Master's Thesis in Graduate School of
Library, Information and Media Studies

Participatory Digital Archives and Issues Relating to Their Contextualization: A Dialogue with the Archival Discourse

March 2021

201921658

ZHANG JIUPEI

Participatory Digital Archives and Issues Relating to Their Contextualization: A Dialogue with the Archival Discourse
参加型デジタルアーカイブスとその文脈化に関する諸問題
―アーカイバル・ディスコースとの対話―

Student No.: 201921658
氏名：張　九沛
Name: ZHANG JIUPEI

Unofficial archiving projects that embrace user-contributed material as the main content are not a stranger these days. As the outcome of these archiving projects, the participatory digital archives (hereinafter referred to as PDA) approach diversifies narratives and democratizes archives by allowing common users to contribute to archival collections. However, the practice of managing these archival materials from over-dispersed provenance does not loyally follow the core principles of traditional archival science such as the principle of *respect des fonds* and the principle of provenance. Admittedly, from a postmodern perspective, the primary goal of collecting archival materials in PDA projects is to diversify the narratives. However, from a traditional archival science perspective, the gathering procedures and organizing methods of PDA result only in fragmented narratives and a lack of contextual information, leading to difficulties in proving their authenticity and erratic evidential value.

Through case studies of several PDA projects and literature review of participation within archival discourse, this study explores the effects the PDA projects' activities have on archival materials, including patterns of participation, personnel composition, composition and presentation/visualization of contextual information, etc. The research findings affirmed the advantages of in-person activities in mobilizing participants and gathering contextual information. The research also acknowledged archivists' value in the archiving activities and concretized archivists' efforts in enhancing the evidential value of archival materials. A critical issue identified in this study is that the administrators in participatory digital archiving projects perceive archival materials' provenance in a different manner, because all those responsible for recording/archiving, collecting/solicitation, and digitization may be seen as creators. This problem affects the organization of archival collections and causes difficulties for users to understand archival materials. The author suggests that the archiving projects' administrators formalize the collecting procedures and take responsibility for and only for the authenticity of the archival materials, in order to promote the evidential value of the archival materials. Furthermore, inspired by the practices in the case studies and by the postmodern archivists' discussions, the author argues that it is necessary to broaden the scope of contextual information in PDA to encompass the entire archival process, as well as all the management, curation, and reuse activities.

Principal Academic Advisor: Tetsuya SHIRAI
Secondary Academic Advisor: Eduard BARYSHEV

Participatory Digital Archives and Issues Relating to Their Contextualization: A Dialogue with the Archival Discourse

ZHANG JIUPEI

Graduate School of Library, Information and Media Studies

University of Tsukuba

March 2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

A huge number of digital collections that allow users to participate in the process of soliciting, preserving, and disseminating documental information have appeared since the beginning of this century. Thematically, they are represented by such diverse topics like nostalgia (e.g. WISEArchive[1], Queens Memory[2], Mass. Memory Road Show[3]), natural disaster (e.g. Hurricane Digital Memory bank[4], Center for Remembering 3.11[5]), human-made disaster (e.g. the September 11 Digital Archive[6], Our Marathon[7]), politics (e.g. The Baltimore Uprising Archive Project[8]), history (e.g. Hiroshima Archive[9]), education (e.g. Digital Archive of Literacy Narratives[10]), etc. Common users can use this content to reinforce their individual identity by exploring family roots and strengthening their group's collective memory by sharing information about a social group's history with other people (冯惠玲, 2015). Furthermore, since most of those content belongs to an unofficial discourse, these resources can be utilized for political mobilization purposes, i.e., in the struggle for human rights (Caswell, 2014). Linguists use this online documental information to trace the changes in pragmatics (Girdharry, 2019); teachers who deal with this content may pursue knowledge for educational purposes. Historians also acknowledge the value

[1] https://www.wisearchive.co.uk/
[2] https://queensmemory.org/
[3] https://openarchives.umb.edu/digital/collection/p15774coll6
[4] http://hurricanearchive.org/
[5] https://recorder311.smt.jp/
[6] https://911digitalarchive.org/
[7] https://marathon.library.northeastern.edu/
[8] https://baltimoreuprising2015.org/
[9] http://hiroshima.archiving.jp/
[10] https://www.thedaln.org/

of these digital collections that document the lives of "grassroots communities" as a supplemental material for historical research (金光耀, 2015). Besides, in some cases, the digital collection project's activities appear to be significantly meaningful as a post-trauma therapy tool allowing people to share their individual traumatic experience with others. (Arthur, 2009; Carlton, 2016).

Notably, a number of such digital collections include in their names the term "archives" and have significant similarities with so-called "conventional archives" when they function as a social memory tool. However, obviously that by the origin and nature they are decisively different from the latter. Due to this fact, until now, these digital archives had been excluded from the archival science's scope. Taking a glance at "digital archives," everyone finds that that derives from "an archive" as a computer science term, widely accepted by the public due to the development and distribution of information and communication technologies in the last decades. Importantly, this new kind of archives tends to be vaguely perceived as everything that is preserved for non-current purposes. (Breakell, 2008).

On the contrary, "archives" in archival professionals' vocabulary is the natural residue of an administrative or executive transaction, that is, the impartial by-product of official affairs that have paramount evidential value. But Schellenberg's (1956) life-cycle theory of archives (especially the idea of archival appraisal) establishes that archival appraisal is oriented to future researchers' potential needs when selecting which records to be preserved, thus making archives no longer natural residues. The proactive involvement of archivists in shaping the mainstream's collective memory through archival appraisal brings archives' reliability into question. However, traditional archivists remain committed to maintaining the authenticity of archives through a diplomatic approach, with Luciana Duranti (1994) and the InterPARES Trust[11] project she has been leading as a director for more than 20 years appears to represent symbolically this traditional archival discourse in the age of postmodernism.

Be noted, even in this circumstance, the phenomenon of community archives of minority groups born under the influence of postcolonialism and answering the call for social justice, has already become an object of the archival discourse. Certainly, the discussion on those practices just barely exceeds the boundaries of so-called "collective archives," and the methods like crowdsourced submission, incorporated

[11] https://interparestrust.org/

in participatory digital collections, are still considered too radical to professional archivists. As a result, Flinn and Sexton (2019) describe the current status of participatory digital archives (hereinafter referred to as PDA) as "archival activities and spaces that occur outside traditional archive structures."

1.2 Problem Statement

The critical problem that prevents archival professionals from including the PDA into their sphere is that these documental collections attempt to organize materials that belong to multiple sources/creators. Naturally, it is impossible to adopt the provenance principle, the very cornerstone of archival science, to these materials. Traditionally, every archival fonds is represented by one or several rather constant creators, and the items in such an archive, being records of specific continuous activity, are highly interrelated to each other that forms a meaningful context. However, since PDA have numerous creators/contributors and lack biographical information on them, their provenance cannot be established in a concrete and authentic way. Many of such user-contributed items are standalone and merely co-related, which makes the narrative of the whole collection highly fragmented. What is more, PDA projects always encourage diverse narratives and storytelling that causes more difficulties when organizing such an archive thematically (Carlton, 2016, p. 14). As a result, the PDA lack features that furnish documental materials with evidential information of high quality.

Notably, in his introduction to the postmodern archives Ketelaar referred to the following expression of Niek van Sas: "postmodernism has not so much been the relativizing of truth but rather the multiplication of perspective." (Ketelaar, 2001, p. 132) These words appear to explain skillfully why the diversity of representations had replaced objectivity in PDA's selection criteria when deciding what to be preserved and why the evidentiality of those collections might be remained beyond question (Carlton, 2016, p. 14). Thus, the traditional archival discourse and the postmodern archival discourse have been developing without a significant interaction between them, looking for entirely different goals. Gilliland's (A. J. Gilliland, McKemmish, & Lau, 2016) archival multiverse theory has settled the divergence between the traditional and the postmodern archival paradigm on a certain level, but it did not encourage the convergence of the two. However, suppose we back to the first wave of deconstruction in the archival discourse and try to bridge it with

contemporary participatory practices; there might be an altered interpretation of provenance and contextual information, leading to a high evidentiality participatory digital collection. Huvila gives his experience on the construction of two digital historical archives that incorporate some levels of user participation and articulates that extending the contextualization of records management to "both records and the entire archival process" is crucial (Huvila, 2008). The most interesting point in Huvila's argument is *contextualization*, a term not commonly used in archival literature. *Context* in archival discourse is described as the background information related to the circumstances surrounding the archival process and the creators of the records that serve to identify the authenticity of the records. In conventional archiving processes, such information is passively collected, and archivists are habitually using terms such as keeping or preserving the context, rather than using the term contextualization, which implying archivists' agency, to describe the collection of contextual information. Huvila's study explains that actively collecting contextual information is necessary for participatory archives, which gives a good idea of rethinking the interpretation of context in participatory archives. The author will follow his concept but will adapt it to rather radical participatory practices. Therefore, contextualization in this study should be understood as all efforts that actively or passively enrich the context of a collection of archival materials.

Chapter 2

Literature Review

2.1 The Contested Archives

Despite the meaning that refers to the archival institution and the archives' generalized perception by the public, archives itself is a contested word in archival discourse. The dispute on the nature of archives – is it primarily evidence or memory? – have been lasting for decades, and there are no grounds to think that it is closing to the end. Famous British archivist Sir Hilary Jenkinson, for instance, gave the following definition in his *A Manual of Archive Administration*:

> A document which may be said to belong to the class of Archives is one which was drawn up or used in the course of an administrative or executive transaction (whether public or private) of which itself formed a part; and subsequently preserved in their own custody for their own information by the person or persons responsible for that transaction and their legitimate successors(Jenkinson, 1937, p. 11).

In his definition, Sir Jenkinson emphasized that the organic structure and the natural accumulation processes of archives are keys to judge whether a document belongs to archives or not. Particularly, he argued that archives is the natural residue of official transactions and forms an undividable whole by themselves. Keeping the organic links between the documents and their origins was another concern of Sir Jenkinson, who as his European predecessors stood against archivists' interference into the archiving process. However, the fundamental changes in socioeconomic

and political conditions since the beginning of the 20 c. questioned this traditional approach: the technological progress of office equipment (typewriter, copywriter, teleprinter, etc.) and the modernization of administration caused the unprecedented proliferation of paperwork in administrative affairs. Due to the huge amount of newly created official documents, it simply became impossible to preserve all of them. Therefore, archivists were forced to decide whether to keep or dispose documents transferred to their custody from governmental agencies. Schellenberg argued that archives had primary and secondary values, and the secondary values is not to serve their origins but the other researchers, especially historians (Tschan, 2002). In his definition of archives, he did not stress the processes of accumulation but the value of records instead (Schellenberg, 1956, p. 16):

> Those records of any public or private institution which are adjudged worthy of permanent preservation for reference and research purposes and which have been deposited or have been selected for deposit in an archival institution.

Hereby, Schellenberg gives his criteria of archives: it is a value for reference and research purpose. In other words, he admits the role of an archivist as not a passive custodian, but as a professional who actively selects documentary materials for future users. As Cook (2013) states, due to the introduction of such a new vision, a general perception of archives had also been changed from "Evidence" to "Memory". To be noticed, Schellenberg's definition limits the scope of archives to documentary materials from institutions or deposits in archival institutions. This definition excluded most grassroots social movements that aimed to preserve history and memories. That is why, the relevance of this definition had been frequently questioned by social activists since the 1970s.

Terry Cook (2013) inclusively summarized the evolution of archival studies in his archival paradigm shift theory that metaphorically divided the process into 4 phases: Evidence, Memory, Identity, and Community. In his opinion, archivists' participation in the social movements that pursued justice and equality in 1970s has converted archivists into radical social activists. As a result, archives began to be utilized in reinforcement of individual and collective identity. Thus, the most controversial topic of archives has changed to identity. After the 1990s, thanks to the development of information communication technologies, the cost and technical barrier to solicit and preserve information was dramatically reduced. Many communities (in

real-world or cyberspace) spontaneously started their own archiving activities and shortly introduced them to the public. Of course, those archiving activities have been noticed by archivists, and come into one of the top topics in archival discourse recently (figure shows the growing literature of community archives in major journals of archival science). However, archival professionals also noticed that those archival collections are not archives that comply with whether Jenkinson's or Schellenberg's definitions. As a result, the controversy of archives is still ongoing.

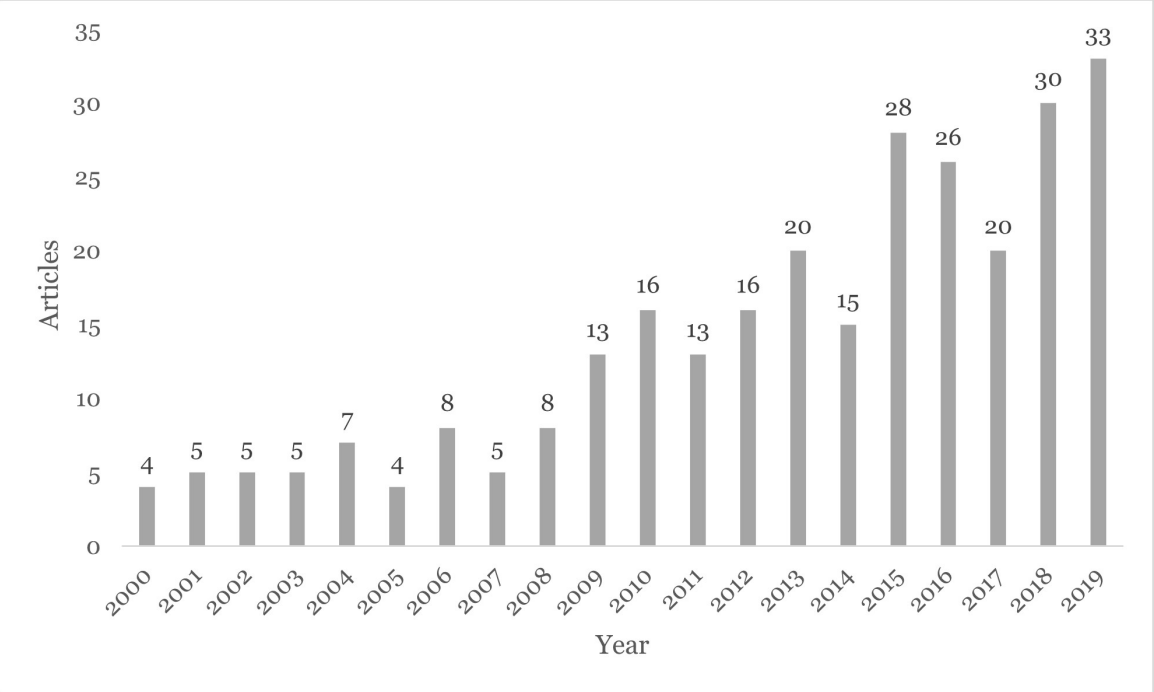

Figure 2.1: Articles on community archives in 4 major periodicals

Cook's perspective on emerging archival phenomena, which is still in a nascent stage like community archives indicating us to embrace a definition of archives other than evidence. Feng (2006) issues that at an abstract level, the essence of archives is solidified information. "Solidified" means here that the information once captured, its integrity cannot be undermined. To circular newly germinated archival phenomena into the scope of archival science, a practical definition of archives based on institutional archives seems not feasible to make any sense; thus, in this study, archives were defined as follow:

> Clear and definitive solidified information that directly generated and captured in the past human activities(冯惠玲 & 张辑哲, 2006, p. 6).

