The Value and Complexity of Collection Arrangement for Evidentiary Work

Ciaran B. Trace
Assistant Professor, School of Information, University of Texas at Austin, 1616 Guadalupe Street, Suite 5.202, Austin, TX 78701-1213. E-mail: cbtrace@ischool.utexas.edu

Luis Francisco-Revilla
Texas Advanced Computer Center, University of Texas at Austin, 10100 Burnet Road (R8700), Austin, TX 78758-4497. E-mail: revilla@tacc.utexas.edu

Archives are invaluable resources for those interested in understanding past activities and events. What makes archival collections valuable as evidence is that they are organized in a way that connects the materials to their creator(s), their associated activities, and recordkeeping systems. In order to create this organization archivists engage in a complex process that involves arranging materials physically and intellectually. This arrangement focuses on documenting the original order of the materials and making them evident to secondary users. The time-consuming nature of this process has resulted in a massive backlog of unprocessed (and thus unavailable) collections, which represents an impediment to the investigative activities of scholars and researchers. To help alleviate such a situation, this article presents a novel approach to archival arrangement, using tabletop computers and digitized images. The article explains the project design and implementation and discusses the evaluation results. This work provides several major contributions to the field, including: a new system that allows archival collections to be arranged digitally, new methods and metrics for evaluating archival arrangements, a detailed analysis of the steps involved in archival arrangement and how they correlate with the final outcomes of the process, and a method for analyzing arrangements based on the topologies created by processing archivists.

Introduction

Many scholars visit archives to find material that can be used as *evidence* of a particular event, story, or scenario. This work involves locating primary sources and then interrogating their external and internal characteristics, so as to make inferences and draw conclusions. Digital technologies such as databases and search engines allow users to access large amounts of archival material online. Users generally like online access because it allows them to have direct access to the materials without having to travel to the archive. However, general-purpose databases and search engines are seldom the best match for the needs of users working with primary sources (Audenaert & Furuta, 2010; Kemman, Kleppe, & Scagliola, 2013; Vajcner, 2008). First, it is not always possible to trace the origin of the documents, or track down how a particular item has been documented and handled through time. Second, for most archive users, search engines are a black box, as users lack the expertise required to understand their algorithms (Kemman et al., 2013). When presenting their results, search engines typically rank them, selecting the order in which items are shown (or not). This automatic selection and ordering often comes at the price of de-contextualizing the materials, since the presentation of the results does not reflect how the creator produced and used the items, nor does it reflect how all items relate to each other. Finally, interactions with search engines and databases represent only a small subset of the complex ecosystem of inquiry used by scholars and researchers. The result is that interpreting the documents becomes increasingly more difficult. Furthermore, the lack of control over their provenance and context compromises its long-term value for evidentiary work.

In contrast to general-purpose databases and search engines, the primary aim of archivists, and archival systems, is to enhance the long-term value of existing materials as evidence and support their interpretation by users. For example, in the trial of persons connected with the forced disappearances committed during Guatemala's civil war, administrative documents from the National Police Historical Archive (Archivo Histórico de la Policía Nacional) were

Received January 14, 2014; revised March 4, 2014; accepted March 16, 2014

 • Published online 7 April 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23295

used. For these records to serve as evidence, proof was required of their authenticity (the documents were genuinely produced by the National Police) and their reliability (the contents represented a full and accurate representation of the activities to which they attested). Ensuring the reliability of these records was the direct responsibility of the archival institution, and the archivists involved accomplished it by preserving the records according to their creator's original order and context (Aguirre, Doyle, Hernández-Salazar, & Guatemala, 2013). As Quintana states, "It is the professionals in the archives, the archivists who know how to deal appropriately with the documents, and those who are aware that their value as witnesses can be affected by the manner in which they are treated. Taken out of context, the majority of documents which bear witness to violations of rights, repression or violence can lose a large part of their value" (Quintana, 2009, p. 76). It is crucial, therefore, that archivists curate their materials by selecting, maintaining, preserving, archiving, and adding value to information objects throughout their lifecycle with the goal of documenting what is known about the reliability of records (as received by the archive), and subsequently making them usable and accessible to interested user communities. Developed and refined over nearly a century of practice, today's process of curating archival collections follows a well-established sequence of phases: appraisal and acquisition, processing (arrangement and description), preservation, digitization, and access.[1] Throughout this process, archivists aim to protect the original order of the collection and try to do minimal rearrangement of it.

However, as the volume of archival material being collected continues to grow exponentially, the limitations of the current curation workflow impede the timely and effective delivery of information to many interested user communities. Nowadays, the workflow for archival curation remains highly segmented and linear in nature. As a result, many collections are stuck in the pipeline, forming a huge backlog. A 1998 survey of Association of Research Libraries' (ARL) special collections libraries found that, on average, about 27% of manuscript collections remained unprocessed, and a similar level (31%) was reported in relation to material in university archives (Panitch, 2001). The situation today is not much improved (Dooley & Luce, 2010). Meanwhile the mean number of archival collections has risen by an average of 50% since 1998 (Dooley & Luce, 2010). Having unprocessed collections, or minimally processed collections, is problematic because these are inaccessible and invisible to scholars and other potential users.

This initial delay in getting collections processed also results in a delay in getting archival content digitized and placed online. In effect, this "process first, digitize second" approach means that collections remain in a holding pattern, often for decades, before being put online. The scope of this problem is nontrivial. A 2010 Online Computer Library Center (OCLC) Research Survey of Special Collections and Archives found that half of archival collections have no online presence (Dooley & Luce, 2010). In the digital humanities field, the lack of available digitized content for research use is talked about in terms of "archival silences."[2] The problem inherent in the traditional "process first, digitize second" approach is compounded by the current archival digitization model which favors selective digitization of collections within an archival repository, and selective digitization of content within chosen collections. This practice ignores the importance of presenting digital objects in context, resulting in digital library collections that may lack any real narrative coherence (Dalbello, 2004). The need to challenge this approach is clear given the call from the OCLC to move away from "boutique" digitization projects to projects that digitize collections as accessioned, in their entirety, at scale (Erway & Schaffner, 2007). Scaling up digitization requires new models that encompass "methodologies for selection of appropriate collections, security, safe handling, sustainable metadata creation, and ambitious productivity levels" (Dooley & Luce, 2010, p. 13).

The goal of the Augmented Processing Table (APT) project is to enable archivists to manage increasing volumes of data and, in the process, to continue to facilitate, and indeed augment, those aspects of the curation workflow that support evidentiary work. A key component of this effort involves the ongoing development and evaluation of a specialized web-based workspace application for archival processing and access, which can run on large or small surface computing devices and empowers archivists to arrange archival collections using digitized images instead of physical materials. This article provides an in-depth discussion of a study of archival curation practices involving both paper and digitized images. It not only sheds light on major issues (e.g., clearing backlogs of unprocessed archival collections), but it also proposes new approaches and methods to study and solve many challenges in the field of archival practice.

[1]The glossary of Archival and Records Terminology (http://www2.archivists.org/glossary/terms/c/collection) provides three definitions for "collection": 1) a group of materials with some unifying characteristic; 2) materials assembled by a person, organization, or repository from a variety of sources; an artificial collection; 3) the holdings of a repository. However, Johnson & Robinson note that:

> "The archival community has not traditionally used the term "collection" to label the aggregates of material they typically describe. Archivists make the distinction between an archival fonds, where the items are of known provenance and their arrangement reflects their original working order as the records of an organisation or individual, and an "artificial collection," where the items are associated but lack the coherence of a fonds. The archivist recognises the fonds as the set of items that have been created and accumulated by an identifiable individual body (or bodies). However, it should be emphasised that both these classes of aggregates (the fonds and the artificial collection) are "collections" in the more general sense in which the term is used here. Within an archival fonds, an item can be fully understood only within the context of its relationship with other items and aggregates in the fonds, and descriptive practice reflects this" (Johnston & Robinson, 2002, p. 2).

This paper uses the term "collections" in the general sense, but in reference to organic rather than artificial collections.

[2]Theimer, K. (2012, March 27). Two meanings of "archival silences" and their implications. In ArchivesNext. Retrieved from: http://www.archivesnext.com/?p=2653

Specifically, the article explains how the methodology of human-computer interaction can be used in the field of archives to: better understand the practices, needs and behaviors of the different stakeholders; design sociotechnical solutions; and objectively evaluate their effectiveness. The article discusses the literature related to contextual information and evidentiary work, followed by a description of the methodology used in the study. The subsequent two sections present the results, and the final section concludes by discussing the implications of the findings and the outlining of future work.

Related Work

APT is situated within the literature that pays attention to the contextual and evidential value and meaning of information as it plays out in the realms of information creation, organization, access, analysis, and preservation. It is said that "in the real world, objects never occur in isolation; they co-vary with other objects and particular environments, providing a rich source of contextual associations" (Oliva & Torralba, 2007, p. 520). In particular, the context in which information objects reside encompasses "the setting and intention within which they were created, the persons and activities involved, as well as the time frame and potentially correlating activities that are somehow linked to or influence . . . [an information] object" (Mayer, Neumayer, & Rauber, 2009, p. 1). Thus, context consists of "the relationship that an information object shares with other information items (such as creation time, type, purpose, creators, users, and others), but also of the embedding of the very content itself, that a piece of information is conveying, such as style, genre, facts, references to other documents, etc." (Mayer & Rauber, 2009, p. 1). The context in which objects reside is an important facet in interpreting those objects; this context also helps to establish the authenticity of objects, and facilitates their subsequent retrieval and use (Mayer et al., 2009). Preserving or associating information together with this larger network of relationships is what allows information to serve as evidence. The intersection of information technology and the contextual and evidentiary aspect of information is a particularly fruitful line of enquiry in fields such as:

- Digital Libraries—for example, the study of the types of contexts and their relevance to the information mediation process in digital libraries (Neuhold, Niederée, Stewart, Frommholz, & Mehta, 2005), and the study of ways to (semi-) automatically determine various contexts of the creation and use of digital objects before ingest into a digital library system (Mayer & Rauber, 2009).
- Digital Preservation—for example, identifying the types of metadata that can be utilized within the preservation process so as to re-contextualize material for future use (Beaudoin, 2012).
- Digital Forensics—for example, designing systems for e-mail recovery that visualize the process by recreating the evidentiary environment and data (Tiwari, Samaddar, Singh, & Dwivedi, 2010).
- Computer Supported Cooperative Work—for example, researching the efficacy of collaborative problem-solving in the presence or absence of complete evidence (Balakrishnan, Fussell, Kiesler, & Kittur, 2010).
- Personal Information Management—for example, highlighting the contextual factors that influence how information is organized in the workplace (Kwasnik, 1989; Barreau, 1995; Henderson, 2005).
- Artificial Intelligence—for example, designing interactive educational assessment environments that take an evidence-based approach in order to support particular assessment claims (Zapata-Rivera, Hansen, Shute, Underwood, & Bauer, 2007).
- Evidence-based medicine (EBM) and evidence-based software engineering (EBSE)—for example, designing technological solutions (text-mining based pipelines) for accelerating the creation and updating of accurate and comprehensive medical evidence reports (Cohen et al., 2010).

In particular, building context is said to be the "sine qua non" of historical research (Duff & Johnson, 2002, p. 486). Context is particularly important to historians (and other humanities scholars) because their mode of research is centered not just on seeking and finding information, but also on the act of contextualizing and therefore interpreting that information (Duff & Johnson, 2002). This interpretation relies not only on building knowledge of the people, organizations, and events that brought the information into existence, but what can be gleaned in the way in which the documents were organized and colocated by the creator (Duff & Johnson, 2002; Duff, Monks-Leeson, & Galey, 2011). Here, the idea of evidence is bound up with notions of proof, trustworthiness, justification, and the establishment of facts. Co-opting Buckland's framework on the meaning of the term "information" (Buckland, 1991), we could say that evidence presents itself as both something intangible (evidence-as-knowledge) and as something tangible (evidence-as-thing). In the first scenario, evidence is seen as *that which justifies belief*. Evidence-as-knowledge exists as propositions, experiences, justifications, beliefs, and mental models. In order to share this knowledge, it must be represented in some concrete form. In the second scenario, evidence and information objects are collocated. Evidence, in this sense, is tangible, or "real," information (Furner, 2004). It is information that exists in context—a fingerprint lifted from the crime scene, a newly discovered document in the archive, or the fossil Archaeopteryx found in a Solnhofen limestone formation.

Such evidentiary work would not be possible without the value-added work that information professionals do with representations (information that has been further processed or otherwise transformed) to reveal, highlight, and perpetuate the evidential nature of information objects and collections (Buckland, 1991). In particular, supporting evidentiary work is the *raison d'être* of the archival activity of arrangement (Meehan, 2006).[3] To archivists, maintaining

[3]It is the archival principles of provenance and original order that ensure that context and evidence remain in the forefront of archival thinking.

DOI: 10.1002/asi

the “provenancial properties of records,” and the “relational properties of archives,” creates representations that are “more contextually rich than loosely assembled collections” (Sternfeld, 2011, p. 551). Before making collections of historical records available to interested user communities, archivists follow a process of inference, identifying and documenting the network of relationships between the records and their external and internal context/structure (Eastwood, 2000). In carrying out this work, archivists draw from the idea that “records have an innate hierarchy imposed by the creating agency’s filing practices and position in a bureaucratic hierarchy and by the processes through which the records were created” (Gilliland-Swetland, 2000, p. 18). This innate hierarchy is reflected in a framework of evidentiary layers or “collectivities” (van Ballegooie and Duff, 2006) that are placed around the records—records are linked to their creator and associated functions, events, geography, and chronological periods (external structure), and to the recordkeeping system and material artifacts in use by the creator (internal structure). Modern archival arrangement defines layers by nesting sets of documents into five levels:

- Fonds: all records of an entity/from the same provenance.
- Subgroup: subset of records with a distinct external structure of provenance.
- Series: group of similar records created, received, or used in the same function or activity, and filed accordingly (internal context which captures the documentary context of the material).
- Subseries: a set of documents within a series, distinguished from the whole by a filing arrangement such as type, form, or content.
- File: set of documents related to the same matter or event.

In documenting these networks or contexts in explicit arrangement schemes (captured within descriptive accounts called “finding aids”), archivists augment collections in a way that supports the evidentiary work of others.

