# Enhancing Evidentiary Work through the Lens of Human-Centered Computing

**LUIS FRANCISCO-REVILLA, CIARAN B. TRACE**
**The University of Texas at Austin**

**Abstract**: In order to understand the past, researchers look for evidence in archival collections. To preserve the evidentiary value of this material for scholars, archivists arrange these collections using a process that ties the records to the creator, their activities, and associated recordkeeping systems. Currently, this complex and time-consuming activity is done manually, often resulting in large backlogs. This paper proposes a novel method (Augmented Processing Table) for creating arrangements using tabletop computers and digitized images. A comparative study of both methods (traditional and APT) provided a deeper understanding of the value and intricacies of creating these arrangements, as well as the need to look to both technology and human factors in order to improve this aspect of the archival curation workflow. The study also revealed that the resulting arrangements vary significantly in terms of topology and quality, pointing out the need to improve replicability and quality control in the arrangement process.

## Introduction

Archives allow researchers (e.g., historians, journalists, and lawyers) to inquire about past human activities. For instance, the use of sources as evidence is at the heart of History's research method, and most historians use archives as the principal source for primary source materials.[1] The information behavior of archive users are part of a complex ecosystem of inquiry that goes beyond examining documents in isolation.[2] This is particularly true for humanities scholars who not only seek to understand a specific document in a collection, but also the context in which it was created, transmitted, and used.[3] Furthermore, many of the scholars' inquiries in archival collections do not directly pertain to the documents themselves, but rather to the individuals and the human activities that are responsible for the existence and state of the documents.

Supporting researchers' work with archival materials requires that archivists curate collections prior to their use. In order to help scholars and researchers find answers to their inquiries, in a manner that is supportive of their search for evidence, archivists carry out work that situates the documents in the collection, making apparent the contexts in which the documents were originally created and used. This process of archival arrangement provides a significant added-value that many researchers have come to expect. Consequently, most archivists consider archival arrangement a critical function that must be performed before users can engage with the collections. However, arrangement is a complex and time-consuming task, and as currently practiced, often results in large backlogs of unprocessed collections.

---

[1] Jennifer Rutner and Roger Schonfeld. Supporting the Changing Research Practices of Historians. Final Report from ITHAKA S+R, December 10, 2012. http://www.sr.ithaka.org/research-publications/supporting-changing-research-practices-historians.

[2] Neal Audenaert and Richard Furuta. "What Humanists Want: How Scholars Use Source Materials." In *Proceedings of the 10th Annual Joint Conference on Digital Libraries*, 283–292. JCDL '10. New York, NY, USA: ACM, 2010.

[3] Ibid.

Improving the workflow for processing collections is crucial for modern archives, although, to date, archivists have been slow to innovate.[4] Archivists have not taken full advantage of modern computational resources in respect to archival processing. This paper introduces a project designed to meet this challenge - the Augmented Processing Table (APT) project - and reports on the design and evaluation of the APT prototype system.

**Archival Curation**

The Digital Curation Centre (DCC) defines curation as a process that involves selecting, maintaining, preserving, archiving, and adding value to collections of data throughout its lifecycle.[5] From all the phases of archival curation, arrangement is distinguished as a cornerstone of archival work. In arrangement, archivists identify and document the network of relationships that exist between the records and their creator, associated functions, activities, chronological periods, geography, and recordkeeping systems.[6] Arrangement thus facilitates the work of scholars by directly linking documents to key internal and external contexts and thus allows researchers to treat the materials as evidence. Given the relevance of this curation phase, significant effort has been previously devoted to create best practice guidelines for how to arrange archival materials. Specifically, best practices require that archivists follow the principles of provenance and original order, and in doing so, all materials are intellectually and physically organized into a hierarchy comprised of standard sets or levels: record group, subgroup, series, subseries, file, and item.[7]

In the arrangement process, archivists interact manually with the collection, looking for discernible patterns that exist for groupings of documents. In doing so, archivists look at the external context (information about the creator, their functions and activities, events, geography, and chronological periods – captured at the record group and subgroup level), and the internal contexts of the collection (the recordkeeping system in use by the creator – captured at the series level and below). While this methodology has been around in modern form for over 80 years, there has not been much study of the actual work process itself, nor has there been much understanding of the underlying tacit knowledge that archivists draw upon to arrange collections.[8] This lack of understanding is a hindrance for developing computational solutions that facilitate this activity.