2.2 Community Archives

According to the author's survey, the term "community archives" might refer to the following documentary materials:

1. Historical files of a specific ethnic groups (in history and anthropology)

   The first type of community archives can be found in the early history or anthropology research aimed at specific ethnic groups. For example, Durnbaugh (1959) referred in his article to the Amana Community archive, a documentary collection of the former German colony in Iowa, USA. Pope (1970, p. 92) also described a part of the Aramaic documents left by the Aramean-Jewish military colony in the 5th century BC as a community archive. From this context, it could be concluded that the community archive documents the public, other than a specific family. Since "community" was once widely considered as a shorthand for ethnic minorities (A. Gilliland & Flinn, 2013, p. 4), community archives should be captured as archives of ethnic groups.

2. Official records of local self-governing districts (municipalities)

   In some countries, the bottom-level administrative entities are termed "community". Correspondingly, these administrative entities' official records are called community archives—for example, the Nanaimo Community Archive[1] and the Mission Community Archives[2] in Canada. In China, "community" can refer to the local autonomous government in the urban area; it names "Village" in the rural area (姜纪云, 2017). However, there is a distinction between community archives that refer to the local authority's records and the community archives that refer to community-based archives in the Chinese language. The former is called "社区 (shequ)," societal district; and the latter is called "社群 (shequn)," societal group (裘丽, 2016). It stands for the official discourse, which quite different from the general thinking of community archives as the repository of non-official discourse. Thus, archival scholars are merely not mentioned this kind of archives, especially in the research of community-based archives.

3. Community-based archives

[1] http://www.nanaimoarchives.ca/
[2] https://missionarchives.com/

According to Bastian & Flinn (2019), the first English literature of community archives that involved community members' participation was in 1942, when public libraries encourage residents to record their particular war experience. Still, the public library can be seen as an official backed institution, which means the content of the community archives are still narrated in a mainstream discourse. Affected by the new-liberalism and post-colonialism, the social movements pursuing justice and equality by minority groups like feminists, queers, ethnic minorities have founded a number of community-based archive projects to preserve documentary materials represented by their own discourse, opposed to the mainstream. Affected by the new-liberalism and post-colonialism, the social movements pursuing justice and equality by minority groups like feminists(e.g. the Feminst Archives North[3]), queers(e.g. Lesbian Herstory Archives[4]), ethnic minorities(e.g. South Asia American Digital Archives[5]) have founded a number of community-based archive projects to preserve documentary materials reperesented by their own discourse, opposed to the mainstream. In other words, community archives are propelled through the suppression of the mainstream. For example, documents of the Black Cultural Archives[6], a community archive of African and Caribbean groups in the UK, treat the *New Cross house fire* of 1981 no more and no less than "the New Cross Massacre." In the BCA's case, this community archive is the marginalized discourse's resistance to the official discourse. Notably, Bak (2016) in his critic of the Library and Archives Canada (LAC)'s Trusted Digital Repository standard remarked that the skepticism to the authorities across the minority groups existed "because their records have not been included in the archives". This also explains why minority communities prefer to keep their records under the control of community members.

Flinn (2009) argued that community archives' defining characteristic was the "active participation of a community in documenting and making accessible the history of their particular group" in their own terms. He articulated that participation was crucial to community archives but still left a free space to explain the "participation." Actually, according to the author's literature survey, community archives' levels of participation varied from case to case and hard to judge whether it is community archives or not from this aspect. However, every step the community members take (including participation) in

[3]https://feministarchivenorth.org.uk/
[4]https://lesbianherstoryarchives.org/
[5]https://www.saada.org/
[6]https://blackculturalarchives.org/

their archiving activity can be interpreted as the efforts to achieve the ultimate goal, that is, to narrate their story in their own independent discourse. Hereby, in this study, the author argues that the unofficial discourse is an inherent character of community archives.

During the survey, some overlapping across these three groups of "community-related archives" was founded. Aboriginal/indigenous archives (Ormond-Parker & Sloggett, 2012) or folks archives of ethnic minorities (田丽媛, 2016) can be given as an example of such overlapping.

The discussion above shows that "community archives" is still a fluid, complex concept in the archival discourse. In this study, while observing the community archives in the literature, the author primarily paid attention to their connection with participatory archives. He has also found that the most relative concept in this relation is "community-based archives." Henceforth, the term "community archives" in this paper is referred mainly to community-based archives characterized by unofficial discourse and active participation.

2.3 Participatory Archives

In 2008, Huvila firstly articulated the participatory archives' concept defining it in the following keywords: "Decentralized curation, radical user orientation, and contextualization of both records and the entire archival process (Huvila, 2008)." He has worked out this concept based on two digital archive projects, which he dealt with and which allowed non-professional users to contribute their knowledge of the historical materials and share the authority of control or curation with non-specialists on some levels. In 2010, in his another paper, Huvila stated that "participation is very limited if it is conceived as letting some others to play with (some of) my toys in my sandbox and not reaching out for the toys of other people in their own sandboxes (Huvila, 2010)." This metaphor seems to divide participatory archives by the authority of control. At the Society of American Archivists Annual Meeting 2011, Kate Theimer gave her definition of participatory archives:

An organization, site or collection in which people other than archives professionals contribute knowledge or resources, resulting in increased un-

> derstanding about archival materials, usually in an online environment (Theimer, 2011).

Theimer's definition is precise and practical, circled a specific area of the participatory archives, which can be accepted by traditional archivists. In Theimer's definition, participation is limited to contributing resources other than content, such as social tagging and commenting, or crowdsourced transcription and description. Of course, there is no sharing authority of control or curation in this pattern of participation. Moreover, Theimer focused on the participation seen in an online environment only.

The disagreement between Huvila and Theimer represents the current state of things in regards to the treatment of participatory archives in archival discourse. Huvila acknowledged that Theimer's definition was practical and realistic but insisted that archivists should "look beyond the actual things that are being done in the present participatory archives and to at the processes and principles that might be about to change (Huvila, 2011)." In 2015, Huvila reported on the review of archival literature, separating the discussion of participatory archives into three parts: Management, Empowerment, and Technology (Huvila, 2015). Under "management" cases that allowed participators to manage the contents of the archival material were represented; under "empowerment" there were cases that would enable participators to contribute their resources in order to empower the meaning or understanding of the original archival material; under "Technology" he put technical discussions around participatory archives. Through the report, the two models (management and empowerment) of participation are discriminated by whether the participators engage in the authority of control of the archival material. To be noticed, there has been a slight shift between the concept of "decentralized curation" and the concept of "management." The term "decentralized curation" means in participatory archives that the participators would be involved with the curation of archival materials and conduct some control on it. However, suppose we perceive curation/control as management. In that case, it is not contrary to empowerment since the point of view of empowerment is primarily judging whether participatory archives are user-contributed or not. Fortunately, Theimer's definition lines the boundary of participatory archives that is already incorporated in the traditional archival discourse. We can take the counterpart as the radical activism (Flinn & Sexton, 2019) of participatory archives, which is not admitted by traditional archivists.

2.4 Digital Archives

In 2015, the Encyclopedia of Archival Science listed four meaning that the phrase "digital archives" have: 1) collections of born-digital records, 2) websites that provide access to collections of digitized materials, 3) websites featuring different types of digitized materials around one topic, and 4) web-based participatory collections (Duranti & Franks, 2015, p. 157-160). Especially, the last understanding of digital archives is highly related to participatory archives. Web-based participatory collections were described as "collections that actively solicit online-contribution," "such collections may be comprised entirely or in part of user-contributed material." Compare to the radical activism of participatory archives, "web-based participatory collections" is a full subset of "participatory archives," which emphasized online as the method of solicitation. Identically, the two cases in Carlton's thesis (2016), the 911 Digital Archive[7] and the Our Marathon[8], are both belong to the category of web-based participatory collections.

In 2011, Kasaba pointed out that "digital archives" in Japan might refer to: digitized archives, digitized collections of cultural heritage (digital cultural heritage), archives of distributed digital data, and born-digital records (笠羽晴夫, 2011). In 2018, Tamura et al. used PDA to describe their digital collection of the Hiroshima Atomic Bombing (田村賢哉, 秦那実, 井上洋希, & 渡邉英徳, 2018). In this project, the solicitation happened in an offline environment and finally contributed to the digital collection, which can be freely accessed from the internet.

Look at the usage of the phrase "Participatory Digital Archives," You will find that the word "digital" only describes a collection's type, not the process of accumulating materials or solicitation environment. Therefore, based on the description of the web-based digital collection, PDA in this study can be defined as follow:

> Digital collections, comprised entirely or partly of user-contributed content, actively solicit archival material from the people who identify themselves as stakeholders of the archive's content.

[7] https://911digitalarchive.org/
[8] https://marathon.library.northeastern.edu/

Chapter 3

Aims of the Study

3.1 Aims of the Study

This study primarily aims to investigate the efforts for contextualizing the collections of PDA and explore the literature related to the participation in the archival discourse, thus aiming to extend the interpretation of context and, eventually, achieve a better practice of contextualizing those collections.

3.2 Research Questions

To archive the research objectives, the following research question must be clarified previously:

1. How did the archival materials have been accumulated?

   The accumulation of archival materials includes the methods of solicitation, patterns of participation, and the participation events' details.

2. What kind of contextual information is primarily represented?

   Drill on the existing contextual information of participatory archives to evaluate and analyze.

3. How did the metadata strategy and implementation affect those collections?

   Metadata is considered the most critical technological aspect that affects electronic records' evidentiality in archival science. Participatory digital archives,

which are also based on electronic media, should not be excluded.

4. What role do archivists play in the archiving activities?

   This question connect to the model of participation and the transformation of archival professionals.

5. How PDA should be positioned in archival science?

   Clarify relating concepts in archival science and introduce approaches of archival science to PDA.

3.3 Significance

This study will encourage future PDA projects to improve the contextualization of archival materials in participatory collections and may result in a promotion of PDA in evidentiality. This study also attempts to bridge archival activism and the guiding principles of archival science, the principle of provenance, and the principle of *respect des fonds*.

Chapter 4

Methodology

4.1 Research Method

This research examines the given cases through website surveys and literature analysis. The research questions will be discussed based on the case study and archival literature.

Website survey

1. Understand the structure and provenance of archival materials;
2. Clarify the technical composition and accessibility as aspects of digital curation;
3. Evaluate the metadata.

Document analysis

1. Review the archival literature relate to participation to position this study in archival science;
2. Review the literature related to the cases in this study to get the detail of archiving activities.

This study will primarily approach the objectives by observing the cases of PDA, open to access from the internet. Thus, through a website survey, it is rather

feasible and efficient to grasp the basic information of an archival collection. The first step will be to visit the digital collection webpage and check in detail their introduction/about page, the category, file list, metadata scheme, etc. The next step will be collecting metadata from the digital collections by using crawlers under the behavior of fare use. This metadata can be used to analyze archival materials' provenance, community involvement in the archiving activity, and metadata quality.

Literature analysis has two different goals. Firstly, it should be noted that archiving activities are usually not shown on their web page of such projects but in news reports and academic papers. Naturally, that literature is crucial to understand the process/procedures of solicitation, guiding policy, personnel status, and decision-making details during the project. Secondly, this kind of analysis aims to review archival literature related to participation in digital archiving and connected topics. The literature itself is represented mainly by the four major periodicals of archival science - *Archival Science, American Archivist, Archivaria, and Manuscripts and Archives,* which can stand for the majority of the archival discourse.

4.2 Limitations

As a result, the reviewed materials will be limited mostly by English language texts that may not cover all aspects of community archives, participatory archives, and archival activism in archival literature. For example, archival science in China was also influenced by postmodernism from the 1990s, and topics like oral history, community archives, and participatory collections have also been discussed in Chinese literature (Lian, 2016, p. 116-117). However, due to the unique political, societal, cultural backgrounds and archival traditions in China, the practices and discussions in Chinese literature are of different character that cannot be compared with the West's archival discourse. There are also some obstacles to retrieving relevant Japanese materials because of the language barrier that becomes obvious when translating abstract concepts such as "archival activism" or "archival multiverse". Therefore, this study will be primarily anchored in English literature to avoid potential conflict between different languages.

The research method relies on existing literature to conduct analysis, which limited the selection of cases and resulting in a bias on the average "quality" of PDA in this study. Generally, prominent participation practices always get more atten-

tion across the public and left more recordings of themselves than others. Since the amount of related literature is used as a criterion for selecting cases, it is inevitable that the quality of cases introduced in this study will, without a doubt, exceed the average level. However, the problem that the selected cases can not represent the majority of PDA will not obstacle the research because this study aims to make future participatory digital collections better, not to advise on the contextualization of existing collections. This study's conclusion based on the prominent samples of PDA will be illustrative when instructing future archiving projects to avoid precedent mistakes. For the same reason, other cases of PDA might be mentioned in the discussion chapter.

4.3 Cases in this Study

Based on the research objectives and the author's understanding of PDA (which discussed in chapter 2), the salient points in selecting cases were: characterizing the community involvement (participation) in archiving activity; a certain level of public influence (mass media exposures, awards, etc.); availability of massive literature, sufficient for a meaningful analysis of the archiving activity.

1. Mass. Memory Road Show[1]

   The Mass. Memory Road Show (hereinafter referred to as MMRS), initialed by archivists and public historians in University Archives and Special Collections in the Joseph P. Healey Library at the University of Massachusetts in Boston, is a statewide event-based participatory archiving program that documents people, places, and events in Massachusetts history through family photographs and stories. Since its very beginning in 2004, the MMRS has held 57 archiving events (including 55 in-person events and two online events) and collected 12,351 items in total.

2. Center for Remembering 3.11[2]

   The Center for Remembering 3.11 (commonly known as recorder311) is a participatory archiving program that aims to disseminate information as well as record the restoration and recovery process after the Great East Japan Earth-

[1] https://openarchives.umb.edu/digital/collection/p15774coll6
[2] https://recorder311.smt.jp/

quake. The Sendai Mediatheque (which is a public cultural facility affiliate with the Sendai Municipal Government) proposed this project on March 22nd,2011, and the first in-person event was held on May 3rd,2011. This project's unique participation pattern is that participators are invited to an in-person event and learn some basics of making records. After participators are ready to submit their documentary materials, there is an acceptance dialogue between the project staff and the participator before their contributions have persevered. The Center for Remembering 3.11 currently holds 573 items, including mixed text/picture contents, videos, and sounds.

3. Our Marathon[3]

   Our Marathon is a crowdsourced archive of pictures, videos, stories, and social media related to the Boston Marathon Bombing on April 15th,2013, which was initially founded by the faculty and students affiliated with the College of Social Science and Humanities, Northeastern University. This project solicits documentary materials related to the bombing primarily via online crowdsourced submissions and relies on in-person events to involve more participators. Our Marathon ended active solicitation in April 2014; after migrated to the Digital Repository Service (provided by the Northeastern University Library), this project ended all solicit activity and turned into a static digital collection, including 22 collections, 7876 items in total.

