Modeling the Process of Information Encountering Based on the Analysis of Secondary Data

Tingting Jiang[1,2(✉)], Shiting Fu[1], Qian Guo[1], and Enmei Song[1]

[1] School of Information Management, Wuhan University, Wuhan, Hubei, China
tij@whu.edu.cn
[2] Center for Studies of Information Resources, Wuhan University, Wuhan, Hubei, China

Abstract. The critical incident technique (CIT) has been applied extensively in the research on information encountering (IE), and abundant IE incident descriptions have been accumulated in the literature. This study used these descriptions as secondary data for the purpose of creating a general model of IE process. The grounded theory approach was employed to systematically analyze the 279 IE incident descriptions extracted from 14 IE studies published since 1995. 230 conceptual labels, 33 subcategories, and 9 categories were created during the data analysis process, which led to one core category, i.e. "IE process". A general IE process model was established as a result to demonstrate the relationships among the major components, including environments, foreground activities, stimuli, reactions, examination of information content, interaction with encountered information, valuable outcomes, and emotional states before/after encountering. This study not only enriches the understanding of IE as a universal information phenomenon, but also shows methodological significance by making use of secondary data to lower cost, enlarge sample size, and diversify data sources.

Keywords: Information encountering · Process model · Secondary data analysis · Grounded theory approach

1 Introduction

In contrast to active and purposive information seeking, information encountering is finding unexpected information passively [1, 2]. Despite the variety of terminologies for describing such phenomenon, "information encountering (IE)" is used consistently throughout this paper to avoid possible confusion. How IE occurs has been one of the most important themes of existing related studies, and a number of empirical models have been established to demonstrate the process of IE [2–7]. These models, however, are mostly constrained to specific contexts (e.g. information retrieval, social media, and work-related) or specific user groups (e.g. professionals, students, and researchers). They were built upon the qualitative data provided by individuals who had experienced IE. The absence of a general model that reveals the process of IE as a universal

N. G. Taylor et al. (Eds.): iConference 2019, LNCS 11420, pp. 41–49, 2019.
https://doi.org/10.1007/978-3-030-15742-5_4

information phenomenon, to a great extent, can be attributed to the shortage of time and budget, which would inhibit any single study from achieving larger data size or greater variety of data sources.

As a result, this study introduced secondary data analysis to the investigation of IE process. Secondary data analysis is a cost-efficient way to make full use of data that are already collected by previous studies to address potentially important new research questions [8]. A total of 279 IE incident descriptions were extracted as secondary data from 14 IE studies published since 1995. These studies were chosen because they relied on the combination of self-report methods and the critical incident technique (CIT) for data collection. The grounded theory approach was employed to systematically analyze the descriptions in QSR Nvivo 11, which gave birth to a general IE process model.

2 Literature Review

2.1 Existing IE Process Models

According to the earliest model by Erdelez [9], IE is embedded within a high-level process of information seeking. A typical IE episode contains five functional steps, i.e. noticing, stopping, examining, capturing, and returning. A further development is the integrated model of online IE which provides a global view of the three phases respectively accommodating, the pre-, mid-, and post-activities of IE. Specifically, IE may happen during online browsing, searching, or social interaction; noticing an information stimulus and examining the information content are both indispensable to acquiring interesting or useful information; and the encountered information may be explored further, used immediately, saved, and/or shared [6].

There are four components in Cunha [7] model of serendipity process: searching for a solution for problem A, precipitating conditions, a bisociation between previously unconnected pieces of information, and an unexpected solution to problem B. Mccay-Peet and Toms [4] modified this model by adding trigger as a necessary element for activating the bisociation as well as an unexpected solution to Task A. More recently, they consolidated several previous models into a new one that consists of trigger, connection, follow-up, valuable outcome, unexpected thread, and perception of serendipity [5].

In the perceptual model of serendipity, Lawley and Tompkins [10] indicated that the happening of an unplanned and unexpected event is preceded by a prepared mind and followed in sequence by recognizing the potential of the event, seizing the moment, amplifying the benefit of the event, and evaluating the effects. This model provided a basis for a later study that established an empirically-grounded model: the serendipity process begins with a mental connection; then there is a cyclic sub-process including forward-facing projections, connection exploiting, and backward-facing reflections; and finally, the whole experience is considered as serendipity given both the value of the outcome and the involvement of insight [3].

