



A Review of Research Methodologies Employed in Serendipity Studies in the Context of Information Research

Xu Sun¹, Xiaosong Zhou², Qingfeng Wang^{3,*}, Sarah Sharples⁴, Yaorun Zhang⁵

¹Faculty of Science and Engineering, University of Nottingham, Ningbo, China,
xu.sun@nottingham.edu.cn

²School of Mechanical and Electrical Engineering, Guizhou Normal University, Guizhou, China,
ivzhouxiaosong@163.com

³University of Nottingham Business School Ningbo China, University of Nottingham, Ningbo, China,
Qingfeng.wang@nottingham.edu.cn

⁴Human Factors Group, University of Nottingham, Nottingham, UK,
Sarah.Sharples@nottingham.ac.uk

⁵Faculty of Science and Engineering, University of Nottingham, Ningbo, China,
yaorun_zhang@yahoo.com

Abstract

Background: *The concept of serendipity has become increasingly interesting for those undertaking serendipity research in recent years. However, serendipitous encounters are subjective and rare in a real-world context, making this an extremely challenging subject to study.*

Methods: *Various methods have been proposed to enable researchers to understand and measure serendipity, but there is no broad consensus on which methods to use in different experimental settings. A comprehensive literature review was first conducted, which summarizes the research methods being employed to study serendipity. It was followed by a series of interviews with experts that specified the relative strengths and weaknesses of each method identified in the literature review, in addition to the challenges usually confronted in serendipity research.*

Results: *The findings suggest using mixed research methods to produce a more complete picture of serendipity and contribute to the verification of any research findings. Several challenges and implications relating to empirical studies in the investigation of serendipity have been derived from this study.*

Conclusions: *This paper investigated research methods employed to study serendipity by synthesizing finding from a literature review and the interviews with experts. It provides a methodological contribution to serendipity studies by systematically summarizing the methods employed in the studies of serendipity and identifying the strengths and weakness of each method. It also suggests the novel approach of using mixed research methods to study serendipity. This study has potential limitations related to a small number of experts involved in the expert interview. However, it should be noted that the nature of the topic is a relatively focused area, and it was observed after interviewing the experts that new data seems to not contribute to the findings owing to its repetition of comments.*

Keywords: Serendipity, Information Research, Research Methodology.

Citation: Sun, X., Zhou, X., Wang, Q., Sharples, S., & Zhang, Y. (2021). A Review of Research Methodologies Employed in Serendipity Studies in the Context of Information Research. *Pacific Asia Journal of the Association for Information Systems*, 13(2), 72-90. <https://doi.org/10.17705/1pais.13204>

Copyright © Association for Information Systems

Introduction

Serendipity studies have suggested that although serendipity is a rare phenomenon, it has become more common in today's information societies, even though we do not always recognize it as such (e.g., Erdelez, 2004; Makri & Blandford, 2012a, 2012b). Serendipity research has garnered an increasing amount of attention in information science, and in recent decades, numerous researchers have studied the phenomenon in the context of information behavior research. Makri and Blandford (2012a, 2012b) unearthed examples of having come across information serendipitously in their research or everyday life. Agarwal (2015) noted that serendipity might occur in purposeful/active or non-purposeful/passive information searches.

Despite its importance, the concept of serendipity remains challenging to define. The word serendipity was originally coined by Horace Walpole, who defined it as the making of "discoveries, by accidents and sagacity, of things which you were not in quest of"¹. To date, there remains a lack of consensus on a formal definition of serendipity (Zhou et al., 2018). In information science, the term "serendipity" is an abstract concept that contains the meaning of accidental discovery of information. New "technical" terms such as information encountering (Erdelez, 2004), information source encountering (Miwa, 2000), and serendipitous information retrieval (Toms, 2000), have been proposed to improve the preciseness of the popularly used term "serendipity". Scholars have tried to define serendipity from different perspectives. In science, serendipity has been considered an essential tool to aid the discovery process, and one for which preparation is perceived as possible, to some extent (Rosenman, 1988). Seifert et al. (1994) suggested that creativity originates in a prepared mind that enables subsequent recognition of the serendipity when it is encountered. However, serendipity in the humanities may also have a role in revealing hidden connections, enabling creative connections to develop (Cory, 1999). In recommender systems, Kotkov et al. (2016) suggested that serendipitous items are those items that are relevant, novel, and unexpected to users. Serendipity is an important concept across disciplinary areas for its role in discovery, creativity and connection building (Foster & Ford, 2003). While in the context of information research, Andel's (1994) description of serendipity suggested that it is "the art of making an unsought finding". McCay-Peet and Toms (2015) defined it as "an unexpected experience prompted by an individual's valuable interaction with ideas, information, objects, or phenomena" (p. 392). The unexpectedness, novelty, and relevance are widely accepted as the key dimensions that form the contours of serendipity (e.g., Dantonio et al., 2012; Maksai et al., 2015; Zhou et al., 2017).

There are various degrees of serendipitous encounters, representing a continuum covering the entire spectrum of different degrees of unexpectedness and meaningfulness, from scientific discoveries to mundane information behavior, from paradigm shifts to memes. Makri and Blandford (2012b) proposed a framework for subjectively classifying the degree of serendipity. According to their framework, the greater the degree of unexpectedness, the greater the degree of insight, and the greater the degree of value, the purer is the degree of serendipity.

As a subjective and rare phenomenon, serendipity can mean different things to different people in different situations, thus posing serious methodological challenges for studying this concept (Makri & Blandford, 2012b). Indeed, Foster and Ellis (2014) argued that the very nature of the serendipitous event of having a large "accidental" component raises significant challenges for systematic data collection. Furthermore, Makri and Blandford (2012a) also highlighted the fact that serendipity involves an element of unexpectedness and, therefore,

¹ This is quoted from Walpole's letter to Sir Horace Mann.

cannot simply be studied on demand.

Through a systematic literature review, this paper aims to summarise the existing research methodologies applied in understanding serendipity and investigate the strengths and weaknesses of each method by conducting a series of interviews with experts in the field. To the best of our knowledge, no existing literature has discussed serendipity from the perspective of methodology. This paper fills this gap by identifying the challenges and providing insights to address the emerging difficulties and problems regarding understanding the concept of serendipity. The implications include the use of mixed research methods to observe serendipitous experiences.