[3]https://marathon.library.northeastern.edu/

Chapter 5

Survey Outcomes

5.1 Mass. Memory Road Show

5.1.1 General Description

The Mass. Memories Road Show derives from the "Massachusetts Studies Project" (MSP)[1], which provides educational materials in Massachusetts's history and culture. Therefore, the MMRS is naturally aiming at "collecting digital surrogates and personal annotations of locally held primary sources that document people, places and events in Massachusetts; and developing a searchable online repository of sources that could be used for educational purposes at all levels." However, the organizers discovered that their events had profoundly promoted the community members' connections throughout the whole region and incorporated community building as their key goal of the project promptly.

The project has been launched in 2004 and has already held 57 events in total by now. In principle, the MMRS only accepts materials from the in-person events, which require participators to fill out the submission forms under the event staff's advice, share the story of the material they submitted on video camera, and take a "keepsake photo" to record their participation. Then they will get a consultation from archivists and historians on preserving family photographs or identifying their photographs (MMRS, 2016). Nevertheless, the COVID-19 Pandemic halted all the in-person events, and it forced the MMRS to turn to online solicit. However, the

[1]See about page: http://blogs.umb.edu/massmemories/about-the-mass-memories-road-show

MMRS still ask participators to fill out the submission form (on the webpage) and have a short virtual meeting with the MMRS staff[2].

A typical workflow of a Road Show starts from the application. The local community that wants to have an event must contact the MMRS staff. Once accepted, a series of meetings will be held to finish the preliminary works: select a date and location, recruit volunteers and participators, find local partners, train all the volunteers, etc. On the event day, the MMRS team will bring technical equipment and staff: videographer(s), archivist(s), historian(s)/reference librarian(s), and their experienced volunteers. They will work together with the local volunteers. Participants need to fill out the registration forms, including a personal information form and a declaration to grant the right that allows MMRS to use the materials they provided in the event. The volunteers will then help participants complete a form about their materials, title, location, time or age, people who associated, description, and whether the participant needs help in identifying their materials. The materials will be sent to the scanning station and digitized; photos will be scanned, and materials that cannot be scanned (e.g., handcrafts) will have their photographs be taken. After digitization, the originals will be returned to the contributor immediately. The participants will then be interviewed to share their stories about the materials they provided, and the video record of the interview also will be preserved. The following step is to take a high-resolution keepsake photograph of the participants to record their presence, and the paperwork of the keepsake photograph will also be done simultaneously. By far, all the information needed by the digital collection should have been collected. The MMRS event provides additional value for the participants. Professional archivists will be present at the event, and participants can consult with them on issues like how to preserve photographs or other records in the family. Experienced historians and reference librarians will also be ready to help participants to identify the materials they shared. After the event day, the MMRS will collect all the records during the event planning and execution, as well as the materials the contributors provided. Those records will be permanently preserved by the University Archives & Special Collections at Joseph P. Healey Library, UMass Boston. The user-contributed materials will be input by the professionals in the host institution and be available on the website in 3 months from the event day.

Until now, the whole collection was comprised of 58 groups (this number exceeds 57 because the Hebrew Senior Life event was held two days and organized sepa-

[2]http://blogs.umb.edu/massmemories/stuck-at-home-show-pilot-project

rately), 12351 items in filetypes of tiff, jpg/jpeg/, and mpeg. Details refer to the appendix.

5.1.2 Personnel and Partners

The thesis gives the survey outcomes in two parts: the personnel and partners information for the whole project and a single MMRS event.
The MMRS team in the University Archives & Special Collections of the Joseph P. Healey Library consists of archivists (whether permanent or interim position), project managers, and student assistants. The team is responsible for the digital collection, including inputting the user-contributed materials and create metadata for digital objects. The library primarily supports the operation of the MMRS project and the digital archives. The project also receives financial supports from the Patricia C. Flaherty'81 Endowed Fund at UMass Boston, Institute of Museum and Library Services[3], LYRASIS Catalyst Fund[4], and the National Endowment for the Humanities[5].

The MMRS team and volunteers staff the MMRS events. The volunteers recruit by both the local community and the MMRS team. The MMRS team will provide certain train sessions for all the volunteers before the event day. Historians from the local historical society and librarians from public libraries may staff in the Road Show. Other partners like TV local stations can help advertise the event, and the local government or senior center can provide transportation to events (especially for seniors)(MMRS, 2016).

5.1.3 Structure

The MMRS digital collection is organized in a plain structure. Users can access the collection of a specific event from the navigation page, and the hyperlink will lead to a search, which keyword is the Road Show name, across the whole collection by the built-in search engine. Users can also search for keywords like date, location, name to get related results.

[3]https://www.imls.gov/about/mission
[4]https://www.lyrasis.org/about
[5]https://www.neh.gov/about

5.1.4 Provenance

The items in the MMRS digital collection can be divided into two parts by provenance. Firstly, user-contributed materials, the provenance of those items are their contributors. The MMRS set a limit that each contributor can only submit up to 3 items on an event day; thus, the whole collection's provenance is distributed and complicated. Secondly, the items that were captured on the event day, the interview videos, and the keepsake photographs. The provenance of those items is the MMRS team, which exists constantly and continuously.

5.1.5 Technical Composition

The MMRS digital collection is hosted on CONTENTdm[6], a software that allows people to build, preserve, and showcase their digital collections. CONTENTdm is provided as SaaS[7] by the Online Computer Library Center[8]. Therefore, the MMRS team doesn't need to have a physical server or allocate hardware resources for the digital collection; CONTENTdm also performs data backup periodically, making a robust infrastructure for the digital collection. CONTENTdm using XML to store metadata and a built-in text-based search engine to perform searches. XML provides extreme flexibility in metadata support; text-based search engine (same as web search engines) which enabled search across all metadata fields gives users a no-barrier experience in information retrieval. CONTENTdm can hold documents, images, videos, and audio files in all formats and incorporate content on third-party providers, such as Vimeo. In the MMRS project, the scanned files are stored in TIFF images, the photographs are stored in JPG/JPEG images, and the videos are stored in MPEG, which host on Vimeo[9].

5.1.6 Metadata

Metadata fields in the MMRS digital collection show in Table 5.1.

[6]https://www.oclc.org/en/contentdm.html
[7]Software as a Service
[8]Commonly OCLC, a major provider of Library Information System, see https://www.oclc.org/en/about.html
[9]Video hosting & sharing service provider, see https://vimeo.com

Table 5.1: Metadata Fields of MMRS Digital Collection

Metadata Fields	Notes
Title	Provide by the contributor; free-text
Description	Provide by the contributor; free-text
Contributor Name	Provide by the contributor; free-text
Date of Original	Optional; provide by the contributor;free-text
Decade	Respect to Date of Original; fixed format
Location Names	Optional; provide by the contributor; controlled vacabulary; fixed format
People	Optional; provide by the contributor; free-text
Road Show Name	Respect to the event; free-text
Type	Controlled Vacabulary; Image/Video
Format	Controlled Vacabulary; tiff/jpg/mpeg
Publisher	University of Massachusetts Boston, Joseph P. Healey Library
Rights	Copyright declaration; free-text
Identifier	Assigned by the MMRS team; fixed format
Road Show Date	Respect to the event; free-text

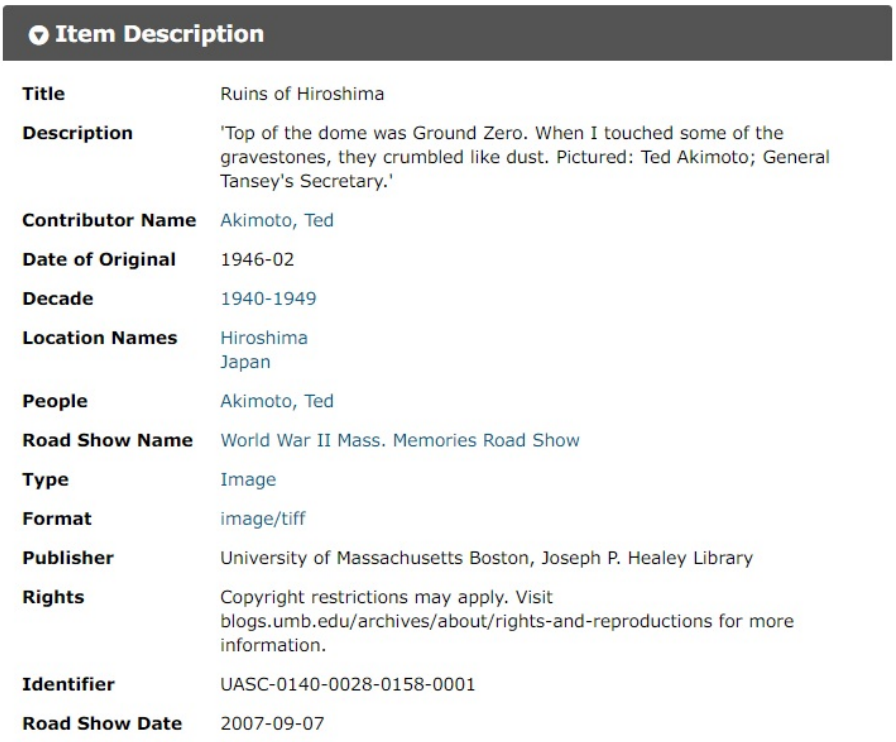

Figure 5.1: Example of MMRS's metadata fields

5.2 Center for Remembering 3.11

5.2.1 General Description

On March 11th, 2011, the most powerful earthquake ever recorded in Japan struck the Pacific coast of Tohoku. The earthquake and concomitant tsunami caused thousands of casualties and a level 7 nuclear accident [10]. This disaster shocked everyone in this area, including Kai, a Sendai Mediatheque staff. Three days later, he proposed a plan to make a place for ordinary people who experienced this tragedy to record and disseminate their feelings, words, and imagery through multiple media forms (佐藤知久, 甲斐賢治, & 北野央, 2018, p. 75-79). On March 22nd, the proposal was approved internally and turned to carry out. This project aims to disseminate information and the restoration and recovery process after the Great East Japan Earthquake, which is on the original concept's trajectory. The Center for Remembering 3.11 (also known as recorder311) opened on May 3rd, 2011, with an interim physical sphere on the second floor of the Sendai Mediatheque. Since the 7th floor, which was damaged by the earthquake, was still under repair, the recorder311 staff arranged a media studio (called Wasuren! Studio) on the second floor's open space to welcome participants (佐藤知久 et al., 2018, p. 86-87). The workflow of the recorder311 is quite different from most participatory collections, which solicit historical material. The recorder311 asks participants to record the restoration and recovery process after the earthquake, which does not happen yet.

To get their records accepted by the recorder311, participants need to write an application form, including personal information and their activity plan on recording (佐藤知久 et al., 2018, p. 117-125). Once the application was accepted, the participants will be gathered at the Wasuren! Studio to learn some necessary information about the recorder311 program: missions, guidelines, potential supports from the recorder311 staff, etc. After finishing the paper forms about copyrights, the participant officially registered as a recorder311 participant and can start recording activity under the support from the recorder311 staff. During the recording period, participants will keep in contact with the staff. Once both the participant and the staff acknowledged that currently recorded materials are qualified to submit, the recorder311 team will hold an acceptance hearing to grasp the contextual information of those materials: date, location, person, motivation to recording, and other backgrounds of the recording activity (佐藤知久 et al., 2018, p. 91-94). After the

[10] https://en.wikipedia.org/wiki/2011_Tohoku_earthquake_and_tsunami

acceptance hearing, the submitted records, contextual information, and provenance information will be input into the Sendai Mediatheque's database system, which has already served the Sendai Mediatheque for years (佐藤知久 et al., 2018, p. 97).

The Sendai Mediatheque also hosted serval parallel projects around the recorder311 program (佐藤知久 et al., 2018, p.216-219). Project Kangaeru Table aims to offer people a place to exchange their thinking on the restoration, local community, and other topics. It locates on the first floor's open square, an open space that can be see-through from the entry. Two programs were hosted at the Kangaeru Table: Tetsugaku Café and Shiga Rieko Lecture. The major topic of the Tetsugaku Café is the recording activity of the recorder311 program; participants gathered to review their activities and discuss future plans to make their recording better. This program started in June 2011 and still ongoing (佐藤知久 et al., 2018, p. 220-224). Shiga Rieko Lecture is a lecture that taught techniques about interviewing and photographing to help the recorder311 participants improve their practices in recording activities. The lecture started in June 2011 and ended in March 2012 (佐藤知久 et al., 2018, p. 225-229).

The Wasuren! Studio also hosted other programs. Salon de Wasurennu, which started in July 2011, aims to promote the connection between participants and relieve the mental stress accumulated during the recording activity by creating a sense of belonging and reducing loneliness. Kiokubu was derived from the Salon de Wasurennu, which shares mutual participants (佐藤知久 et al., 2018, p. 229-232). Kiokubu focuses on the editing technology of videography (佐藤知久 et al., 2018, p. 232-233). Kiokubu also held the first reuse event of the digital collection, a rehearsal screening "Hoshi to Michi" in October 2011 (佐藤知久 et al., 2018, p. 234).

The regular reuse events "Hoshi to Michi" (from March 2012) and the following "Koe Cinema" (from August 2012) are screening parties held in the theater, which locates on the 7th floor. The prominent part of the recorder311 program is that the staff try to capture all the in-person events related to the program. All the events mentioned forehead are documented in some forms, might be an event report or video record of the event, even the schedule of reuse events can be found in the collection. It shows that the recorder311 is not a static collection but a dynamic, ever-evolving participatory collection.

By July 7th, 2020, the recorder311 digital collection holds 573 items, containing

251 mixed text image articles, 288 videos, and 34 sound files. Part of those items was organized in 28 series by provenance.

5.2.2 Personnel and Partner

The Sendai Mediatheque's employees primarily staff the recorder311 program, and there is no archivist in the team (佐藤知久 et al., 2018, p. 108). Satou et al. (2018, p. 112-113) identified five major roles that are indispensable in the staff, which is:

- One who has video & audio recording/editing equipment & software skills that can support the recording technology in the project;
- One who is experienced in information dissemination via websites, blog, and social network service that able to in charge of reuse activities;
- One who can organize and categorize records into a context that may lead the compilation/curation of user-contributed materials;
- One who is competent to carry out the repeated & detailed work of compilation/curation under the leader's instruction;
- One who is skilled in communication that can facilitate community members' participation and maintain public relations.

Additionally, staff and participants' boundary is not clear since they are not mutually exclusive but always overlapping and exchangeable. The staff themself is a community member who experienced the earthquake and concomitant tsunami, and they also desire to narrate and record their feelings.

The Sendai Mediatheque fully endorsed this project; they provided personnel, event venue, technical support in the database system and web server, part of the technical equipment[11], advertising, and information dissemination (佐藤知久 et al., 2018, p. 81-82); external partners from academic institutions and industrials provided technical equipment; the local newspaper helped in advertising this project to the local community (佐藤知久 et al., 2018, p. 130); the Kingdom of Netherlands sponsored the bilingual program that translates their digital collection to English.

[11] https://recorder311.smt.jp/support_member/

5.2.3 Structure

The collection was primarily deposited in the database by format and also can be accessed by activities. The records of a series of recording activities formed a subcollection, whose name is "series." Users can browse the series page to locate records of a specific event or program. Especially, the recorder311 website provides intuitive methods to access their materials: by visualizing the records on maps and timelines.

5.2.4 Provenance

For all 573 items in the digital collection, 557 of them were labeled with at least one creator, which completeness is 97.2%. Across the 557 items, 147 creators were identified, including the recorder311 staff, individual participants, and the member of juridical person participants. Besides, 399 items were labeled with at least one contributor, and a total of 76 contributors were identified.

However, the provenance of the recorder311 collection can not be easily covered by metadata. For example, in an item of the Kagaeru Table event on September 22nd, 2012, the "creator" labeled in metadata doesn't accurately refer to the one who directly created the record but may refer to the one who originally collected the material[12]. This situation also recurs in other items, especially when the contents are transcriptions of conversation or discourse.

By observing the metadata field "creator", the recorder311 staff contributed to 194 items, exceeding one-third of the labeled items. Considering the situation that the creator of some records was wrongly labeled as the original collector, it can be asserted that the most prominent provenance of the collection is the recorder311 staff.

5.2.5 Technical Composition

The digital collection was preserved on the legacy database system of the Sendai Mediatheque, the file server equipped with RAID 5 hard disk array, and the data in

[12]https://recorder311.smt.jp/blog/27342/

the file server have a fully redundant backup, which is also equipped with RAID 5 hard disk array. What's more, the data have another backup in tape drives, which means it is triple secured. Besides, an uninterruptible power supply (UPS) system was adopted for the file server and profoundly strengthened the hardware's stability (佐藤知久 et al., 2018, p. 97).

The website was a WordPress[13] customization, which enabled features like metadata, keyword, and geolocation. To incorporate the contextual information collected in the acceptance hearing, the recorder311 staff didn't adopted a fixed template for records. The widely used template, which is standalone record plus metadata, was intentionally abolished by the recorder311 staff (佐藤知久 et al., 2018, p.99).

5.2.6 Metadata

The metadata scheme and the status is shown in Table 5.2. The author calculated the weighted average IDF to quantify the level of informativeness of the metadata filed, especially in search/retrieval. IDF stands for inverse term frequency, used in informatics studies to measure that if a term suit for search/retrieval keyword (Manning, Raghavan, & Schutze, 2009). The equation of IDF is shown below, in which N is the total number of documents, df is the number of documents that the term appeared.

$$idf = \log_{10} \frac{N}{df_t} \tag{5.1}$$

In this case, the author calculates all the terms that appeared in the metadata fields and the average IDF of terms to evaluate the metadata field's informativeness; the calculation is defined below.

$$idf_{wa} = \frac{\sum_{t=1}^{v} \log_{10} \frac{N}{df_t} \times df_t}{\sum_{t=1}^{v} df_t} \tag{5.2}$$

[13] https://wordpress.com/

Table 5.2: Metadata Status of Center for Remembering 3.11

Metadata Field	Type	Total	Valid Total	Max Freq	Min Freq	Ave Freq	Ave IDF	Completeness
キーワード	625	4806	4806	384	1	7.7	1.3	100.0%
きろくしゃ	147	573	557	96	1	3.8	1.8	97.2%
シリーズ	26	573	246	70	1	9.5	1.5	42.9%
さんかしゃ	76	573	399	61	1	5.3	1.6	69.6%
きろくび	367	573	555	24	1	1.5	2.5	96.9%
きろくばしょ	244	573	554	108	1	2.3	2.0	96.7%
タイトル	573	573	573	1	1	1.0	2.8	100.0%
URL	573	573	573	1	1	1.0	2.8	100.0%

Figure 5.2: Example for Recorder311's Metadata Fields

5.3 Our Marathon

5.3.1 General Description

The Boston Marathon began in 1897, is an annual marathon race hosted by several cities in greater Boston metropolitan region. It is traditionally held on Patriots' Day set in the commemoration of the American War of Independence[14]. On April 15th, 2013, two Chechen Kyrgyzstani-Americans detonated two homemade pressure-cooker bombs near the race's finish line[15]. The terror attack killed three (including a nine-year-old boy) and injured 264 people[16]. During the following manhunt, the terrorists killed an MIT policeman and severely wounded two police officers, one of whom died a year later.

Elizabeth M. Dillon and Ryan Cordell, faculty of the College of Social Science and Humanities (CSSH), Northeastern University (NEU), proposed a crowdsourced participatory archive that aims to mend and strengthen the fabric of their community to the college for project funding right after the bombing. In May 2013, the team received the grant and hired several graduate students from the English and History Department. The team also found local TV & radio stations and newspapers as partners to advertise this project and help the solicitation. A particular oral history project was initiated based on the co-operation with the WBUR[17], which is experienced in oral history projects. At the same time, the team released an Omeka[18] website and allowed crowdsourced submissions via the internet (McGrath & Peaker, 2018, p. 20-21). The crowdsourcing submission process described below:

- Select the file type of the submitting record;
- Choose the location of the record took place;
- Choose the date of the record;
- Fill out personal information (optional);
- Final consent on reuse permission.

[14]https://en.wikipedia.org/wiki/Boston_Marathon
[15]https://en.wikipedia.org/wiki/Boston_Marathon_bombing
[16]https://www.ketv.com/article/boston-honoring-marathon-bombing-victims-with-acts-of-kindness/32154489
[17]Local radio station, see https://www.wbur.org/inside/highlights-history
[18]An open-source web publishing software, see https://omeka.org/

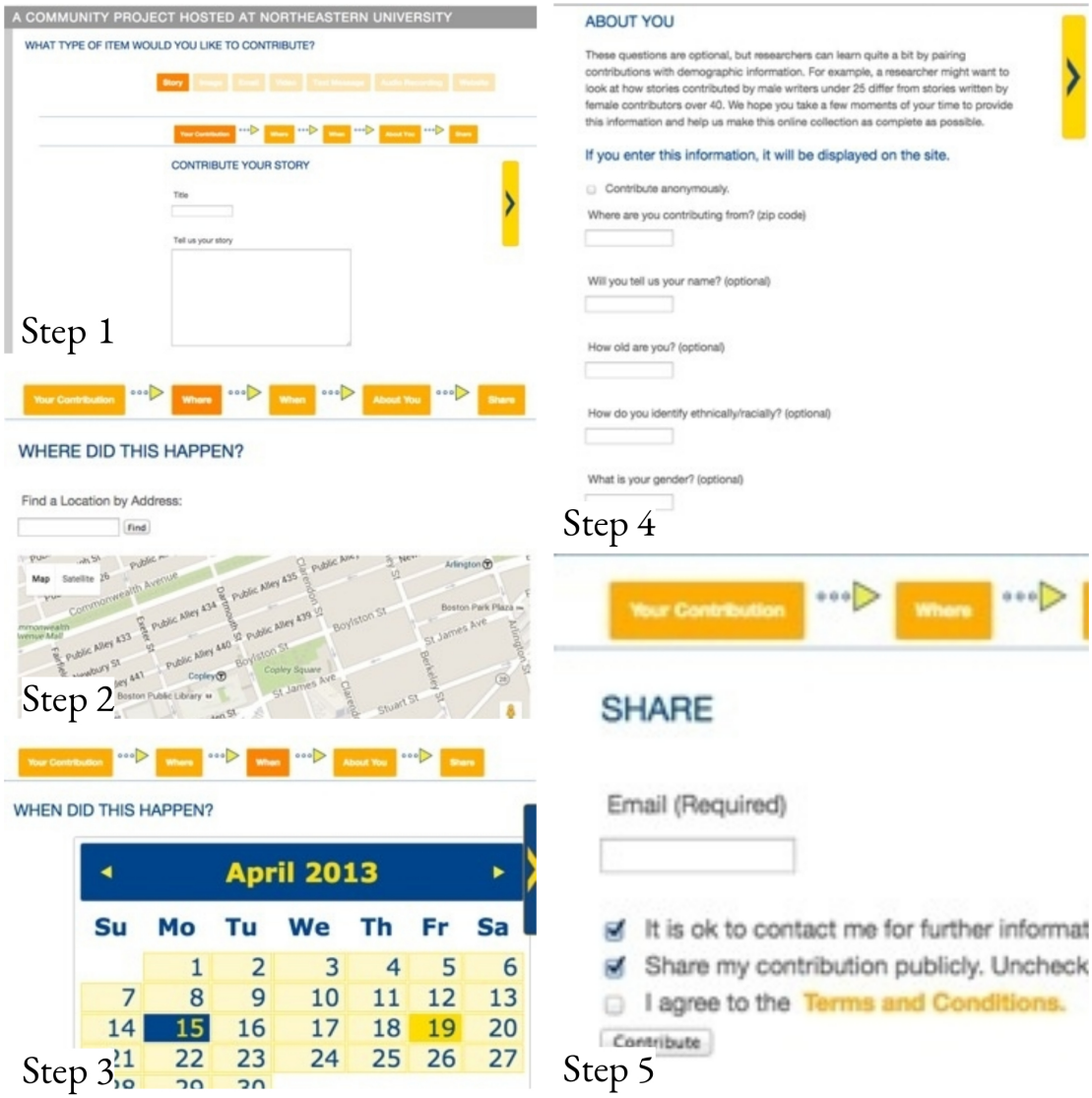

Figure 5.3: Our Marathon Crowdsourced Submission Page

With the help from the Archives & Special Collections division of the NEU Library, the team hosted an internship program for students in the Simmons College's MLIS program[19]. In June 2013, under the support of the NEU Library's metadata specialist, the team concluded the metadata scheme for Our Marathon and started metadata cleaning for collected materials. Simultaneously, the co-operation programs with the WBUR and the WCVB[20] were set in motion to preserve audio and video records, and the Boston Globe[21] opened an entry to the crowdsourced submission on their website (McGrath & Peaker, 2018, p. 24-25). At the early time

[19]Current *Simmons University*, a private university located in Boston, Massachusetts
[20]Local television station, see https://www.wcvb.com/
[21]Local newspaper, see https://www.bostonglobe.com/

of the project, the team is anxiously trying to populate the records; soon enough, the team realized that they might have to outreach their crowdsourced practice out of the digital form and into the real world. In October 2013, the six-month memorial exhibition and the first "Share Your Story" event held at the NEU Library (McGrath & Peaker, 2018, p. 24). The "Share Your Story" event has held an exclusive space in the exhibition area, with plenty of donated or borrowed laptops. The visitors can view the Our Marathon sites, talk with staff, and share their stories via the submission page using the exhibition space's laptop. This event archived great success and was conducted 14 times across the greater Boston metropolitan region from January 2014 to April 2014. In April 2014, the Boston City Archives donated the digital copy of two collections: Boston Marathon Bombing response mail[22] and Boston Marathon Bombing Memorial Collection[23], and Iron Moutain[24] supported the digitization of those collections. The records from the Boston City Archives were quickly incorporated in the first-anniversary exhibition in April 2014.

In September 2014, Our Marathon project stopped active solicitation. Soon enough, the Digital Scholarship Group[25] and the Archives and Special Collections interfered with the long-term preservation planning of the Our Marathon's digital assets. In fall 2017, Our Marathon closed the crowdsourcing submission page, and then the Digital Scholarship Group helped Our Marathon migrated to the Digital Repository Service[26]; the new Our Marathon website, which is powered by WordPress, was opened in April 2018[27].

Currently, the digital collection has 7876 items in total, including 5495 images, 1570 zipped files, 628 text files, 156 sound files, and 27 videos. Those items were organized into 22 subcollections by provenance. Details see Table 5.3.

Table 5.3: Our Marathon Subcollections Review List

Collection Name	Image	Audio	Video	Text	Zip	Items
WBUR Oral History Project	34	137	0	2	0	173

Continued on next page

[22] *City of Boston Archives, 0247.003* See https://archives.cityofboston.gov/repositories/2/resources/92
[23] *City of Boston Archives, 0247.004* See https://archives.cityofboston.gov/repositories/2/resources/727
[24] A private company who dedicated to preserve information assets, see https://www.ironmountain.com/about-us
[25] Formed in January 2014, offering technical support for NEU's faculty and students, see https://dsg.northeastern.edu/home/about/
[26] Developed by Digital Scholarship Group, a data system for faculty and staff to protect their data, see https://dsg.neu.edu/services/drs/
[27] see about page: https://marathon.library.northeastern.edu/home/about

Collection Name	Image	Audio	Video	Text	Zip	Items
Boston Marathon Temporary Memorial (Boston City Archives)	109	0	0	0	5	114
Letters to the City of Boston (Boston City Archives)	2722	0	0	0	1469	4191
No Story Too Small Crowdsourced Submissions	375	7	19	269	0	670
Your Story Crowdsourced Submissions (Boston Globe's GlobeLab)	0	0	0	289	0	289
"LOCKDOWN" Police Scanner Recordings	0	9	0	1	0	10
Boston Medical Center Cards and Meaages of Support	158	0	0	29	96	283
Strong Medicine The Response to the 2013 Marathon Bombing	24	0	0	26	0	50
Photos by Nabila Abuljadayel	38	0	0	0	0	38
Photos by James Schmidt	108	0	0	0	0	108
Photos by Sarah W.	16	2	0	0	0	18
Photos and Stories by NUPR	43	0	0	0	0	43
ONE RUN	1106	0	1	0	0	1107
Marathon Daffodils	47	0	0	0	0	47
Stoneham Strong 5K	187	0	0	0	0	187
ONE RUN for Boston Relay	204	0	1	0	0	205
Salute for Our Heroes	23	0	0	0	0	23
Reactions to Rolling Stone's Dzhokhar Tsarnaev Cover	27	0	0	0	0	27
2013 Boston Marathon Internet Memes and Other Digital Content	258	1	1	0	0	260
WCVB-TV Boston Marathon Footage	0	0	5	0	0	5
Our Marathon Lesson Plans	0	0	0	5	0	5
Our Marathon Student Digital Exhibits	16	0	0	7	0	23
Total	5495	156	27	628	1570	7876

5.3.2 Personnel and Partners

The core team of Our Marathon is comprised of the faculty of the College of Social Science and Humanities, NEU, Ph.D. candidates and graduate students in the English and History Department, and graduate students in the Simmons College MLIS program. After August 2013, the control of the project was delegated to Ph.D. candidates. The students carried out almost all the concreted work, including metadata creating & cleaning, in-person event planning, and project management (McGrath & Peaker, 2018, p. 23). The NEU Library provided technical support in developing the original Omeka website, consultation in metadata scheme designing. Digital Scholarship Group of the NEU Library, which formed in January 2014, had supported the long-term preservation planning and digital assets migration. The Digital Scholarship Group also operates the Digital Repository Service, which preserved Our Marathon's digital assets. Outside the NEU, the local public libraries and university libraries provided event venue for exhibiting and soliciting. The librarians also helped the Our Marathon team in communicating with local community members (McGrath & Peaker, 2018, p. 24-26).