2.2 CIT-Based IE Research

According to Flanagan [11] definition, the critical incident technique (CIT) is "a set of procedures for collecting direct observations of human behavior in such a way as to facilitate their potential usefulness in solving practical problems and developing broad psychological principles". An incident refers to any observable human activity that allows for inferences and predictions to be made about the subject of the activity. The critical incident data can be collected through interviews, questionnaires, and written records, which engenders functional descriptions of people's experiences as reflected in the integration of time, places, persons, conditions, and activities [12]. Such qualitative data then enters an inductive analysis process in which incidents are sorted and grouped into categories for important themes and patterns to surface.

The CIT aims to elicit an accurate and in-depth description of the event from the participants and ensures that the incidents are meaningful to participants instead of researchers, which particularly attends to occasional or rare incidents such as IE. IE researchers tended to apply the CIT in interviews or diaries. It should be mentioned that some studies did not indicate explicitly that the CIT was applied, but their data collection processes did reflect the procedure of the CIT.

CIT-based interviews focus on individual events and attach great importance to the details of the interviewees' behavior and mind. Proper guide from the researcher side is often indispensable to effective face-to-face communication [3, 5]. For example, Makri and Blandford [3] invited 28 interdisciplinary researchers to participate in their semi-structured CIT-based interview. The participants firstly needed to talk about their understanding of IE; then an interview guide with several questions was used loosely to help them recall and recount in detail the events in their work and life that they deemed IE incidents; and lastly they were asked about their attitudes towards and opinions on IE. Each interview took about 50 min on average. The 28 participants contributed 46 IE incidents in total.

CIT-based diaries ask participants to record the events as soon as they happen. It is necessary to provide the participants with a recording tool, such as electronic questionnaires and mobile applications, to create diaries by themselves, sometimes following specific instructions or requirements [13, 14]. For example, Rahman and Wilson [15] recruited 14 active Facebook users to fill in a diary entry daily to report their interactions with a search engine. The participants needed to provide an open-ended description of any IE experience. Illustrated examples of their queries were presented to assist recollection. Then they were asked to explain why they clicked on or did not click on highlighted serendipitous results. The diary study lasted 11 days and collected 57 IE incidents.

3 Data Collection

Thanks to the extensive application of the CIT in IE research, abundant descriptions of IE incidents have been accumulated in the literature. Such qualitative data was collected and analyzed in the original studies as primary data to address various research questions. This study, instead, used these descriptions as secondary data for the purpose of creating a general model of IE process. The foremost advantage of reusing the data

in published studies is the decreased investment in time, money, and manpower for data collection. Second, the validity and reliability of secondary data have been ensured during the publication of original studies. Last but not least, resorting to multiple data sources enabled this study to base the investigation on much larger and more widely distributed samples.

This study conducted several rounds of searches on Google Scholar which provides more complete coverage of IE studies, with a series of queries consisting of "critical incident technique" and "serendipity", "information encountering", "opportunistic/incidental acquisition of information", "incidental information acquisition", or "accidental/opportunistic discovery of information", between July 8th and July 12th, 2018. The time span of the searches was set to "1995 ~ now". The returned publications were further screened according to three criteria: (1) IE-related research papers written in English; (2) the CIT adopted for data collection; and (3) original data provided. As a result, 14 papers were selected as data sources, including Erdelez [2], Makri and Blandford [3], Mccay-Peet and Toms [5], Jiang et al. [6], Sun et al. [13], Pontis et al. [14], Rahman and Wilson [15], Foster and Ford [16], Makri and Warwick [17], Dantonio [18], Yadamsuren and Erdelez [19], Miwa et al. [20], Yadamsuren and Heinström [21], Makri, Ravem, and Mckay [22]. They contained 279 IE incident descriptions collected for exploring the characteristics, process, factors, and value of IE. These descriptions were found in different sections of the above papers, such as data analysis, results, and appendix. A number of example descriptions are provided in the Appendix. Although the 14 studies were conducted at different times and/or in different contexts, the participants provided their IE experiences in similar ways, which made it possible to analyze the descriptions in a uniform framework.

4 Data Analysis and Results

The 279 IE incident descriptions were transcribed without any changes into text files, which were then imported into QSR Nvivo 11 for data analysis. Since this study intended to derive a new theoretical model from a large amount of qualitative data, the grounded theory approach was the most appropriate method. The idea was to extract important concepts from original data and integrate them into categories in a bottom-up fashion [23]. Specifically, the data analysis followed a three-step process: open coding for identifying categories and subcategories, axial coding for relating the categories to one another, and selective coding for determining a core category that represents the central phenomenon.

In the open coding step, the researchers read each IE incident description carefully and annotated the critical statements. Take the four statements in Table 1 for examples. Conceptual labels were created for them in the first place. Upon the completion of the annotating for all the descriptions, similar labels were merged into subcategories. The open coding of the 279 descriptions engendered 230 conceptual labels and 33 subcategories in total. The top 10 subcategories were "textual stimuli" (N = 79), "useful" (N = 69), "positive emotions after encountering" (N = 54), "information content" (N = 45), "online environments" (N = 32), "purposeful searching" (N = 32), "interesting" (N = 31), "negative emotions before encountering" (N = 30), "verbal stimuli" (N = 25), and "negative emotions after encountering" (N = 24).

Table 1. Open coding examples

Original statements	Conceptual labels	Subcategories
"…searching for the library catalogue for a specific book…"	Searching for specific books	Purposeful searching
"…an eBook with a similar title caught her eye…"	eBook-triggered	Textual stimuli
"…decided it was a better match for her information need than original target…"	Satisfying one's need	Useful
"…helping her to write her literature…"	Helpful to one's research	Satisfy their own needs

The axial coding involved the constant comparisons of the 33 subcategories deriving from the open coding. The researchers focused on discerning the particular aspect of the IE phenomenon reflected by each subcategory. If two subcategories reflected the same aspect, they belonged to the same category. For example, both "positive emotions after encountering" and "negative emotions after encountering" are pertinent to the emotional state one might have after IE occurred. It is therefore reasonable to relate them to each other and incorporate them into a higher-level category "emotional state after encountering". The 33 subcategories were incorporated into 9 main categories, including "stimuli" (N = 131), "foreground activities" (N = 102), "examination of information content" (N = 97), "emotional state after encountering" (N = 78), "valuable outcomes" (N = 56), "environments" (N = 55), "interaction with encountered information" (N = 55), "emotional state before encountering" (N = 53), and "reactions" (N = 22).

The final step of data analysis, i.e. selective coding, was devoted to detecting the connections among the 9 main categories. The researchers traced these categories back to the original statements and found that they were mentioned in specific sequence in the descriptions. For instances, "foreground activities" was often mentioned in the beginning, and "stimuli" preceding "examination of information content". Therefore, the 9 main categories were unified around a core category, i.e. "IE process". The resulting theory is shown in the general model of IE process (Fig. 1). It comprises the 9 components (categories) that each is enriched with all the possible situations (subcategories). The overall process can be further divided into three phases, i.e. pre-encountering, encountering, and post-encountering.

There are two major types of environments in which IE occurs, i.e. offline and online. The former refers to physical places such as homes, schools, libraries, stores, cafes, and movie theaters, etc., while the latter virtual places on the Internet such as social media, digital libraries, search engines, E-commerce platforms, Internet forums, and so on. These environments are built for different purposes which support users' foreground activities, including typical online information behavior like purposeful searching, purposeless scanning, exploratory searching, and browsing, as well as everyday routines (e.g. listening to music and shopping), social networking (e.g. chatting), and work and study (e.g. doing research and attending conferences). When engaged in the foreground activities, users may feel positive (e.g. happy and excited) or negative emotions (e.g. bored and frustrated).

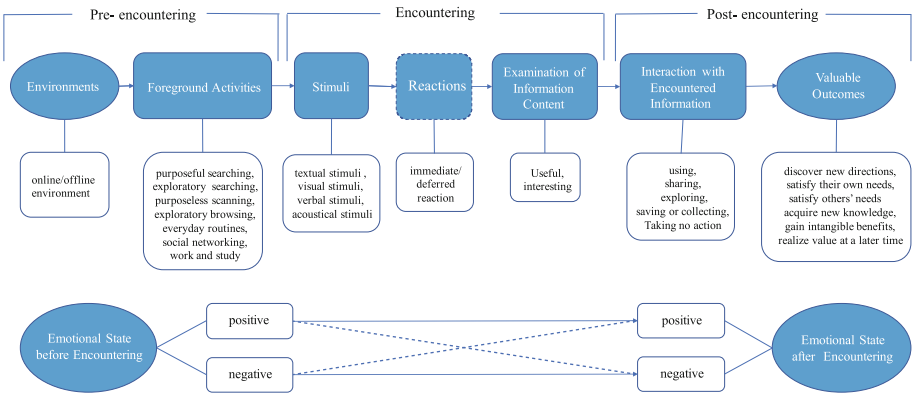

Fig. 1. The general IE process model

IE occurs when a stimulus in the environment that is irrelevant to the foreground activity attracts one's attention. Although written text is the most common type of stimuli, IE can be also triggered by visual (e.g. images, videos, and TV programs), verbal (e.g. conversation), or acoustical (e.g. music and radio) stimuli. After noticing the stimuli, users may make an immediate or a deferred reaction, depending on the urgency of their foreground activities, and examine the information content represented by the stimuli. The content may be determined as useful (i.e. solving an existing problem) or interesting (i.e. matching one's interests or leading one to new domains).