System Design and Development

The APT project forms part of the larger study of the role of interactive technology in environments where people are specifically engaged in information-intensive work. In the professional realm, Pierre Wellner at EuroPARC developed the first tabletop computer system in the 1990s (Wellner, 1991). Wellner’s DigitalDesk explored the use of more natural interactions in the physical world that could also be supported and augmented by the computer’s functionality. Subsequently, Wigdor, Perm, Ryall, Esenther, and Shen (2007) expanded this research by evaluating the viability and effectiveness of using multitouch tabletop computers in lieu of the traditional office computer. While Wigdor et al. focused on single user environments, Scott et al. (2010) report on of the use of tabletop interfaces in support of collaborative planning and decision making in military maritime operations. Similarly, Clifton et al. (2011) studied how digital tabletops can be used by engineers and designers in the collaborative design process, and Nebe, Klompmaker, Jung, and Fischer (2011) looked at how processes in disaster control management can be optimized using multitouch and other tangible interaction techniques. The intersection of multitouch interfaces and information work is also being looked at in combination with sophisticated applications that facilitate such information activities as document review in e-discovery (O’Neill, Privault, Renders, Ciriza, & Bauduin, 2009) and collaborative web search among colocated group members (Morris, Lombardo, & Wigdor, 2010).

Following in these footsteps, the design process for APT has iterated through several cycles of prototyping, testing, analyzing, and refining. In the fall of 2011, the researchers (Francisco-Revilla & Trace) began to explore the problem space by studying the viability of creating evidentiary layers by taking a “digitize first, process second” approach (Crow, Francisco-Revilla, Norris, Shukla, & Trace, 2012). While the viability of the “digitize first, process second” was validated by the first prototype, a number of key questions remained unanswered, including whether archivists create different evidentiary layers when using digitized images as opposed to when using paper materials, and how well the APT system actually supported evidentiary work, from a usability standpoint. The second APT prototype (APT 2.0) is based on the following principles:

- Support the creation of evidentiary layers even when some or all traces of the original order have been lost.
- Base the visualizations of the evidentiary layers on the standard hierarchical structure used by archives (i.e., subgroup, series, subseries, files, and items).
- Use high-quality digitization that allows users to easily infer the actual physical characteristics of the materials.

The functionality of APT 2.0 facilitated:

- Reviewing all documents within a collection (facilitated through dispersal and scrolling functions, as well as the extensible nature of the workspace).
- Grouping of materials using different tools (virtual clips, piles, and containers).
- Entering metadata and taking notes (facilitated at the document and workspace level).
- Matching and comparing pages (including a zoom function at the document level and functionality that renders a document transparent when it is moved and placed over another document).

System Design

Specifically, APT 2.0 is a $5' \times 5'$ standing-height interactive tabletop computer with a $47'' \times 28''$ interactive area

Provenance refers to the idea that records which come from an entity (organization, family, or individual) form an organic unit and must be kept together. Original order refers to the idea that the way that the creator originally filed their records provides important context about the activity that generated the records in the first place, and therefore this original order must be preserved.

DOI: 10.1002/asi

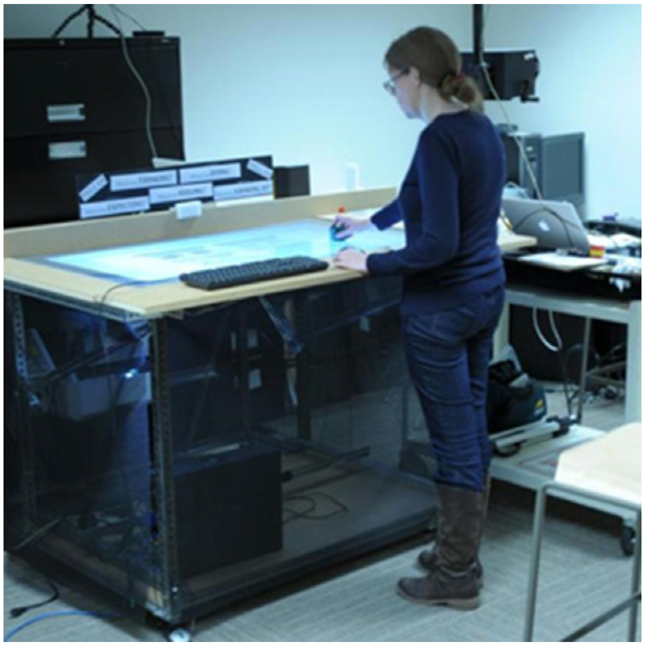


FIG. 1. APT standing-height interactive tabletop computer. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

running WARP (see Figure 1). APT uses the rear Diffused Surface Illumination (DSI) technique to provide multitouch support. DSI involves shining nonvisible light through a clear surface (in this case a pane of acrylic) that shows machine-readable "blobs" where someone touches the pane. Webcams located underneath the touch screen capture these blobs, which are converted to TUIO (Table-Top User Interfaces Objects) protocol data using the Community Core Vision (CCV) software, allowing the user's touch to be understood by the software platform. The client side of the application is written in JavaScript and HTML, utilizing the jQuery and Fabric.js frameworks, running in a touch-enabled web browser. It supports both touch and mouse input, although touch is the primary method of input. The server side of the application is a Java Servlet, and is responsible for storing and retrieving document workspace states that are rendered in the client. The state of the workspace is saved incrementally, giving the ability to replay the history of system actions in whole or in part.

The APT prototypes use a visual workspace in order to facilitate the curation task of archival arrangement. The workspace uses a Spatial Hypermedia interface that provides "a big 2D space (a canvas) for sorting information" (Grønbæk, Kristensen, Ørbæk, & Eriksen, 2003, p. 11). In spatial hypermedia all objects are placed in a space and have explicit visual characteristics such as horizontal and vertical position, background color, and font size. A distinctive feature of spatial hypermedia is its ability to represent both explicit and implicit relationships. Explicit relationships can be represented in various ways. For example, objects can have explicit links between them, just like traditional webpages. Another way to represent explicit relationships is by using hierarchical spaces (box-within-a-box metaphor). The explicit object characteristics (e.g., position and appearance) make it possible to express relationships implicitly. For example, a group can be expressed implicitly as a "pile." Spatial hypermedia supports several types of implicit groups, including horizontal lists and vertical lists. By manipulating the visual characteristics and spatial positioning, users can express and adjust fine nuances (e.g., arranging objects as vertical lists to represent an ordered sequence, or using piles to represent unordered groups). More important, the manipulation of objects allows users to reflect-on-action and reflect-in-action as they analyze and develop an understanding of the overall structure and organization of the objects. APT 2.0 was developed by using and extending a Web-based Spatial Hypermediaplatform called WARP (Web-based Multi-Model Adaptive Spatial Hypermedia), which was previously created by Francisco-Revilla (Francisco-Revilla & Shipman, 2004).

The APT system is built around the concept of a workspace with a direct manipulation interface (see Figure 2). When document images are first imported into APT they are presented as a pile. Users can choose to work with them one-by-one, or they can tile them across the workspace to get an overall impression of the size of a collection. In the initial state, all items are located at the root level of the workspace hierarchy. Users can move and scale items freely in 2D. APT facilitates matching document activities (e.g., identifying all the pages in a handwritten letter) by making documents semitransparent when moved over other documents. APT's function to digitally clip documents together supports collectional and collocational actions. APT uses the visual metaphor of a container to represent the action of creating evidentiary layers (subgroup, series, subseries, files). Icons on a toolbar allow users to create these virtual containers. Each type of container is associated with a different color to help differentiate the evidentiary layers. Users create, designate, and title the groups to which they can add or remove items. In turn, these containers can be sized and resized, and related to each other through the nesting of these virtual containers (series, for example, can be nested within subgroups). These containers can be redesignated through the use of a properties box, and can also be destroyed (although not the contents in them) if the grouping is of a temporary nature. Users can view and edit the metadata associated with each document and with each virtual container through a contextual menu. Currently, the metadata fields capture basic information about a document such as the title, date(s), and genre/form. The ability to add additional comments and annotations to documents and any of the evidentiary layers is facilitated through the use of digital post-its. At the workspace level, a search feature allows documents to be identified and highlighted, based on information recorded in the metadata fields. The workspace itself is extensible—users zoom in or out as needed. This allows documents, and aggregations of documents, to be viewed true-to-scale, as well as at other ratios, as required.

FIG. 2. APT interface. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Research Process and Evaluation

Research Objectives and Questions

The study considered two main research objectives. The first included two research questions (RQ) directed toward understanding the process of archival arrangement and the creation of evidentiary layers. The second included three research questions directed toward measuring the level of success of APT in terms of improving the process of arrangement and improving the outcomes of the process.

Objective 1: Understanding the process of archival arrangement.

RQ. 1.1. What is the nature of the arrangement process?
RQ. 1.2. Do archivists create similar or different evidentiary layers (arrangements) based on variables such as media type?

Objective 2: Measuring the success of APT.

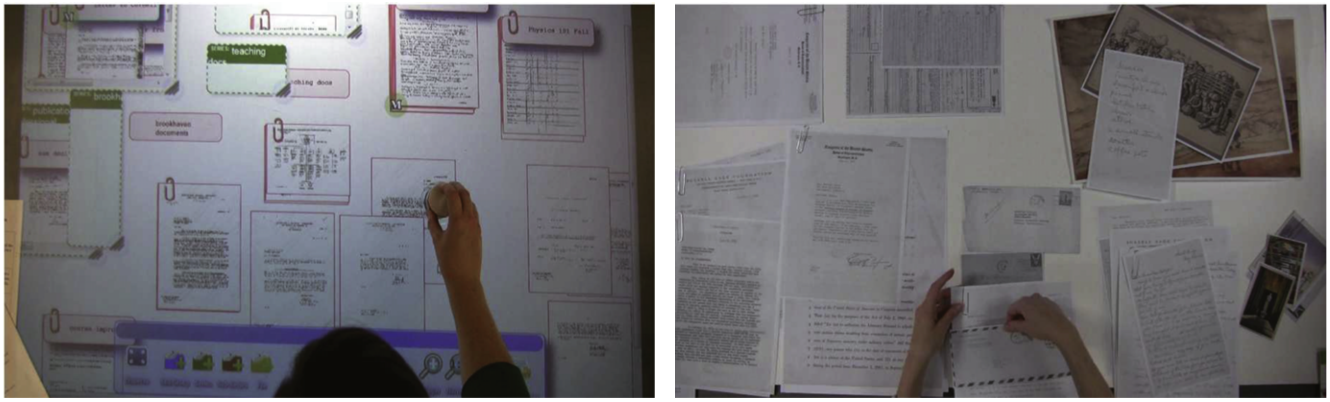
RQ. 2.1. Is APT an effective tool for processing a collection?
RQ. 2.2 Is APT an efficient tool for processing a collection?
RQ. 2.3 Is APT the preferred tool for processing a collection (preferred over the baseline)?

The first two questions were examined in terms of process (phases and styles) and the topologies (structure) of the resulting arrangement schemes. For the remaining questions, effectiveness was measured by task success (error rates and overall scores for the assigned task) and quality of the outcomes (i.e., how good were the arrangements). Efficiency was measured using time (total and normalized) and looking at correlations between time and all other effectiveness metrics. Preference was measured using a Likert scale questionnaire at the end of the experimental sessions.

Research Procedures

In order to address the research questions, the project evaluated APT 2.0 in relation to a baseline system, which for this study was the current best practice in archival science for processing collections of historical materials. The baseline system consists of a large horizontal workspace that allows archivists to work with physical materials, arranging them into recognized groupings (evidentiary layers), and take notes on a separate medium (such as notepads and laptops). In contrast, the APT system consists of a large multitouch table computer that allows archivists to interact with digitized materials, group them virtually, and take notes.

In order to compare the effects of creating evidentiary layers using digitized images or paper materials, the study sought out a group of participants to work with a collection



a) Digital condition b) Paper condition

FIG. 3. Experimental conditions for arranging materials. (a) Digital condition. (b) Paper condition. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of records and create an evidentiary layer that captured their provenancial, procedural, and documentary contexts. Since APT ideally should be usable by anybody who has a minimum knowledge of the archival method of arrangement, and since new entrants to the archival profession are often sought out as processing archivists, the study recruited 16 graduate archival students (ages 18 to 40) who knew the method, even though they were not experts. The participants were mostly female (15 out of 16), 18 to 39 years old, and had a variety of educational backgrounds (science/engineering, information science, art history, literature/languages and history, and other humanities disciplines). As part of the study, participants were asked to conduct two tasks: to process a collection of historical materials on each system (both APT and baseline). In order to ensure that the experimental collections were truly representative, the study selected two publicly available archival collections: Estelle Ishigo Papers and Samuel Goudsmit Papers. Because of the size of the collections, it was necessary to select a subset from each, which participants could process in one session. Randomized selection was not appropriate, as there was no way to guarantee that the resulting subset would have the evidentiary relationships required for the study. Hence, the researchers had to assemble the test collections manually, using the existing finding aids and biographies to ensure that they provided evidence of different aspects of the life of the creators. Most traces of the original order (original filing system) were removed from the experimental collections. While knowing the original order of a collection is a great aid for creating an evidentiary layer that captures its provenance and recordkeeping, APT aims to support this creation even when this knowledge is not directly available. The final experimental collections included 22 items for Ishigo (55 pages), and 24 items for Goudsmit (47 pages).

The processing of the two collections took place on separate sessions, separated by a minimum of 1 day and a maximum of 4 days. As Figure 3 shows, in each session

TABLE 1. Participant groups.

	Task 1		Task 2	
Group A	Baseline	Ishigo	APT	Goudsmit
Group B	Baseline	Goudsmit	APT	Ishigo
Group C	APT	Ishigo	Baseline	Goudsmit
Group D	APT	Goudsmit	Baseline	Ishigo

participants created an evidentiary layer (arrangement scheme) for a collection, once while using APT (digital) and once while using paper (baseline).

The experiment was balanced to account for factors such as fatigue, learning, and innovation. Half of the participants used APT first and the baseline second, and the other half used the baseline first and APT second. As Table 1 shows, the study performed a within-participant comparison that considered eight conditions (2 systems $\times$ 2 collections $\times$ 2 sessions). Participants were assigned randomly to one of four groups (A, B, C, or D). In order to account for the lack of familiarity with the APT table, the APT sessions began with a short tutorial given by a member of the research team (Francisco-Revilla), in which the participant was shown APT's functionality using a subset of the test collection.

