# Dispatches, Digests and Doodles:
# Exploring the Significant Properties of Field Notebooks

**Andrea K. Thomer**
thomer2@illinois.edu

**Nicholas M. Weber**
nmweber@illinois.edu

**Center for Informatics Research in Science and Scholarship**
**Graduate School of Library and Information Science**
**University of Illinois at Urbana-Champaign**

## Abstract

Whether a first-person narrative, a structured observational data source, or simply a diary of doodles, field notebooks are *the* material instantiation of most site-based social and natural sciences. In this poster, we present preliminary work that explores the significant properties of field notebooks across three diverse disciplines: Paleontology, Archaeology, and Field Biology. This work is meant to supplement existing norms in digitization processes, and begin moving many of those projects towards a more general, interoperable approach to preserving the rich content held within field note collections.

*Keywords:* digital preservation, field notes, natural history, museum informatics, digitization.

## Introduction

### Beyond the silos of the LAMs: making field data broadly accessible

Field notes – the recording and cataloguing of observations in site-based field-work – often contain a huge amount of data structure in their writing, yet libraries, archives and museums typically do not have the resources or domain expertise to reveal that structure. Creating a typology of this structure and making it machine-readable will only become more important as more notebooks are not just digitized but also transcribed. Currently efforts to make that structure discoverable through annotation or text markup are nascent-to-nonexistent (Thomer et al., 2012).

### What's significant about significant properties?

Much has been written about the preservation of digitized cultural objects from a curatorial standpoint, including a differentiation between *representation properties*, which describe the data content of a digital object, and *significant properties,* which describe the informational content of a digital object (CEDARS, 2002). Our use of significant properties in this poster is similar to Knight and Pennock, and Wilson's definition: "...significant properties are defined as the characteristics of an information object that must be maintained to ensure that object's continued access, use, and meaning over time as it is moved to new technologies" (2011, p.163; 2007). Here however, we want to emphasize the importance of maintaining an intelligible transfer of the content from a recorded observation found in a field notebook to the digital instantiation of that content, which may be viewed as both a surrogate of the original analogue text and a data point (for a more thorough discussion see Thomer et al., 2012).

**Methods.** We performed a comparative analysis of individual field notebooks from three different scholarly disciplines (paleontology, archaeology, field biology), and reviewed domain-specific literature describing best practices to help ensure that our analysis was commensurate with field practices of that discipline in a general sense, as opposed to field practices specific to a particular person or expedition.

---

In assessing each field notebook for its significant properties, we followed Grace and Knight's typology of significant properties: content, context, rendering, structure and behavior (2008). Below we map those significant property types onto questions we answered in analyzing the field note collections from our three chosen disciplines:

- What temporal division is most frequently recorded, and how (e.g. day, field campaign, year)? **Context**
- Is there a cataloguing or pagination scheme in place? **Structure**
- How are external sources of data included or cited (e.g. sketched, pasted in, cited)? Are external databases referenced? **Content**
- What structured data is collected and how is that structure represented? **Rendering**
- What and how is quantitative data collected?  **Behavior**
- What events are recorded? **Content**