The partnership with the local mass media supported the project in many facets. Without a doubt, the media helped advertise this project and involved many people in the local community in sharing their reflections on the bombing. They also have other collaborations with the Our Marathon team. The WBUR started an oral history project and provided specialists and technical equipment; the WCVB participated in reporting and recording the memorial events and donated their footage; the Boston Globe opened their website to solicit crowdsourced submission and eventually donated it to Our Marathon.

5.3.3 Structure

The digital collection can be accessed from both Our Marathon website and the Digital Repository Service website. Our Marathon website provided an index of subcollections that make it possible to browse by subcollection. However, the primary arrangement of the collection doesn't have any hierarchical structure. An identifier can directly locate every single item from the root directory. On the Digital Repository Service website, the digital collection can only be accessed by 22 subcollections. While a user browses an item, the subcollection's name will show in a

format like "Our Marathon > Subcollection Name > Item Name." Be noted that the Our Marathon website actually requests digital assets from the Digital Repository Service by a Word Press plugin, which means the curators of Our Marathon website intentionally hid the original hierarchical structure.

5.3.4 Provenance

As shown in Table 5.3, 4305 items came from the Boston City Archives collections. 670 items were collected through the Omeka website's crowdsourced submission, including spontaneous submission via the internet and submission from in-person events. Boston Globe also contributed 289 crowdsourced submissions. The WUBR Oral History project provided 173 items. Especially, there are 28 items collected from the reuse activity of the digital collection. The rest items have come from individuals or institutional donators.

5.3.5 Technical Composition

Our Marathon originally released a customized Omeka website based on Omeka ver. 2.0.3. With the website service hosted on the NEU Library's server, the library's specialists made it possible to remotely manage the records. After migrating to the Digital Repository Service, the new website was customized from Word Press. The Digital Scholarship Group provided a WordPress plugin, CERES Exhibit Toolkit[28], to help researchers curate their digital assets. This plugin will send requests to the Digital Repository Service and receive JSON data when users browse an item. The website and the database were distributed, making the collection secure and robust.

The original Omeka website implemented a metadata scheme referred from Dublin Core, and the metadata cleaning had finished before migration. However, the Digital Repository Service is using Metadata Object Description Schema (MODS)[29], the migration made many errors in metadata fields, and some of them had been abandoned.

[28] https://dsg.neu.edu/ceres/

[29] A schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. see https://www.loc.gov/standards/mods/mods-schemas.html

5.3.6 Metadata

The original Omeka website is no longer available now; therefore, this survey is based on the current WordPress website and the MODS style metadata. The survey result is shown in Table 5.4.

Table 5.4: Metadata Status of Our Marathon

Metadata Field	Type	Total	Valid Total	Max Freq	Min Freq	Ave Freq	Ave IDF	Completeness
Referenced by	85	7876	1484	107	1	17.5	2.4	18.8%
Digital origin	2	7876	7867	4725	3142	3933.5	0.3	99.9%
Map data	1117	7876	3854	179	1	3.5	2.7	48.9%
Biographical or Historical	435	7876	602	19	1	1.4	3.6	7.6%
Publisher	3	7876	49	43	1	16.3	2.4	0.6%
Source note	57	7876	6132	4178	1	107.6	0.7	77.9%
Type of Source	4	7876	7876	7102	27	1969.0	0.2	100.0%
Related item	101	11418	11418	7880	1	113.0	0.6	100.0%
Contributor	72	7876	1610	1096	1	22.4	1.3	20.4%
Use and reproduction	15	7876	7790	3354	1	519.3	2.8	98.9%
Title	3380	7876	7876	341	1	2.3	2.9	100.0%
Date created	175	7876	7876	2308	1	45.0	0.5	100.0%
Language	4	7876	4991	4974	1	1247.8	0.2	63.4%
Subjects and Keywords	1025	37601	37601	7876	1	36.7	1.8	100.0%
Abstract and Description	3079	7876	7525	1190	1	2.4	2.5	95.5%
Our Marathon Omeka ID	3	7876	6	2	2	2.0	3.6	0.1%
Notes	641	7876	788	137	1	1.2	3.5	10.0%
Genre	37	7876	7875	4476	1	212.8	0.5	100.0%
Creator	586	7876	3999	380	1	6.8	2.3	50.8%
Item text	4	7876	13	10	1	3.3	3.1	0.2%
Format	3	7876	7286	7102	27	2428.7	0.1	92.5%

Chapter 6

General Discussion

6.1 Patterns of Participation and Procedures of Solicitation

Based on the author's investigation of the above cases of participatory digital archiving, the participation can be divided broadly into two types: online and offline.

6.1.1 Online Participation

Our Marathon project was grounded at the outset in collecting people's reflections on the Boston Marathon bombings through online crowdsourcing. It sought to do so in a way that would achieve the original goal of repairing and strengthening the fabric of the community that had been damaged by the bombing. The MMRS collection, on the other hand, was entirely driven by in-person events. Still, the COVID-19 pandemic halted the tradition of their practices and forced the suspension of in-person events. As a result, two projects that been already in the planning process had to be replaced by participation through online methods, namely the STUCK AT HOME event in 2020[1]. The MMRS also differs from the Our Marathon in the following point: the former restricts participants to specific communities, i.e., the two communities where in-person events are planned before the pandemic; in contrast, the latter does not restrict participants at all. MMRS required participants to complete the same online forms as those completed at the in-person event under the guidance of volunteers, as well as to be interviewed by staff via virtual meeting software to obtain additional information of interest that should be included in

[1] http://blogs.umb.edu/massmemories/stuck-at-home-show-pilot-project/

the description of the user-contributed material. In contrast, Our Marathon only requires participants to complete the procedures described in Figure 5.3.

The author attributed this distinction to the following two reasons. The first point lies in the mission of the project. On the one hand, MMRS is dedicated to collecting and preserving regional historical materials or providing educational resources for teachers and students in the region, and also plays a role in facilitating (regional) community building. These goals all emphasize the regional characteristic of the MMRS, defines the scope of the MMRS collection and the content organization and curation activities directed to serve the local residents. The participation method used by this project should promote communication among community members. These practices of offline collecting activities demonstrated their effectiveness in the past, and the STUCK AT HOME events are destined to be just a temporary solution during the pandemic. On the other hand, Our Marathon's mission is concretely aimed at remedying the illusions of safety and stability that community members have been shattered by the bombing, exchanging or sharing individual traumatic experiences and reflections on the bombing with community members, and reinforcing the broken community fabric with expressions of support. The Great Boston Metropolitan region is extraordinarily cosmopolitan[2], with many foreigners living here, thus giving the local community a greater density of international connections, making it inevitable that concerns about the Boston Marathon Bombing have extended to a worldwide scale. For example, one of the three victims was a Chinese student and thus the tragedy received a lot of attention in China[3]. On the other hand, the activism of Islamist extremism around the world, the very background of the terrorist attack, has also led to empathy for the local community among the people all over the world who also face the threat of terrorism.

The second aspect is the technical composition of each project. The CONTENTdm system used by MMRS is a digital collection system provided by OCLC for libraries and other cultural institutions. It requires administrators to use a browser or client to manage the content and does not support crowdsourced submissions or decentralized curation. The MMRS staff at the STUCK AT HOME event utilized WordPress to create a temporary submission page that could not interact with the digital collections in the CONTENTdm system, and the materials submitted by participants on the page needed to be manually inputted into the CONTENTdm system after being handled by the staff, just like the materials collected in previous

[2]See *New Bostonians Demographic Report*: https://www.cityofboston.gov/newbostonians/pdfs/dem_report.pdf
[3]*CAS Mourns the Loss of Graduate Student Lu Lingzi*, see http://www.bu.edu/cas/lu-lingzi/

in-person events. On the other hand, the Omeka system used by Our Marathon project at the beginning is an open-source digital collection system that supports crowdsourcing submission and decentralized curation for both individuals and institutions. Without a doubt, that was crucial for the project's speedy response to crowdsourcing collecting right after the bombing. The Omeka-based website was open to crowdsourcing submissions between its launch in May 2013 and its migration to DRS in early 2018, while the new DRS system did not support crowdsourcing submissions. The literature suggests that the Operation team plans to hold new collecting events for the fifth (2018) and tenth (2023) anniversaries of the tragedy (McGrath & Peaker, 2018). As of late 2020, the author has not found new entries in Our Marathon's digital collection after migrating to the DRS.

The notable weakness of online participation is insufficient power to mobilize community members to participate in the event. The initial efforts to boost the number of materials in the Our Marathon project did not yield the expected results, but after the great success of the in-person event in October 2013, the Our Marathon team held 14 more in-person events to exhibit their collections and to engage community members to participate as contributors to the project (McGrath & Peaker, 2018, p. 24). MMRS demonstrated the weakness of online events in terms of mobilization in another way: during two STUCK AT HOME events only a few dozen records have been collected in comparison to the hundreds of records that were routinely collected at previous in-person events (see Appendix).

6.1.2 Offline Participation

The offline participation observed in the case studies can be classified into two groups: in-person events and entrusted recordings. Event-based participation is typified by MMRS, and local communities are encouraged to participate through online promotions, phone calls, and flyers, with the MMRS team and other partners' help. On the event day, participants bring their materials to share to the event venue, where with staff and volunteers' support describe their submissions and participate in keepsake photo and videography to gather the background information of the contributors and the materials they submitted. The MMRS team will permanently preserve the materials collected from in-person events in a library along with the records generated from the planning of that event (including meeting minutes, official notes, copies of materials used to advertise the event, etc.) (MMRS, 2016), and will

organize and curate the digital collection with respect to the specific in-person event. Our Marathon's in-person events mixed exhibiting and gathering purposes, and the gathering at the in-person events was also conducted through the Omeka system. Participants could communicate with Our Marathon staff, volunteers, or librarians, or oral history scholars from partners at the event venue, being invited and supported to participate in crowdsourcing submissions or oral history recordings (McGrath & Peaker, 2018). However, Our Marathon does not prioritize keeping the records of the in-person events activity. Crowdsourced submissions collected at different in-person events are neither distinguished from each other in the Omeka system, nor from those spontaneously submitted via the Internet.

As mentioned in the previous chapter, Recorder311 collects materials, primarily through entrusted recording. Recorder311's in-person events, in principle, do not aim at gathering records but rather at providing the chance of training and communication for the participants. However, one of Recorder311's fundamental objectives is to record and disseminate information that participants shared at the events. As a result, many in-person events have been recorded in the digital collection. For example, Kangaeru Table and Wasuren! Studio's event records and the participants' narratives about the materials they collected are captured at these events. In particular, Recorder311 preserves the records of the events it uses, including the records of the screenings and other events, as well as the reflections of the participants on the events. Importantly, if the event-based approach is taken, an event organizer has a leading authority over all the collecting activities. At the same time, common users can just contribute to the collection but are not involved actively into the organization and management of recording materials.

Collecting through in-person events has two advantages compared to online collecting. First, it involves more community members in the event, and, second, it facilitates the communications between participants to achieve the overall objective of the participatory digital archiving project. However, offline and online activities are essentially the same, except for recording the in-person event. The collecting in Our Marathon project's in-person events should be understood as an extension of the online collecting, and MMRS's online collecting serves as a mock-up of the in-person events. There is hardly any difference between the two in terms of characteristics of gathered materials, but in the case of MMRS and recorder311 projects, the recording of the in-person events themselves enriches the contextual information of the materials gathered in the corresponding events, and thus a contextualization

of the collecting activities in a macro sense is being achieved.

Entrusted recording is a primary method in Recorder311 project. Recruited participants (natural or legal persons) each independently conduct gathering activities and transfer the collected materials (usually a series of related records) to the Recorder311 project as a whole. The acceptance hearings held at the time of transfer can provide additional contextual information that is brought to the user by the staff in curation. The entrusted recordings in Our Marathon also includes organizations and individuals, such as television stations and photographers (McGrath & Peaker, 2018, p. 22). Through these participants, records of many post-bombing commemorative events were captured. The records collected through entrusted recordings generally have exact provenance and can be traced to specific recorders. Of more importance is that they are usually a continuous series of documentary materials that were generated directly from a recording activity with a definite purpose and can be contextualized internally between the records. Those involved in entrusting recordings are commonly so-called expert users, or, in the case of Recorder311, they train and support the participants through intensive offline activities. Records from entrusted recordings are routinely divided by provenance upon acceptance and are regarded as an organic whole in preservation, organization, and curation, where the wills of the participants are respected preserved.

The author argues that the presence of super-participants can threaten the diversity of content in PDA, and that is the problem keenly related to entrusted recordings. As Owens (2012) has pointed out, the most successful crowdsourcing projects have nothing to do with the majority of contributors, and in the cases examined in this study, it has also been proved that a few people contributed the vast majority of materials. The MMRS project skillfully avoided this problem in terms of results by limiting the number of items allowed to be submitted by one contributor to three, because of the event time limit (MMRS, 2016). In recorder311 project, the most active participants excluding staff contributed 73 records from a total of 557 records, while the average of this number was 3.8. This represents one-fifth of the 363 records that excluded staff-contributed material. Similarly, the ONE RUN sub-collection accounts for 30% of our Marathon's total, except for materials from the City of Boston Archives. A possible solution would be to proactively coordinate the presentation of crowdsourcing submissions in curation to show the diversity, as was done in the former Our Marathon before the migration (Barney, 2018). But this approach would undoubtedly confront critiques of impartiality, and the retrieval

techniques used for digital collections of PDA are simple flat structures that allow users to bypass curatorial presentations without barriers.

In this section, the author attempted to answer the research question on how archival materials are accumulated in PDA. Through comparative analysis, the author has concluded that there is no essential difference between online and offline crowdsourcing submissions, but in-person events have great advantages in mobilizing participants and facilitating community communication. The records of in-person events can achieve a contextualization of the collecting activities and serve as background information for the user-contributed materials collected during the activities, improving the authenticity of the materials. Entrusted recordings can provide record groups with exact provenance and solid internal relations, but an involvement of expert users may put at risk the balance of diversity in PDA.

6.2 Metadata in Contextualization

In the archival discourse, a context is broadly understood as "the organizational, functional, and operational circumstances surrounding materials' creation, receipt, storage, or use, and its relationship to other materials[4]." More concretely, the context includes information about an archival creator (usually an institution or a representative of an institution), organizational structure of a creator, and functions and activities of creators. Specifically, in the electronic environment, archives' context is defined as "the framework in which a record is created, used, and maintained[5]." Notably, all these definitions are associated with diplomatics, literally "the study of documents," which is known as one of the origins of archival science. Diplomatics determines the authenticity of a document by verifying its internal and external characteristics, such as a type of paper, ink, typeface, phrasing conventions, and so on. The concern with the context in archival science is also focused on domains related to authenticity and evidentiality.

Traditional archival discourse discovers the origins of reliability of records in their forms and procedures of creation (Duranti, 1995, p. 6), Archivists have known a long time ago that a perfect authenticity of records does not guarantee a reliability of their content, especially if these records have been accumulated outside of institutional

[4] *SAA Archives Terminology:* https://dictionary.archivists.org/entry/context.html
[5] *InterPARES Trust Terminology:* https://interparestrust.org/terminology/term/context

archives. Moss (2008, p. 76) used the following example to illustrate such a situation. A painter James McNeill Whistler, in a letter to his mother written on September 26/27, 1876, informed her sorrowfully that he was too busy paying off his debts to make it back to visit her. However, Whistler was found to have hosted a grand dinner party at his home on his mother's birthday, since the dinner party's menu has survived. In the light of the methodology of diplomatics, a mail written and sent by the very person cannot be considered a forgery. Archivists who adhere to the conventional archival thinking may attribute this to the fact that creators are not trusted, but the problem is that such "untrusted" creators and contributors are overwhelmingly present when you explore under the participatory discourse. Moreover, Bak (2016, p. 377) points out that trust does not exist in a simple binary with distrust, and that traditionally "trusted" governmental and social-cultural institutions are not inherently immune to distrust. As Bak implied, unilaterally establish a legitimate official archives and claims it has been trusted are not feasible in current diversified society. There is no common sense on what is trusty across all society members, so we have to consider that who dose a PDA needs to be trusted by?

While the author argues that the authenticity and evidentiary nature of conventional institutional archives primarily serve judicial and accountability purposes, PDA share the objective of empowering their communities through the construction and maintenance of collective memory. In other words, PDA primarily serve a given group of people with shared interests. Considering that the origins of community archives are distancing from official discourse, the author argues that being trusted in such groups, a sense of belonging and closeness to narratives and discourses would be the most critical factors. Up to this point, the preceding discussion seems to affirm the postmodernist postulate in regards to diversity of perspectives to approach the truth that was stated in Chapter 1 as bringing about evidentiality problems. Although the case studies' outcomes show that the evidentiality problems are real, it is possible to discover alternatives for refinement in the traditional archival discourse by comparison.

In PDA, additional information from users during submitting archival materials is usually recorded in the metadata. At the same time, contextual information also exists elsewhere that will be discussed in the next section. The question what kind of contextual information exists in the metadata can be approached by analyzing metadata fields, as shown in Figure 6.1. Through the investigation of the metadata

of the above mentioned cases, the author has discovered the following problems:

1. creation and archiving of records were being confused;
2. insufficient description of collecting activities;
3. overlapping or misplacing of functions between metadata and content.

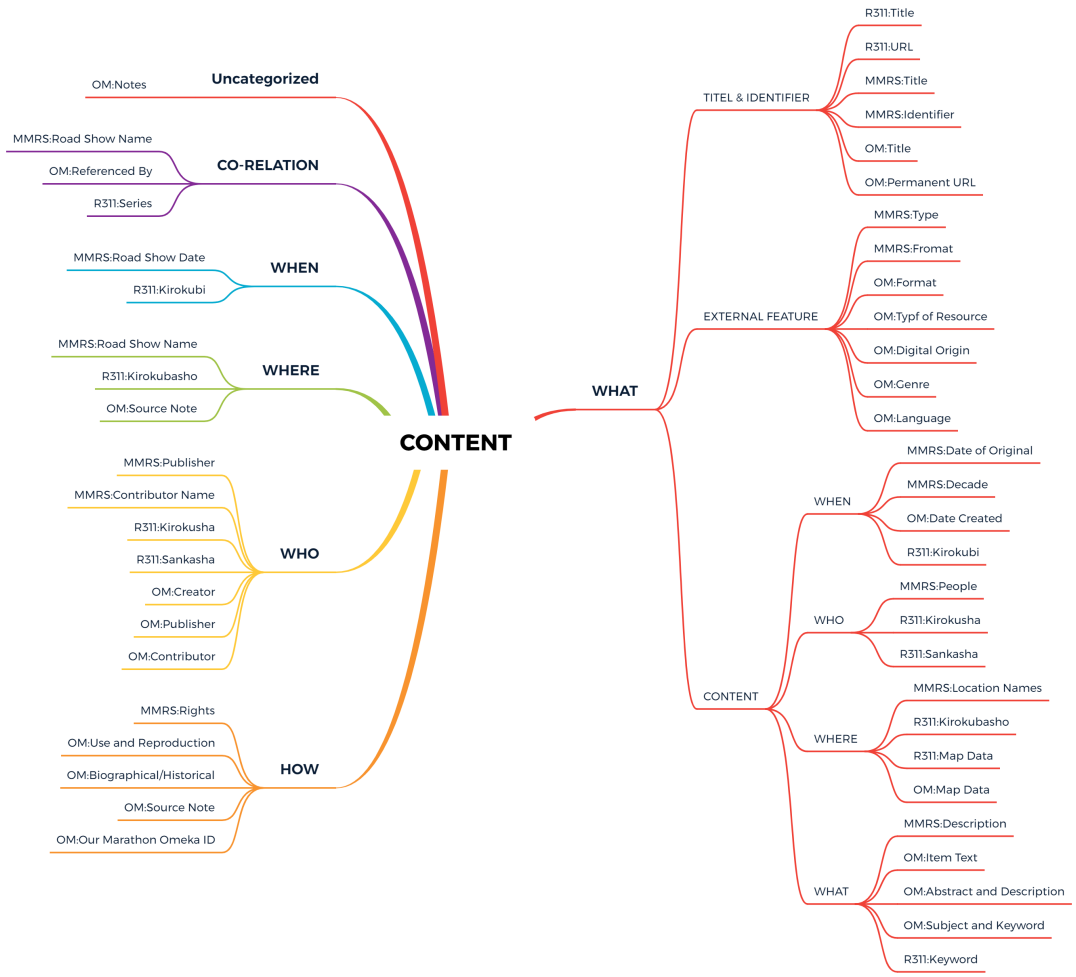

Figure 6.1: Contextual Information in Metadata Fields

Compared to ordinary archives, descriptions of creation and archiving in the metadata of PDA are the most confusing aspects. Descriptions of people and dates, which are considered the essential contextual information in conventional archives, are problematic in PDA' metadata. "Archiving" has different meanings among the various participatory methods mentioned in Section 1, and in some participatory methods, "archiving" is even equivalent to "creation." "Archiving" may

be understood as the activity by which a participant initially captures a record, as the activity of submitting a record to the organizer, or as the activity of creating a digital object in a digital collection system.

This issue was not found in MMRS digital collections where archivists were involved. MMRS used the Dublin Core-based metadata schema, and the Dublin Core Metadata Initiative defines the creator, contributor, publisher, date of creation (in MMRS, the date of creation of the digital object), date of submission, and date of acceptance. The single collection method of MMRS also reduces the risk of ambiguity. Our Marathon initially planned a metadata schema based on Dublin Core with metadata specialists' help, but the metadata collected by the author was a combination of MODS metadata after the migration to DRS. In addition, the existence of metadata fields from the EAC-CPF[6] was also identified in Our Marathon. Our Marathon's metadata cleaning did not involve archival specialists. The tagging of photos from photographers in Our Marathon was mixed with the creator and contributor tags, which might cause users to doubt the meaning of the metadata fields and the accuracy of the metadata. Besides, Our Marathon only has a date of creation that describes the date when a digital object was created and does not use the date of submission and date of acceptance. There is also ambiguity in capturing digital source objects; for example, the description on the date of creation of a screenshot from a web page or instant messaging software contains two understandings, the date of creation of the content and the date of taking the screenshot.

Recorder311 project, which does not adopt any standardized metadata schema or archival specialists' intervention, is entirely chaotic in regards to this issue. For example, in an instance mentioned in Chapter 4, a transcription of an interview with a participant at the Kangaeru Table event about her recordings, Kirokubi (date of recording) is labeled as the time period in which the interviewee performed her recording activity, Kirokusha (recorder) is labeled as the interviewee, Kirokubasho (location) is labeled as the location of the interviewee's recording activity, and the non-required field Sankasha (means participant) is also labeled as the interviewee, which means that the person, responsible for editing the metadata, may interpret a material as a record of the interviewee's recording activity which talked in the interview, instead of a record of the interview. As it is seen in this instance, in which a content section presents an interview's time and place, an interviewer and an interviewee are not shown in the metadata. Similarly, information on interviewees in the

[6]Encoded Archival Context for Corporate Bodies, Persons, and Families is a supplemental metadata schema for encoding archival materials, see https://eac.staatsbibliothek-berlin.de/

oral history project Wasuren! Stories also does not shown in the metadata, but only in the content. The author supposes that the possible reason for this is that the Recorder311 project staff mistakenly (indeed, from the archival metadata perspective) regarded participants and contributors as equivalent. Literally, participants do contribute materials to the Recorder311 digital collection, but at the very first, participants were the direct creators of those materials. Therefore, participants in Recorder311 should be defined as creators of records. Those who contribute to the materials for which the participant recorded, such as interviewees for oral histories, are contributors in the standard metadata schema.

For conventional archives, archiving is, perhaps, the holiest moment: the moment when the document says goodbye to current business and enters a continuous and uninterrupted chain of preservation that begins to enjoy a guaranteed safety. As mentioned before, the forms and procedures of creating archives are crucial to their reliability. The fact that records of PDA are not comparable in forms to those created in institutions highlights the significance of the collecting procedures. The documentation of the collecting activity is an evidence of the birth of the materials contributed by users as "archives", and its value for the authenticity and evidentiality of archival materials cannot be overestimated. The MMRS staff has collected and preserved a complete set of records related to in-person event planning from the very beginning. Still, this part is not presented in the digital collection. The digital collection records were organized and presented by events, ensuring that they could be traced back to the event's specific location and date that is essential background information. The digital collection also includes videos and keepsake photos of participants being interviewed, thus proving the participants' presence. Recorder311, on the other hand, recorded in-person events of the participants but did not actively record specific collecting procedures, such as acceptance hearing. Our Marathon, however, unfortunately, left no record of offline collecting activity and failed to label offline activities in the digital collection.

Metadata of conventional archives are expected to contain valuable additional information, absent in the format of the content, but in the free-form PDA, this expectation is frustrated or even reversed. Conventional archives typically have a persistent creator with a relatively stable organizational structure, shouldering specified functions, and performing specific activities. These constants become self-explanatory contextual information in identifying and understanding the archive, and simply tagging the creator in the metadata is sufficient. However, in the records

of PDA, the creator's name alone is not meaningful enough to help users understand and judge the records. The metadata for these records, either provided directly by participants (MMRS) or created by curators based on interviews with participants (Recorder311), provide information that in no way exceeds the boundaries of the total information contributed by participants through various means, which, in most cases, can be obtained by browsing the content. The role of metadata in PDA has more to do with helping to presents contextual information than providing additional information. In particular, Our Marathon converts text-type records into images for preservation after migration. It logs the full text of the content in the description field, creating an abnormal reverse of content being a subset of metadata. These issues inevitably call into question both the capability and the value of metadata to provide additional information in PDA.

In response to these problems, the author suggests that an organizer should be shown as a provenance, and that efforts should be made to formalize collecting procedures and documentation of collecting activities, thereby enriching contextual information that evidences the collected records' authenticity. Without a doubt, metadata will be crucial in the presentation of contextual information that helps in conveying the efforts that make users "trust" the archival materials in PDA. In terms of reliability, it is important to anchor the provenance, i.e., an organizer, as a constant entity, based on the project's mission to curate a consistent and continuous narrative framework, which is oriented towards a group of participants, in order to obtain the community's trust.

6.3 Presentation/Visualization of Contextual Information

In the survey, the author discovered the existence of contextual information beyond metadata in PDA. Roeschley et al.'s (Roeschley, Kim, & Zavalina, 2020) analyzed the status of the presence of content and context in the description fields of user-contributed materials in one of the MMRS events, inspired by their work, an analogous approach is used in this section to investigate the status of contextual information in digital collections. However, the total number of cases in this study would be too large to be exhaustive; thus, random sampling was adopted for each sub-collection in the 3 cases in order to cover all the various collection methods. The number of samples was based on the sub-collection volume, which was usually no less than the number of formats and no more than five.

The primary justification for the division of contextual information for investigation is Sheridan's (2018) comment on the contextual information that is required by historians (users) and archivists in the digital environment. He asserts that the purpose of a historian is to understand events and places in the record through contextual information in order to clarify the chronological sequence of events, while the archivist is thinking about the function of the creator as well as the influence of the records, the environment in which the records were created and the custodial history of the records. The researcher also argues that in PDA, the records of reuse activity also essential contextual information. Postmodern archivists (Ketelaar, 2001; Cook, 2001) have already pointed out that archives are "records of infinite activation(Ketelaar, 2001, p. 136)." In other words, each new interaction with the archives (including reuse) contributing new meaning to the archives. The purpose of activities in PDA is demonstrated in reuse activities, such as the content screened in Recorder311 or the educational exploit's responses in Our Marathon. The *Tacit Narratives* (Ketelaar, 2001) in the reuse activity are the contextual information for all recording activities in the project and part of the project's storytelling framework as a whole. The division of contextual information based on the above assumptions is shown in Figure 6.2.

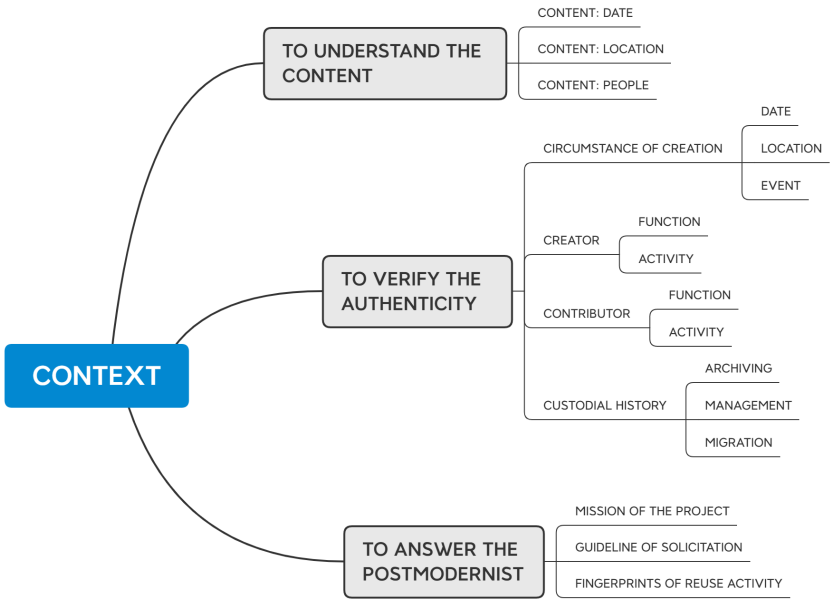

Figure 6.2: Division of Contextual Information

The result of the investigation of the status of the contextual information in the digital collection is shown in Table 6.1. It is remarkable that none of the three cases

neglected to represent the contextual information of the records' content, which was found in both the metadata and the content. The efforts to provide the context of the content are in harmony with Huvila's assertion that participatory archives are "radically user-oriented."