The remainder of the paper is organized as follows. Section 2 presents a systematic literature review of the research methods applied to study serendipity and the comparisons of each method. The expert interviews are outlined in Section 3, exploring the strengths and weaknesses of the research methods identified in the literature review. Section 4 discusses the use of mixed research methods to produce a complete picture of serendipity, and presents several challenges and implications relating to empirical studies in the investigation of serendipity in the context of information research.

Systematic Literature Review

Methods

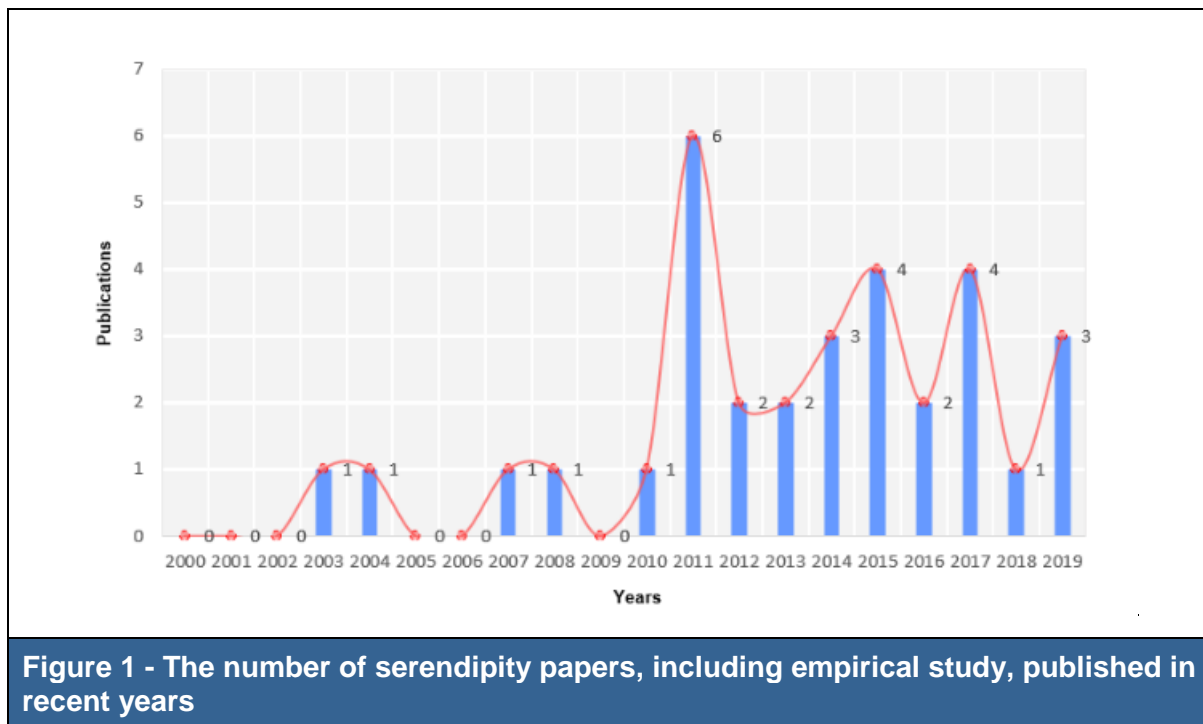
A literature review was conducted, analyzing current studies of serendipity in the context of information research, focusing on the research methodologies applied in understanding the esoteric nature of this phenomenon. Five library databases (IEEE Explore, ScienceDirect, Web of Science, Scopus & Elsevier) were used for an initial search of articles published from January 2000 to December 2019, with the use of multiple combinations of key terms, such as serendipity, information encountering, chance encounters, incidental information acquisition, and opportunistic acquisition of information. By doing so, we hope the terms used were both comprehensive and accurate. The following inclusion criteria were applied (refer to Table 1).

The initial search identified 678 potential articles for the literature review. These papers were carefully selected based on their relevance to the topic, source of publication and originality of findings. During the evaluation process, 653 articles were excluded after reading the titles, keywords, and abstracts, mainly because of the following two reasons: 1) the articles were not in the context of information research; 2) empirical studies were not within the scope of understanding serendipity (Abualigah et al., 2018a; Abualigah et al., 2018b; Agarwal, 2015; Björneborn, 2017; Foster & Ellis, 2014).

Table 1 - Evaluation Criteria	
Publication Language	English
Journal	Only peer-reviewed journals
Author	Articles of popular authors irrespective of ranking, but must be peer-reviewed
Setting	Empirical studies (e.g., interviews, controlled laboratory study).
Research Area	Information Science
Date of publication	2000-2019

Additional articles were identified by reviewing the reference lists of already included publications and a secondary web search (using Google Scholar) with multiple keywords to ensure that we did not miss any pertinent studies. The secondary search process led to the identification of seven additional articles. A total of 32 articles were ultimately included in the literature review.

Figure 1 shows the years in which the articles were published, each of which involves empirical studies relating to the research topic of serendipity. It is evident that before the year 2011, empirical studies regarding serendipity research were very limited, but an increasing number of such studies have been published ever since. A possible reason for this phenomenon is that a special issue of “Information Research — on the opportunistic discovery of information” was published in 2011, where the role of serendipity was highlighted in the context of information research.



Results

Table 2 lists the methodologies employed in the 32 journal articles. It is evident from the table that various research methods have been used to study serendipity, including interviews, surveys/questionnaires, observation, think-aloud, diary-based studies, online ethnography, the Wizard of Oz approach, selective blog mining, selective Twitter data mining, narrative, and network analysis. Some of the studies were conducted using a single method (e.g., interviews, questionnaires, or selective blog mining), while others employed mixed methods to gain an understanding of serendipity.

Collected data type: qualitative vs quantitative

Qualitative methods are often employed when researchers are uncertain about what they are trying to find out when there is no predetermined hypothesis. Qualitative approaches deal with abstract characteristics that cannot be measured numerically and remain open to all possibilities expressed in the data.

The qualitative research method is the mainstream method used to study abstract concepts,

such as creativity and emotion. Researchers conduct qualitative research to uncover creativity in natural settings, and produce descriptive data collected by interviewing, observing, and studying documents (Meador et al., 1999). The self-report method (e.g., interviews), is commonly accepted in the studies of emotions (Sailunaz & Alhaji, 2019). This method relies on asking participants to describe the nature of their experience. Eye-tracking technology and physiological sensors have been used in a recent study investigating the impact of emotion on perceiving serendipitous information encountering (Sun et al., 2021). This study highlighted the potential of using physiological sensors to assess human information behavior. Physiological measures could provide objective data in relation to user tasks, environment, and status. In addition, they are not constrained to limitations inherent in self-reported subjective measurements that rely heavily on a subject's memory to remember the experience. However, there are concerns with physiological measures that they are intrusive in nature.