## Analysis

**Paleontology (Simpson, 1930; Whitmore, 1975; Walcott, 1879).** In paleontology, field books are labeled with expedition name, date range, book number within a larger set (e.g. "Book 1"), and author name.
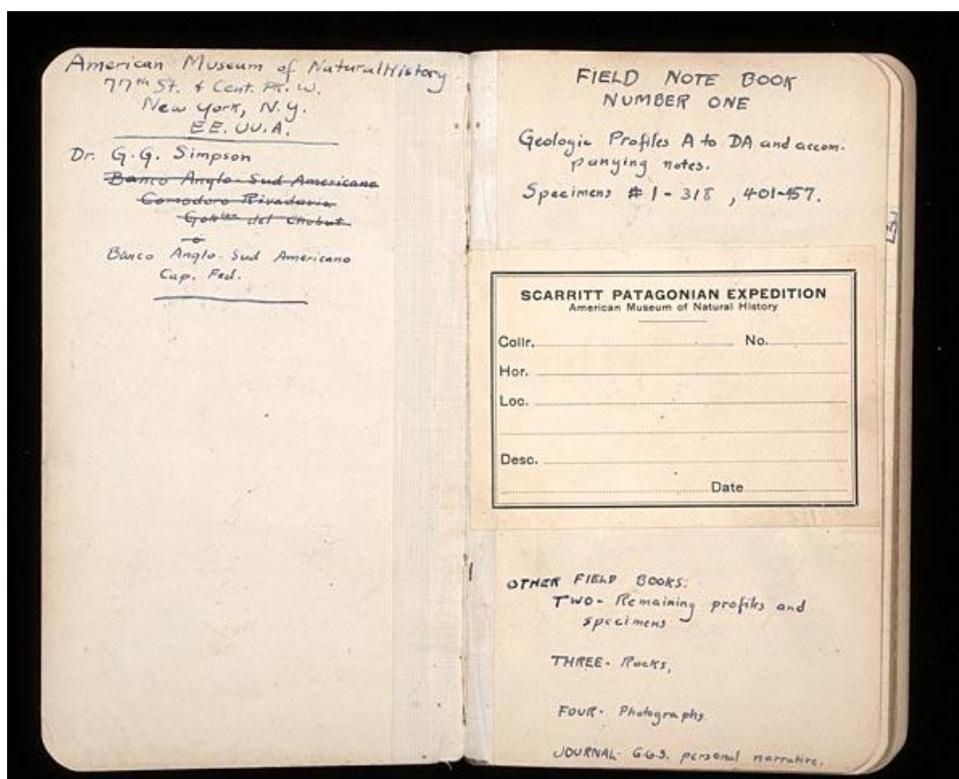


*Figure 1.* Index to George G. Simpson's notebook, 1930. Available from
http://research.amnh.org/paleontology/notebooks/simpson-1930a/01.jpg

Entries are divided by day, but the books are divided by both subject and campaign (that is, each expedition is recorded in a new series of notebooks). Pages are hand numbered and are often continuous from one book to another within a series (e.g. in Simpson's case, book two starts on page 167). The inside cover of the first book in a series includes a list of specimen numbers that are referenced in the book (this list would be added after the book was 'completed'), as well as an index to *other field books* describing other aspects of the expedition (Figure 1) – particularly the geologic context surrounding each collected specimen -- thus creating a kind of database-of-notebooks all related to one particular

938

expedition or site. While not every paleontologist keeps this same system of notebooks, many have similar practices; "locality" ledgers containing detailed stratigraphic and geologic data are common (if not necessary).
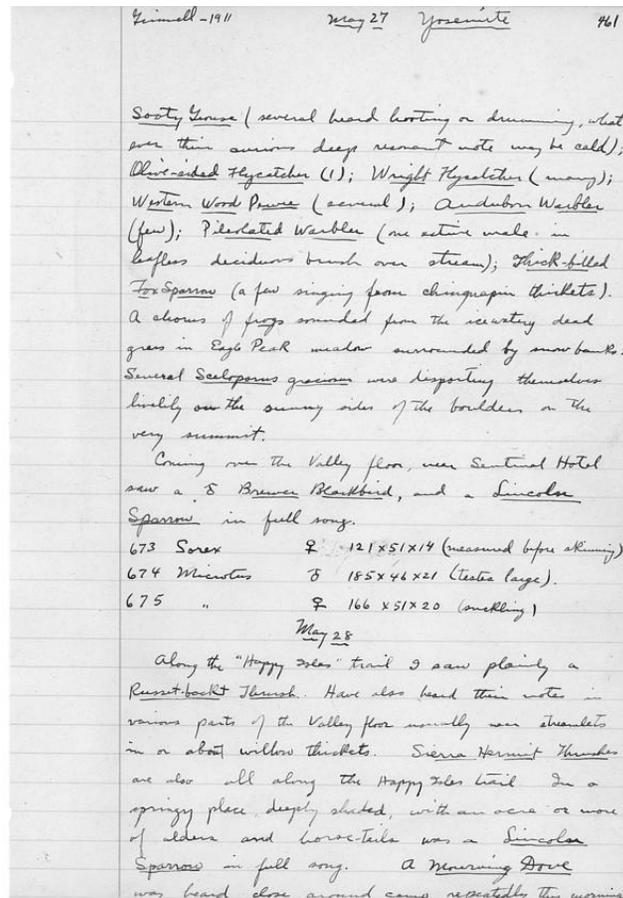


*Figure 2.* A page from Joseph Grinnell's notebook (1911) featuring free text narrative and structured data. Available from http://bit.ly/TQ1YOe