After encountering, users may interact with the encountered information in various ways, including saving or collecting, using, sharing, and exploring etc., or take no further action. It is common for users to adopt multiple ways for their interaction. The value of IE takes different forms to different users. The encountered information may help them discover new directions, satisfy their own needs, satisfy others' needs, acquire new knowledge, and gain intangible benefits (e.g. opportunities), or has its value realized at a later time. The IE experience may maintain or reverse users' emotional states or strengthen or weaken their initial emotions.

5 Discussion and Conclusions

A comprehensive understanding of how IE occurs is indispensable to taking advantage of encountering for more effective information acquisition. Despite the variety of existing IE process models, this study made a special effort to enrich this stream of research by contributing a general model that provides a panoramic view of this universal information phenomenon and combines the behavioral and affective components of IE processes. The significance of this study not only consists in the identification or clarification of some essential components of an IE process which have been ignored in previous models, but also the introduction of secondary data analysis to IE research.

5.1 The Value of the General IE Process Model

The model established in this study is characteristic of a multi-dimension (i.e. behavior and emotions) and multi-phase (i.e. pre-encountering, encountering, and post-encountering) demonstration of the process of IE. It provides important improvement for or addition to the previous understanding.

First and foremost, the general IE process model recognizes the roles of emotions. Previous models either failed to take the affective aspect into consideration [5, 9] or just deemed positive emotions after encountering a kind of valuable outcome [3, 25]. In contrast, the new model reflects the transition of emotional states as aroused by encountering. It is widely believed that emotions have an impact on humans' information seeking behavior [24]. As IE behavior is almost effortless, it is more likely to be emotionally charged [21]. Therefore, it is possible to facilitate or hinder IE by affecting users' emotions. The interaction between the behavioral and emotional dimensions during an IE process is a promising direction for future research.

Second, environments are treated as an independent component. They were usually mixed with foreground activities as the contexts of IE [5, 14]. It has been found that certain environmental characteristics are conducive to IE [26, 27]. Focused research on this component will generate more practical implications for designing IE-friendly environments.

Third, an inclusive categorization has been engendered for the stimuli that trigger IE. Although researchers have noticed the importance of stimuli [5, 6], there still lacks an in-depth understanding of the triggering mechanisms, let alone taking different types of stimuli into consideration. Besides, this model adds the possible reactions to the stimuli, suggesting the underlying selection between processing stimuli and proceeding with foreground activities.

Last but not least, the post-encountering phase is critical to realizing the value of IE. The phase was originally referred to as "capturing" [28]. The rise of online IE has greatly enriched the modes of dealing with encountered information. It is desirable that more powerful support is provided to encourage the immediate using of encountered information that can be easily forgotten.

5.2 Limitations

This study relied on secondary data analysis, which involves the reuse of the existing qualitative data from published studies, to explore the process of IE. The benefits of secondary data mainly consisted in the lower cost, larger samples, and diversified sources. However, this study was also limited to some extent by the data collection method. Since the 14 studies collected the IE incidents for their own research objectives, some of them failed to provide complete and detailed descriptions as desired in this study. There might lack certain important components in some descriptions, such as "environments" and "interaction with encountered information". The absent components might have not been collected or have been omitted in the papers. In addition,

not every CIT-based IE study disclosed their original data. Therefore, research data sharing should be encouraged to facilitate reuse as long as it does not invade the participants' privacy. In addition, the researchers plan to test the validity of the general IE process model which was generated in an inductive process, with real users in case studies.

Acknowledgement. This research has been made possible through the financial support of the National Natural Science Foundation of China under Grants No. 71774125 and No. 71420107026.

Appendix

The 3 representative types of IE incident descriptions are provided as follows:

Complete Incident: "While booking train tickets from Brussels to London on the Eurostar website, P5 noticed a carousel advertising free entry to London museums with the purchase of a train ticket. As she likes to visit museums, P5 clicked for more information and discovered 'you can go to galleries that I didn't even know existed in London. So it was pretty interesting to know.' The page mentioned a few attractions she would like to visit, but she was not sure if she would have time; she made a mental note of these attractions and booked her train tickets. The information she had encountered had the potential to be useful but this potential had not yet been realized" [22, p. 283].

Partial Incident: "R14 stated that she visits the Yahoo! Site because 'they have just a ton of random links'. She said she likes just 'clicking on something' and finding 'interesting things' that she 'wasn't intending to read about'" [19, p. 5].

Psychological Activities: "R20 explains that her negative feelings about incidental exposure to online news often were triggered by her thinking that she mostly finds unpleasant or doubtful information on the Internet" [21, p. 486].