Table 6.1: Contextual Information in Digital Collections

Fields	MMRS	Recorder 311	Our Marathon
Content: Date	metadata; content	metadata; content	metadata; content
Content: Location	metadata; content	metadata; content; map	metadata; content
Content: People	metadata; content	metadata; content	metadata; content
Date	metadata	metadata; content	metadata
Location	metadata	metadata; content; map	/
Event	metadata	content	metadata
Creator: Function	about page	associated biography page	collections' description, only on DRS
Creator: Activity	about page	content	collections' description, only on DRS
Contributor: Function	associated interview video	content	content
Contributor: Activity	associated keepsake photography	content	content
Archiving	/	/	/
Management	metadata	metadata	metadata
Migration	/	/	metadata
Mission of the project	about page	about page	about page
Guideline of solicitation	/	/	/
Fingerprints of reuse activity	/	associated records	associated records

Another factor that plays a significant role in the presentation and visualization of contextual information is spatial information. For instance, Recorder311 marks the location for most of the records and represents them on a map, allowing users to browse records through the map, as shown in Figure 6.3[7]. On the other hand, Our Marathon uses high-precision floating-point numbers for the map data values in their metadata fields. The latter, in the author's opinion, is probably caused by the digital collection system's migration. The search for the legacy version of the Our Marathon website on Archive-it has proved that the records of Our Marathon had been once represented on a map when the Omeka system functioned. The legacy version of the website is shown in Figure 6.4[8].

Figure 6.3: Recorder311 Map View

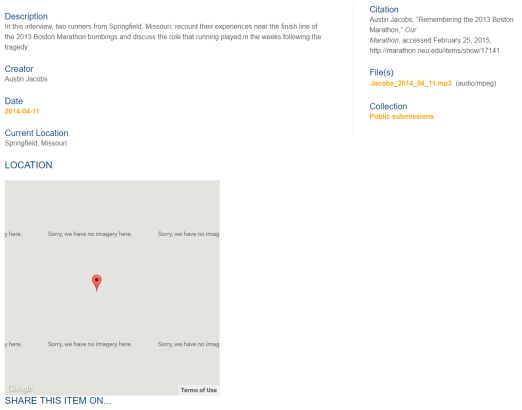

Figure 6.4: Our Marathon Legacy Version with Omeka

Watanabe argues that PDA faces users who lack both knowledges on the composition of the digital collection and specialized retrieval skills. As a consequence, in his opinion, structured categories can impede cross-searching and discourage the correlations between the records (渡邉英徳, 2013, p. 155, p. 160). That is why, it was concluded that tree-like structures were unsuitable for organizing such archival materials. His project, Hiroshima Archive, abandoned using the categorized index, building a PDA that curated the whole collection based on geolocation with Cesium[9] (shown in Figure 6.5). The author concludes that for regional PDA projects, the utilization of the digital map technology in curation can dramatically enhance the understanding of the content by local users. Through map information, the user's pre-existed knowledge of their region is translated into contextual information about the record upon viewing. Basically, it is the same mechanism as links to a creator in perusing institutional archives.

[7] https://recorder311.smt.jp/mapview/
[8] https://wayback.archive-it.org/all/20150226001020/http://marathon.neu.edu/items/show/17141
[9] A 3D Geospatial platform, see https://cesium.com/

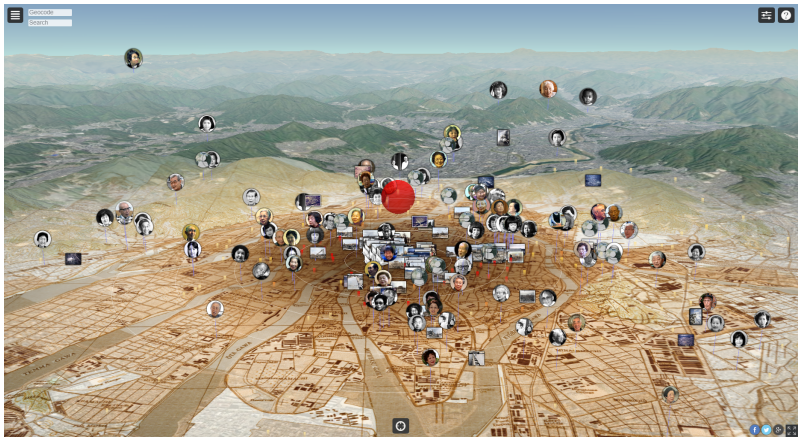

Figure 6.5: Cesium Adpotion in Hiroshima Archive

6.4 Archival Professional's ommission

Unquestionably, an involvement of archivists significantly contributes to the implementation of PDA projects. In terms of cases surveyed, archivists' involvement in MMRS activities started from metadata scheme planning and was seen through in-person event planning and the entire process of registration in the digital collection system. The professionalism of the archivist is demonstrated in the following aspects during the archiving activity's process: the preservation of the activity records reflects the archivist's instinct for records directly generated from activities; the handling of the records from multiple-day event reflects the archivist's sensitivity of provenance; the almost fulfilled metadata reflects the archivist's requirement for consistency in the form of records; the accurate capturing for each entity appeared in the creation process in the metadata labeling reflects the archivist's professional competence. While not featured in the case studies, archivists' capability to help non-expert teams capture information in a digital environment has also confirmed by Flinn (2008, p. 123).

Recorder311, a project without any archivist involvement at all, correspondingly reflects some limitations. The metadata fields are relatively few and not accurately described. However, due to the objective of information dissemination, Recorder311 staff accidentally gathered and preserved considerable contextual information from in-person events. That information is distributed among the content, metadata, biographical pages, introduction pages of series, or records to which they have any references. Unfortunately, it does not well described within metadata, which

means it failed to provide an intuitive presentation for the records' authenticity. Our Marathon project has been launched and also fulfilled without archivists' involvement. Its team neglected the necessity of collecting records of activities and did not demonstrate a proper respect to provenance in the management of archival materials. The metadata schema was designed with consultation from archivists, but the metadata cleaning, or describing, had many preventable errors. Repeated titles and descriptions in different records, mislabeled contributors and creators, placing sources in the notes field, etc. The most grievous one was that a part of the collection had to be abandoned when migrating from Omeka to the DRS using MODS (Heilbrun, 2018). Obviously, MODS has been used primarily to describe libraries' resources, and that is why it cannot be seen as the most appropriate metadata scheme for PDA. For sure, the absence of archivists must have caused more than just these regrets.

Postmodernism was the context in which these archival projects emerged, and it likewise supports the attendance of archivists in their activities. In 2001, Cook (2006, p. 170) emphasized in his paper that the archivist's job included making social and historical memories, thus obliging archivists to think about what kind of materials should be preserved for the future. By "thinking," he apparently means thinking in a postmodernist environment (Cook, 2001, p. 6). Ketelaar's (2002, p. 237) commentary on the power of archivists suggests that archivists can be either an "accomplice of oppression and torture" or a "friend of liberation and justice." In particular, Schwartz & Cook (2002, p. 17) stated that archivists should "search thoroughly for the missing voices." These perspectives encouraged archivists to step out of the traditional custodial paradigm and engage in the archiving of minorities and marginalized groups. Flinn's (2007, p. 170) argument jumps out from responding to the call for social justice by arguing that the absence of archivists in these grassroots archives may result in missing "the opportunity to work to ensure that the national archival heritage (inside and outside the archive walls) reflects the whole of society." He also remarks that the relationship between the mainstream and the marginal is not fixed but can be mutually transformative. In order to provide future historians with a broader scope of vision, archivists should not judge the worthiness for preservation by what is mainstream or marginal, success or failure in contemporary discourse (Flinn, 2008, p. 110-111).

Today, even traditional archival institutions have been slowly transforming under the postmodernist discourse (Wang, 2017, p. 123), but the community archives

paradigm predicted by Cook (2013, p. 113-116) has not yet fully emerged. Unlike many community archives that share commonalities with collecting archives, PDA that collect archival materials through a radical crowdsourcing practice, as earlier, remain outside the archival discourse (Flinn & Sexton, 2019, p. 173). The author has to admit that archivists, certainly, primarily serve their employers. However, archivists working in public social and cultural institutions should also be responsible for the collective memory of society. In current diversified society, legally recognized as an official repository cannot make one the only source of future information, the days of only authority has passed (Bak, 2016, p. 394). As this survey demonstrates, archivists' skills can be brought out in PDA, and these practices, already emerged, are waiting to be embraced by archivists.

6.5 "Broader Contextualization" in Participatory Digital Archives

PDA, along with other practices of archival activism, are facing more intrinsic interrogation than the defects of authenticity. Matthews (2016) argues that post-modern archival theory and contemporary archival activism based on deconstructionism (p. 214) are grounded in a misreading of Derrida's archival fever (p. 213, p. 225, p. 256). He notes that "the call for justice involves ethical and political commitments that must be justified but cannot be justified by deconstruction (p. 224)" because "deconstruction has no ethico-political commitments. (p. 224)" Thus, archival activism practices, and the political decisions that archivists made in those practices, must be critiqued and interrogated by deconstruction as well. Gilliland & Flinn (2013, p. 5) argue that community archives are "inherently political," and this is the point the author agrees with, because the motivations of community archives projects inevitably include creating records in their own narratives that mean engaging in the struggle for the discourse of power.

PDA that are facing the challenge of deconstructionism too has the option to defend themselves in the same manner as traditional institutional archives: transparency. Such a justification relies on a definitive and coherent narrative framework, such as the one referred at the end of Section 2. A concrete embodiment of the whole project's macro-narrative framework is the broader contextualization of its archival material. Huvila (2008) summarizes his practice and suggests that it is necessary to extend the domain of contextual information to the entire archival process. Com-

pared to conventional archives, PDA, on the one hand, lack correlations between records, but, on the other hand, the factors associated with the creation are more sophisticated. Such entities as original creators, submitters, contributors, and curators rely on more contextual information to identify and describe. A variety of collecting methods and relatively flexible procedures inevitably impact the consistency of records' forms. Here again, it requires more contextual information to account for the consequences that happen to the records' forms for adopting various methods and procedures.

However, extending contextualization to the entire archival process has yet to respond to postmodern archival theory's articulation of the archive as the "infinite activation of the record." Ketelaar (2001, p. 136) explains that creation of records is influenced continuously by changing social, cultural, political, and even religious factors that must be incorporated into the understanding of archives. Across a greater time span, these changes in the external circumstances are reflected by the reuse of the archival materials. The fingerprints of these reuse activities may add new meanings and values to the archival materials. In PDA, the significance of reuse activities contributes new meanings to the archival materials and becomes the context for the whole project, thus enhancing the narrative framework of the whole project.

If the purpose of conventional archival preservation is to serve future historians' uncertain demands, the purpose of PDA projects is obviously somewhat oriented. Therefore, there is no better way than reuse activities to state the project's purpose and stance. PDA projects usually do not instruct or imply the aims and solicitation guidelines to encourage participants to speak freely (佐藤知久 et al., 2018, p. 90). Yet, the censorship of content remains true (佐藤知久 et al., 2018; Carlton, 2016, p. 108, p. 239). This unstated contextual information can be perceived within the reuse activities to understand the record's creation process's circumstances in more comprehensive ways. Thus, the author argues that the broader contextualization in PDA projects means that the management, curation, and reuse activities are equally critical with the entire archival process.

Chapter 7

Conclusion

Start from the evidentiality problem of the archival materials in PDA, this paper depicts the basic portrait of PDA's activities through the several case studies. The author has reviewed the literature related to PDA in terms of postmodern archival theory and analyzed the factors in PDA's archiving activities that may contribute to the better contextualization of archival materials under the current archival discourse. The investigation the author conducted allows to make answers to the research questions settled at the beginning:

1. How did the archival materials have been accumulated?
   Archival materials have been collected directly from contributors by project teams via crowdsourced submissions or received from trusted recorders. Crowdsourced submissions may occur on the Internet or in the real world, and offline collecting activities generally can mobilize more participants and capture more contextual information. Materials from trusted recorders can be regarded independently as an organic whole with stronger authenticity. However, the high proportion of archival materials transferred to the collection by expert users may threaten the diversity of narratives that the PDA projects intentionally or unintentionally pursuing.

2. What kind of contextual information is primarily represented?
   The contextual information in the PDA can be roughly divided into contextual information about the record itself and contextual information about the record's content, which are both distributed in the metadata, content, biographical pages, about pages, and associated records. Apparently, PDA focuses more

on representing contextual information about the records' content, thus reflecting their user-oriented nature. The author argues that the implementation of the digital map technology can significantly enhance the user's understanding of the records' content.

3. How did the metadata strategy and implementation affect those collections?
   The metadata planning and implementation can dramatically affect the metadata's status, including its completeness, consistency, and accuracy. High-quality metadata help users understand the circumstances in which records were created and provide a more intuitive presentation of contextual information about records. However, unlike conventional archives, metadata of PDA usually do not provide additional value beyond the original information submitted by participants. In the light of this, the author argues that metadata has no decisive function in contextualizing PDA.

4. What role do archivists play in the archiving activities?
   Usually, archivists are not indispensable in PDA projects. An archivist may act as an initiator of the metadata scheme, a recorder of the archiving activities, the archival process's regulator, a person responsible for metadata annotation/archival description, or a person in charge of archival management. Undeniably, such involvement of archivists can dramatically improve the overall authenticity and reliability of PDA.

5. How is PDA positioned in archival science?
   Along with other archival activism practices, PDA have been excluded from traditional archival discourse. Even in postmodern archival theory, the PDA model is rather marginalized. The records continuum model allows to incorporate user participation (Eveleigh, 2017, p. 303), but this model is quite different from the PDA in this paper, basically composed of user-contributed materials as content. Perhaps, PDA will remain ignored by archival discourse for a long time due to their unique and deviant collecting methods.

This paper's original contribution to the PDAs theory is in expansion of the "broader contextualization" proposed by Huvila (2008). In full accordance with the postmodern archival theory of "infinite activation of the record" (Ketelaar, 2001, p. 136), the author argues that it is necessary to broaden the scope of contextual information in PDA to encompass the entire archival process, as well as all the management,

curation, and reuse activities.

In this study, the author consciously neglected "archival multiverse" theory that acknowledges the diversity of archival beings as well as methodologies (A. J. Gilliland et al., 2016). Admittedly, archival multiverse theory quells the identity crisis of archival professions that Cook (2013, p. 109-113) warned us about. Still, it also justifies us to stay in our comfort zones, where exciting cross-domain collisions and encounters are no longer encouraged (Piggott, 2020). As archival multiverse theory has been gradually incorporated into archival education (李子林, 2018, p. 141), the author became more sensitive to this concealed problem. It is undeniable that PDA's brutal practice is also part of what Richard Pearce-Moses (2007, p. 22) calls "our social mandate of preserving the cultural record." The author also believes that archivists should proactively break the compromise (which archival multiverse theory made) and embrace these radical archiving activities with our professional skills.

Acknowledgement

I wish to express my gratitude first and foremost to my supervisors, Assistant Professor Eduard Baryshev and Professor Tetsuya Shirai, for their enlightening guidance and persistent encouragement, without which this research would have no chance to be finished. I would additionally like to show my gratitude to all the seminar members in Shirai Laboratory, whose sharp commentaries helped me improve this study.

I sincerely thank my friends who support me to overcome the difficulties in life and study, especially under the current pandemic. My stay in Tsukuba with them will become one of my most precious recollections.

Finally, I would not achieve anything without my parents' and sister's continuous support and kind counseling at all times. I would like to extend my deepest gratitude to them for their love, understanding and support.

References

Arthur, P. (2009). Trauma online: Public exposure of personal grief and suffering. *Traumatology*, *15*(4), 65-75.