**Field Biology (Grinnell, 1911; Rafinesque, 1818).** Many biologists write in their notebooks in a manner that is similar to, but less geologically-oriented than, paleontologists: many maintain a pagination scheme that stretches over many notebooks (possibly over the duration of their entire careers), and they mix personal narrative with semi-structured descriptions and lists of the animals and plants observed while travelling. Biologists conducting field work often follow a variation of the "Grinnell Method" of data collection: plant and animal names are underlined; their number or even absence are noted; and collected specimens are recorded in a table including a catalog number, their gender, a map reference, and additional notes describing their appearance or behavior (Perrine & Patton, 2011). Grinnell himself did not seem one to sketch (though many other biologists do sketch their subjects), but he did on occasion paste annotated maps into the front or back pages of his journals detailing where, exactly, he did his work.

**Archeology (Strong, 1933; Smith, 1938).** Archaeological field books tend to be temporally bound by the length of an expedition, and the 'entries' are recorded in frequency by day. In some of the books we surveyed, we noted a high amount of marginalia; one author (Strong) seems to have summarized some of his narrative into short parenthetical statements to the left of the main text. Archaeological journal entries are also less systematically laid out than in other disciplines; the author often relies on the pre-printed page numbers in the upper corners of each page instead of creating his own numbering system, and if artifacts were collected, their catalog numbers are not as prominently or systematically recorded.

While these notebooks do contain data in the form of maps and detailed drawings of pottery sherds, these diagrams are primarily used to augment the author's daily narrative describing his travels, meals, and conversations with other people. More extensive and structured data collection often happens outside the field notebook and in the context of formal excavations.
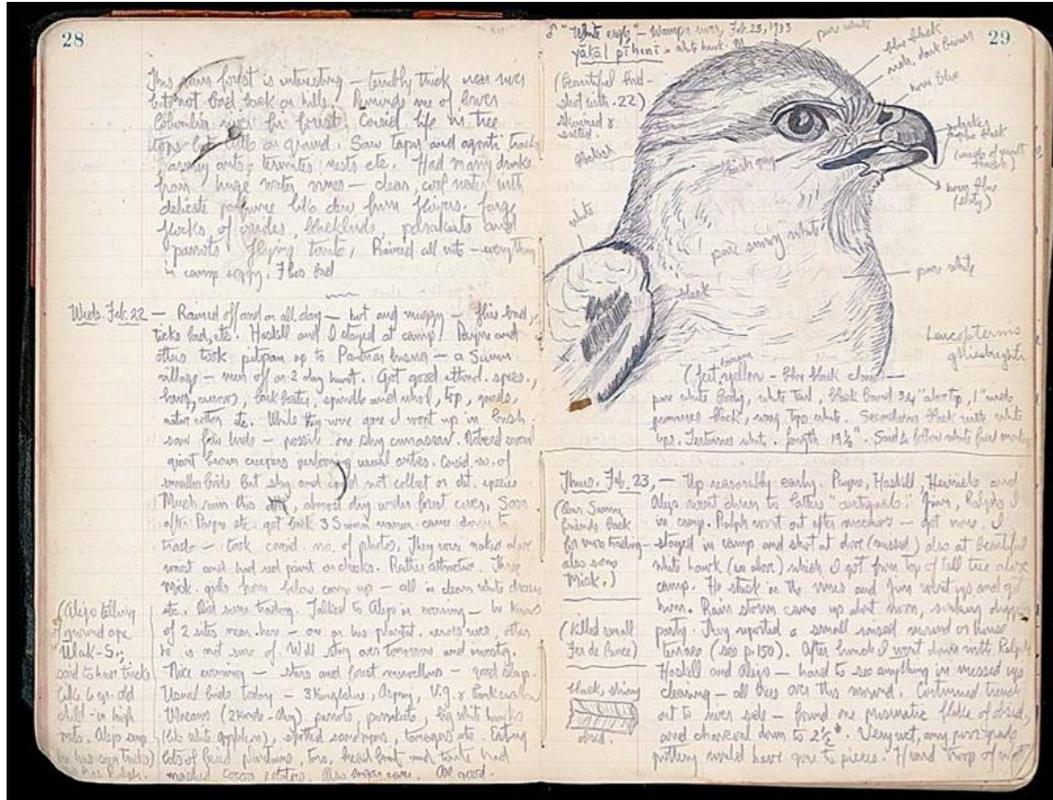