The Augmented Processing Table (APT) project aims to tackle some of these issues by doing baseline studies of the activity of archival arrangement, redesigning the archival workflow to take a 'digitize first, arrange second' approach, and by using large scale surface computing devices to augment and facilitate the arrangement process. The first APT prototype, a spatial hypermedia application, served as a proof of concept for this approach.[9] A pluralistic walkthrough evaluation of the system determined that archival processing is amenable to be conducted digitally using interactive surfaces. The study also revealed key interaction phases internal to the arrangement process, as well as the need to perform additional tasks, such

---

[4] An exception is the recent workshops on technology and archival processing sponsored by the Radcliffe Institute for Advanced Study at Harvard University.

[5] "What is Digital Curation?," Digital Curation Center, accessed April 3, 2013, http://www.dcc.ac.uk/digital-curation/what-digital-curation

6 Terry Eastwood, "Putting the Parts of the Whole Together: Systematic Arrangement of Archives," *Archivaria* 50 (2000): 94–116.

[7] A related term for the record group is "fonds."

[8] For studies that have examined the nature of the archival arrangement process see Crow et al. (2012) and Lemieux (2013). Jeff Crow, Luis Francisco-Revilla, April Norris, Shilpa Shukla, and Ciaran Trace, "A Unique Arrangement: Organizing Collections for Digital Libraries, Archives, and Repositories." In Proceedings of the International Conference on Theory and Practice of Digital Libraries (2012): 335-344. Victoria Lemieux, "Visual Analytics, Cognition and Archival Arrangement and Description: Studying Archivists' Cognitive Tasks to Leverage Visual Thinking for a Sustainable Archival Future," Archival Science (2013).

[9] See Crow et al., "A Unique Arrangement," 335-344.

as matching pages of two-sided pages, when arranging collections using digitized images. Based on these results, a second prototype was developed - APT 2.0.

**Problem Statement**

APT extends traditional archival arrangement methods to work with digitized images. Specifically, APT provides archivists with a large interactive surface where they can work with digitized images similarly to how they manipulate paper materials in a traditional workspace. This allows archivists to take advantage of their prior experiences working with paper, as well as the new functionalities that digital technology can afford. APT is designed to help archivists process at the item level. While file level processing is more common, the APT team began with item level processing in prototype 1 and 2 in order to tackle the activity of arrangement at the most granular level. Figure 1 shows the second APT prototype.
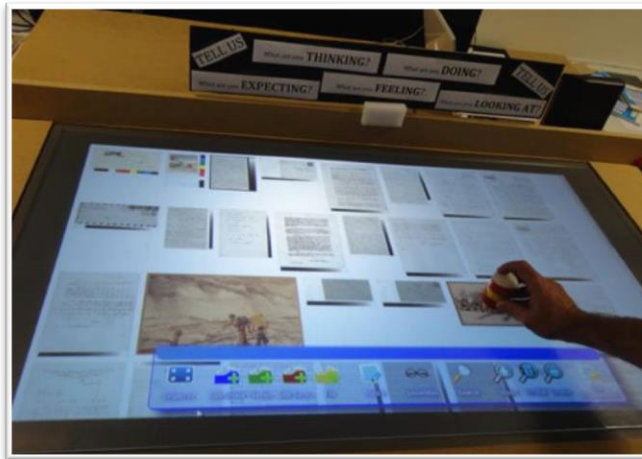


**Figure 1. APT 2.0 Prototype**

The second APT prototype remains a large interactive tabletop computer (47" x 28" interactive area), but the software uses a Web-based spatial hypermedia platform (WARP)[10]. The new functionality 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/or placed over another document)

APT 2.0 allowed the research team to investigate a number of key research questions.

- What is the nature of the arrangement process?
- Do archivists create similar or different evidentiary layers (arrangements) based on variables such as media type?
- Is APT an effective tool for processing a collection?
- Is APT an efficient tool for processing a collection?
- Is APT the preferred tool for processing a collection (preferred over the baseline)?

