

# Understanding Methodological Diversity in Born-Digital User Communities

Dr Adam Nix

University of Birmingham

# Introduction

Born-digital archives present both challenges and opportunities for users

Existing and new users come with diverse methods and motivations

Relevant to AI-supported preservation and access needs



# EMCODIST

EMAIL CONTEXTUALISATION DISCOVERY TOOL

[HOME](#)  
[ABOUT](#)  
[CONTACT](#)



Welcome to EMCODIST, a research tool for email archives and datasets. This tool is currently in a prototype phase, and has been configured for use on a single collection of organizational emails known as the Enron Email Dataset. To begin exploring Enron through these emails, please select from the version options listed below.

**Choose a version to begin**

## EMCODIST BASIC >

- ✓ Free text search of ENRON email corpus
- ✓ AI-generated topic models
- ✓ Filter results by date or topic
- ✓ Word cloud of results

## EMCODIST PLUS >

**Everything in EMCODIST  
Basic, plus:**

- ✓ Ordering results by relevance
- ✓ Semantic matching of search terms

# User diversity and marginalisation

## User diversity can be categorised variously:

- Geographically
- Culturally
- Intersectionally
- Professionally

## Users categorised by method:\*

- Analytical approach (data vs. source)
- Knowledge motivation (empirical vs. theoretical)

\*Method: As formally or informally understood.



## Conceptualising methodological diversity among born-digital users: insights from the garbage can model

Adam Nix<sup>1</sup> · Stephanie Decker<sup>1</sup> · David A. Kirsch<sup>2</sup>

Received: 28 August 2024 / Accepted: 4 February 2025  
© The Author(s) 2025

### Abstract

The benefits of AI technologies in archival preservation are well recognised, though questions remain about their integration into existing processes. AI also shows promise for enhancing user experience and discovery in accessing born-digital materials. However, a limited understanding of the diverse methodological needs surrounding born-digital access risks the creation of one-size-fits-all solutions that suit certain approaches and research questions better than others. This article reviews current efforts in born-digital access and applies the Garbage Can Model from organisation theory to conceptualise the challenge of developing AI-based tools for multiple user types, highlighting the iterative and often decentralised nature of multi-stakeholder decision-making. We address this challenge by creating four born-digital archival user types—the aggregator, the synthesiser, the fact finder, and the narrator—each with distinct motivations and research approaches. Finally, we identify some new opportunities for stakeholders to inform how AI-based tools can be developed to better meet the variety of methodological needs that exist in relation to born-digital archives.

**Keywords** Archival users · Garbage can model · User perspective · Archival methods · Archival access

### 1 Introduction

Born-digital preservation and access have received considerable scholarly and professional attention. Indeed, recent multi-stakeholder initiatives like the Digital Preservation Coalition (DPC), Unlocking our Digital Past with Artificial Intelligence (LUSTRE), and the National Digital Stewardship Alliance in the US have mobilised a community of interest focused on archiving and accessing the digital past. Amongst the various issues tackled are the new and changing ethical and legal risks that digital sources represent (Fisher 2020; Jaillant and Caputo 2022; Zhang 2012), the scale and resource intensity of appraisal and processing tasks when faced with large digital accessions (Bunn and Higgins 2016; Hawkins 2022; Tzouganatou 2022), and the role that machine-learning technologies may have in assisting these challenges (Decker et al. 2022; Jaillant and

Rees 2022). Here, the increased availability of mainstream AI tools like Chat GPT has further enhanced the potential impact of such technologies but added new transparency and authenticity challenges for researchers and information professionals (Budhwar et al. 2023; Hannigan et al. 2024). Aligning with the sentiment that access is ultimately what is being preserved (Talboom and Underdown 2019), there has also been growing interest in understanding the current and future users of born-digital collections (Jaillant and Aske 2024; Welland 2023). In this regard, born-digital not only represents a shift in the formats represented within archives but also in how and why users engage with them.