Data Collection and Analysis

The study used multiple research techniques. At the beginning of the first session, a survey collected demographic data including age, gender, technological background, and experience of arranging collections. Data from the sessions included video recordings (1.27 terabytes of data from 32 usability sessions of ~2 hours each, running two high-definition video cameras), researchers' notes taken when the sessions were underway, participants' notes, and HTML and WARP files that captured the final arrangement

DOI: 10.1002/asi

of the collections on the APT table. At the end of each session, each participant also produced a final arrangement scheme in written form. At the end of the second session, an exit survey collected the participants' assessments of using both systems. The exit survey started by asking about the participant's perception of time and learning outcomes vis-à-vis the two systems. The survey also asked participants what they liked and did not like about processing collections using APT, and whether or not they would like to use APT in the future. Finally, a free-form interview at the end of the second session asked participants for additional comments and suggestions (when transcribed this resulted in ~131 pages/55,000 words of transcribed text).

The study included two analysis stages. In the first stage, the researchers analyzed the results of each research technique independently, using quantitative and qualitative methodologies as appropriate. For example, the initial and exit surveys were analyzed quantitatively, and the interviews were analyzed qualitatively. In the second stage, the researchers integrated and correlated the results of all research techniques. The analysis of the video recordings used a combination of qualitative and quantitative methodologies. Following a grounded theory approach, two team members marked each video recording, using labels (codes) and timestamps, to indicate what the participants did throughout the arrangement task. Then, one of the PIs (Trace) reviewed the emerging coding schemes of both versions in order to identify variations and discrepancies, and to unify the choice of words for the coding. Mostly, the two versions were quite similar, with only slight variations in the timestamps. The few discrepancies in the interpretation of the participants' activities were resolved by discussing them as a group. At the end, this analysis provided a unified codebook and an annotated timetable for every experimental session. The researchers analyzed these timetables, looking for patterns in the order of activities. This helped to answer Research Objective 1. The researchers compared the timetables for the APT and baseline conditions in order to detect potential effects of the different activities on the quality of the final arrangement and the total time required to complete it. This helped to answer Research Objective 2.

The analysis of the final arrangement schemes consisted of three parts: quantifying errors, comparing the shape and features of the final hierarchy, and assessing the overall quality of the final arrangement schemes produced by the research participants. Assessing the overall quality was particularly challenging, as the archival profession currently has no metrics in place to judge the quality of archival arrangement. The researchers used a holistic evaluation approach that allowed clear metrics to be devised for judging the quality of an arrangement. The process was as follows:

1. Rank the arrangements, one by one, based on their quality according to the metric being considered.
2. Verbalize and discuss the reasons why one arrangement was placed above or below another.
3. Capture all the arguments and rationale used for the judging.
4. Discuss and refine the rationale that provides a base for the current metric.
5. Identify whether or not the rationale includes aspects belonging to different dimensions.
6. Create a more precise definition that helps to assess the arrangements based on that dimension.
7. Iterate by going back to step 1, (re)evaluating the arrangements based on the new metric.
8. Come up with a final ranking or absolute score for the arrangements.

In order to define these dimensions explicitly, the panel of experts focused on the following questions:

- What affordances do we value in an arrangement?
- Why do we value that affordance of the arrangement?
- How does an arrangement provide this affordance?

After two sessions, the expert panel generated five metrics for evaluating arrangements. Three of the metrics resulted in a ranking: *Narrativity* (whether the arrangement accurately reflected the life of the creator); *Recordkeeping* (whether the arrangement reflected the recordkeeping practices of the creator); and *Materiality* (whether the arrangement accurately expressed the types of material contained in the collection). The two metrics that resulted in an absolute value or score were *Perspective* (whether the arrangement explicitly reveals the meta-level context of the materials), and *Coherence* (whether the arrangement follows the archival rules and established a coherent structure).

Results

Process (Phases, Styles, Time)

As a process, archival arrangement has become "naturalised and internalised" (Cook & Schwartz, 2002, p. 181) to the point that "explanations of the archival reasoning process, by and large, are absent from the archival literature" (Lemieux, 2012, p. 7). To our knowledge, only one other published research study has provided an analysis of the process that archivists go through in arranging archival collections. Researchers at the University of British Columbia used a think-aloud cognitive task analysis method to discover what two archivists were thinking, feeling, and looking at when team-processing a collection (Lemieux, 2012). This research established that archival processing to the file level involves a "three-stage sense-making process" (Lemieux, 2012, p. 9). Stage one involved the archivists analyzing and developing a mental model of the overall structure of the collection and creating a draft arrangement. The second phase involved confirming and refining the arrangement structure. In the third stage, the archivists documented the final arrangement in an archival finding aid. The researchers observed that the work involved in stage one (to figure out the timeline and context of the collection)

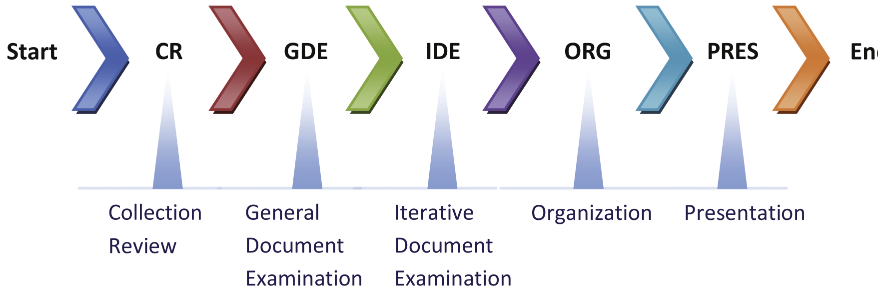


FIG. 4. Arrangement phases. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Phase	(Optional) 1	2	3	4	5	Participant count
Style 1	CR Collection Review	GDE General Document Examination	IDE Iterative Document Examination	ORG Organization	PRES Presentation	11
Style 2		GDE	IDE/ORG	ORG	PRES	4
Style 3	CR	GDE/ORG	IDE/ORG	ORG	PRES	12
Style 4		GDE/ORG	IDE/ORG	ORG/PRES		1
Style 5		GDE/ORG		ORG	PRES	2

FIG. 5. Arrangement styles. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

created a significant cognitive load, with participants reporting feeling more fatigued during this phase of the process (Lemieux, 2012).

The analysis of our video recordings corroborated the fact that the task of archival arrangement has clearly distinguishable phases. Furthermore, we found that while there were slightly different processing styles, most participants followed the same process of sense-making, independently of whether they used paper or digital images. Figure 4 shows the five phases of the arrangement process.

- Collection Review [CR]
  - Objective: becoming familiar with materials, preserving order as received.
  - Involves quickly surveying the material in order to get a sense of what is in the collection. From all phases, this one was the only one “optional” (not everybody used it).
- General Document Examination [GDE]
  - Objective: becoming familiar with material, and inferring structure of collection.
  - Involves a systematic and/or methodical look at all the documents in the collection, and results in the creation of basic piles (groups).
- Iterative Document Examination [IDE]
  - Objective: refining structure.
  - Involves going back through the material after everything has been looked at once.
  - Results in the creation of refined piles (groups) that designate sets of related documents.
- Organization [ORG]
  - Objective: designating structure.
  - Involves manipulating sets of material (physically and intellectually) as they are assigned a defined archival level (file, series, subseries, and subgroup).
  - Results in at least a draft arrangement scheme.
- Presentation [PRES]
  - Objective: cementing structure.
  - Involves writing the final arrangement, during which the participant envisions himself/herself as a reader. In the process of writing the final arrangement, a final refinement may take place.
  - Results in the final arrangement.

In general, participants followed the sequential process shown in Figure 4. However, participants sometimes merged two phases into one. This gave rise to the five processing styles shown in Figure 5.

In Style 1, each of the four or five phases took place in a regimented fashion. In Styles 2 through 5, the organizing phase (that phase where groups form highly defined sets) runs in parallel with other phases. This means that participants were making determinations about what constitutes a clearly defined set at the same time as they were getting to know the collection and were refining the preliminary groupings. The separate and distinct phase for formalizing sets (Style 1) was the preferred style for paper collections (particularly in the case of paper Ishigo). On the other hand, a more integrated organization phase (Style 3) was the preferred style for collections processed on APT. This suggests that the functionality of the APT system prompts participants to conceive and formulate defined sets earlier in the arrangement process.

In addition to identifying the phase-composition of each processing style, the analysis studied the absolute and relative time spent on each phase. Figure 6 shows the absolute processing time for all experimental conditions.

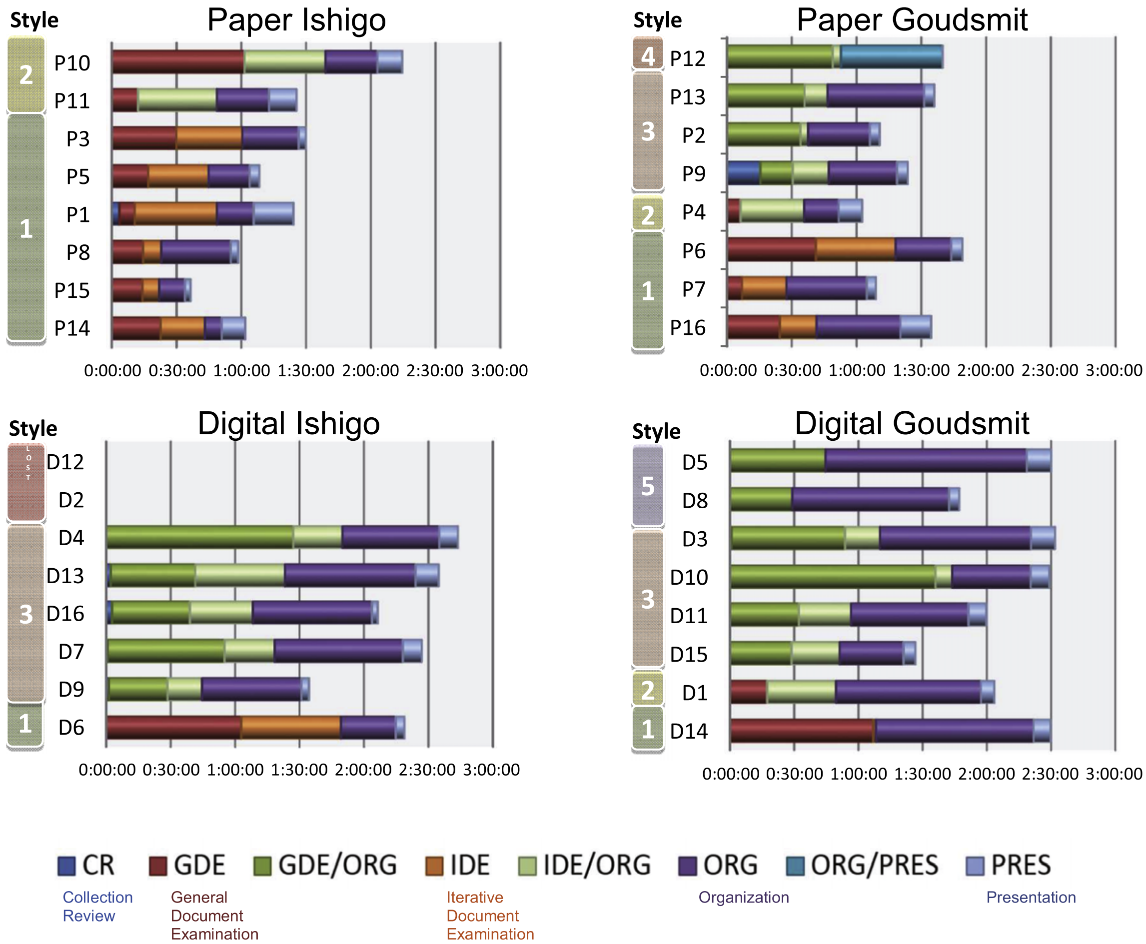


FIG. 6. Absolute processing time. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The processing style for each sample is indicated by the style marks left of the vertical axis. Two samples of the digital Ishigo condition could not be properly quantified due to an unexpected system crash that disrupted the processing task.

The results show a large variation in the time required to arrange the collections, even when following the same processing style. At the aggregate level, there were no statistically significant differences between collections. However, as Figure 7 shows, participants processed the paper collections significantly faster than they did the digitized collections on APT ($p < .01$). This difference in processing time was due to several factors. First, participants were more familiar with paper, and had limited experience with APT. Second, the experiment revealed some bugs, glitches, and problematic interactions in APT that slowed the overall performance (e.g., clipping function). Third, there was an uneven workload between the two conditions. Specifically, participants using the digital version had to process more documents, since in the digitization process two-sided documents become two separate and unique images. While these findings highlight some problems, they provide important insights for the design of the next prototype.

	Ishigo	Goudsmit
All	Average = 1:43:40 St. Dev. = 51:50	Average = 51:50 St. Dev. = 3.22
Paper	Average = 1:33:49 St. Dev. = 22:57	Average = 1.875 St. Dev. = 2.17
Digital	Average = 2:31:21 St. Dev. = 1:10:54	Average = 5.375 St. Dev. = 3.25

FIG. 7. Average absolute processing times.

In order to study and compare the potential impact of the different phases and alternative processing styles, the analysis normalized all the times, making them proportional to the total time required by each participant to complete the arrangement. Figure 8 shows the relative processing times.