**Methodology**

---

[10] Luis Francisco-Revilla and Frank M. Shipman, "WARP: A Web-Based Dynamic Spatial Hypertext." In Proceedings of the 15th Conference on Hypertext and Hypermedia (2004): 235-236.

In order to compare the two systems, the study sought out a group of participants to work with collections and to create an arrangement 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 study performed a within-participant comparison that considered eight conditions (2 systems x 2 collections x 2 sessions). Participants were asked to take part in two sessions, each one lasting 90 to 180 minutes and separated by one to four days. In each session, each participant processed and created an arrangement scheme for a collection, once using APT (digital) and once using paper. Two collections were used in order to ensure that participants would work with a previously unknown collection in each session. The order of systems and collections was balanced to account for fatigue and learning effects. Participants were randomly assigned to one of the four groups (A, B, C, D).

| | 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 |

**Figure 2. Participant Groups**

In order to ensure that the experimental collections were truly representative, the researchers selected two publically available archival collections: Estelle Ishigo Papers[11] and Samuel Goudsmit Papers.[12] Due to the size of these collections, it was necessary to select a subset of materials from each so that participants could process the collection in one session. Assembling these subsets was a careful qualitative process. The use of randomized selection was not appropriate because there was no way to guarantee that the resulting subset would have the evidentiary relationships that the study required. Hence, the researchers had to manually assemble the test collections, using the existing finding aids and biographies to ensure that they provided evidence of different aspects of the life of the creators. Obvious traces of the original order (filing system) were removed from the experimental collections because, while knowing the original order of a collection is a great aid for creating an arrangement that captures its provenance and record keeping, APT aims to support this creation even when this knowledge is not easily or directly available. The final experimental collections included 22 items for Ishigo (55 pages), and 24 items for Goudsmit (47 pages).

At the beginning of the first session, a survey collected demographic data including age, gender, technological background and experience arranging collections. During the sessions the study produced data in the form of video and audio recordings, researchers' notes, and participants' notes. At the end of each session every participant produced a written final arrangement scheme. At the end of the second session, an exit survey collected the participants' assessments of using the two systems (paper and APT). Finally, a free form interview asked participants for additional comments and suggestions.


**Results**

---

[11] "Estelle Ishigo Papers," Online Archive of California, accessed April 3, 2013,
   http://www.oac.cdlib.org/findaid/ark:/13030/tf409nb2b5/
[12] "Samuel A. Goudsmit Papers," American Institute of Physics, accessed April 3, 2013,
www.aip.org/history/nbl/collections/goudsmit

The study produced qualitative and quantitative data, which was analyzed using visualizations, coding schemes, descriptive statistics, and ANOVA. 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.