Within this article, digital archives focusing on user experience and access can support more detailed consideration of user base and expectations remain. Of course, users of digital archives have many motivations and many respects, the user experience has traditional

✉ Adam Nix  
a.nix@bham.ac.uk

<sup>1</sup> Birmingham Business School, University of Birmingham, Birmingham, UK

<sup>2</sup> Robert H. Smith School of Business, University of Maryland, College Park, MD, USA



# Typology of Born-Digital Archival Users

Knowledge  
motivation

*Theoretical*

Aggregator

Synthesiser

*Empirical*

Factfinder

Narrator

Analytical approach

*Data*

*Source*

*A couple of caveats...*

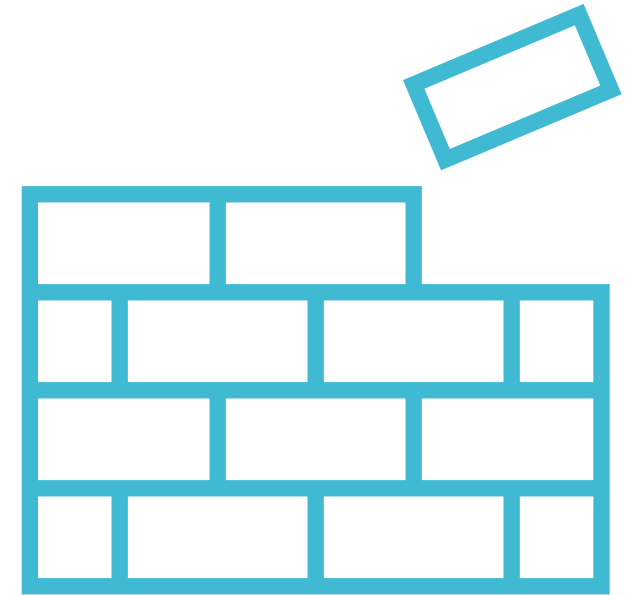


## Aggregator

- Uses data to produce generalisable, quantitatively derived insights
- Values completeness, consistent structures, and extractability
- Raw data is “rendered” or “wrangled” into an analysable dataset
- Examples: economists, corpus linguists, statisticians

# Synthesiser

- Uses sources to build thematic or exploratory insights
- Values richness and volubility over consistency or completeness
- Often seeks complex explanations with broader implications
- Examples: historians, sociologists, social commentators





## Fact Finder

- Uses data to confirm or refute a version of reality
- Values accuracy and comprehensive recall
- Seeks specific, well-defined information
- Examples: journalists, lawyers, policy researchers



# Narrator

- Seeks to elaborate on specific people, places, or events
- Values volubility and traceability.
- Often motivated by personal or public relevance
- Examples: genealogists, historians, biographers



# Implications for Access and AI Tools

How users interact with records changes based on their method and motivation



Changes their needs in terms of AI-supported preservation and access



Different types of AI are likely to suit certain users better than others

|                    | Interfaces                                  | Search                                                              | Processing                                               | Preservation                                               |
|--------------------|---------------------------------------------|---------------------------------------------------------------------|----------------------------------------------------------|------------------------------------------------------------|
| <b>Aggregator</b>  | Data dashboard, programmatic, bulk download | Inclusion/exclusion, completeness, replicability                    | Aggregation and extraction, automated dataset generation | Raw data and metadata, data volume                         |
| <b>Synthesiser</b> | Contextual overview, thematic/open querying | Semantic and contextual search, thematic relevance, recommendations | Summarising, cross-referencing, analytical tagging       | Original structure and context                             |
| <b>Fact Finder</b> | Structured querying, referential provenance | Comprehensive recall, filtering, relevance                          | Cross-referencing, context interpretation                | Completeness and auditability                              |
| <b>Narrator</b>    | Content overview, relational navigation     | Actor/event tracing, biographical reconstruction                    | Information collation, rich-form exporting               | Granular detail, multi-level records, personal information |

## Closing remarks

- Methodological plurality in born-digital usage
- One-size-fits-all solutions support some but not all
- AI offers new capacity for plural forms of access