The study looked at correlations between style, specific phases, and total time. The researchers found no correlation between style and total time, nor between total time and IDE (Iterative Document Examination), IDE/ORG, or PRES

DOI: 10.1002/asi

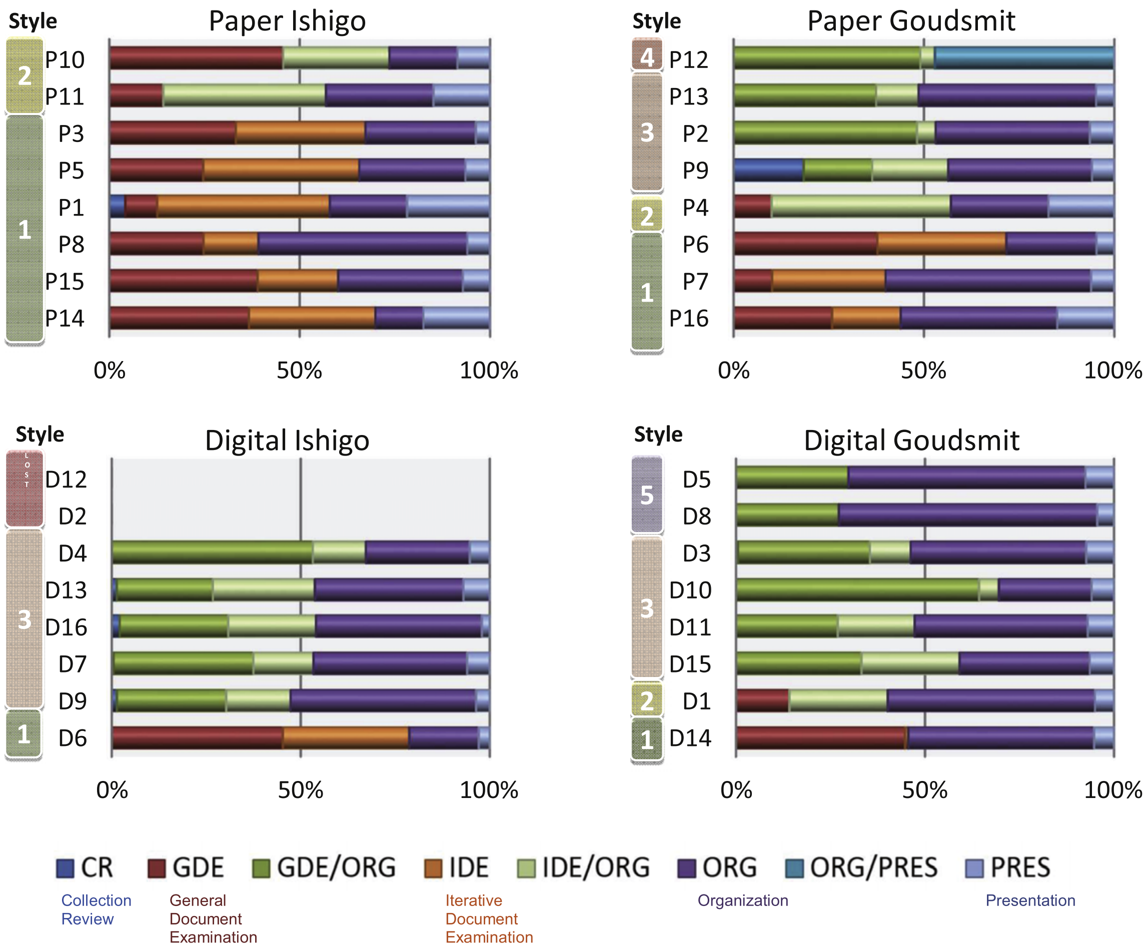


FIG. 8. Relative processing time. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

(Presentation) phases. However, the more time a person spent on the GDE (General Document Examination) and ORG (Organization) phases (or combined GDE/ORG phases), the more total time was spent processing the collection. This implies that future versions of APT should include functionality aimed at reducing the time required to complete these two phases. Finally, participants who included a CR (Collection Review) phase spent less time on the ORG phase, and less total time processing the collection. While this is a small sample, this observation provides motivation to extend the support of APT for an initial Collection Review in the next prototype.

Outcome (Topology, Errors, and Overall Quality)

Based on the participants' data, the researchers generated a composite interpretation for each system (digital and paper). This revealed hidden patterns, and provided valuable insights about the overall nature of the arrangements and the effects of using digitized images or paper materials. The analysis compared the current practices with APT in terms of replicability, processing errors, and overall quality. Specifically, replicability refers to the ability of both methods to let different archivists produce similar arrangements for a given collection. Processing errors refers to mistakes committed by the participants when trying to identify all the pages belonging to a single item. Overall quality refers to the degree that the final arrangement adds value to the collection. All of these aspects speak to the issue of quality control in archival work, and the ability of researchers to trust that the outcome of the work of archival arrangement will be similar across multiple conditions.

Looking at the replicability aspect required examining the structure of the different hierarchies produced by all participants. This paper refers to the structure of the hierarchies as the *arrangement topology*. The following subsection (A) presents the topological features common to all arrangements and the variations associated with different experimental conditions. In addition to identifying and quantifying mistakes objectively, the issue of processing errors required studying the factors that influence the propensity for human errors in the resulting arrangement schemes, such as hard documents and the platform used (paper or digital). This is discussed in subsection B. The assessment of the overall quality of the final collection is discussed in subsection C.

DOI: 10.1002/asi

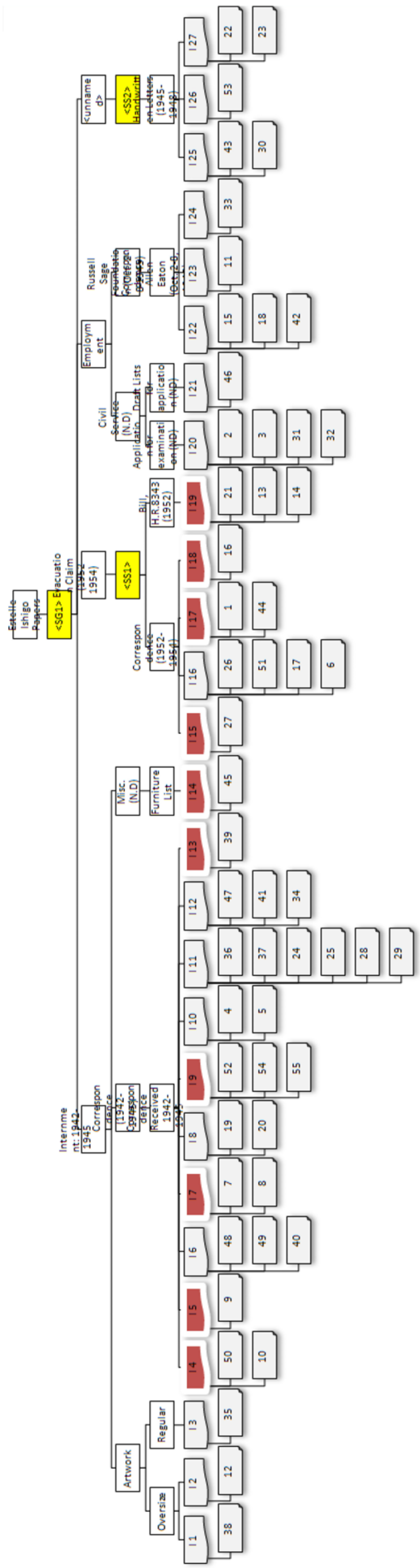


FIG. 9. Hierarchical tree visualization. This facilitates browsing one arrangement at a time. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

A. Topological characteristics of arrangements. The basis for the topological analysis was the written arrangement schemes that each participant produced after processing a collection. Each arrangement scheme laid out the sets or levels (subgroup, series, subseries, file, etc.) that the participant had deduced in processing the collection. The analysis compared topologies across experimental conditions and aggregated them within experimental conditions. This was facilitated by creating visualizations that made it easy to examine the topology as a single (hierarchical tree [Figure 9], or to compare multiple trees [Figures 10 and 11]).

The topology of the arrangements refers specifically to how nodes in the tree are connected to each other. In order to compare trees, it was necessary to normalize their topologies such that every tree had the same depth and included all levels. Creating phantom nodes (shown in yellow in Figures 9 to 11) in all the empty levels (the levels that the participants did not include) made it easier to conduct a side-by-side visual comparison (the red nodes signify matching errors).

The researchers studied the arrangements using large high-resolution displays that showed an integrated visualization of all the experimental conditions simultaneously. However, to show the integrated visualization here would be impractical due to the limited space available. Therefore, the visualization has been divided into four figures, 12 to 15, which show the trees for Groups A, B, C, and D. In all figures, each column shows the two arrangements created by a given participant. The background shading makes it easier for researchers to quickly distinguish arrangements with specific characteristics. For example, purple means arrangements that do not follow all the principles of archival arrangements (e.g., nesting errors). In total, there were only two cases of nonstandard arrangements. In six cases, we detected a bug in the APT system that complicated the interaction with some documents (orange color). One sample was lost completely, due to a system crash late in the task. While much of the task was completed and videotaped, the participant was unable to write the final arrangement (gray color in Figure 15). Therefore, this sample was removed from the analysis.

Figures 12 to 15 illustrate how researchers visually compared all the arrangements by participant, and across experimental conditions. While the text is not readable at this scale, it is possible to see which conditions had more errors and more phantom levels. The analysis revealed two important observations:

1. Participants showed a great deal of flexibility when it came to using optional levels.
2. Participants only partially followed the accepted method (i.e., arrangement schemes were missing expected levels such as subgroups).

While the number of possible topologies was shaped by the constraints of the archival method and the

DOI: 10.1002/asi

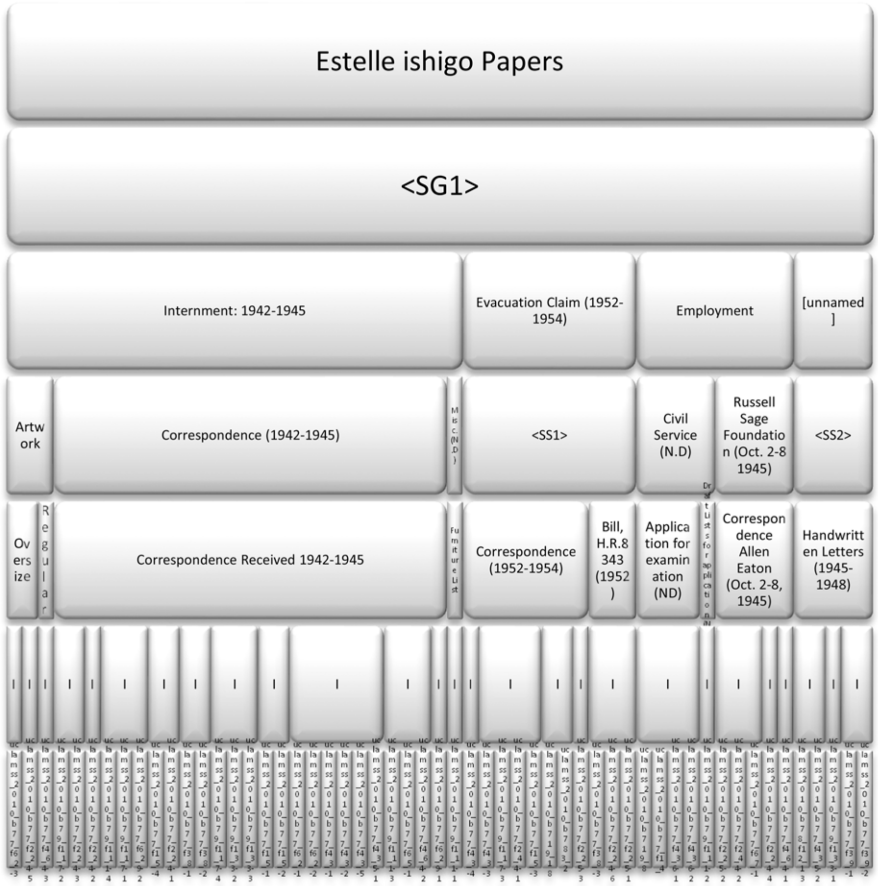


FIG. 10. Square visualization. This facilitates side-by-side horizontal comparisons.

experimental conditions, the results were unexpected. The initial hypothesis was that the (seemingly prescriptive) nature of the archival methodology would produce similar arrangements. However, as Figures 12 to 15 show, there were marked topological variations across all participants.

In order to determine if these individual variations would "average out" when considered as a group, the analysis aggregated all the trees in each condition using an agglomerative hierarchical clustering (AHC) algorithm. Figure 16 shows that for both collections, the aggregate trees for paper or digitized images have clear differences at the subgroup, series, and subseries levels.

Based on the topology of the majority of the trees in a set, AHC builds an aggregate tree, one level at a time, starting at the pages and finishing with the fonds, for example, AHC connects pages P1 and P2 at the item level, if and only if, P1 and P2 have the same parent in more than half of the trees.

In order to better understand the variations shown in Figure 16, the analysis quantified, aggregated, and analyzed several topological features including real and phantom nodes. Figures 17 and 18 show the real, phantom, and total (sum of real and phantom) nodes in paper and APT experimental conditions. The average count is shown as a solid line, green for Ishigo, and blue for Goudsmit. The maximum and minimum counts are shown as green rectangles and red rectangles, respectively.

The analysis revealed statistically significant differences between arrangements constructed using paper or digitized images. Specifically, using APT, with the associated digital images, resulted in:

- More real and total subgroups ($p < .02$).
- More phantom series ($p < .02$).
- Less phantom subseries ($p < .04$).

These findings indicate that the arrangement schemes created under the APT condition do a better job of highlighting the narrative elements of the arrangement (typically encoded at the subgroup level). That is, the arrangements more often capture the level (subgroup) that links the material to the life and activities of the creator. Conversely, the arrangements created under the APT condition place less emphasis on the level (series) that references the material aspects of the collection (the genre and format of the documents in the collection). One reason for these findings may be that the virtual containers provided by the APT system led participants to pay more attention to the contextual layers that exist at the intellectual level (i.e., linking

DOI: 10.1002/asi

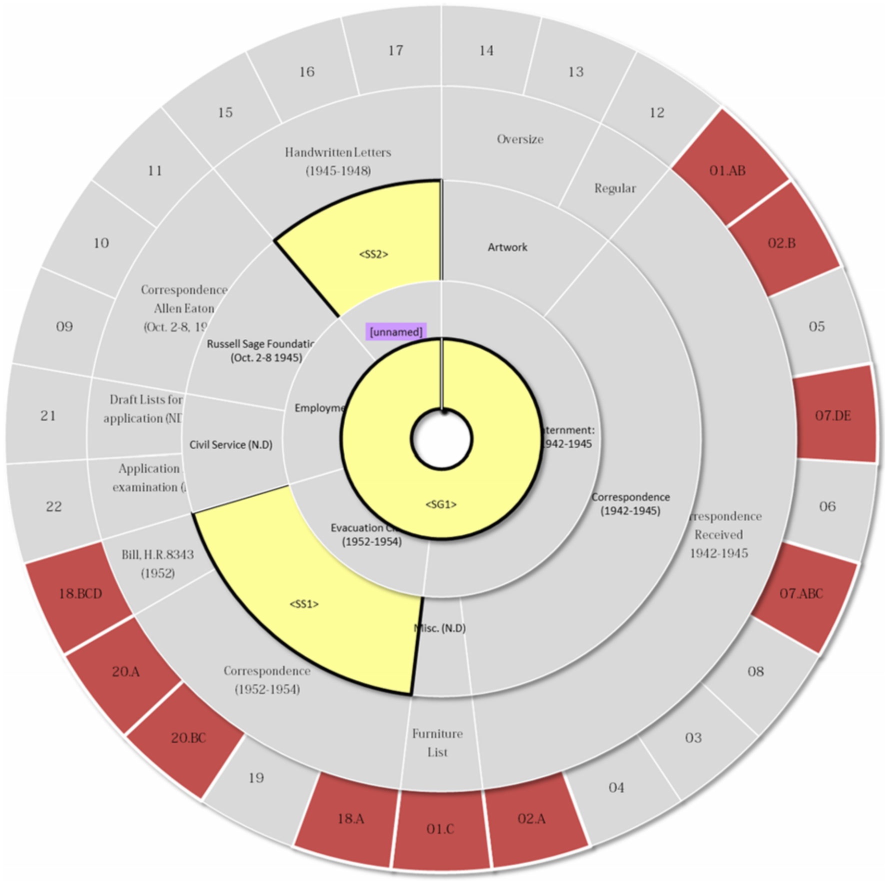


FIG. 11. Circular visualization. Facilitates visual comparisons in any direction. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

