RETROSPECTIVE, SUBJUNCTIVE, PROSPECTIVE: PROVENANCE CHALLENGES ACROSS TIME

Rhiannon Bettivia

Simmons University
USA
bettivia@simmons.edu
0000-0003-4593-562X

Yi-Yun Cheng

Rutgers University
USA
yiyun.cheng@rutgers.edu
0000-0001-6123-7595

Michael R. Gryk

University of Illinois
USA
gryk2@illinois.edu
0000-0002-3483-8384

Abstract – This panel will explore provenance: as theory and practice; as a tool for sustainability; and as a space of shared struggle and challenge for digital preservations and those in fields ranging from archives to cluster computing. In digital preservation, provenance tells us where an object has come from, the myriad preservation actions we could take to care for it, and where we predict the object will need to go in future. This panel is intended for anyone who is interested in the world of provenance: defining it, understanding it, modeling it, addressing the vague dissatisfaction practitioners often have when researching and documenting it. Provenance is more about the journey than the destination: this panel aims to surface a variety of experiences with provenance and to facilitate discussion and a community of practice around the relationship between digital preservation and provenance.

Keywords – provenance, authenticity, evidence, archival values, information theory

Conference Topics – SUSTAINABILITY: REAL AND IMAGINED; WE'RE ALL IN THIS TOGETHER; FROM THEORY TO PRACTICE

I. INTRODUCTION

The OED defines "provenance" as origin, source, and ownership, tied tightly to the ability to determine authenticity. Provenance can be used to describe what did happen (retrospective provenance), what could happen (subjunctive provenance), and what will happen (prospective provenance). Provenance has many faces in different fields: the tree of life in phylogeny; ancestry of families in genealogy; layers of sediments in stratigraphy. Provenance transcends disciplines. In digital preservation, the custodial chain, audit trails, iteration reports, and change logs are building blocks for establishing authenticity in the face of managed change over time. Measurable properties and questions about the identity of digital objects engender challenges in modeling and recording different stages of computational projects. The lack of provenance information for born-digital objects in each stage of a research pipeline can reduce the transparency, trustworthiness, and reproducibility. Xu et al. [6] state that reliance on process has actually changed and expanded traditional uses of the term provenance:

"The notion of provenance has been adopted and extended in the field of Computer Science and applied to concepts such as data, computation, user interaction, and reasoning. In this context, provenance is no longer limited to origin or history, but also includes the process and other contextual information."

Technologies like blockchain are bound up with procedural provenance in their very form and function [5,3]. Provenance stories will play an increasing role as AI artifacts may impede the ability of archival materials to accurately represent the historical records [4].

We define provenance broadly as how something has come to be, and we incorporate the following key concepts into our exploration of provenance [1,2]:

- Provenance is fluid and transcends time;
- Creating provenance descriptions is both a conceptual modeling, a metadata

iPRES 2023

recording exercise, and a persuasive exercise;

- Working with provenance is both a ubiquitous and field-agnostic act.

As they became more commonly encountered in archives and other information institutions, digital records destabilized commitments and assumptions of traditional preservation. The lack of provenance information for born-digital objects in each stage of a research pipeline can reduce the transparency, trustworthiness, and reproducibility. Archival concerns over the mutability of digital records eventually gives way to a realization that their particular affordances may support additional techniques for ensuring provenance than physical records, such as embedded metadata, blockchain technologies, and digital forensics approaches.

Provenance becomes an intellectual and moral concern as collections of digital objects are managed through their life cycles, migrated, emulated, and remediated in new formats and interfaces, such as virtual and augmented reality. How does our conceptualization of provenance adapt to these new conditions and ensure that we can continue to trust the authenticity and integrity of copies over time?

II. Panel Structure

We have assembled a panel to have an exploratory discussion about the concept of provenance. This 90-minute panel brings together participants from across the ASIS&T community to represent concerns from information organization, research data management, metadata, cultural heritage, archives, digital curation, data curation, and digital preservation.

A. Part I: Establishing a Baseline

B. Each panelist will share their thoughts on provenance and how it intersects with their work.

C. PART II: Interactive Q&A

The panel will address topics posed by a moderator. These topics include:

-PREMIS as a provenance standard

-Workflows as prospective provenance

-Subjunctive provenance as a mediator with future audiences

III. Panelists and Coordinators

Karin Bredenberg (invited panelist) is the Metadata Strategist at the Kommunalförbundet Sydarkivera, a local federation of 37 municipalities in Sweden.

Dr. Alexandra Chassanoff (invited panelist) is an Assistant Professor at the School of Library and Information Sciences at North Carolina Central University.

Dr. Zack Lischer-Katz (invited panelist) is Assistant Professor in Digital Curation and Preservation at University of Arizona's School of Information.

Dr. Mike Twidale (invited panelist) is a Professor and the PhD Program Director at the iSchool at the University of Illinois, Urbana-Champaign.

Dr. Rhiannon Bettivia (coordinator) is faculty at Simmons University in Library and Information Science.

Dr. Yi-Yun (Jessica) Cheng (coordinator) is faculty at SC&I at Rutgers University.

Dr. Michael R. Gryk (coordinator) is Associate Professor of Molecular Biology and Biophysics at UCONN Health.

1. REFERENCES

[1] Bettivia, R., Cheng, Y., & Gryk, M. (2022a). Documenting the Future: Navigating Provenance Standards. Springer Nature.

[2] Bettivia, R., Cheng, Y., & Gryk, M. (2022b). Storied Past, Bright Future: A Prov Jam Session. ASIS&T 2022 Annual Meeting, Pittsburgh, PA.

[3] Dell, N., Perrier, T., Kumar, N., Lee, M., Powers, R., & Borriello, G. (2015, February). Digital Workflows in Global Development Organizations. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 1659-1669).

[4] Prelinger, R. (2021, April 28). NFTs and AI Are Unsettling the Very Concept of History. Wired.

[5] Woodall, A., & Ringel, S. (2020). Blockchain archival discourse: Trust and the imaginaries of digital preservation. new media & society, 22(12), 2200-2217.

[6] Xu, K., Ottley, A., Walchshofer, C., Streit, M., Chang, R., & Wenskovitch, J. (2020, June). Survey on the analysis of user interactions and visualization provenance. In Computer Graphics Forum (Vol. 39, No. 3, pp. 757-783).