Since APT aims at enhancing the processing of archival collections, it was important to compare the current practices with APT. The first aspect that was compared that will be described here is outcome replicability, namely the ability of both methods to let different archivists produce similar arrangements for a given collection. This speaks 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 notion of replicability required examining the structure of the different hierarchies produced by all participants (the structure of the hierarchies is referred as the *arrangement topology*). The following subsection presents the topological features common to all arrangements and the variations associated with different experimental conditions. The next subsection discusses how the propensity for human errors in the resulting arrangement schemes depends on the platform used to create them (digital or paper).
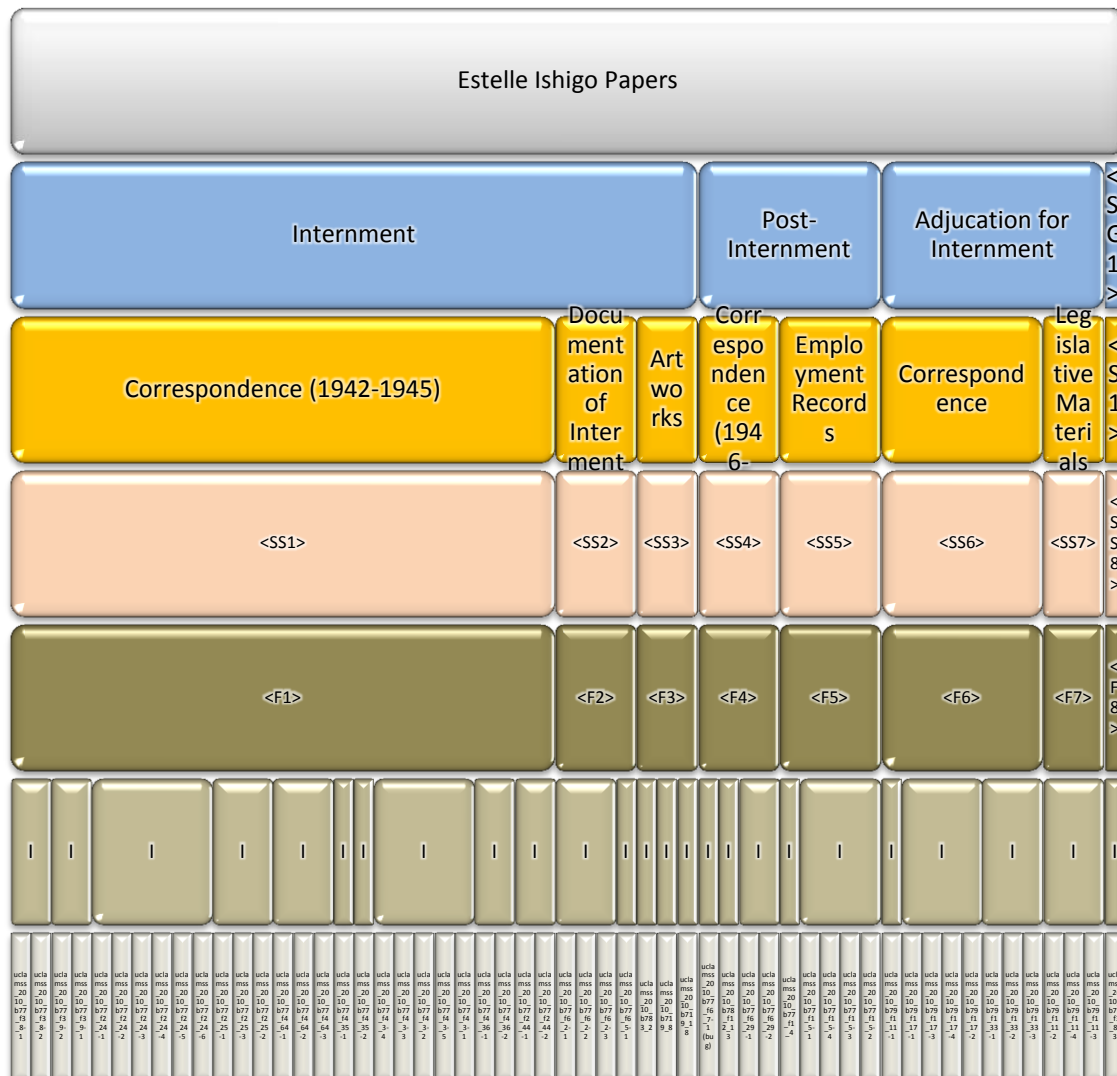
**Figure 3. Arrangement topology.**

*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 of a single tree, or to compare multiple trees. Figure 3 shows an example of an arrangement topology visualized in a way that facilitates side-by-side comparisons between different participants. Each row in Figure 3 represents an archival arrangement level. The top represents the overall collection level, followed by sub-groups, series, sub-series, files, and items at the bottom-level. In order to facilitate comparison across participants, all collections were normalized. As a result, missing levels are shown between brackets, e.g. when a participant put an item directly inside a series without nesting it inside a file and a subseries, the *phantom* subseries and file are denoted as <SS1> and <f1> respectively.

The analysis also quantified and compared the number of elements at each level. This revealed statistically significant differences between arrangements constructed using paper or digitized images. The means that the arrangement hierarchies and use of standard sets/levels (the outcome of the arrangement process) differs under the two conditions. Specifically, using APT, with the associated digital images, resulted in:

- More real and total subgroups ($p < 0.02$)
- More phantom series ($p < 0.02$)
- Less phantom sub-series ($p < 0.04$)

These findings indicate that the arrangement schemes created under the APT condition do a better job of highlighting the narrative elements of the arrangement. That is, the arrangements more often capture the level that links the material to the life and activities of the creator (subgroup). 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 attention to the contextual layers that exist at the intellectual level (i.e. linking 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).

*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 matching process 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 in the same item. Merges occur when a person puts two or more unrelated pages in the same item. Figure 4 shows the error count for every item in both collections.

In general, the results show that the number or errors for a given item are lower when using paper. This indicates 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 also found that there is a propensity in working with digital images to split things that look different (e.g. envelopes from letters). Results also indicate that overall, page matching was harder to do with the Ishigo Collection, possibly due to the greater diversity of genre formats in this collection.
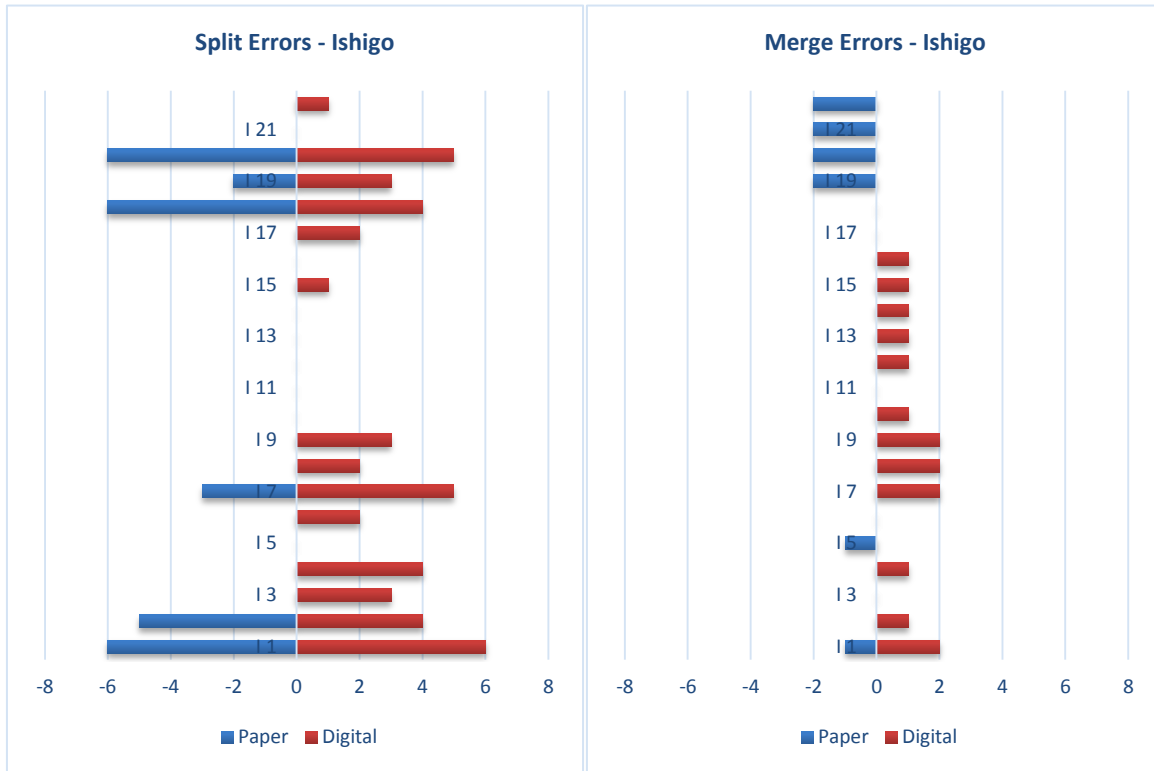
**Figure 4. Error Count in Final Arrangements.**

The splits results indicated that some items are just hard to match (they had high error counts regardless of the system being used). The analysis looked at identifying which items are 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 results show that paper and digitized images present different challenges for page matching, although merge errors are more common in the digital environment. Unlike the split errors, the number of merge errors was too small to identify clear differences between collection and media types.

| | Paper | Digital |
|---|---|---|
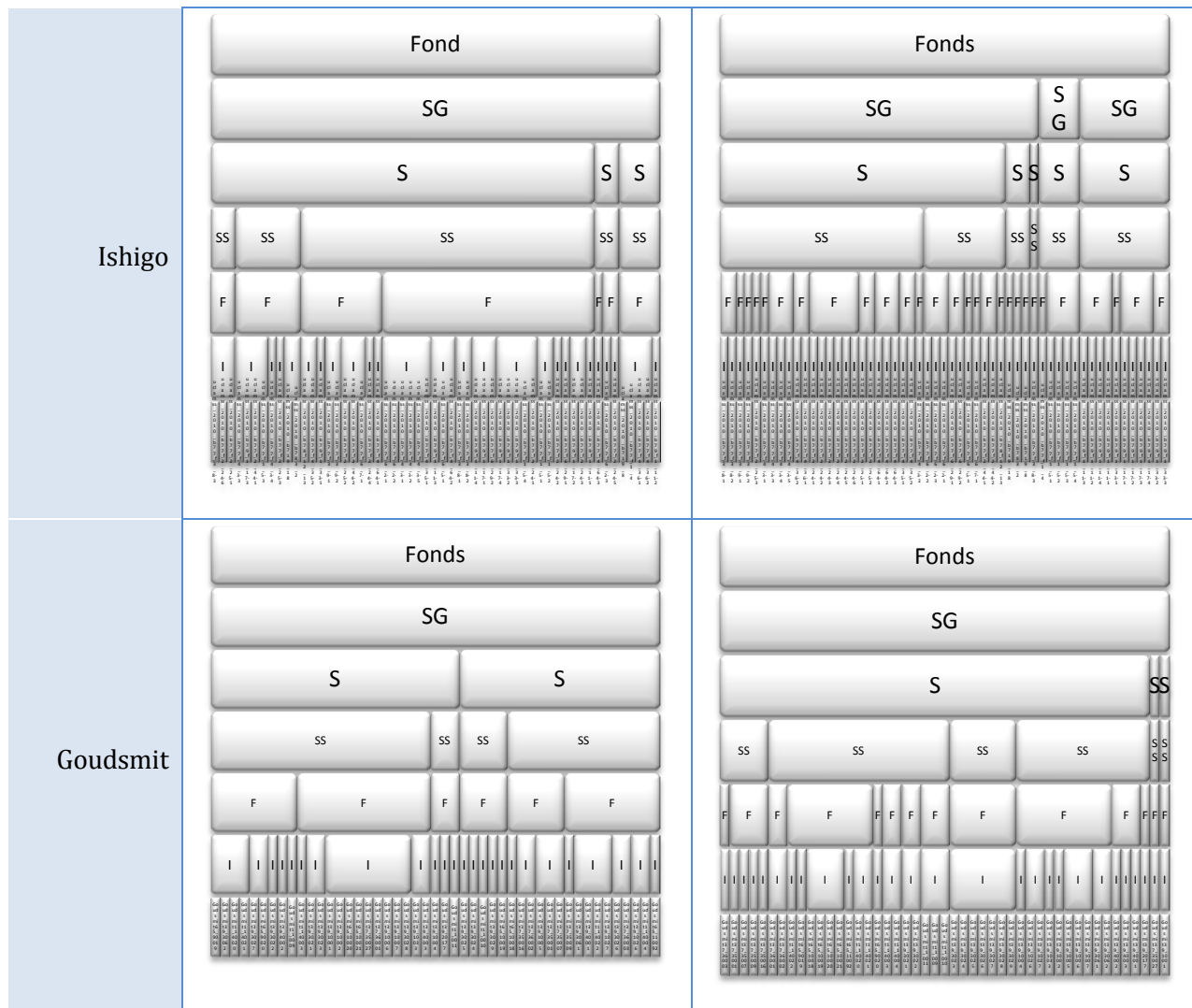
**Figure 5. Average Arrangement Topologies.**

## Findings

In terms of replicability, the study revealed that the current archival methodology does not guarantee the creation of similar arrangements. Whether using APT or paper, the final arrangements presented significant variations. These variations are not correlated with archival experience (the amount of experience a participant had in processing collections), or the quality of the topology.

An important finding was that significant topological variations occur at the levels of the arrangements that encode the provenance of the records and the filing system that reflects how records were created, received or used in the same activity. Moreover, when looking at groups of arrangements, individual variations do not average out. Figure 5 shows the average arrangement topology for each experimental condition.

The average topologies show that the alternative approaches result in different variations, both in terms of topology and propensity for matching. These findings demonstrate that archivists may privilege certain levels/sets within the arrangement scheme resulting in certain aspects of the collection being emphasized or deemphasized.

**Discussion and Future Work**

The possibility of variation between arrangements raises questions about the replicability and usefulness of arrangements. With a degree of flexibility being shown in the arrangement process, the question remains as to why archivists sometimes deviate from the arrangement guidelines already in place and what, if anything, to do about it. This leads to a broader question of whether it makes sense to think that there is, or should be, one 'right' arrangement scheme for each collection.

The study clearly established that matching errors are possible when processing collections and that, as a result, documents may be placed in the wrong context (file, series, subgroup, etc.). From a researcher's perspective, a potential outcome of this scenario is that a document could be misunderstood or misinterpreted. Another result of matching errors is that documents, or parts of documents, may not be discovered at all, and therefore fail to become a part of the researcher's accumulated body of evidence.

A logical follow-up study based on these findings would be to look at whether the current topology (hierarchical tree), along with the associated sets/levels, actually provides the best possible support for those doing evidentiary work within diverse research communities (historians, journalists, lawyers, etc.). 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 how that network of relationships should be represented. It may be that the traditional hierarchical arrangement schemes with levels representing the creator, their activities, and associated recordkeeping system, may stand up to further scrutiny. However, if communities demonstrate unique research needs, the workflow and the outcome of the arrangement process may need to change to reflect this.

The next iteration of APT (APT 3.0) will build in additional functionality for processing at the file level. This will allow for greater scalability in terms of the size of collections that can be processed using APT. Based on findings from the user study, the design will also continue to evolve, with a particular eye on providing additional support for matching documents at the item level, and for linking documents to the larger functions and activities of the creator. Such solutions will 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. The foregrounding of information about the creator within the APT workspace (through the use of a timeline feature, for example) would also help to ensure that the external contexts of the collection are not overlooked in the arrangement process.

**Resources**

American Institute of Physics. "Samuel A. Goudsmit Papers." Accessed April 3, 2013.
    http://www.aip.org/history/nbl/collections/goudsmit/

Audenaert, Neal and Richard Furuta. "What Humanists Want: How Scholars Use Source Materials." Proceedings of the 2010 Joint International Conference on Digital Libraries (2010): 283-292.

Crow, Jeff, Luis Francisco-Revilla, April Norris, Shilpa Shukla, and Ciaran B. Trace. "A Unique Arrangement: Organizing Collections for Digital Libraries, Archives, and Repositories." Proceedings of the International Conference on Theory and Practice of Digital Libraries (2012): 335-344.

Digital Curation Center. "What is Digital Curation?" Accessed April 3, 2012.
    http://www.dcc.ac.uk/digital-curation

Eastwood, Terry. "Putting the Parts of the Whole Together: Systematic Arrangement of Archives." Archivaria 50 (2000): 93-116.

Francisco-Revilla, Luis, and Frank M. Shipman. "WARP: A Web-Based Dynamic Spatial Hypertext." Proceedings of the15th Conference on Hypertext and Hypermedia (2004): 235-236.

Lemieux, Victoria. "Visual Analytics, Cognition and Archival Arrangement and Description: Studying Archivists' Cognitive Tasks to Leverage Visual Thinking for a Sustainable Archival Future." *Archival Science* (2013).

Online Archive of California. "Estelle Ishigo Papers." Accessed April 3, 2013. http://www.oac.cdlib.org/findaid/ark:/13030/tf409nb2b5/

Rutner, Jennifer, and Roger C. Schonfeld. Supporting the Changing Research Practices of Historians. Final Report from ITHAKA S+R (December 10, 2012). http://www.sr.ithaka.org/research-publications/supporting-changing-research-practices-historians.