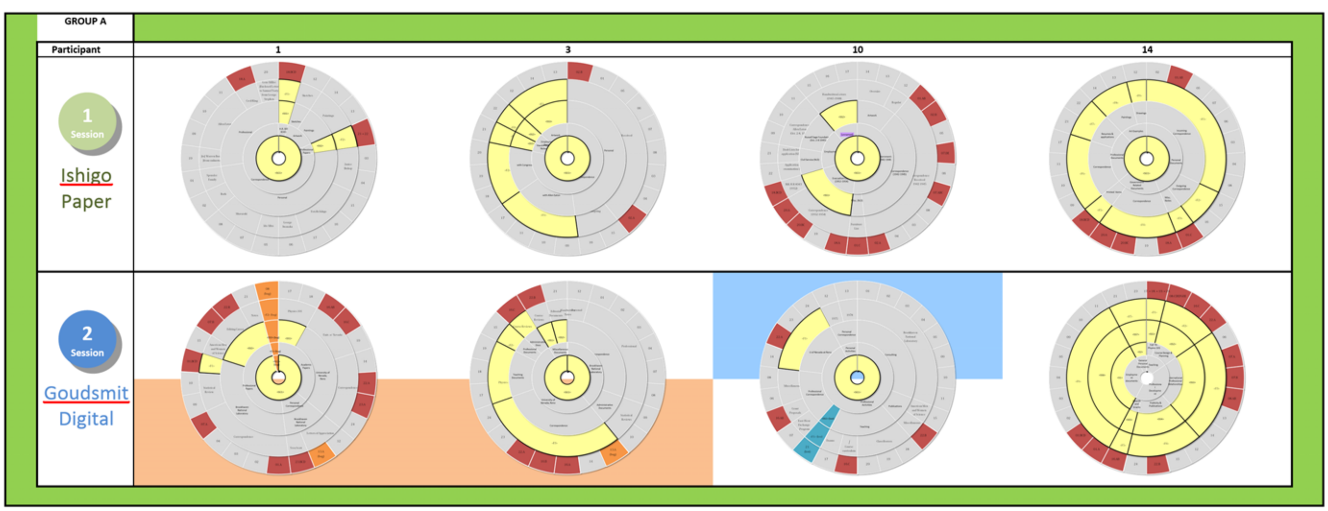


FIG. 12. Topological variations for group A. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

documents to external contexts such as creators and associated functions and activities). The lack of materiality of the processing experience with APT (i.e., the inability to touch and hold the documents in the manner that paper collections make possible) may also suggest why participants created fewer series (series are where the actual recordkeeping system is manifested in the arrangement hierarchy).

DOI: 10.1002/asi

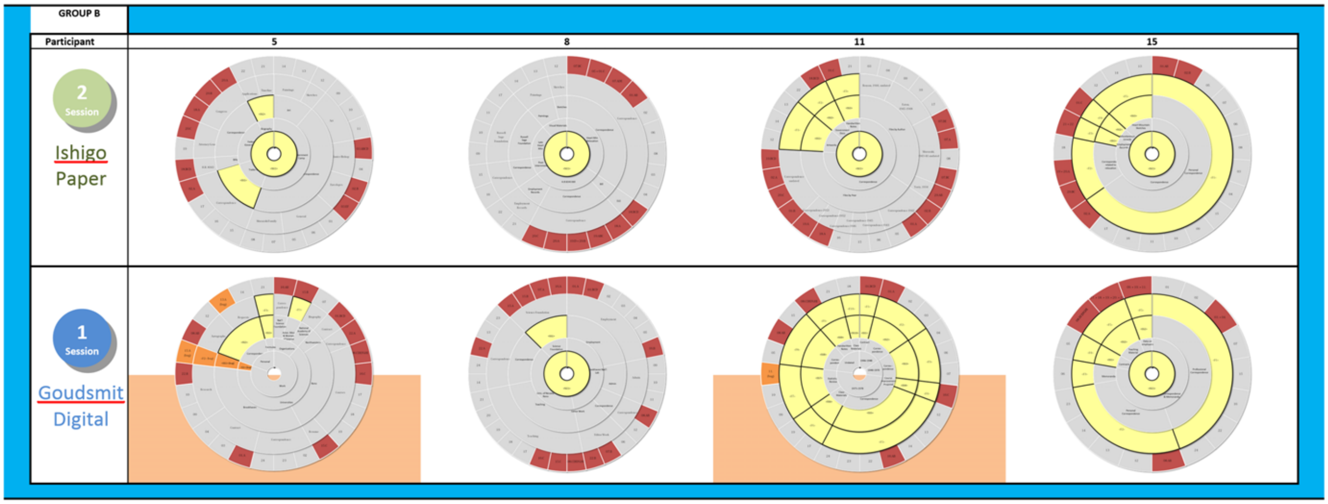


FIG. 13. Topological variations for group B. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

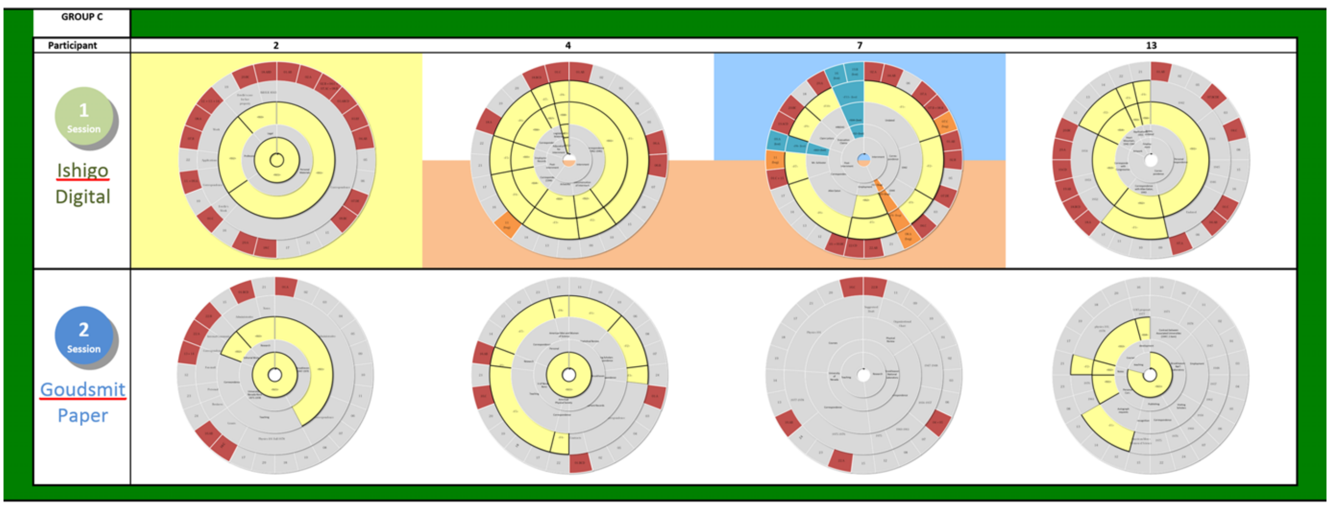


FIG. 14. Topological variations for group C. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

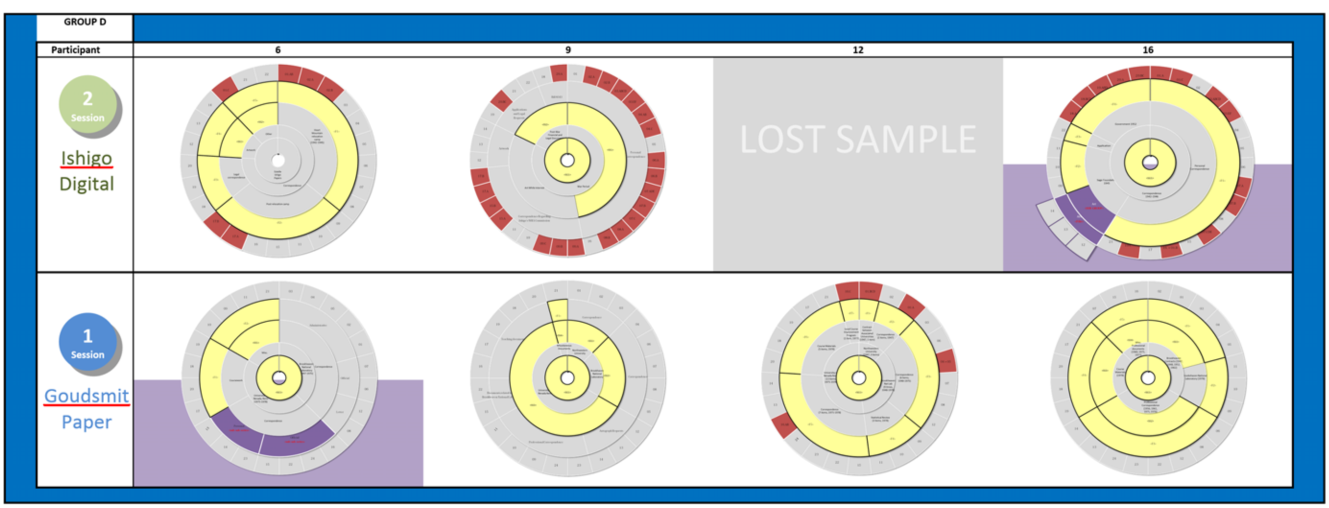


FIG. 15. Topological variations for group D. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

DOI: 10.1002/asi

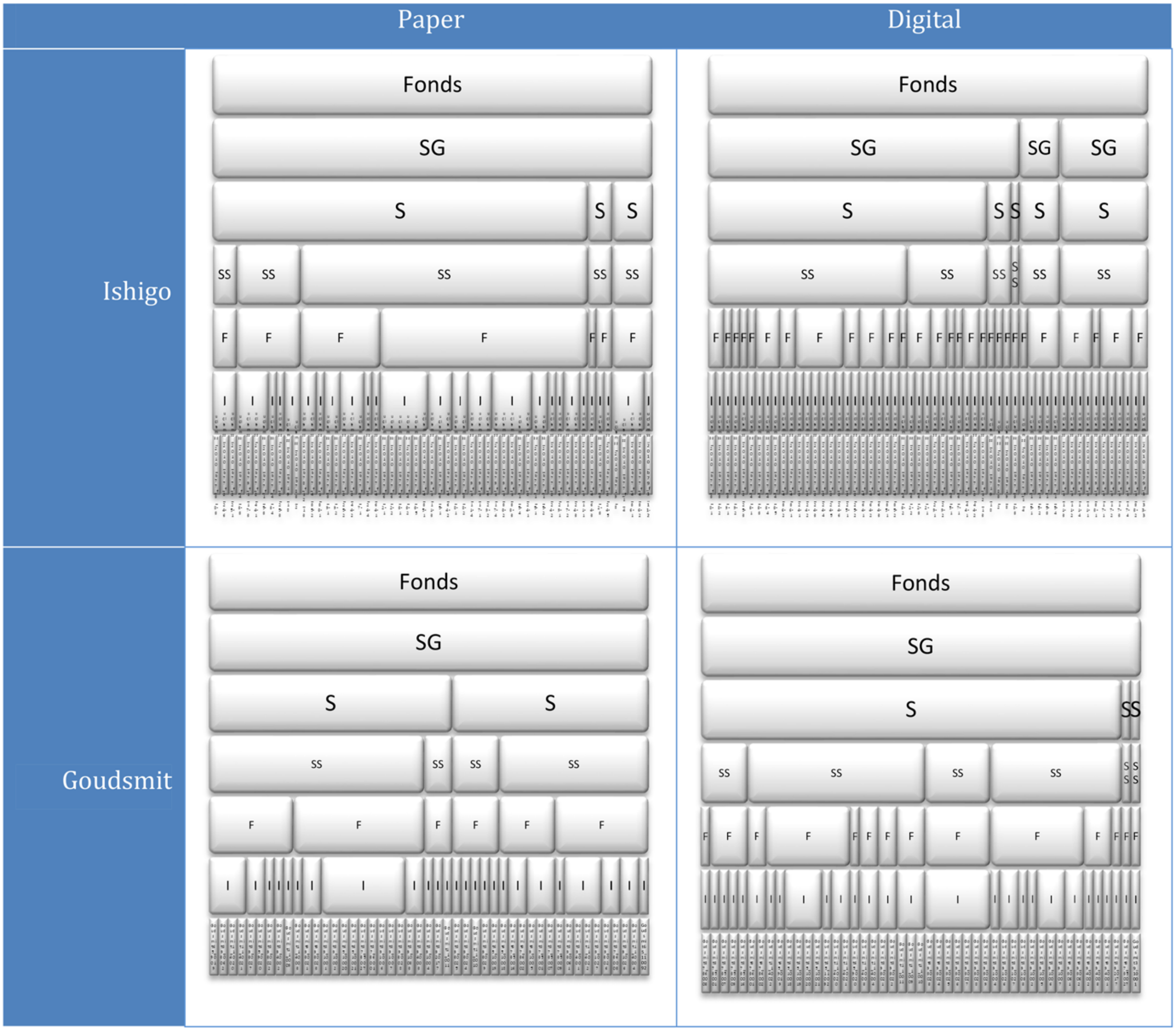


FIG. 16. Average arrangement topologies. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