Bak, G. (2016, 12). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, *16*(4), 373-402.

Barney, M. (2018). *Our marathon 1.0 to Our Marathon 2.0: Reflections on digital platforms.* Retrieved 2020/12/03, from `https://dsg.neu.edu/our-marathon-1-0-to-our-marathon-2-0-reflections-on-digital-platforms/`

Bastian, J., & Flinn, A. (2019). *Community Archives, Community Spaces: Heritage, Memory and Identity.* Facet Publishing.

Breakell, S. (2008). Perspectives: Negotiating the archive. *Tate Papers*, *9*(3).

Carlton, P. (2016). *From ashes to ashé: Memorializing traumatic events through participatory digital archives* (Doctoral Dissertation). University of Central Florida.

Caswell, M. (2014). Toward a survivor-centered approach to records documenting human rights abuse: Lessons from community archives. *Archival Science*, *14*(3-4), 307-322.

Cook, T. (2001). Archival science and postmodernism: New formulations for old concepts. *Archival Science*, *1*(1), 3-24.

Cook, T. (2006). Remembering the future: Appraisal of records and the role of archives in constructing social memory. In F. X. Blouin & W. G. Rosenberg (Eds.), *Archives, documentation, and institutions of social memory: Essays from the sawyer seminar* (p. 169-181). University of Michigan Press.

Cook, T. (2013). Evidence, memory, identity, and community: Four shifting archival paradigms. *Archival Science*, *13*(2-3), 95-120.

Duranti, L. (1994). The concept of appraisal and archival theory. *The American Archivist*, *57*(2), 328-344.

Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria*, *39*.

Duranti, L., & Franks, P. (2015). *Encyclopedia of Archival Science* (1st ed.). London: Rowman & Littlefield.

Durnbaugh, D. F. (1959). Johann Adam Gruber, Pennsylvania-German Prophet and Poet. *The Pennsylvania Magazine of History and Biography*, *83*(4), 382-408.

Eveleigh, A. (2017). Participatory archives. In H. MacNeil & T. Eastwood (Eds.), *Currents of archival thinking* (2nd ed., p. 299-325). Santa Barbara, CA, United States: Libraries Unlimited.

Flinn, A. (2007). Community histories, community archives: Some opportunities and challenges. *Journal of the Society of Archivists*, *28*(2), 151-176.

Flinn, A. (2008). Other ways of thinking, other ways of being. Documenting the margins and the transitory: What to preserve, how to collect. In L. Craven (Ed.), *What are Archives? Cultural and Theoretical Perspectives: A Reader* (1st ed., p. 109-128). Farnham, Surrey, England: Ashgate Publishing Company Burlington, VT.

Flinn, A., & Sexton, A. (2019, 01/30). Activist participatory communities in archival contexts: Theoretical perspectives. In E. Benoit & A. Eveleigh (Eds.), *Participatory archives: Theory and practice* (p. 173-190). London, United Kingdom: Facet Publishing.

Flinn, A., Stevens, M., & Shepherd, E. (2009). Whose memories, whose archives? Independent community archives, autonomy and the mainstream. *Archival Science*, *9*(1-2), 71.

Gilliland, A., & Flinn, A. (2013). Community archives: What are we really talking about. In *CIRN Prato Community Informatics Conference.*

Gilliland, A. J., McKemmish, S., & Lau, A. J. (Eds.). (2016). *Research in the archival multiverse* (1st ed.). Clayton, Victoria, Australia: Monash University Publishing.

Girdharry, K. (2019). A personal essay on digital community engagement. *A Journal of Community-Engaged Writing and Rhetoric*, 195.

Heilbrun, D. (2018). *Crowdsourced Metadata for the Our Marathon Digital Archive.* Retrieved 2020/12/03, from `https://dsg.neu.edu/augmentation-and-enhancement-of-crowdsourced-metadata-for-the-our-marathon-digital-archive/`

Huvila, I. (2008). Participatory archive: Towards decentralised curation, radical user orientation, and broader contextualisation of records management. *Archival Science*, *8*(1), 15-36.

Huvila, I. (2010). *What is a participatory archive?* Retrieved 2020/07/06, from `http://istohuvila.eu/what-participatory-archive`

Huvila, I. (2011). *What is a participatory archive? for real(?).* Retrieved 2020/07/06, from `http://www.istohuvila.eu/what-participatory-archive-real`

Huvila, I. (2015). The unbearable lightness of participating? Revisiting the discourses of "participation" in archival literature. *Journal of Documentation*, *71*(2), 358-386.

Jenkinson, H. (1937). *A Manual of Archive Administration* (2nd ed.). London, Great Britain: P. Lund, Humphries & co., ltd.

Ketelaar, E. (2001). Tacit narratives: The meanings of archives. *Archival Science*, *1*(2), 131-141.

Ketelaar, E. (2002). Archival temples, archival prisons: Modes of power and protection. *Archival Science*, *2*(3-4), 221-238.

Lian, Z. (2016). A history of archival ideas and practices in China. In A. J. Gilliland, S. McKemmish, & A. J. Lau (Eds.), *Research in the archival multiverse* (1st ed., p. 96). Clayton, Victoria, Australia: Monash University Publishing.

Manning, C. D., Raghavan, P., & Schutze, H. (2009). *Inverse Document Frequency.* Retrieved 2020/12/03, from `https://nlp.stanford.edu/IR-book/html/htmledition/inverse-document-frequency-1.html`

Matthews, R. J. (2016). Is the archivist a "radical atheist" now? Deconstruction, its new wave, and archival activism. *Archival Science*, *16*(3), 213-260.

McGrath, J., & Peaker, A. (2018). Our Marathon: The Role of Graduate Student and Library Labor in Making the Boston Bombing Digital Archive. In *Digital humanities, libraries, and partnerships: A critical examination of labor, networks, and community* (1st ed., p. 19-29).

Hull, United Kingdom: Chandos Publishing.

MMRS. (2016). *The Mass. Memories Road Show Project Handbook.* Retrieved 2020/07/06, from https://openarchives.umb.edu/digital/collection/p15774coll6/id/1327/

Moss, M. (2008). Opening pandora's box: What is an archive in digital environment? In L. Craven (Ed.), *What are Archives? Cultural and Theoretical Perspectives: A Reader* (1st ed., p. 71). Farnham, Surrey, England: Ashgate Publishing, Ltd.

Ormond-Parker, L., & Sloggett, R. (2012, 06/01). Local Archives and Community Collecting in the Digital Age. *Archival Science*, *12*(2), 191-212.

Owens, T. (2012). *The crowd and the library.* Retrieved 2020/12/03, from http://www.trevorowens.org/2012/05/the-crowd-and-the-library/

Pearce-Moses, R. (2007, 08). Janus in Cyberspace: Archives on the Threshold of the Digital Era. *The American Archivist*, *70*(1), 13-22.

Piggott, M. (2020). Community archives and the records continuum. In J. A. Bastian & A. Flinn (Eds.), *Community Archives, Community Spaces: Heritage, Memory and Identity* (1st ed., p. 41-60). London: Facet Publishing.

Pope, M. H. (1970). Archives from Elephantine: The Life of an Ancient Jewish Military colony. *Journal of Biblical Literature*, *89*(1).

Roeschley, A., Kim, J., & Zavalina, O. L. (2020). An exploration of contributor-created description field in participatory archives. In A. Sundqvist, G. Berget, J. Nolin, & K. I. Skjerdingstad (Eds.), *Sustainable Digital Communities* (pp. 638–648). Cham: Springer International Publishing.

Schellenberg, T. R. (1956). *Modern Archives: Principles and Techniques* (4th ed.). Chicago, IL, United States: The University of Chicago Press.

Schwartz, J. M., & Cook, T. (2002). Archives, records, and power: The making of modern memory. *Archival Science*, *2*(1-2), 1-19.

Sheridan, J. (2018). *Digital Archiving: Context is Everything.* Retrieved 2020/12/03, from https://blog.nationalarchives.gov.uk/digital-archiving-context-everything/

Theimer, K. (2011). Exploring the participatory archives. In *Presentation at the 2011 Society of American Archivists annual meeting.* Retrieved 2020/07/06, from http://www.slideshare.net/ktheimer/theimer-participatory-archives-saa-2011

Tschan, R. (2002). A comparison of Jenkinson and Schellenberg on appraisal. *The American Archivist*, *65*(2), 176-195.

Wang, Y. (2017). Different record different past: The rise and theoretical influence of oral history archives. *Archives Science Study*, *30*(5), 40-44.

佐藤知久, 甲斐賢治, & 北野央. (2018). **コミュニティ・アーカイブをつくろう！　せんだいメディアテーク「3がつ *11* にちをわすれないためにセンター」奮闘記**. 日本: 晶文社.

冯惠玲. (2015). 当代身份认同中的档案价值. **中国人民大学学报** (2015 年 01), 96-103.

冯惠玲, & 张辑哲. (2006). **档案学概论** (2nd ed.). 北京: 中国人民大学出版社.

姜纪云. (2017). “千村档案”建设问题研究——以浙江省 “千村档案” 建设为例. **档案学研究** (6), 68-70.

李子林. (2018). 国外档案多元论的发展及其启示. **档案学研究** (6), 24.

渡邉英徳. (2013). **データを紡いで社会につなぐ：デジタルアーカイブのつくり方** (1st ed.). 講談社.

田丽媛. (2016). **苗族民俗文化建档保护研究** (Master's Thesis). 云南大学.

田村賢哉, 秦那実, 井上洋希, & 渡邉英徳. (2018). ヒロシマ・アーカイブにおける非専門家による参加型デジタルアーカイブズの構築. **デジタルアーカイブ学会誌**, *2*(4), 370-375.

笠羽晴夫. (2011). デジタルアーカイブの 15 年; それが意味するもの. **アーカイブズ学研究**, *15*, 28-37.

裘丽. (2016). 英国社群档案建设及其启示. **档案与建设** (4), 21-23.

金光耀. (2015). 口述历史与城市记忆. **现代声像档案**, *17*(2), 92-93.

Appendices

Table 1: MMRS Subcollections Review List

Road Show name	Year	Total Items	tiff	jpg/jpeg	mpeg
Norwood Mass. Memories Road Show (2004)	2004	106	100	6	
Dorchester (Codman Academy) Mass. Memories Road Show (2005)	2005	7		7	
Dorchester (Codman Square Great Hall) Mass. Memories Road Show (2005)	2005	10	10		
Roxbury (Hibernian Hall) Mass. Memories Road Show (2005)	2005	22	22		
Dorchester Mass. Memories Road Show (2006)	2006	141	110	18	13
Dorchester (Codman Square) Mass. Memories Road Show (2006)	2006	27	17	3	7
Dorchester (Lower Mills) Mass. Memories Road Show (2006)	2006	66	46	12	8
Deerfield Mass. Memories Road Show (2007)	2007	46	29	9	8
Dorchester (Kit Clark Center) Mass. Memories Road Show (2007)	2007	22	13	7	2
Grub Street Project Mass. Memories Road Show (2007)	2007	13	13		
Hebrew Senior Life Mass. Memories Road Show (Day 1, 2007)	2007	85	32	32	21
Hebrew Senior Life Mass. Memories Road Show (Day 2, 2007)	2007	99	50	27	22
Quincy Mass. Memories Road Show (2007)	2007	215	158	35	22

Continued on next page

Road Show name	Year	Total Items	tiff	jpg/jpeg	mpeg
World War II Mass. Memories Road Show (2007)	2007	204	126	41	37
Brewster Mass. Memories Road Show (2008)	2008	93	61	16	16
Natick Mass. Memories Road Show (2008)	2008	81	45	22	14
Remembering Home: Boston Public Housing (2008)	2008	112	82	16	14
Falmouth Mass. Memories Road Show (2009)	2009	167	146	21	
Norwell Mass. Memories Road Show (2009)	2009	203	158	26	19
Project ALERTA Mass. Memories Road Show (2009)	2009	45	44	1	
Reading Mass. Memories Road Show (2009)	2009	162	102	31	29
Truro Mass. Memories Road Show (2009)	2009	216	146	42	28
Danvers Mass. Memories Road Show (2010)	2010	168	120	30	18
Duxbury Mass. Memories Road Show (2010)	2010	307	204	56	47
Stoneham Mass. Memories Road Show (2010)	2010	482	370	75	37
Boston Harbor Islands Mass. Memories Road Show (2011)	2011	217	130	54	33
New Bedford Mass. Memories Road Show (2011)	2011	369	268	55	46
Sharon Mass. Memories Road Show (2011)	2011	195	148	25	22
Waltham Mass. Memories Road Show (2011)	2011	605	435	123	47
Halifax Mass. Memories Road Show (2012)	2012	179	136	26	17

Continued on next page

Road Show name	Year	Total Items	tiff	jpg/jpeg	mpeg
Lowell Mass. Memories Road Show (2012)	2012	254	157	70	27
Milton Mass. Memories Road Show (2012)	2012	369	280	52	37
Peabody Mass. Memories Road Show (2012)	2012	103	81	14	8
Roxbury (Dudley Square) Mass. Memories Road Show (2012)	2012	47	30	10	7
Irish Immigrant Experience Road Show (2013)	2013	310	209	54	47
Lexington Mass. Memories Road Show (2013)	2013	376	214	130	32
Provincetown Mass. Memories Road Show (2013)	2013	399	276	84	39
Stoughton Mass. Memories Road Show (2013)	2013	387	249	93	45
Allston Brighton Mass. Memories Road Show (2014)	2014	224	150	41	33
UMass Boston 50th Anniversary Mass. Memories Road Show (2014)	2014	284	209	44	31
Wayland Mass. Memories Road Show (2014)	2014	408	324	50	34
West End (Boston) Mass. Memories Road Show (2014)	2014	125	66	34	25
Hingham Mass. Memories Road Show (2015)	2015	451	336	66	49
Martha's Vineyard Mass. Memories Road Show (2015)	2015	242	162	61	19
Hyde Park (Boston) Mass. Memories Road Show (2016)	2016	380	320	34	26
Spencer Mass. Memories Road Show (2016)	2016	171	144	15	12

Continued on next page

Road Show name	Year	Total Items	tiff	jpg/jpeg	mpeg
Eastham Mass. Memories Road Show (2017)	2017	346	280	34	32
Marshfield Mass. Memories Road Show (2017)	2017	200	146	34	20
Nahant Mass. Memories Road Show (2017)	2017	213	162	28	23
Wilmington Mass. Memories Road Show (2017)	2017	475	400	43	32
Amesbury Mass. Memories Road Show (2018)	2018	287	229	31	27
Chinese American Experiences Mass. Memories Road Show (2018)	2018	223	162	41	20
Show Em Whatcha Got Mass. Memories Road Show: The Hip-Hop Edition (2018)	2018	338	254	42	42
Winchester Mass. Memories Road Show (2018)	2018	394	311	55	28
Brockton Mass. Memories Road Show (2019)	2019	242	185	30	27
Plymouth Mass. Memories Road Show (2019)	2019	363	299	41	23
Bellingham Mass. Memories STUCK-AT-HOME Show (2020)	2020	20		20	
Malden Mass. Memories STUCK-AT-HOME Show (2020)	2020	56		56	
Total		12351	8956	2123	1272