As can be seen from Table 2, most of the data collected from the studies are qualitative in nature, and of the various research methods employed, the interview method was most widely used (Sun & May, 2014). Nine of the 32 studies were conducted using interviews alone. For example, Foster and Ford (2003) carried out open-ended interviews with academic researchers and then interpreted the notion of serendipity in the context of information seeking. They found serendipity was widely experienced among inter-disciplinary researchers, and it played an important role in their information encountering and the subsequent development of new ideas. Makri and Blandford (2012a, b) conducted semi-structured interviews with interdisciplinary researchers based on their memories of serendipitous encounters in their research or daily life. They put forward a memorable process model of serendipity. Further, they presented a framework to identify serendipity, where "unexpectedness", "insight", and "value" were used to evaluate the strength of a serendipitous incident. Through the cases of online information encountering collected from interviews, Jiang et al. (2015) proposed an integrated model of online information encountering based on McBirnie's (2008) consideration of the process-perception duality of serendipity. The model divides the information encountering process into pre-, mid- and post-activities, while the element of perception constitutes the three clusters of the user, information and environment, where each cluster is comprised of constant factors and dynamic factors.

Compared to qualitative data, quantitative data is primarily collected using surveys/questionnaires (Braun & Clarke, 2006; Kefalidou et al., 2014; Wichary et al., 2005). Heinström (2006) discussed the psychological aspects of serendipity (which were identified as incidental information acquisition) through three survey studies, covering participants with ages ranging from 12 to 53. Heinström (2006) showed that an energetic personality, high motivation levels, and positive emotions could enhance the possibility of experiencing serendipity. In contrast, low levels of motivation, stress, and insecurity reduce the likelihood of such experiences. The online survey study performed by Maccatrozzo et al. (2017) which involved a total of 187 participants, also suggested that individuals with higher levels of curiosity and coping potential are inclined to experience more serendipitous results. McCay-Peet and Toms (2015) conducted a web-based survey of 289 participants. They found those environments that can enable connections and lead to unexpectedness are more conducive to serendipity. Lutz et al. (2017) surveyed 1,173 German internet users. They argued that the backgrounds and results of serendipity encountering were only associated with user satisfaction in the context of social network sites. Grange et al. (2019) suggested that, in the context of online shopping, serendipitous experiences are affected by both environmental factors (i.e., website design) and internal factors (i.e., individual attitudes towards uncertainty). In one of their experimental studies, they identified that the integration of social media and electronic commerce could cultivate serendipity. In this research, each individual is conceptualized as a social actor. The serendipitous system design constructs information not only through individual needs but also through social interactions. This is consistent with the suggestion of McCay-Peet and Toms (2015) who highlight that social networking implies unplanned and unstructured opportunities for the accidental coming together of ideas.

We identified eight of the 32 studies employed surveys/questionnaires together with other methods to understand the process of information encountering (Björneborn, 2008; Erdelez, 2004; McCay-Peet et al., 2014; McCay-Peet & Toms, 2011; Miwa et al., 2011; Pontis et al., 2016; Saadatmand & Kumpulainen, 2013; Waugh et al., 2017; Yadamsuren & Heinström, 2011).

Miwa et al. (2011) conducted a controlled laboratory study during which an eye-tracker was employed to investigate information encountering. Their collected data includes quantitative data, such as the number of viewed web pages, the number of eye-gaze points, and the duration time on different web pages. Based on the collected data, they argued that participants could better recall their feelings and thoughts at the very moment of information encountering with the help of the objective quantitative data. And, they found that participants acted differently in well-defined tasks and exploratory searches, which may have led to a “reconsideration of the definition of information encountering.”

Table 2 - Research methods employed in understanding serendipity

Employed Methodology	Number	Research Setting	Articles
Single method			
Interview	8	Naturalistic	Foster & Ford (2003); Jiang et al. (2015); Makri et al., (2014); Makri & Blandford, 2012a, 2012b; McCay-Peet & Toms (2015); Nutefall & Ryder (2010); Pálsdóttir (2011).
Survey/questionnaire	6	Naturalistic	Grange et al. (2019); Heinström (2006); Lutz et al. (2017); Maccatrozzo et al. (2017); McCay-Peet et al. (2014, 2015); Stewart & Basic (2014).
Selective Blog Mining	1	Naturalistic	Bogers & Björneborn (2013); Rubin et al. (2011).
Survey/questionnaire	2	Controlled Laboratory	Erdelez (2004); McCay-Peet & Toms (2015).
Interview	1	Controlled Laboratory	Makri et al. (2015).
Mixed methods			
Primarily interview-based	2	Naturalistic	McCay-Peet et al. (2014); Yadamsuren & Heinström (2011).
Primarily observation-based	3	Naturalistic	Björneborn (2008); Makri et al. (2019); Waugh et al. (2017).
Primarily diary-based	4	Naturalistic	Kefalidou & Sharples (2016); Makri et al. (2017); Sun et al. (2011); Zhou et al. (2018).
Primarily quantitative descriptive	1	Naturalistic	Srirahayu et al. (2019).
Primarily online Ethnography	1	Naturalistic	Saadatmand & Kumpulainen (2019).
Primarily Wizard of Oz	1	Naturalistic	Pontis et al. (2016).
Primarily questionnaire-based	1	Controlled Laboratory	Miwa et al. (2011).

Naturalistic setting vs controlled laboratory setting

Under a naturalistic setting, researchers typically observe study subjects in their everyday setting (Given, 2008). It involves collecting data from the subjects of interest within their daily living or working environment. Serendipity has usually been studied in a naturalistic setting using interviews (Makri & Blandford, 2012a; Makri & Blandford, 2012b; McCay-Peet & Toms, 2011). Other researchers also suggest that it can be difficult or impossible to observe serendipity in controlled settings, advocating the need to study it in more naturalistic settings (e.g., André et al., 2009; Björneborn 2017). For instance, André et al. (2009) argued that “because serendipity is inherently rare, it is hard for researchers to capture or induce it for study and experimentation” (p. 307). Björneborn (2017) further reasoned that “we cannot design environments always leading to serendipity as serendipity is a highly subjective and situational phenomenon” (p. 20).