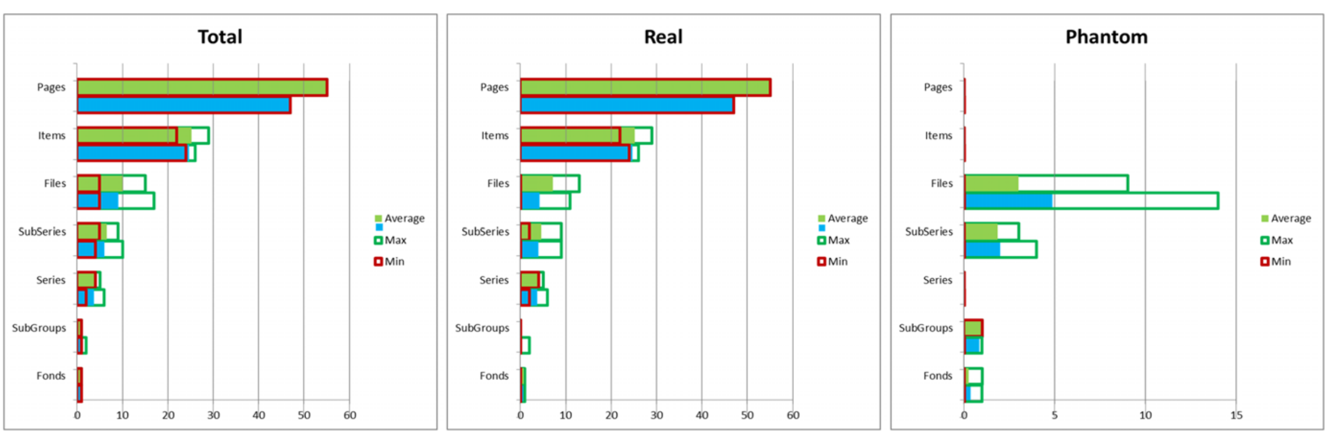


FIG. 17. Total number and per level in paper arrangements. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

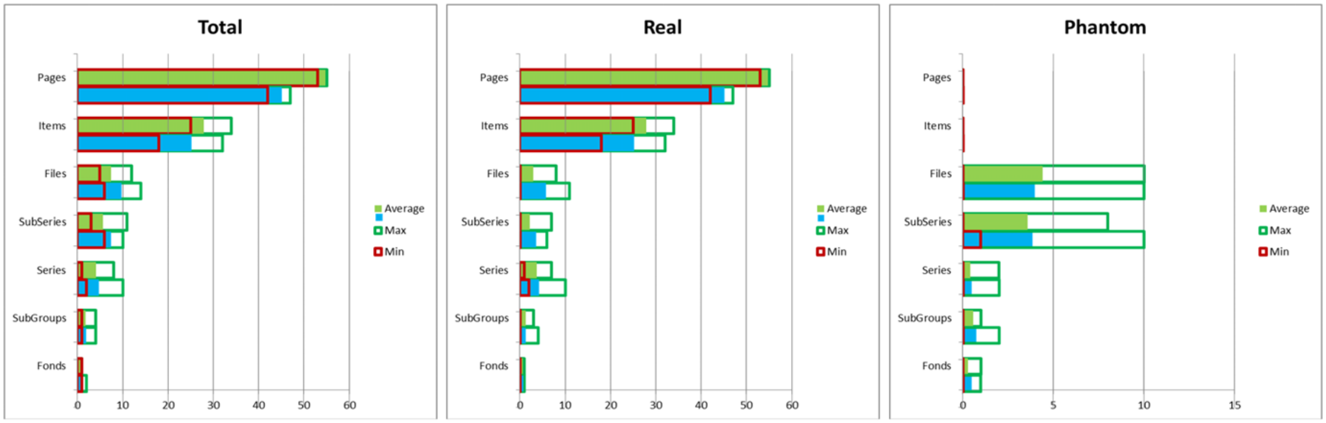


FIG. 18. Total number and per level in digital arrangements. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

B. Effects of processing methods in the propensity for errors. As noted earlier, APT aims to support the creation of arrangements even when some or all traces of the collection's original order have been lost. When the creator's original filing system for the documents has been disturbed, this raises the challenge of page matching, where archivists have to review every page (paper or image) and decide if they are part of the same item. The user study indicated that documents are matched using two main methods:

- Using information intrinsic to the document (e.g., date, name, content, phrase/sentence fragment).
- Using information extrinsic to the document (e.g., size, color, script).

The level of human intervention in the process of matching introduces the possibility that errors occur during processing (it is important to note, however, that the current archival literature does not address this possibility). Specifically, the analysis focused on two error types: splits and merges. Splits occur when a person fails to put two related pages together that belong to the same item. Merges occur when two or more unrelated pages are grouped in the same item. Figure 19 shows the average error counts for each type of error and the overall affected items (an affected item might have two splits and a merge).

At the aggregate level, the analysis revealed significant differences between the two collections. Figure 20 shows the total error counts (splits and merges).

Looking at the first row in Figure 20, the average of matching errors was higher for Ishigo (5.73) than Goudsmit (3.63). Although this difference was not statistically significant ($p < .078$) for the overall sample, it was significant for the paper condition ($p < .03$). However, there was no significant difference in the case of the digitized images ($p < .28$). Since the Ishigo collection included more types of documents, it was initially hypothesized that the variation in materials would make it easier to sort and match the pages, particularly for paper. A possible factor might be that the Ishigo collection had more items with mixed materiality such as items with enclosures (e.g., an item contained a typewritten letter, an envelope, and a handwritten note). However, the fact that this significant difference was observed only in the paper condition raises the question of whether the physical instantiation of the materials was distracting or biasing.

Consequently, the analysis looked for possible correlations between items and errors. Figure 21 charts the number of errors for every item in each collection (I_1-I_{22} for Ishigo, G_1-G_{24} for Goudsmit). The left and right bars represent the number of errors in paper and APT (digital), respectively. In general, the larger error bars on the right indicate that page matching is harder with digitized images than with paper. One reason for this is the greater work involved—what is a two-sided document on paper becomes two separate images when digitized. Another reason to consider is that digitized images do not provide the same access to the extrinsic features of documents. The physical and tactile affordances that come into play when people are matching paper documents are not present with digitized images. We found that there is a propensity when working with digital images to split things that look different (e.g., envelopes from letters). The results also indicate that, overall, page matching was harder to do with the Ishigo Collection, possibly due to its higher diversity of genre formats. The splits charts indicate that some items are just hard to match. These included:

- Items with pages with specialized content (such as equations and formulas).
- Items with pages that were not all uniform in genre, size, and appearance.
- Items with pages where the internal structure and the makeup of the document differed from the expected norm.

There were fewer merge errors in comparison to split errors. The merges charts show that paper and digitized images present different challenges for page matching, although merge errors are more common in the digital

Averages		Paper	Digital
Ishigo	Splits	1.272727	2.045455
	Merges	0.454545	0.727273
	Affected Items	4.25	7.428571
Goudsmit	Splits	1.125	3.625
	Merges	0.75	1.75
	Affected Items	1.875	5.375

Splits	Paper	Digital
Ishigo	1.272727	2.045455
Goudsmit	1.125	3.625

Merges	Paper	Digital
Ishigo	0.454545	0.727273
Goudsmit	0.75	1.75

Affected Items	Paper	Digital
Ishigo	4.25	7.428571
Goudsmit	1.875	5.375

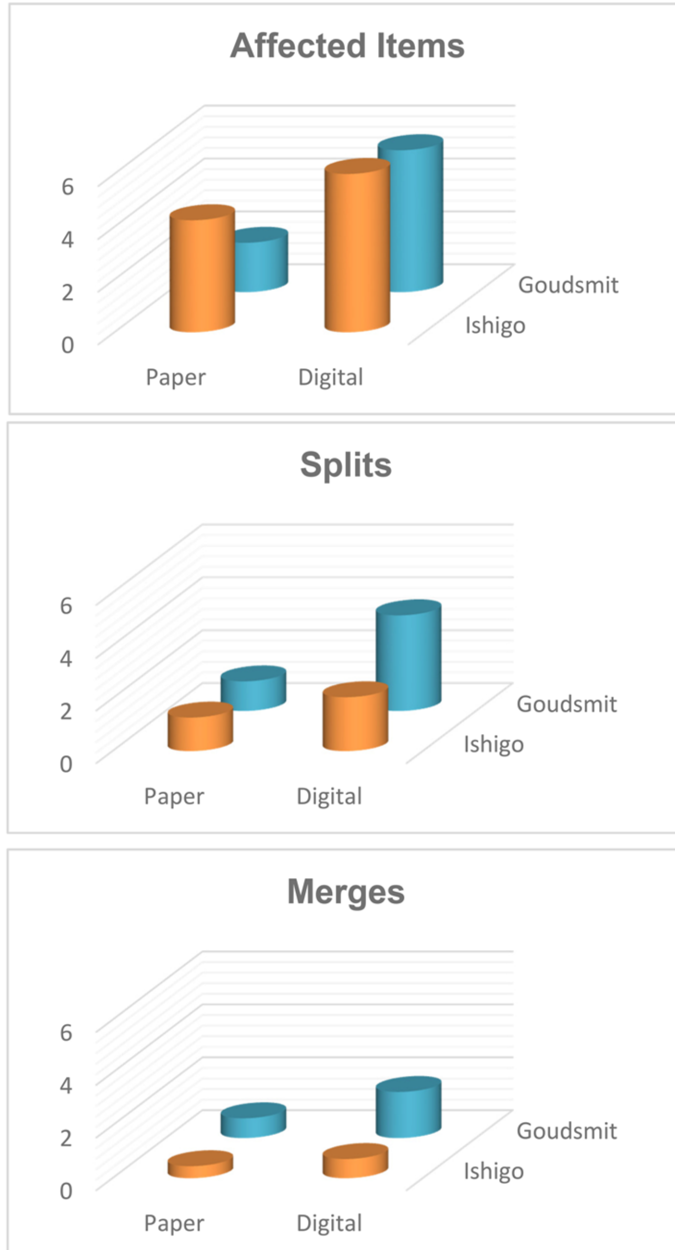


FIG. 19. Error counts. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

environment. Unlike the split errors, the number of merge errors was too small to identify clear differences between collection and media types.

C. Overall quality. In order to gauge the effectiveness (RQ2.1) and efficiency (RQ2.2) of APT, the study compared the overall quality of the final arrangements across all the experimental conditions based on five metrics: narrativity, recordkeeping, materiality, perspectives, and coherence. Figure 22 shows the overall quality results and Table 2 explains what it is measured by each metric, and why each metric is important.

	Ishigo	Goudsmit	p-value
All	Average = 5.73 St. Dev. = 3.20	Average = 3.63 St. Dev. = 3.22	0.08
Paper	Average = 4.25 St. Dev. = 1.83	Average = 1.875 St. Dev. = 2.17	0.03
Digital	Average = 7.43 St. Dev. = 3.69	Average = 5.375 St. Dev. = 3.25	0.28

FIG. 20. Average error counts.

DOI: 10.1002/asi

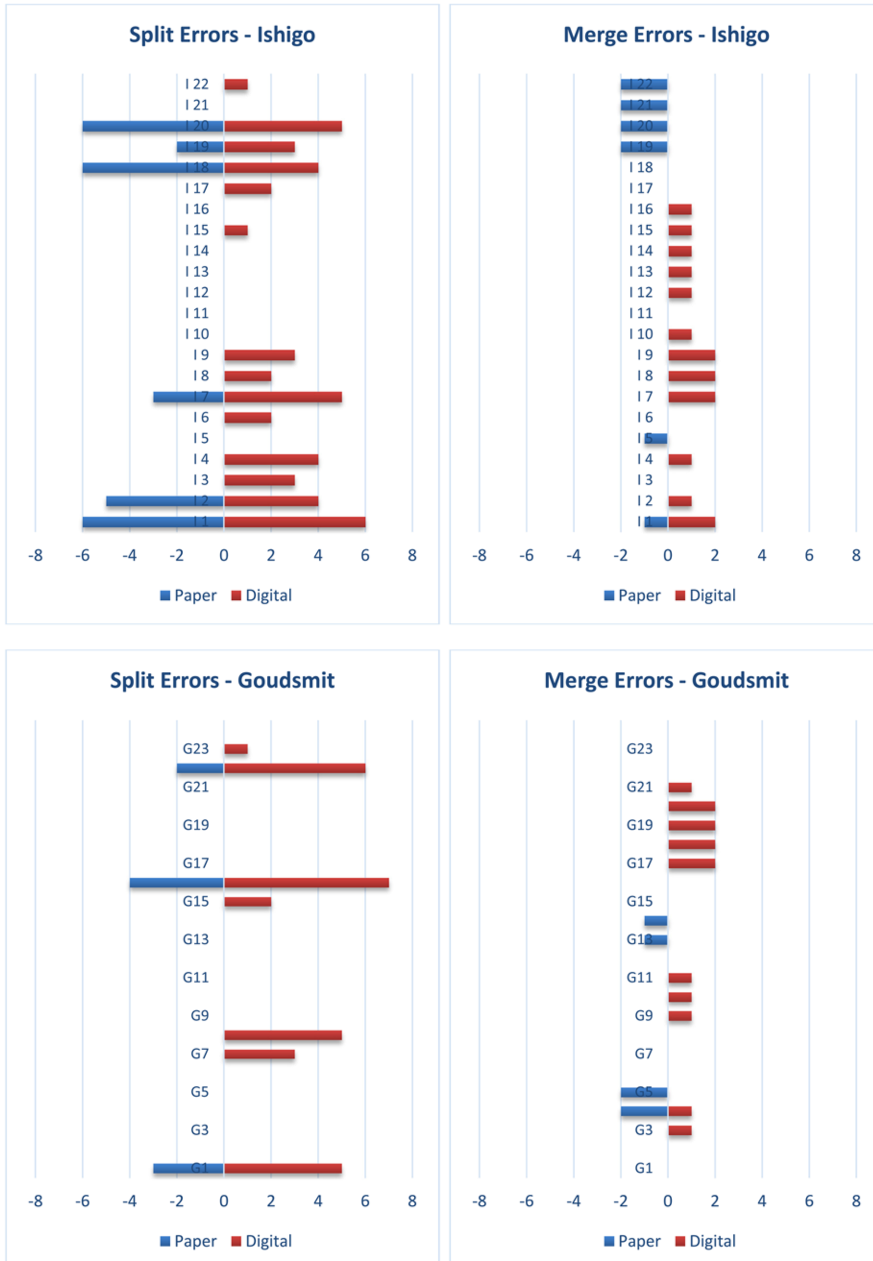


FIG. 21. Errors by items. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The rank 1 is given to the best arrangement. The Overall Ranking shows the average ranking of a document across the five metrics. The second session of participant 12 was lost due to a system crash. It is shown in Figure 22 only for completeness, but it has been removed from the data analysis.

In order to study the contribution of each metric to the overall quality assessment of the arrangements, the study analyzed the correlations between different metrics (see Figures 23–25).

