

DH Benelux Conference – 2023

Belgium at KBR – the Royal Library of Belgium, Brussels, from 31 May -2 June, 2023

“If data is king, context is the crown”: making sense of digitized sources with Tropy
and paving the way for integrative source processing

Presenters:

Anita Lucchesi¹

Douglas McRae²

Sean Takats³

This paper introduces [Tropy](#), a free, open-source tool designed specifically for today's scaled-up archival research in an open and multilingual research environment. Alongside the presentation of Tropy's ability to manage large numbers of digital images, we will discuss the integrative potential of the tool for transforming the research data lifecycle. The presentation will refer to concrete user cases and build on the perspective of *scholars as collectors* and *data curators* (Cooper & Rieger, 2018).

Over the past two decades, digitized sources have become increasingly prevalent in the historian's craft. The sheer volume of digital data, however, can be overwhelming and make it difficult to understand and analyze. Any historian who has worked with digitized primary sources knows that establishing a consistent mechanism for organizing and making sense of large sets of research photos can be a daunting and time-consuming task. For his 2020 paper, historian Ian Milligan conducted a survey among academic historians working in Canada, and only 3.17% of the 253 respondents reported *not* taking digital photographs. According to the same survey, 90% received no professional training in this field. When asked how many photos they took for their most recent substantive project, more than a third of respondents (39.74%) said more than 2000 (Milligan 2020). This may be just discrete evidence, but it demonstrates a significant shift in historians' research practices as a development of digitization projects as well as the introduction of digital cameras in archives reading rooms (Rutner & Schonfeld 2012, 12). Notably, in the last few years, historians are spending more time taking photos than intellectually engaging with the sources during their archive trips. In such a new research landscape, historians may frequently find themselves with multiple folders full of images, each with its own naming convention and date stamp, but these fruits of hard work risk becoming meaningless without the proper metadata attached to them.

¹ Lucchesi's ORCID ID: <https://orcid.org/0000-0002-8523-111X>

² McRae's ORCID ID: <https://orcid.org/0000-0003-0989-5122>

³ Takats' ORCID ID: <https://orcid.org/0000-0002-7851-5069>

Given the complexity, abundance, and fuzziness of data today, we should remember that if content, or data, is king, context is its crown (Snickars 2012; Fickers 2013). This is where Tropy comes in: as a digital tool designed specifically for managing and annotating digital images, it provides a means of adding context to digitized sources. Tropy was initially developed at the Roy Rosenzweig Center for History and New Media (RRCHNM) and was first introduced as a beta version in 2016. Currently, its development is a joint effort between RRCHNM, the Luxembourg Centre for Contemporary and Digital History (C²DH), and Digital Scholar, a non-profit corporation that operates several data management projects such as Zotero, Omeka, and Sourcery. The development these tools were, somehow, catalyzed and driven by the growing demand for support with information management practices and technologies; a major dilemma in the Library and Information Science realms (Trace & Karadkar, 2016).

Tropy is an asset for researchers who need to manage a large number of digital images. Through tags, lists, notes, and metadata templates, the tool helps researchers organize, describe, and annotate their research photos. For the sake of consistency, these features allow researchers to customize their data management process while sustaining standardization through customized metadata templates and controlled vocabularies. Furthermore, the search function enables quick queries encompassing all the project information, including metadata, notes, and tags.

In this presentation we will show how Tropy's features improve the contextualization of historical artifacts and facilitate information retrieval – two of the most explicit benefits of using Tropy to make sense of digitized sources. A further aspect we want to emphasize is that well organized and annotated sources might be interesting not only to the individual researcher, but also to the surrounding community that could benefit from them. In fact, another goal of this presentation is to highlight Tropy's implicit and collective benefits as an integrative research data management engine. After demonstrating how Tropy works, we will discuss how cross-platform and data reuse emerged in different user cases, based on how researchers explored Tropy's integrative perks to bring their data elsewhere. In general, after curating their Tropy projects and writing about their sources, users could consider the work done for that specific dataset. But with Tropy, users can also take a further step by exporting their research data to other platforms for remixing, reusing or reshaping, either by themselves or by other researchers and audiences. Tropy currently allows exporting in several formats such as CSV, JSON-LD and PDF, or through direct integration with Omeka via plugin.

By assisting researchers shift away from individualized and isolated work and toward more collaborative endeavors, Tropy paved the way for a more integrated source processing where the collective can benefit from the tremendous amount of work historians use to do alone. Consider how much time and resources could be saved if individual efforts to collect and organize archival material could be seamlessly shared with others. Furthermore, what could be the added value for a broader community of researchers? On the one hand, we argue, the tool can help individual historians safely climb their *Himalaya of Data* (Snickars 2014), finding a consistent system to bring logic and meaning to their stack of research photos; on the other, it can contribute to the development of an open science culture by facilitating the establishment of flexible yet systematic workflows that make it easier to adhere to values such as scientific research transparency, collaboration, and reproducibility.

Whether they work alone or in groups, researchers can use Tropy to fulfill some of the most important requirements for moving digital scholarship toward open science. Take, for example, reproducibility and replicability as timeless components of scientific research. With today's array of digital tools, researchers undoubtedly have abundant (even free) means to publish their data, code, and findings. However, not all tools are simple to use for non-technologists and not all of them take interoperability into account. With Tropy, all researchers can create a metadata-rich environment using its user-friendly interface, maintaining fine-grained control over their research data that is essential for advancing the research cycle.

References:

- Cooper, Danielle, and Oya Y. Rieger. 2018. "Scholars ARE Collectors: A Proposal for Re-Thinking Research Support." *Ithaka S+R*, November 28
<https://sr.ithaka.org/publications/scholars-are-collectors/>
- Fickers, Andreas. 2013. "Veins filled with the Diluted Sap of Rationality: A Critical Reply to Rens Bod". *BMGN - Low Countries Historical Review* 128 (4): 155–63.
- Milligan, Ian. 2020. "AHA 2020 - Becoming a Desk(top) Profession". Conference paper presented in the American Historical Association Annual Meeting, New York, janeiro 5. <https://www.ianmilligan.ca/talk/aha-2020/>.
- Rutner, Jennifer, e Roger Schonfeld. 2012. "Supporting the Changing Research Practices of Historians". *Ithaka S+R*. <https://doi.org/10.18665/sr.22532>
- Snickars, Pelle. 2012. "If Content Is King, Context Is Its Crown". *VIEW Journal of European Television History and Culture* 1 (1): 34–39.
<https://doi.org/10.18146/2213-0969.2012.jethc006>.
- . 2014. "Himalaya of Data". *International Journal of Communication* 8: 13.
- Trace, Ciaran B., and Unmil P. Karadkar. 2017 "Information Management in the Humanities: Scholarly Processes, Tools, and the Construction of Personal Collections." *Journal of the Association for Information Science and Technology* 68, (2): 491–507.
<https://doi.org/10.1002/asi.23678>