Nevertheless, Erdelez (2004) argued that challenges to the study of the opportunistic acquisition of information in controlled environments “can be overcome with very careful planning, high attention to detail, and ongoing adjustments in the development and execution of a research design” (p. 1023). Erdelez (2004) designed a laboratory setting in which ten participants were assigned to a web-searching task related to online shopping for a surfboard as the foreground task. A trigger for information encountering was embedded within the information-seeking activities to induce respondents’ experiences of information encountering. After the task was completed, a post-survey was conducted to investigate participants’ perceptions of information encountering. This study also gave an initial insight into observing participants’ reactions to information encountering in a controlled laboratory setting and demonstrated several methodological possibilities for observing serendipity in a controlled laboratory research setting. In addition, in the laboratory setting, Erdelez (2004) used eye-tracking technology to capture the subtle changes in users’ attention. These sensitive data could facilitate further research efforts in information encountering.

McCay-Peet and Toms (2011) invited 123 individuals to participate in their study in an office environment, where they were asked to conduct web browsing in a wikiSearch system for twenty minutes. This system involves different interfaces that can give feedback to participants regarding their search tasks and provide a list of suggested items that are somewhat related to the searched results. The participants were then invited to a survey regarding serendipity scales drawn organically from Björneborn’s (2008) ten dimensions of the physical library. Their study ultimately identified five factors (enabled connections, introduced the unexpected, presented variety, triggered divergence, and induced curiosity) as the core elements that can facilitate serendipity in a digital environment.

Makri et al. (2015) is another example of a controlled laboratory study. They recruited 45 participants to perform self-selected searching tasks in three different digital environments: digital libraries, e-commerce sites and online news sites. Over at least 30 minutes, these participants were asked to bookmark and/or screenshot the information they considered useful. Participants were asked to think aloud while using digital information environments. After the tasks had been completed, a post-interview was conducted with each participant, focusing on understanding the two dimensions of “unexpectedness” and “usefulness” in respect of their stored bookmarks or screenshots. Participants were not informed of the purpose of the research until the end of the study. This study demonstrates that “with a carefully-considered approach, serendipity-related information interaction behavior can be directly observed” (Makri et al., 2015, p. 1).

Based on our examination of the 32 empirical studies, it was evident that there are two major concerns in current serendipity studies in information research. The first is that, to date, it remains the conventional research methods (e.g., interviews, surveys, diary-based studies, and observation) that play a primary role in understanding serendipity. The second concern is

related to the collected data type. Most of the discussed studies only collected subjective data, relying heavily on individuals' perceptions of serendipity. The pros and cons remain unclear for applying these different methods when studying serendipity, and what challenges may be confronted when employing these methods. A series of interviews with experts were then conducted to gain insight and understanding of these concerns and challenges.

Interviews with Experts

Participants

Nine experts were interviewed, each of whom had at least one year's research experience into serendipity. The participants were recruited through printed and electronic advertisements on notice boards at the workshop of "The Serendipity Factor: Evaluating the Affordances of Digital Environments (SEADE)", which was organized under the ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR) 2016, where a group of researchers were gathered to discuss ongoing serendipity research. The participants comprised seven females and two males, with experience in serendipity study ranging from one to eight years (see Table 3).

Table 3 - Research methods employed by the different experts				
Expert	Years Researching Serendipity	No. of Relevant Publications	Research Setting	Methodology Used in Serendipity Studies
P1	8	14	Naturalistic	Interview, diary-based study, Wizard of Oz, group interview
P2	6	7	Naturalistic	Interview, online survey
P3	5	17	Controlled Laboratory	Web-data scanning, interview, online survey, stories
P4	2	9	Naturalistic	Diary-based study, Wizard of Oz, interview
P5	1	5	Naturalistic, Controlled Laboratory	Interview, focus group, diary-based study, think-aloud
P6	2	3	Naturalistic, Controlled Laboratory	Interview, survey
P7	1	4	Naturalistic	Survey
P8	1	1	Naturalistic	Interview
P9	1	2	Naturalistic	Interview, survey, think-aloud

Methods

A structured interview was conducted with each expert. Each interview lasted approximately half an hour and focused mainly on methodological issues when performing serendipity studies. See Table 4 for the detailed interview protocol.

Table 4 - Expert Interview Protocol

Q1: How long have you been researching serendipity?
Q2: What are the challenges or problems of applying conventional research methods (e.g., interviews, surveys, focus groups) in studying serendipity?
Q3: Scenario: Ann was concerned about an important piece of literature for her research, but she could not get hold of the paper. One day, when Ann was searching for the paper online in her spare time, a web link popped up, which indicated that the paper could be freely downloaded. Ann was curious, so she opened the link and searched for the paper on the website, and although she could not download the paper, she found the author's contact information. She then sent an email via the website and asked if the author could send her the paper. To her surprise, the next day, she received the author's reply via the website with the paper attached. In this scenario, what are the interesting aspects you may look into, and how will you collect data relating to the same?
Q4: What research methods have you used previously to conduct research into serendipity?
Q5: What was your user population in your previous study?
Q6: What was your research context in your previous study?
Q7: What are the advantages of applying your research methods in the study of serendipity?
Q8: What are the disadvantages of applying your research methods in the study of serendipity?
Q9: Having discussed the pros and cons of these research methods, how do you suggest balancing the methods to conduct good research into serendipity?

Data collection and analysis

All interviews were voice recorded with the permission of the participants. The collected data were qualitative in nature. A thematic analysis was conducted to analyze the interview data (Sun & May, 2014). First, based on the questions on the interview protocol, a top-bottom thematic analysis was conducted. Next, various topics, such as the challenges and problems caused by conventional methods, scenario-related questions, qualitative and quantitative research methods, strengths and weaknesses of the methods, etc., were coded into different themes. Subsequently, a bottom-up coding process was conducted further to categorize any similar sub-themes to a new theme. The coding was performed using Nvivo 11.

Results

Table 3 lists the research methods employed by the experts. It can be seen from the table that they have used most of the research methods discussed in the literature review section.

Interviews

The interview is the most widely used method employed in studying and understanding serendipity. The advantage of using interviews in these studies is that they can provide rich and detailed data from participants' perceptions of serendipity.

"Interviews can get the qualitative data, collecting detailed information of user perceptions of serendipity which is really useful" (P2)

Researchers can identify the details of the context of participants' serendipitous encountering retrospectively based on the interview data. They can look deeper into their thoughts and how they understand serendipity as and when it happens.

"Interviews can provide quite a lot of information from people in terms of the context of relevant interactions or experiences, even though you obtained the context retrospectively". (P1)

However, the main disadvantage of interviews is that the collected data is based on participants' memories, and it is retrospective data rather than instant data from the moment when an episode of serendipity occurs.

"You don't get the information when it happens, so there's a bit of time delay from the participants." (P2)

"It relies on people's instant opinions." (P3)

As the definition of serendipity remains elusive, thus far, participants have different perceptions of serendipity, which often lead to a perception gap between researchers and participants, challenges to interpret participants' responses and to judge whether an episode can be considered as serendipity.

"I found one big challenge is interpreting what [participants] are saying, and several times when I was doing research, [participants] didn't talk a lot like us in the field of serendipity research. They said things like 'I had an accidental find' or 'I was searching for a thing' or 'I saw it on the Internet', so it's like picking out what they are actually saying and interpreting that as serendipity." (P8)

Group interviews

Compared to individual interviews, group interviews can collect data from a group of participants simultaneously, and it is flexible enough for researchers to set different research settings to study serendipity. One expert mentioned that he could collect excellent feedback from group interviews regarding serendipity research design strategies.

"It is good for reaching concerns or design decisions in relation to the design digital tools to facilitate serendipity." (P5).

Group discussion can support interactivity, and such dynamic situations can encourage participants to raise ideas that may not be identified from a one-to-one interview. Researchers can also collect interaction data during the study process and different data layers by distinguishing between participant groups.

"It gives you different layers of richness or information when it is a focus group." (P1)

"You can have multiple interactions between different participants, so you can find people, and elicit information that cannot be obtained from a single individual." (P1)

There are also disadvantages of employing group discussions, where one individual's opinion can often have an influence on others. For example, one participant's understanding of serendipity may impact the other participants in a group. In addition, the social interaction process may put pressure on participants.

"Participants might appear to be under pressure owing to the social interaction, and this then influences their responses in the focus group." (P1)

"Sometimes individual opinions can affect the opinions of others." (P5)

Survey

Surveys are also widely used in conducting serendipity-related studies. McCay-Peet and Toms (2015) used a survey to validate their qualitative findings. Another use of surveys is to evaluate serendipity by employing Makri and Blandford's (2012b) framework. The survey

questions can be easily distributed to participants (e.g., emails, survey websites, paper). An obvious advantage of applying a survey is that it can collect a large number of samples at a relatively low cost. However, some experts pointed out that it is not easy to enroll a sufficiently representative group of participants in a serendipity study.

“It’s hard with surveys to get representative people who have experienced serendipity to fill out questionnaires and also it is difficult to seriously look into their thoughts especially when you are doing web-based surveys” (P 2).

Moreover, surveys cannot examine participants’ experience about a serendipitous encounter in real-time.

“There is a time dimension for normal research methods.” (P7)

Diary-based Study

A diary study often lasts at least for several consecutive days. It can record the occurrence of serendipity in the first place (e.g., time, mood, responses in a serendipitous episode). In serendipity studies, the diary method is often combined with a post-interview (e.g., Sun et al., 2011; Kefalidou & Sharples, 2016). Thus, any information provided about incidents can help participants to better recall the context of when serendipity occurred. In addition, serendipity is a user-defined notion, and the diary method permits researchers to define the data that are not easily extracted by other methods. For example, by identifying the frequency of serendipitous encountering, researchers can distinguish between user groups, ranging from super-encounterers to non-encounterers (Erdelez, 1997).

The drawbacks of the diary method are that it is time-consuming and capturing the data can be tedious. Since the diary method is often associated with a post-interview to understand participants’ different episodes of serendipitous encountering, this often leads to a relatively lengthy interview.

“[Diary study] makes it difficult to obtain large numbers of sample participants, and generalizing the context is also difficult, as the contexts people came up with were really varied, and often they were quite unusual in their personal life and things like this, so it was quite difficult to identify generalizable findings like those which could be obtained in a laboratory type of context. Some were not sure they would particularly associate themselves with the examples that we obtained.” (P4)

Another drawback of applying the diary study method is that participants are not always introspective in their daily lives and may have trouble recording episodes in a diary entry. Lazar et al. (2017) showed that “it is often difficult to strike a balance between a sufficiently frequent series of diary entries and infringement on daily activities (user participation may then trail off)”.

Think-aloud

Some experts reported that they used think-aloud, often in controlled laboratory studies, to better understand participants’ experiences of serendipity. Using think-aloud in the study of serendipity can help participants remember specific cues that appeared during the study. Researchers can also understand participants’ ongoing cognitive processes relating to their serendipitous experience. A particular benefit of this method for studying serendipity is that during the post-interviews, the collected data can greatly help participants recall the context during which the episode occurred, and researchers can collect more detailed and accurate data during the interview sessions.

“Especially in the digital environment, I find it is very useful to use the think-aloud approach, [where] it helped participants to remember certain cues and other things in the environment, because when I began the interviews, they had different memories of serendipitous experiences online, but they had to remember exactly what had caused the serendipitous sparks. They can easily point out to me certain parts of triggers, even though the content is changing.” (P9)

However, it should also be noted that the more a think-aloud is used during a study, the more it is possible to interrupt a participant’s cognitive flow, and this will often lead to more time being required to finish the designated tasks. It is also challenging for a researcher to use this method when some participants do not feel comfortable speaking aloud, especially when alone.

Wizard of Oz

One expert employed the Wizard of Oz approach in studying serendipity. This approach is often associated with a low cost and can be completed in a short period of time, where a human plays the role of a system to provide related information to participants. On most occasions, participants are not aware that the information they find has been designed purposely, but consider the information to be natural feedback from the system. Therefore, this method helps explore a user’s perception of serendipity before a real system is designed, or when the appropriate technology does not yet exist. For example, one expert developed a prototype of a notebook tool, and based on this tool, s/he employed a Wizard of Oz approach by sending push text suggestions to participants according to details obtained before the study concerning their experiences and interests. Participants’ immediate responses to particular kinds of information were collected, and sometimes these included serendipitous information.

“[Wizard of Oz] provides information about how people respond to particular kinds of tailored information that you have provided to them; studies are by ground theory, and at the same time, wizard settings can provide prompt responses from participants when they experience serendipity.” (P1)

The limitation of this method comes mainly from errors made by researchers. It is possible that the wizard’s role, which a human plays, could make errors when listening to dictation or when typing words. One expert explained that the original aim of adopting a Wizard of Oz approach in her study was to collect data from participants’ immediate responses to a serendipitous encounter, but it was also often the case that participants cannot give feedback promptly on the context, which means the researcher fails to capture the entire storyline and context surrounding the serendipitous encounter: *“While [the Wizard of Oz approach] could provide timely information as it happens, you may not obtain the full context surrounding an incident or events, and that could be more of a disadvantage.” (P1)*

Discussion and Conclusion

Based on our literature review and the interviews with experts, it is evident that various research methods have been employed in the study of serendipity. While there have been multiple efforts to remove the veil of serendipity in recent decades, this paper provides a methodological contribution to serendipity studies in the context of information research.

Methodological challenges in studying serendipity

Observing serendipity in real-time

As reported by most experts, a significant challenge is a difficulty in observing serendipity in

real-time. Most had conducted related studies in a naturalistic setting using interviews, surveys, group discussions, etc. Our results showed these data collection methods are dependent on participants' memories, resulting in the interviewers being concerned about the challenge of observing serendipity in real-time. Three experts had conducted controlled studies trying to capture serendipity, including the collection of a participant's log file data on a website, and as one expert pointed out, *"It's not easy to directly observe serendipity, but I think these problems can be overcome with more research in this area."* (P4)

Foster and Ellis (2014) argued that there should be a means by which the researcher may learn when a serendipitous recognition occurs, either at the moment when the serendipity occurs or at some later point. Many controlled laboratory studies demonstrated the possibilities of observing serendipity under a controlled laboratory setting. Such lab-based studies may be a solution for capturing more detailed data at the moment when serendipity occurs. Erdelez (2004) proposed the use of eye-tracking technology to capture subtle changes in users' attention shifts. Miwa et al. (2017) eye-tracking data provided strong evidence that "information encountering sometimes led to a temporary deviation from the initial task, but may have resulted in a change of the topic of the paper and/or destination of the trip" (p. 1). In addition to eye-tracking technology, it is also possible to employ other HCI methods to capture the moment of serendipity in the lab-based studies, such as muscular and skeletal position sensing, motion tracking, and using physiological tools to collect physiological data (e.g., electrodermal activity, cardiovascular signals, respiration, brain activity, and muscle tension) (Peng et al., 2020). Hence, by careful planning and detailed laboratory study design, researchers can capture sensitive data at the moment when serendipity happens.

Definition of serendipity

Several experts discussed how the lack of consensus on what serendipity is has contributed to a significant obstacle to carrying out their studies. This leads to a perception gap between participants' recognition and a researcher's identification of serendipity. Participants typically do not directly report their encountering an episode of serendipity, and they may describe this phenomenon using other similar words, such as accident or chance, or may even do so without any description. Researchers need to identify the notion of serendipity based on the varying responses from different participants. This interpretation process inevitably results in a potential challenge for researchers to not introduce bias to participants' original understandings of serendipity.

Despite the lack of consensus on the definition of serendipity, several proposed frameworks showed a promise to facilitate researchers to identify serendipity. For example, Makri and Blandford (2012b) proposed a framework to classify serendipitous cases. They determine whether serendipitous cases occurred based on the level of "unexpectedness", "insight", and "value" of these cases. Their framework has been adopted in recent studies to determine whether participants have experienced serendipity (e.g., Kefalidou & Sharples, 2016; Pontis et al., 2016). Zhou et al. (2018) further extended Makri and Blandford's (2012b) framework. All these frameworks emphasize the properties of serendipity rather than redefining what serendipity is. By doing so, they provide researchers with a solution to identify serendipity.

Individual differences

The participants also reported individual differences present challenges to the study of serendipity. First, the ability to recall information is different for all individuals. For example, one expert revealed that some participants could describe their experience of serendipity very well during interviews, while others could not remember the details of an encounter. Another challenge is that individuals may experience a different level of serendipity. Erdelez (1997) divided people who encounter serendipity into five categories, ranging from super-encounterers to non-encounterers. One expert (P5) stated that she has a colleague who

claimed that she had never experienced serendipity, while the expert considered herself to be a frequent encounterer of serendipity in her daily life. Such individual differences make it difficult for researchers to recruit a representative sample of participants to perform their studies. A possible solution is to employ online crowdsourced studies. According to Lazar et al. (2017) “crowdsourcing studies use online platforms to collect data from participants over the web, usually through the use of web software designed to enroll participants, provide training, and complete relevant tasks” (p. 429). Selective blog mining, or what the expert called “web scanning”, is typically used in a crowdsourced study. These methods are used to search and analyze blogs to generate insights that might otherwise not be found by examining an individual blog. This kind of method allows researchers to collect data from large samples, and thus it is possible to draw more reliable statistical conclusions than from a small selection of participants.

Mixed research methods

Based on the literature review and expert interviews, it is shown that a variety of research methods have been employed in the study of serendipity. However, there is not a single all-purpose method that can be optimally applied to study serendipity as each comes with its own set of strengths and weaknesses. For example, an interview is a good solution for looking into a participant’s deep thoughts about a serendipitous episode, but it is dependent on their memory. Mixed research methods are suggested to produce a more comprehensive understanding of serendipity and contribute to more reliable findings. For example, combining diary studies or controlled laboratory studies with pictures, videos, or log files to collect any instances of serendipitous encountering can better help participants recall the context in which serendipity occurred. In addition, a think-aloud session, especially conducted in a controlled laboratory setting, would bring significant benefits for researchers wanting to identify participants’ cognitive processes and provide possible clues for any post-interview or post-survey elements of a study.

It is well acknowledged that there are individual differences in experiencing serendipity (Erdelez, 1997). To address this concern, a survey could be conducted before an experimental study to categorize participants into super-encounterers and super- non-encounterers. This could be helpful to avoid the selection of non-encounterers as participants.

This paper investigated research methods employed to study serendipity by synthesizing finding from a literature review and the interviews with experts. It provides a methodological contribution to serendipity studies by systematically summarizing the methods employed in the studies of serendipity and identifying the strengths and weakness of each method. It also suggests the novel approach of using mixed research methods to study serendipity. This study has potential limitations related to a small number of experts involved in the expert interview. However, it should be noted that the nature of the topic is a relatively focused area, and it was observed after interviewing the experts that new data seems to not contribute to the findings owing to its repetition of comments.

Acknowledgements

This study received financial support from the National Natural Science Foundation of China program (7207010273) and Ningbo Hefeng Creative Industrial Park.

References

- Abualigah, L. M. Q. (2019). *Feature Selection and Enhanced Krill Herd Algorithm for Text Document Clustering*. Berlin: Springer.
- Abualigah, L. M., Khader, A. T., & Hanandeh, E. S. (2018a). Hybrid clustering analysis using improved krill herd algorithm. *Applied Intelligence*, 48(11), 4047-4071.
- Abualigah, L. M., Khader, A. T., & Hanandeh, E. S. (2018b). A new feature selection method to improve the document clustering using particle swarm optimization algorithm. *Journal of Computational Science*, 25, 456-466.
- Agarwal, N. K. (2015). Towards a definition of serendipity in information behaviour. *Information Research*, 20(3), paper 675.
- Andel, P. V. (1994). Anatomy of the unsought finding. serendipity: Origin, history, domains, traditions, appearances, patterns and programmability. *The British Journal for the Philosophy of Science*, 45(2), 631-648.
- André, P., Schraefel, M. C., Teevan, J., & Dumais, S. T. (2009, October). Discovery is never by chance: designing for (un) serendipity. In *Proceedings of the seventh ACM Conference on Creativity and Cognition*, (pp. 305-314).
- Björneborn, L. (2008). Serendipity dimensions and users' information behaviour in the physical library interface. *Information Research*, 13(4), paper 370.
- Björneborn, L. (2017). Three key affordances for serendipity: Toward a framework connecting environmental and personal factors in serendipitous encounters. *Journal of Documentation*, 73(5), 1053-1081.
- Bogers, T., & Björneborn, L. (2013). Micro-serendipity: Meaningful coincidences in everyday life shared on Twitter. In *Proceedings of iConference* (pp. 196-208).
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Cory, K. A. (1999). Discovering hidden analogies in an online humanities database. *Library Trends*, 48(1), (pp. 60-71).
- Dantonio, L., Makri, S., & Blandford, A. (2012). Coming across academic social media content serendipitously. In *Proceedings of the American Society for Information Science and Technology*, 49(1), 1-10.
- Erdelez, S. (1997, August). Information encountering: A conceptual framework for accidental information discovery. In *Proceedings of an international conference on information seeking in context* (pp. 412-421).
- Erdelez, S. (2004). Investigation of information encountering in the controlled research environment. *Information Processing & Management*, 40(6), 1013-1025.
- Erdelez, S., Basic, J., & Levitov, D. D. (2011). Potential for inclusion of information encountering within information literacy models. *Information Research*, 16(3), paper 489. <http://InformationR.net/ir/16-3/paper489.html>
- Foster, A., & Ellis, D. (2014). Serendipity and its study. *Journal of Documentation*, 70(6), 1015-1038.
- Foster, A., & Ford, N. (2003). Serendipity and information seeking: An empirical study. *Journal of Documentation*, 59(3), 321-340.
- Given, L. M. (2008). *The Sage encyclopedia of qualitative research methods*. Sage Publications.

- Grange, C., Benbasat, I., & Burton-Jones, A. (2019). With a little help from my friends: Cultivating serendipity in online shopping environments. *Information & Management*, 56(2), 225-235.
- Heinström, J. (2006). Psychological factors behind incidental information acquisition. *Library & Information Science Research*, 28(4), 579-594.
- Jiang, T., Liu, F., & Chi, Y. (2015). Online information encountering: Modeling the process and influencing factors. *Journal of Documentation*, 71(6), 1135-1157.
- Kefalidou, G., Skatova, A., Brown, M., Shipp, V., Pinchin, J., Kelly, P., Dix, A., & Sun, X. (2014, September). Enhancing self-reflection with wearable sensors. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services* (pp. 577-580).
- Kefalidou, G., & Sharples, S. (2016). Encouraging serendipity in research: Designing technologies to support connection-making. *International Journal of Human-Computer Studies*, 89, 1-23.
- Kotkov, D., Veijalainen, J., & Wang, S. (2016). Challenges of serendipity in recommender systems. In *Proceedings of the 12th International Conference on Web Information Systems and Technologies, Vol 2*, (pp 251-256).
- Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.
- Lutz, C., Pieter Hoffmann, C., & Meckel, M. (2017). Online serendipity: A contextual differentiation of antecedents and outcomes. *Journal of the Association for Information Science and Technology*, 68(7), 1698-1710.
- Maccatrozzo, V., van Everdingen, E., Aroyo, L., & Schreiber, G. (2017). Everybody, more or less, likes serendipity. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 29-34).
- Makri, S., & Blandford, A. (2012a). Coming across information serendipitously—Part 1 - A process model. *Journal of Documentation*, 68(5), 684-705.
- Makri, S., & Blandford, A. (2012b). Coming across information serendipitously—Part 2 - A classification framework. *Journal of Documentation*, 68(5), 706-724.
- Makri, S., Blandford, A., Woods, M., Sharples, S., & Maxwell, D. (2014). Making my own luck: Serendipity strategies and how to support them in digital information environments. *Journal of the Association for Information Science and Technology*, 65(11), 2179-2194.
- Makri, S., Bhuiya, J., Carthy, J., & Owusu-Bonsu, J. (2015). Observing serendipity in digital information environments. In *Proceedings of the Association for Information Science and Technology*, 52(1), 1-10.
- Makri, S., Chen, Y. C., McKay, D., Buchanan, G., & Ocepek, M. (2019). Discovering the unfindable: The Tension between findability and discoverability in a bookshop designed for serendipity. In *IFIP Conference on Human-Computer Interaction* (pp. 3-23). Springer, Cham.
- Makri, S., Ravem, M., & McKay, D. (2017). After serendipity strikes: Creating value from encountered information. In *Proceedings of the Association for Information Science and Technology*, 54(1), 279-288.
- Maksai, A., Garcin, F., & Faltings, B. (2015, September). Predicting online performance of news recommender systems through richer evaluation metrics. In *Proceedings of the 9th ACM Conference on Recommender Systems* (pp. 179-186).

- McBirnie, A. (2008, November). Seeking serendipity: The paradox of control. In *Aslib Proceedings: New Information Perspectives*, 60 (6) pp. 600-618. Emerald Group Publishing Limited.
- McBirnie, A., & Urquhart, C. (2011). Motifs: Dominant interaction patterns in event structures of serendipity. *Information Research*, 16(3), <http://InformationR.net/ir/16-3/paper494.html>.
- McCay-Peet, L., & Toms, E. G. (2015). Investigating serendipity: How it unfolds and what may influence it. *Journal of the Association for Information Science and Technology*, 66(7), 1463-1476.
- McCay-Peet, L., Toms, E. G., & Kelloway, E. K. (2015). Examination of relationships among serendipity, the environment, and individual differences. *Information Processing & Management*, 51(4), 391-412.
- McCay-Peet, L., Toms, E. G., & Kelloway, E. K. (2014). Development and assessment of the content validity of a scale to measure how well a digital environment facilitates serendipity. *Information Research*, 19(3), paper 630.
- McCay-Peet, L., & Toms, E. (2011). Measuring the dimensions of serendipity in digital environments. *Information Research: An International Electronic Journal*, 16(3), paper 483.
- Meador, K., Hunsaker, S. L., & Kearney, K. (1999). Qualitative Research Methods for Studying Creativity. In A. S. Fishkin, B. Cramond, & P. Olszewski-Kubilius (Eds.), *Investigating creativity in youth: Research and methods* (pp. 239-261). Cresskill, NJ: Hampton.
- Miwa, M. (2000). *Use of human intermediation in information problem solving: The users' perspective*. Syracuse, NY: ERIC Clearinghouse on Information and Technology.
- Miwa, M., Egusa, Y., Saito, H., Takaku, M., Terai, H., & Kando, N. (2011). A method to capture information countering embedded in exploratory Web searches. *Information Research*, 16(3), 3-16.
- Nutefall, J. E., & Ryder, P. M. (2010). The serendipitous research process. *The Journal of Academic Librarianship*, 36(3), 228-234.
- Pálsdóttir, Á. (2011). Opportunistic discovery of information by elderly Icelanders and their relatives. *Information Research*, 16(3), paper 485.
- Peng, X., Huang, Z., & Sun, X. (2020). Building BROOK: A multi-modal and facial video database for Human-Vehicle Interaction research. *arXiv preprint arXiv:2005.08637*.
- Pontis, S., Kefalidou, G., Blandford, A., Forth, J., Makri, S., Sharples, S., Wiggins, G., & Woods, M. (2016). Academics' responses to encountered information: Context matters. *Journal of the Association for Information Science and Technology*, 67(8), 1883-1903.
- Rosenman, M. F. (1988). Serendipity and Scientific Discovery. *Journal of Creative Behavior*, 22(2), 132-138.
- Rubin, V. L., Burkell, J., & Quan-Haase, A. (2011). Facets of serendipity in everyday chance encounters: A grounded theory approach to blog analysis. *Information Research*, 16(3), paper 488.
- Saadatmand, M., & Kumpulainen, K. (2013). Content aggregation and knowledge sharing in a personal learning environment: Serendipity in open online networks. *International Journal of Emerging Technologies in Learning (IJET)*, 8 (S1), 70-77.
- Sailunaz, K., & Alhaji, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, 36, paper 101003.

- Srirahayu, D. P., Irfana, M. S. A., Mannan, E. F., & Anugrah, E. P. (2019). Serendipity on information searching behavior in use e-journal collection. *Library Philosophy and Practice*, paper 2608.
- Stewart, K. N., & Basic, J. (2014). Information encountering and management in information literacy instruction of undergraduate students. *International Journal of Information Management*, 34(2), 74-79.
- Sun, X., & May, A. (2014). Design of the user experience for personalized mobile services. *International Journal of Human-Computer Interaction*, 5(2), 21-39.
- Sun, X., Sharples, S., & Makri, S. (2011). A user-centred mobile diary study approach to understanding serendipity in information research. *Information Research*, 16(3), 3-16.
- Sun, X., Zhou, X., Wang, Q. & Sharples, S. (2021). Investigating the impact of emotions on perceiving serendipitous information encountering. *Journal of the Association for Information Science and Technology*, forthcoming.
- Thomas, D. R. (2003). A general inductive approach for qualitative data analysis. *American Journal of Evaluation*, 27(2), 237-246.
- Toms, E. G. (2000). Serendipitous information retrieval. A position paper for the European Research Consortium for Informatics and Mathematics. Workshop conducted at the *First DELOS network of excellence workshop on information seeking, searching and querying in digital libraries*, Zurich, 11–12 December. Available: <http://www.ercim.org/publication/ws-proceedings/DelNoe01>
- Waugh, S., McKay, D., & Makri, S. (2017, March). 'Too Much Serendipity' The tension between information seeking and encountering at the library shelves. In *Proceedings of the 2017 conference on conference human information interaction and retrieval*, (pp. 277-280).
- Wichary, M., Gunawan, L., Van den Ende, N., Hjortzberg-Nordlund, Q., Matysiak, A., Janssen, R., & Sun, X. (2005). Vista: Interactive coffee-corner display. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems*, (pp. 1062-1077).
- Yadamsuren, B., & Heinström, J. (2011). Emotional reactions to incidental exposure to online news. *Information Research*, 16(3), paper 486.
- Zhou, X., Xu, Z., Sun, X., & Wang, Q. F. (2017, July). A new information theory-based serendipitous algorithm design. In *Proceedings of the International Conference on Human Interface and the Management of Information*, (pp. 314-327).
- Zhou, X., Sun, X., Wang, Q., & Sharples, S. (2018). A context-based study of serendipity in information research among Chinese scholars. *Journal of Documentation*. 74(3), 526-551.

About the Authors

Dr Xu Sun is an Associate Professor and course director for Product Design and Manufacture at the University of Nottingham, Ningbo, China. Her current research focuses on intellectual property rights in the design, digital adaptive educational game, and creative computer supported cooperative work, transportation design, and sustainable product design.

Dr Xiaosong Zhou is a lecturer at the Department of Mechanical and Electrical Engineering, Guizhou Normal University. His current research focus is on human-computer interaction and additive manufacturing.

Dr Qingfeng Wang is an Associate Professor in Economics and Quantitative Methods at the University of Nottingham Business School Ningbo China, University of Nottingham, Ningbo, China. His current research focuses on intellectual property rights and population/labour economics.

Sarah Sharples is a Professor of Human Factors in the Department of Mechanical, Materials and Manufacturing Engineering and Pro-Vice-Chancellor for Equality, Diversity and Inclusion at the University of Nottingham. She completed her PhD in 1999 and has been a researcher, research manager or grant holder on a number of industrial, government and EU funded projects, including a long term programme of research for Network Rail examining implications, design and implementation of novel interfaces for railway control and use of rail simulation for human factors research.

Yaorun Zhang is a PhD candidate at the Faculty of Science and Engineering, University of Nottingham Ningbo, China.