 10 No. 8, pp. e0134826

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., Manoff, M. and Frame, M. (2011), "Data sharing by scientists: practices and perceptions", *PloS One*, Vol. 6 No. 6, pp. e21101

Thompson, R.L., Higgins, C.A. and Howell, J.M. (1991), "Personal computing: toward a conceptual model of utilization", *MIS Quarterly: Management Information Systems*, Vol. 15 No. 1, pp. 125-142.

Tohidinia, Z. and Mosakhani, M. (2010), "Knowledge sharing behaviour and its predictors", *Industrial Management & Data Systems*, Vol. 110 No. 3/4, pp. 611-631.

Van House, N.A., Butler, M.H. and Schiff, L.R. (1998), "Cooperative knowledge work and practices of trust: Sharing environmental planning data sets", *Paper presented at the The ACM Conference on Computer Supported Cooperative Work, Seattle, WA*, pp. 335-343.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology: toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.

Vickers, A.J. (2006), "Whose data set is it anyway? Sharing raw data from randomized trials", *Trials*, Vol. 7 No. 1, pp. 15.

Yoon, A. (2015), "Data reuse and users' trust judgments: toward trusted data curation", Ph.D. dissertation, University of North Carolina, Chapel Hill, NC.

Yoon, A. (2016), "Data Reusers' trust development", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 4, pp. 946-956.

Zimmerman, A. (2007), "Not by metadata alone: the use of diverse forms of knowledge to locate data for reuse", *International Journal on Digital Libraries*, Vol. 7 No. 1/2, pp. 5-16.

Zimmerman, A.S. (2008), "New knowledge from old data: the role of standards in the sharing and reuse of ecological data", *Science Technology and Human Values*, Vol. 33 No. 5, pp. 631-652.

Appendix 1

Construct	Items	Sources
Perceived usefulness	Reusing other researchers' data improves the quality of my research Reusing other researchers' data enhances the effectiveness of my research Reusing other researchers' data reduces the time/cost/effort I spend on my research	Davis (1989), Davis *et al.* (1989)
Perceived concern	If I reuse other researchers' data, I worry that I might misinterpret the data If I reuse other researchers' data, I worry that I might cause infringement If I reuse other researchers' data, I worry that I might not publish with that data	Lee (2009), Littler and Melanthiou (2006)
Perceived effort	Reusing other researchers' data requires time and effort to locate data sets Reusing other researchers' data requires time and effort to access (or get permission to use) data sets Reusing other researchers' data requires time and effort to process data sets for a new study	Davis *et al.* (1989), Thompson *et al.* (1991)
Attitude toward data reuse	Reusing other researchers' data is valuable Reusing other researchers' data is desirable Reusing other researchers' data is pleasant	Ajzen and Fishbein (2005), Tohidinia and Mosakhani (2010)
Norm of data reuse	In my discipline, it is expected that researchers could reuse other researchers' data In my discipline, many of researchers currently reuse data In my discipline, reusing other researchers' data is a common practice	Ajzen and Fishbein (2005), Kostova and Roth (2002)
Availability of data repository	In my discipline, data repositories are available for researchers to share data In my discipline, researchers can easily access data repositories to reuse data	Thompson *et al.* (1991), Venkatesh *et al.* (2003)
Metadata standards	In my discipline, metadata is available for researchers to share data In my discipline, researchers can easily use metadata to reuse data	Thompson *et al.* (1991), Venkatesh *et al.* (2003)
Data reuse intention	I am likely to reuse other researchers' data for my future research I intend to reuse other researchers' data for my future research I will try to reuse other researchers' data for my future research	Ajzen and Fishbein (2005), Tohidinia and Mosakhani (2010)

Table AI. Measurement items for research constructs

Appendix 2. Survey Questionnaire

ABOUT YOUR DISCIPLINE

What is your primary and specific research disciplines?

DATA REUSE PERCEPTION

Please indicate to what extent you agree with the following statements. For validation reasons, we may have to ask similar questions. If you cannot answer any question(s) or they are not applicable to you, please leave them blank.