As Figures 23–25 show, Narrativity and Recordkeeping were highly correlated with the overall ranking (0.818 and 0.838, respectively). That means that arrangements that had a good narrative or recordkeeping score also had a good score across the other metrics. Furthermore, these two metrics were also correlated with each other (0.637 for Goudsmit, 0.873 for Ishigo, and 0.728 overall). Perspectives and Coherence were also correlated with overall ranking, but to a lesser degree (0.661 and 0.526, respectively). Individually, Perspectives was correlated with Recordkeeping (0.539). Interestingly, Coherence was not correlated with any other metric. Similarly, Materiality was not correlated with anything. Given that Narrativity and Recordkeeping are typically expressed in the subgroup and series levels,

Participant ID	Session	Medium	Collection	Group	Narrative	Perspectives	Recordkeeping	Materiality	Coherence	Overall Ranking
1	1	Paper	Ishigo	1	7	4	6	4	6	5.400
3	1	Paper	Ishigo	1	11	6	10	7	4	7.600
10	1	Paper	Ishigo	1	2	1	1	1	3	1.600
14	1	Paper	Ishigo	1	12	7	9	2	3	6.600
5	2	Paper	Ishigo	2	5	6	3	4	6	4.800
8	2	Paper	Ishigo	2	6	2	8	7	5	5.600
11	2	Paper	Ishigo	2	8	4	6	5	1	4.800
15	2	Paper	Ishigo	2	9	4	10	8	5	7.200
2	1	Digital	Ishigo	3	10	5	11	3	4	6.600
4	1	Digital	Ishigo	3	1	2	3	6	3	3.000
7	1	Digital	Ishigo	3	[illegible]	3	4	7	5	4.600
13	1	Digital	Ishigo	3	8	2	7	6	2	5.000
6	2	Digital	Ishigo	4	6	2	8	10	3	5.800
9	2	Digital	Ishigo	4	3	6	5	6	6	5.200
12	2	Digital	Ishigo	4						
16	2	Digital	Ishigo	4	11	6	12	9	7	9.000
2	2	Paper	Goudsmit	3	8	3	6	2	6	5.000
4	2	Paper	Goudsmit	3	9	2	4	8	8	6.200
7	2	Paper	Goudsmit	3	1	1	1	3	1	1.400
13	2	Paper	Goudsmit	3	5	5	7	1	7	5.000
6	1	Paper	Goudsmit	4	3	4	5	10	9	6.200
9	1	Paper	Goudsmit	4	4	6	5	11	5	6.200
12	1	Paper	Goudsmit	4	4	3	2	7	3	3.800
16	1	Paper	Goudsmit	4	13	5	10	4	9	8.200
1	2	Digital	Goudsmit	1	11	4	13	9	8	9.000
3	2	Digital	Goudsmit	1	2	4	4	5	4	3.800
10	2	Digital	Goudsmit	1	10	4	4	3	9	6.000
14	2	Digital	Goudsmit	1	12	7	11	6	10	9.200
5	1	Digital	Goudsmit	2	9	2	8	3	2	4.800
8	1	Digital	Goudsmit	2	7	3	3	10	3	5.200
11	1	Digital	Goudsmit	2	6	7	12	6	1	6.400
15	1	Digital	Goudsmit	2	14	8	9	7	3	8.200

FIG. 22. Overall quality results. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE 2. Overall quality metrics.

	What is Measured	Why it is Measured
Narrativity	Degree to which the arrangement accurately reflects the life of the creator	Facilitates understanding the creator and his/hers/its activities
Recordkeeping	Degree to which the arrangement reflects the recordkeeping practices of the creator	The way that records are kept relates the role of the documents to the life of the creator
Perspective	Number of perspectives about the materials' meta-level context as explicitly revealed in the arrangement	Facilitates understanding the life of the creator from various contextual perspectives, such as: time, places, events, institution, functions and activities
Coherence	Number of ways that an arrangement establishes (or fails to establish) an overall coherence, and follows the archival rules	Enforces archival principles, facilitates navigation, and makes it easier to understand all other metrics
Materiality	Number of material types in the collection that the arrangement accurately expresses (e.g., handwritten notes, contracts, etc.)	Helps users understand the types of materials found in the collection

respectively, these results imply that the highest levels of the arrangement are the most relevant in terms of overall quality.

Correlations

In addition to analyzing the correlations within individual metrics, the study analyzed the correlations between all the different measurements. The initial and exit surveys provided insights about the participants' previous experience in processing collections, their comfort with multi-touch technology, and their perceptions after the study. Tables 3 and 4 show some of the questions asked to the participants. Figure 26 shows the correlations for survey questions and errors.

	Narrativity	Perspectives	Recordkeeping	Materiality	Coherence	Overall Ranking
Narrativity	1.000					
Perspectives	0.406	1.000				
Recordkeeping	0.637	0.626	1.000			
Materiality	-0.046	0.187	-0.003	1.000		
Coherence	0.405	0.151	0.246	0.052	1.000	
Overall Ranking	0.800	0.673	0.795	0.333	0.591	1.000

FIG. 23. Overall quality correlations for Goudsmit Collection. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

	Narrativity	Perspectives	Recordkeeping	Materiality	Coherence	Overall Ranking
Narrativity	1.000					
Perspectives	0.585	1.000				
Recordkeeping	0.873	0.445	1.000			
Materiality	0.085	-0.103	0.403	1.000		
Coherence	0.005	0.415	0.161	0.233	1.000	
Overall Ranking	0.842	0.644	0.928	0.479	0.401	1.000

FIG. 24. Overall quality correlations for Ishigo Collection. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

	Narrativity	Perspectives	Recordkeeping	Materiality	Coherence	Overall Ranking
Narrativity	1.000					
Perspectives	0.486	1.000				
Recordkeeping	0.728	0.539	1.000			
Materiality	0.009	0.067	0.157	1.000		
Coherence	0.286	0.238	0.192	0.112	1.000	
Overall Ranking	0.818	0.661	0.839	0.394	0.526	1.000

FIG. 25. Overall quality correlations for all arrangements. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The results show that the actual processing time was not correlated with the participants' perception of time ($PT_{Digital}$ or PT_{Paper}), or the participants' perception that the media facilitated the understanding of archival processing ($PL_{Digital}$ or PL_{Paper}). However, $PT_{Digital}$ and $PL_{Digital}$ were directly

TABLE 3. Initial survey questions.

Question ID	Question
Exp_{Arch}	How many archival collections have you processed (arranged and described) in the past? ○Zero ○One ○Two ○Three ○More ○Decline to answer
Exp_{Tech}	Please select all the surface computing devices you have used in the past? ○Multi-touch Surface ○Multi-touch Wall ○Other ○Decline to answer
$Comfort_{Tech}$	How comfortable are you in using multi-touch technologies like smart phones (e.g., iPhone or Android) or Tablets? ○Very Uncomfortable ○Uncomfortable ○Neutral ○Comfortable ○Very Comfortable ○Decline to answer

correlated with APT Use. At the same time, PT_{Paper} was inversely correlated with APT Use. The choice to use APT in the future was also directly correlated with the participants' technological background (Exp_{Tech}), and Exp_{Tech} was also indirectly correlated with PT_{Paper} and PL_{Paper}. This implies that the participants' perceptions and their technological background play a strong role in their choice of whether to use APT in the future.

There were differences between paper and digitized images. For example, $PT_{Digital}$ was correlated with $PL_{Digital}$ but PT_{Paper} was not correlated with PL_{Paper}. This observation leads us to think that the digitized condition facilitated the understanding of the arrangement process more than did the paper condition.

It is worth noting that the results did not show any correlation between the number of errors (mismatched items) and the Processing Time, or with the participants' perceptions of time ($PT_{Digital}$, $PT_{Digital}$). Furthermore, as Figure 27 shows, the errors were not correlated with any other metric, including the overall quality metrics. This goes against the hypothesis that higher overall quality, or less errors, could be explained by longer processing times.

Discussion and Conclusion

In "identifying, preserving, and communicating the relationships between records and events, archivists select, shape, and situate records such that they can be regarded and used as documentary sources that are capable of serving as evidence of past events" (Meehan, 2006, p. 143). As archivists aim to enhance their support of the evidentiary work of humanities scholars, they need methods that can continue to capture the important network of relationships between creators, activities, records, and recordkeeping systems. The best practice for creating, capturing, and conveying this information is by purposefully arranging collections. This paper improves the understanding of the value and the inherent characteristics of arrangement schemes and the arrangement process. It also reveals important insights concerning the current practices for archival arrangement and the

DOI: 10.1002/asi

TABLE 4. Exit survey questions.

Question ID	Question
$PT_{Digital}$	Processing collections using APT takes less time as compared to processing paper collections. ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer
$PL_{Digital}$	I learned more about processing collections while working on APT as compared to processing a paper collection. ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer
Practice	APT can be used to practice processing collections. ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer
$PT_{Digital}$	Processing paper collections is faster than processing collections on APT. ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer
PL_{Paper}	Processing paper collections helped me better understand archival processing as compared to processing collections using APT. ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer
APT Use	Would you like to use APT in future? ○Strongly Disagree ○Disagree ○Neutral ○Agree ○Strongly Agree ○Decline to answer

	$PT_{Digital}$	$PL_{Digital}$	Practice	PT_{Paper}	PL_{Paper}	APT Use	Exp_{Arch}	Exp_{Tech}	$Comfort_{Tech}$	Mismatched Items	Processing Time
$PT_{Digital}$	1.00										
$PL_{Digital}$	0.69	1.00									
Practice	0.50	0.65	1.00								
PT_{Paper}	-0.64	-0.62	-0.34	1.00							
PL_{Paper}	-0.21	-0.48	-0.42	0.39	1.00						
APT Use	0.50	0.65	0.49	-0.73	-0.42	1.00					
Exp_{Arch}	-0.57	-0.31	-0.22	0.44	-0.15	-0.30	1.00				
Exp_{Tech}	0.11	0.39	0.21	-0.63	-0.50	0.55	0.18	1.00			
$Comfort_{Tech}$	-0.12	0.09	0.15	0.00	0.00	0.30	-0.36	-0.27	1.00		
Mismatched Items	0.20	0.26	-0.18	-0.28	0.08	-0.21	-0.07	0.14	-0.36	1.00	
Processing Time	0.09	0.12	0.24	-0.09	-0.08	0.22	-0.04	0.20	0.10	-0.28	1.00

FIG. 26. Correlations for survey questions and errors. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

quality of its outcomes. Furthermore, this paper establishes the need to look to both technology and human factors in order to improve this aspect of the archival curation workflow.

The study clearly established that matching errors can occur when processing collections, resulting in documents being placed in the wrong context (file, series, subgroup, etc.). From a researcher's perspective, this scenario could lead to a document being misunderstood or misinterpreted. Matching errors may also mean that documents, or parts of documents, may not be discovered at all (due to splitting mistakes), and therefore fail to become a part of the researcher's accumulated body of evidence. On the positive side, this work exposes clear opportunities for dealing with this challenge using sociotechnical approaches. For example, future prototypes of the APT system could automatically identify and highlight relevant metadata in the image (e.g., names, dates, sentence fragments), thus improving the process of document matching. At the same time, archivists could put new usage protocols in place to allow scholars to provide expert feedback and to mark potential matching mistakes.

Furthermore, the results show that more work is needed to guarantee the replicability of the outcomes since, whether

DOI: 10.1002/asi

	Mismatched Items	Processing Time	Narrativity	Perspectives	Recordkeeping	Materiality	Coherence	Overall Ranking
$PT_{Digital}$	0.20	0.09	-0.05	0.07	-0.18	-0.17	-0.01	-0.12
$PL_{Digital}$	0.26	0.12	0.08	0.31	0.09	0.08	0.03	0.16
Practice	-0.18	0.24	-0.03	0.01	0.03	-0.05	-0.11	-0.05
PT_{Paper}	-0.28	-0.09	-0.25	-0.48	-0.31	0.17	-0.12	-0.29
PL_{Paper}	0.08	-0.08	0.09	-0.18	-0.15	-0.21	0.05	-0.10
APT Use	-0.21	0.22	0.06	0.38	0.16	0.09	0.15	0.22
Exp_{Arch}	-0.07	-0.04	0.12	-0.07	0.21	0.32	0.12	0.23
Exp_{Tech}	0.14	0.20	0.16	0.37	0.32	-0.10	0.14	0.26
$Comfort_{Tech}$	-0.36	0.10	0.09	0.20	0.08	-0.01	0.03	0.10
Mismtached Items	1.00	-0.28	0.17	0.20	0.07	0.03	-0.16	0.10
Processing Time	-0.28	1.00	-0.10	-0.12	0.05	-0.03	0.16	-0.01

FIG. 27. Correlations for time, errors, and survey questions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

using APT or paper, the final arrangements presented significant variations. Moreover, when looking at groups of arrangements, individual variations do not average out. The average topologies show that the alternative approaches (APT and paper) result in different variations, both in terms of topology and propensity for matching. Above all, the study found that significant topological variations occur at the levels of the arrangements that encode information about the provenance of the records and the filing system that reflects how records were created, received, or used in the same activity (see section Topological characteristics of arrangements, above). These findings demonstrate that archivists can and do privilege certain levels/sets within the arrangement scheme. A likely outcome is that researchers will interact with collections in which certain networks of relationships have been emphasized or deemphasized. Interestingly, the variations are not correlated with archival experience (the amount of experience a participant had in processing collections), or with the quality of the topology. Instead, the variations seem related to factors such as: who created the arrangement, the system used (e.g., using digitized images changes the propensity for human errors), and the contents (some items are just hard to match).

Overall, the study suggests that the way that archival processing work gets done does not always align with codified principles and practices. Given the degree of flexibility shown in the arrangement process, the question remains as to why archivists may deviate from the arrangement guidelines already in place, and what, if anything, to do about it. From an archival standpoint, the relational and proximate contexts surrounding an information object are said to be just as important as the content inscribed on that object (Horsman, 2011). This contextual information typically relates to “the entities involved in the creation of the records—individuals, organizations and functions—and the nature of the records themselves” (Tibbo, Lee, Marchionini, & Howard, 2006, p. 213). Overall, the study suggests that such contexts are generally downplayed in the arrangement process. Also, while the archival principle of “original order” holds that it is the archivist’s job to arrange the actual records themselves according to the original filing system of the creator, the results indicate that archivists may not always follow this dictate. Instead, the preference seems to be to impose an order on the collection, or parts of the collection, based on assumptions about how researchers would like to engage with the material. The arrangement that the participants seemed to think the user would favor was one based mainly on genre/format. There is no doubt that materiality, or what Lee (2011) calls “form of expression,” provides an important “gateway” to contextual information (Foscarini et al., 2010). Duff and Johnson (2002, p. 485), for example, have found that certain historians seek information from particular kinds of sources or genres (an example is a social historian’s use of psychiatric case files). Yet the merit of a model of context that privileges materiality over narrativity has yet to be ascertained through any kind of formal user study.

The findings raise the larger question as to whether, or to what degree, the contexts represented in the archival arrangement process continue to be useful. From a theoretical perspective it can be argued (the work of postmodern archival theorists such as Brothman, 2002, Nesmith, 2005,

DOI: 10.1002/asi

and Beattie, 2009, notwithstanding), that archivists have traditionally adopted a fairly positivist approach to context. A positivist view of context is one in which context is seen as a representational problem, and where context itself is seen as fixed, rational, and unchanging (see Dourish, 2004). Adopting a phenomenological and critical approach to context (see Dourish, 2004), for example, would allow archivists to expand their understanding of context—including the notion that context is a dynamic, occasioned, and emergent property, and that context itself can be studied as a conditioned or privileged structure. The notion that context is in principle "infinite" (Duff & Harris, 2002, p. 276) has long been accepted in other fields. Digital libraries research, for example, has adopted a more expansive view of context. In one such model, context types are looked at from a total lifecycle perspective that captures the context of the information and its constituent components; the community processes of which the information is a part; as well as the context in which the information is created, interpreted, administered, archived, and used (Neuhold et al., 2005). This model embraces not only the context of the creator, but that of the information professional, and the subsequent audience.

A logical follow-up study based on the findings would be to look at whether the current archival topology (along with the associated sets/levels) or an expanded set of contextual layers provides the best possible support for users within diverse research communities (journalists, scientists, etc.). Such work may run in parallel with research into people's preferences when accessing archival material through digital finding aids (see, for example, Fachry, Kamps, & Zhang, 2008). User studies focused on individual disciplines or areas would help determine the particular network of relationships that need to be captured in any arrangement scheme, and the way that the network of relationships should be represented. It may be that the traditional hierarchical arrangement schemes with levels representing the creator, activities, and associated recordkeeping system stand up to further scrutiny. However, if communities demonstrate unique research needs, the workflow and the outcome of the arrangement process will need to change to reflect this.

Along with these theoretical issues, the study provided a thorough comparison between paper-based arrangement and APT, which leads to a number of practical design implications. First, while the arrangement schemes created under the APT condition do a better job of highlighting the narrative elements of the arrangement (typically encoded in the subgroups) it is necessary to:

- speed-up collection arrangement
- improve overall quality
- enhance outcome replicability

The next iteration of APT (APT 3.0) will build in additional functionality for processing at the file level, thus allowing for greater scalability in terms of the size of collections that can be processed using the APT software. Based on findings from the current user study, the design of the APT system will also continue to evolve, with an eye to providing additional support for processes that participants found challenging in the current prototype.

With APT 2.0, a large part of the task of arranging collections was found to involve space management. Scrolling was found to be too time consuming and it led to people losing track of where they were in the workspace. Design solutions for APT 3.0 may include nested layers of space (see Kirk et al., 2010) which would facilitate people's ability to be selective about what they pay attention to or what they interact with at any particular time. APT 3.0 also needs to provide advanced functionality to help users with pattern recognition and document matching. Some participants adopted the strategy of zooming out the workspace and manipulating the documents at a level where the text was not legible. APT 3.0 design will need to further facilitate document matching, by automatically analyzing and visualizing more document metadata, such as dates, names/entities, and physical characteristics—page size, page color, stains on the documents, etc. Any resolution of this problem will undoubtedly also involve the digitization process itself, including the use of optical character recognition (OCR) technology to help highlight those intrinsic elements of the document that people utilize in the matching process.

In terms of improving overall quality, additional functionality will need to be provided in order to facilitate and encourage better workflows (e.g., conducting a Collection Review—CR phase—at the beginning of the task). It is possible to envision automatic evaluation algorithms that provide an estimation for different quality metrics. Finally, any future design would need to take more account of the multiple affordances of objects. The classic example in the APT study is that of the digital paper clip. Participants used this clip to demarcate temporary as well as permanent relationship between items, and to move items in bulk around the table. Each of these possibilities for action will need to be explicitly designed into the new system.

References

Aguirre, C., Doyle, K., Hernández-Salazar, D., Guatemala, & University of Oregon. (2013). From silence to memory: Revelations of the Archivo Histórico de la Policía Nacional. Eugene, OR: University of Oregon Libraries.

Audenaert, N., & Furuta, R. (2010). What humanists want: How scholars use source materials. In Proceedings of the 10th Annual Joint Conference on Digital libraries (JCDL '10) (pp. 283–292). New York: ACM.

Balakrishnan, A.D, Fussell, S.R., Kiesler, S., & Kittur, A. (2010). Pitfalls of information access with visualizations in remote collaborative analysis. In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10) (pp. 411–420). New York: ACM.

Barreau, D.K. (1995). Context as a factor in personal information management systems. Journal of the American Society for Information Science, 46(5), 327–339.

Beattie, H. (2009). Where narratives meet: Archival description, provenance, and women's diaries. Libraries & the Cultural Record, 44(1), 82–100.

DOI: 10.1002/asi

Beaudoin, J.E. (2012). Context and its role in the digital preservation of cultural objects. D-Lib Magazine, 18 (11/12).

Brothman, B. (2002). Afterglow: Conceptions of record and evidence in archival discourse. Archival Science, 2(3-4), 311–342.

Buckland, M.K. (1991). Information as thing. Journal of the American Society for Information Science, 42(5), 351–360.

Clifton, P., Mazalek, A., Sanford, J., Rébola, C., Lee, S., & Powell, N. (2011). SketchTop: Design collaboration on a multi-touch tabletop. In Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '11) (pp. 333–336). New York: ACM.

Cohen, A.M., Adams, C.E., Davis, J.M., Yu, C., Yu, P.S., Meng, W., . . . (2010). Evidence-based medicine, the essential role of systematic reviews, and the need for automated text mining tools. In T. Veinot (Ed.), Proceedings of the 1st ACM International Health Informatics Symposium (IHI '10) (pp. 376–380). New York: ACM.

Cook, T., & Schwartz, J.M. (2002). Archives, records and power: From (postmodern) theory to (archival) performance. Archival Science, 2(3–4), 171–185.

Crow, J., Francisco-Revilla, L., Norris, A., Shukla, A., & Trace, C.B. (2012). A unique arrangement: Organizing collections for digital archives and libraries. In Theory and Practice of Digital Libraries (pp. 335–344). Berlin, Germany: Springer.

Dalbello, M. (2004). Institutional shaping of cultural memory: digital library as environment for textual transmission. Library Quarterly, 74(3), 265–298.

Dooley, J.M., Luce, K., & OCLC Research. (2010). Taking our pulse: The OCLC research survey of special collections and archives. Dublin, OH: OCLC Research.

Duff, W., & Harris, V. (2002). Stories and names: Archival description as narrating records and constructing meanings. Archival Science, 2(3–4), 263–285.

Duff, W.M., & Johnson, C.A. (2002). Accidentally found on purpose: Information-seeking behavior of historians in archives. The Library Quarterly, 72(4), 472–496.

Duff, W., Monks-Leeson, E., & Galey, A. (2011). Contexts built and found: A pilot study on the process of archival meaning-making. Archival Science, 12 (1), 69–92.

Dourish, P. (2004). What we talk about when we talk about context. Journal of Personal and Ubiquitous Computing, 8(1), 19–30.

Eastwood, T. (2000). Putting the parts of the whole together: Systematic arrangement of archives. Archivaria, 50, 93–116.

Erway, R., Schaffner, J., & OCLC (2007). Shifting gears: Gearing up to get into the flow. Dublin, OH: OCLC Programs and Research.

Fachry, K.N., Kamps, J., & Zhang, J. (2008). Access to archival material in context. In P. Borlund, J.W. Schneider, M. Lalmas, A. Tombros, J. Feather, D. Kelly, & A.P. de Vries (Eds.), Proceedings of the Second International Symposium on Information Interaction in Context (IIiX '08) (pp. 102–109). New York: ACM.

Foscarini, F., Kim, Y., Lee, C.A., Mehler, A., Oliver, G., & Ross, S. (2010). On the notion of genre in digital preservation. In J.P. Chanod, M. Dobreva, A. Rauber, & S. Ross (Eds.), Proceedings of the Dagstuhl Seminar (Vol. 10291).

Francisco-Revilla, L., & Shipman, F.M. (2004). WARP: A web-based dynamic spatial hypertext. In Proceedings of the Fifteenth ACM Conference on Hypertext and Hypermedia (HT'04) (pp. 235–236). New York: ACM.

Furner, J. (2004). Conceptual analysis: A method for understanding information as evidence, and evidence as information. Archival Science, 4(3-4), 233–265.

Gilliland-Swetland, A.J., & Council on Library and Information Resources. (2000). Enduring paradigm, new opportunities: The value of the archival perspective in the digital environment. Washington, DC: Council on Library and Information Resources.

Grønbæk, K., Kristensen, J.F., Ørbæk, P., & Eriksen, M.A. (2003). Physical hypermedia: Organising collections of mixed physical and digital material. In Proceedings of the Fourteenth ACM conference on Hypertext and Hypermedia (pp. 10–19). New York: ACM.

Henderson, S. (2005). Genre, task, topic and time: Facets of personal digital document management. In Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-human Interaction: Making CHI Natural (CHINZ '05) (pp. 75–82). New York: ACM.

Horsman, P. (October, 2011). Wrapping records in narratives: Representing context through archival description. Paper contributed to the "Usability of the Archives of the International Tracing Service Workshop" at the International Tracing Service. Bad Arolsen, Germany.

Johnston, P., & Robinson, B. (2002). Collections and collection description. Collection description focus briefing paper 1. UKOLN.

Kemman, M., Kleppe, M., & Scagliola, S. (2013). Just Google it: Digital research practices of humanities scholars. Studies in the Digital Humanities. arXiv: 1309.2434.

Kirk, D.S., Izadi, S., Sellen, A., Taylor, S., Banks, R., & Hilliges, O. (2010). Opening up the family archive. In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (pp. 261–270). New York: ACM.

Kwasnik, B.H. (1989). The influence of context on classificatory behavior (doctoral dissertation). Newark, NJ: Rutgers State University of New Jersey.

Lee, C. (2011). A framework for contextual information in digital collections. Journal of Documentation, 67(1), 95–143.

Lemieux, V. (2012). Envisioning a sustainable future for archives: A role for visual analytics. Paper contributed to the International Council on Archives Conference, Brisbane, Australia.

Mayer, R., Neumayer, R., & Rauber, A. (2009). Interacting with (semi-) automatically extracted context of digital objects. In Proceedings of the 1st Workshop on Context, Information and Ontologies (CIAO 2009). New York: ACM.

Mayer, R., & Rauber, A. (2009). Establishing context of digital objects' creation, content and usage. In Proceedings of the JCDL Workshop on Innovation in Digital Preservation (InDP 2009).

Meehan, J. (2006). Towards an archival concept of evidence. Archivaria, 61, 127–146.

Morris, M.R., Lombardo, J., & Wigdor, D. (2010). WeSearch: Supporting collaborative search and sensemaking on a tabletop display. In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10) (pp. 401–410). New York: ACM.

Nebe, K., Klompmaker, F., Jung, H., & Fischer, H. (2011). Exploiting new interaction techniques for disaster control management using multitouch-, tangible- and pen-based-interaction. In Human-Computer Interaction. Interaction Techniques and Environments (pp. 100–109). Berlin, Heidelberg: Springer.

Nesmith, T. (2005). Reopening archives: Bringing new contextualities into archival theory and practice. Archivaria, 60, 259–274.

Neuhold, E., Niederée, C., Stewart, A., Frommholz, I., & Mehta, B. (2005). The role of context for information mediation in digital libraries. In Digital Libraries: International Collaboration and Cross-Fertilization (pp. 133–143). Berlin, Heidelberg: Springer.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. Trends in Cognitive Sciences, 11(12), 520–527.

O'Neill, J., Privault, C., Renders, J.-M., Ciriza, V., & Bauduin, G. (2009). DISCO: Intelligent help for document review. In Global E-Discovery/E-Disclosure Workshop—A Pre-Conference Workshop at the 12th International Conference on Artificial Intelligence and Law, Barcelona, Spain.

Panitch, J.M., & Association of Research Libraries. (2001). Special collections in ARL libraries: Results of the 1998 survey sponsored by the ARL Research Collections Committee. Washington, D.C: Association of Research Libraries.

Quintana, A.G. (2009). Archival policies in the protection of human rights. Paris: International Council on Archives.

Scott, S.D., Allavena, A., Cerar, K., Franck, G., Hazen, M., Shuter, T., & Colliver, C. (2010). Investigating tabletop interfaces to support collaborative decision-making in maritime operations. In Proceedings of ICCRTS 2010: International Command and Control Research and Technology Symposium (pp. 22–24).

Sternfeld, J. (2011). Archival theory and digital historiography: Selection, search, and metadata as archival processes for assessing historical contextualization. American Archivist, 74(2), 544–575.

Tibbo, H.R., Lee, C.A., Marchionini, G., & Howard, D. (2006). VidArch: Preserving meaning of digital video over time through creating and capture of contextual documentation. In Archiving Conference (Vol. 2006, No. 1, pp. 210–215). Society for Imaging Science and Technology.

Tiwari, L.K., Samaddar, S.G., Singh, A.K., & Dwivedi, C.K. (2010). Evidentiary usage of e-mail forensics: Real life design of a case. In Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia (pp. 226–230). New York: ACM

Vajcner, M. (2008). The Importance of context for digitized archival collections. History and Computing, 11(1) (2008).

van Ballegooie, M., & Duff, W. (2006). DCC Digital Curation Manual: Installment on Archival Metadata. HATII, University of Glasgow; University of Edinburgh; UKOLN, University of Bath; Council for the Central Laboratory of the Research Councils.

Wellner, P. (1991). The DigitalDesk calculator: Tangible manipulation on a desk top display. In Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology (pp. 27–33). New York: ACM.

Wigdor, D., Perm, G., Ryall, K., Esenther, A., & Shen, C. (2007). Living with a tabletop: Analysis and observations of long term office use of a multi-touch table. In Horizontal Interactive Human-Computer Systems, 2007. TABLETOP'07. Second Annual IEEE International Workshop on (pp. 60–67). IEEE.

Zapata-Rivera, D., Hansen, E., Shute, V.J., Underwood, J.S., & Bauer, M. (2007). Evidence-based approach to interacting with open student models. International Journal of Artificial Intelligence in Education, 17 (3), 273–303.

DOI: 10.1002/asi