References

1. Case, D.O.: Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior. Emerald Group Publishing, Bingley (2012)
2. Erdelez, S.: Information Encountering: An Exploration Beyond Information Seeking. Syracuse University, New York (1995)
3. Makri, S., Blandford, A.: Coming across information serendipitously – Part 1. J. Doc. **68**(5), 684–705 (2012)
4. McCay-Peet, L., Toms, E.G.: The process of serendipity in knowledge work, pp. 377–382 (2010)
5. McCay-Peet, L., Toms, E.G.: Investigating serendipity: how it unfolds and what may influence it. J. Assoc. Inf. Sci. Technol. **66**(7), 1463–1476 (2015)
6. Jiang, T., Liu, F., Chi, Y.: Online information encountering: modeling the process and influencing factors. J. Doc. **71**(6), 1135–1157 (2015)

7. Cunha, M.P.E.: Serendipity: Why Some Organizations are Luckier than Others Universidade Nova de Lisboa, Lisbon (2005)
8. Heaton, J.: Secondary analysis of qualitative data: an overview. ANS Adv. Nurs. Sci. **33**(3), 33–45 (2008)
9. Erdelez, S.: Towards understanding information encountering on the web. In: Proceedings of the ASIS Annual Meeting, pp. 363–371 (2000)
10. Lawley, J., Tompkins, P.: Maximising serendipity: the art of recognising and fostering unexpected potential - a systemic approach to change. In: The Developing Group 2008 (2008)
11. Flanagan, J.C.: The critical incident technique. Psychol. Bull. **51**(4), 327–358 (1954)
12. Gremler, D.D.: The critical incident technique in service research. J. Serv. Res. **51**(1), 65–89 (2004)
13. Sun, X., Sharples, S., Makri, S.: A user-centred mobile diary study approach to understanding serendipity in information research. Inf. Res. **16**(3), 492 (2011)
14. Pontis, S., et al.: Academics' responses to encountered information: context matters. J. Assoc. Inf. Sci. Technol. **67**(8), 1883–1903 (2015)
15. Rahman, A., Wilson, M.L.: Exploring opportunities to facilitate serendipity in search. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval 2015, pp. 939–942. ACM, New York (2015)
16. Foster, A., Ford, N.: Serendipity and information seeking: an empirical study. J. Doc. **59**(3), 321–340 (2003)
17. Makri, S., Warwick, C.: Information for inspiration: understanding architects' information seeking and use behaviors to inform design. J. Assoc. Inf. Sci. Technol. **61**(9), 1745–1770 (2010)
18. Dantonio, L.: Reciprocity and investment: the role of social media in fostering serendipity University College London, London (2010)
19. Yadamsuren, B., Erdelez, S.: Incidental exposure to online news. Proc. Am. Soc. Inf. Sci. Technol. **47**(1), 1–8 (2010)
20. Miwa, M., Egusa, Y., Saito, H., Takaku, M., Terai, H., Kando, N.: A method to capture information encountering embedded in exploratory web searches. Inf. Res. **16**(3), 487 (2011)
21. Yadamsuren, B., Heinström, J.: Emotional reactions to incidental exposure to online news. Inf. Res. **16**(3), 486 (2011)
22. Makri, S., Ravem, M., Mckay, D.: After serendipity strikes: creating value from encountered information. Proc. Assoc. Inf. Sci. Technol. **54**(1), 279–288 (2017)
23. Corbin, J.M., Strauss, A.: Grounded theory research: procedures, canons, and evaluative criteria. Zeitschrift Für Soziologie **19**(6), 418–427 (1990)
24. Case, D.O., Andrews, J.E., Johnson, J.D., Allard, S.L.: Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. J. Med. Libr. Assoc. **93**(3), 353–362 (2005)
25. Rubin, V.L., Burkell, J., Quan-Haase, A.: Facets of serendipity in everyday chance encounters: a grounded theory approach to blog analysis. Inf. Res. **16**(3), 488 (2011)
26. Martin, K., Quan-Haase, A.: "A process of controlled serendipity": an exploratory study of historians' and digital historians' experiences of serendipity in digital environments. Proc. Assoc. Inf. Sci. Technol. **54**(1), 289–297 (2017)
27. McCay-Peet, L., Toms, E.G., Kelloway, E.K.: Examination of relationships among serendipity, the environment, and individual differences. Inf. Process. Manage. **51**(4), 391–412 (2015)
28. Erdelez, S.: Investigation of information encountering in the controlled research environment. Inf. Process. Manage. **40**(6), 1013–1025 (2004)