In this survey, data reuse refers to using other researchers' data for another research purpose (e.g., replicating a study, c omparing to another study, conducting meta-analysis, or asking a different research question, etc.).

[Usefulness]

Reusing other researchers' data:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Improves the quality of my research	1	2	3	4	5
Enhances the effectiveness of my research	1	2	3	4	5
Reduces the time/cost/effort I spend on my research	1	2	3	4	5

[Concerns]

If I reuse other researchers' data I worry that:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I might misinterpret the data	1	2	3	4	5
I might cause infringement	1	2	3	4	5
I might not publish with that data	1	2	3	4	5

[Efforts]

Reusing other researchers' data requires time and effort:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
To locate data sets	1	2	3	4	5
To access (or get permission to use) data sets	1	2	3	4	5
To process data sets for a new study	1	2	3	4	5

DISCIPLINARY CONTEXTS

[Social Norms]

In my discipline:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
It is expected that researchers could reuse other researchers' data.	1	2	3	4	5
Many of researchers currently reuse data.	1	2	3	4	5
Reusing other researchers' data is a common practice.	1	2	3	4	5

[Data Repositories]

In my discipline:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Data repositories are available for researchers to share data.	1	2	3	4	5
Researchers can easily access data repositories to reuse data.	1	2	3	4	5

[Metadata] *Note: Metadata is a set of data that provides information about one or more aspects of the original research data (e.g. Ecological Metadata Language).

In my discipline:	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Metadata is available for researchers to share data.	1	2	3	4	5
Researchers can easily utilize metadata to reuse data.	1	2	3	4	5

[Intention to Reuse Other Researchers' Data]

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I am likely to reuse other researchers' data for my future research.	1	2	3	4	5
I intend to reuse other researchers' data for my future research.	1	2	3	4	5
I will try to reuse other researchers' data for my future research.	1	2	3	4	5

(*continued*)

DEMOGRAPHIC INFORMATION

How many years have you worked in your current research field?
a) Less than 5 b) 6-10 c) 11-15 d) 16-20 e) 21-25 f) 26-30 g) More than 30

What is your age?

a) Under 24	b) 25-34	c) 35-44
d) 45-54	e) 55-64	f) 65+

What is your gender?

a) Male	b) Female

What is your ethnic background?

a) Asian/Pacific Islander	b) Black/African-American	c) Caucasian
d) Hispanic	e) Native American	f) Other/Multi-Racial

What is your highest education so far?
a) Associate Degree b) Bachelor's Degree c) Master's Degree d) PhD/Doctoral Degree

What is your current position?

a) Assistant Professor	b) Associate Professor	c) Full Professor	d) Professor Emeritus
e) Professor of Practice	f) Lecturer/Instructor	g) Post-Doctoral Fellow	h) Researcher
i) Graduate Student	j) Other (Specify)		

Please choose the option most applicable to you.

a) Tenured	b) On Tenure Track But Not Tenured	c) Not on Tenure Track	d) Retired

About the authors

Yeon Kyoung Joo is an Assistant Professor in the Department of Digital Media at Myongji University. She received her PhD degree in communication from Stanford University and the BS and MS degrees in communication from Seoul National University. Her research interests include technologies in scholarly communication, human-computer interaction and methodology in information science. Dr Joo is the recipient of the scholarship of Korea Foundation for Advanced Studies (2007-2012).

Youngseek Kim is an Assistant Professor in the School of Information Science at the University of Kentucky. He completed his PhD at Syracuse University, and he received the Eugene Garfield Doctoral Dissertation Award from the Association for Library and Information Science Education and Emerald/EFMD Outstanding Doctoral Research Award (highly commended) in Information Science. His current research efforts in eScience try to understand the nature of scientists' data practices, focusing on data sharing and reuse, the education of eScience professionals and adoption and use of cyberinfrastructure. He published articles in *Journal of Association for Information Science and Technology*, *Journal of Information Science*, *Library and Information Science Journal* and *International Journal of Information Management*. Youngseek Kim is the corresponding author and can be contacted at: youngseek.kim@uky.edu