*Figure 3.* A page from Strong's notebook. http://www.nmnh.si.edu/naa/features/strong2z.htm

## Discussion

### A typology of not just notes, but also disciplinary practice

Field books differ not just between disciplines, but over time; social and scientific practices evolve from one year to the next, and notebooks within in the same domain vary wildly, for example, from 1850 to 1950. While we did find that the significant properties of field notes across these three disciplines fit broadly within the Grace and Knight's (2008) typology, we also note that the property types have very different implications for the re-use of these records as data sources. For instance, the 'context', 'content' and 'behavior' properties are likely most important for discovery, while the 'structure' and 'rendering' properties are likely more important for cross-disciplinary collecting and digitization purposes.

Categorizing field notes not just by discipline, but by data structure and data collection method may better support later use of field notes while also informing our understanding of how data collection practices evolve over time. Almost every notebook that we looked at contained data that would be usable by another field; paleontological notes contain descriptions of towns, historical events and people that would be of potential interest to historians and geographers, and archaeological field notes often contain detailed drawings and descriptions of local flora and fauna that would be important to biologists.

# Future work

## Making use of undiscovered personal knowledge for public science

This work is meant as a first step in exploring how the similarities in research methods and document structures recorded in field notebooks might be formally translated into a better understanding of the significant properties of digitized field notes. In turn, we believe that a better understanding of these properties is important for the success of digitization projects and the development of preservation standards in libraries, archives and museums. We also believe that working towards a taxonomy of field notes and field data will not only aid future efforts in creating a field note markup/annotation schemas, but will also provide important insights to the varied ways that recorded data inform a process of knowledge production.

# References

CEDARS Project. (2002). *Cedars guide to: Digital collection management.* Retrieved September, 2012 from http://www.leeds.ac.uk/cedars/guideto/collmanagement/

S. Grace and Knight, G. (2008) What are significant properties and why should I care?. Presentation delivered at Digital Curation 101, October, 7 2008. Edinburgh, Scotland.

Knight, G., & Pennock, M. (2009). Data Without Meaning: Establishing the Significant Properties of Digital Research. *International Journal of Digital Curation, 4,* 1.)

Grinnell, J. (1911). Field notes from Yosemite trip. Archives of the Museum of Vertebrate Zoology, University of California, Berkeley. Retrieved from http://bscit.berkeley.edu/mvz/volumes.html; Section 1, 27 May 1911, page 461. Accessed 20 September 2012.

Perrine, J. D., & Patton, J. L. (2011). Letters to the Future. In M. R. Canfield (Ed.), *Field Notes on Science & Nature.* Cambridge, Massachusetts: Harvard University Press.

Rafinesque, C. (1818). Notebook kept by Rafinesque on a trip from Philadelphia to Kentucky, 1818. Smithsonian Institute Archives; RU 007250, Box 1 Folder 3.

Simpson, G. A. (1930). Field notes from Scarritt Patagonia Expedition. American Museum of Natural History Archives. Retrieved from http://research.amnh.org/paleontology/notebooks/simpson-1930a/. Accessed 20 September 2012.

Smith, A. (1938). Field notes from UNM field school at Chaco NHP. Bc 50: 2083. Chaco Culture NHP Museum Archive, University of New Mexico, Albuquerque. Retrieved from http://www.chacoarchive.org/media/pdf/000806.pdf Accessed 20 September 2012.

Strong, W. D. (1933). Selections from the Field Journal of William Duncan Strong (Honduras, 1933). National Museum of Natural History Archives. Retrieved from http://www.nmnh.si.edu/naa/features/strong.htm. Accessed 15 September 2012.

Thomer, A., Vaidya, G., Guralnick, R., Bloom, D., & Russell, L. (2012). From documents to datasets: A MediaWiki-based method of annotating and extracting species observations in century-old field notebooks. ZooKeys, 209, 235–253. doi:10.3897/zookeys.209.3247

Walcott, C. D. (1879). Field notes : October 15-November 3, 1879. Smithsonian Institution Archives. SIA RU007004, Box 32, Folder 1.

Whitmore, F. C. (1975). Notebook, 1965-1980. Smithsonian Institution Archives. SIA Acc. 02-012, Box 1, Folder 3, Env 5.

Wilson, A. (2007). Significant properties report. Retrieved September, 202, from http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf