

# **A COMMUNICOLOGICAL CRITIQUE OF EVALUATIVE NORMS FOR DIGITAL PRESERVATION SUCCESS**

**Stephen Abrams**

**BA (Mathematics)**

**ALM (History of Art and Architecture)**

Submitted in fulfilment of the requirements for the degree of  
Doctor of Philosophy

School of Information Systems  
Faculty of Science  
Queensland University of Technology  
2023



# Keywords

Affordance, Communicology, digital preservation, Expectation-Confirmation Theory (ECT), policy, Qualitative Content Analysis (QCA), semiotics, significant properties, success, trustworthiness

# Abstract

Digital preservation is critical to the successful continuity of memory in the digital age. However, current theory and practice do not encompass operational measures of that success. Nevertheless, this study reveals that important normative attitudes regarding success do permeate domain discourse. In doing so, it illuminates *why* the effective measure of success has remained elusive by identifying the conceptual limitations of those norms as actionable evaluative factors as well as proposing a new multivalent evaluative framework responding to those limitations. The study draws upon Expectation-Confirmation Theory to define success as the degree to which realized outcomes of preservation service-provider intentions satisfy purposive stakeholder expectations. Communicology Theory is used to position determinations of that satisfaction in the context of semiotically-explicable communication that unfolds across time and ever-widening technical and cultural distance. Although acts of preservation-enabled communication are technically mediated, they must be evaluated for final efficacy in terms of their human consequences.

Provider/stakeholder interactions take place, and should be assessed, in the context of the social “contract” of reciprocal commitments and reliances implicitly established by digital preservation policies. The normative positions underlying a representative sampling of policy statements are recovered through Predicate Reduction, a novel Qualitative Content Analysis technique developed specifically for this inquiry. This analysis empirically establishes four primary evaluative norms regarding imperatives for the ongoing integrity, authenticity, accessibility, and usability of preserved digital objects. While the first three are well-established archival concepts, Communicological critique underscores that they are essentially artifactual in scope and explanatory power. That is, they primarily characterize what preserved digital objects *are*, but not necessarily what they enable their consumers to *understand* or *do*. Usability, on the other hand, does directly embrace normative consideration of contingent communicative experience. However, as an evaluative concept, usability currently lacks definition sufficiently detailed to support derivation of actionable metrics. This study formalizes usability in terms of a new semiotic model of the preservation enterprise.

These findings explain the field's current evaluative emphasis on the *trustworthiness* of preservation programs, processes, and outputs rather than the *success* of consequent outcomes. The former provides necessary baseline assurances regarding the persistence of integral, authentic, and accessible information *objects*. The latter, however, is necessary for more comprehensive assessment reflecting the complementary persistence of legitimate information *experiences*. In essence, prior research regarding digital preservation assessment builds upon an underlying synthetic question: What characteristics of digital preservation agency and systems bolster confidence in their ability to perform their obligations? This study, on the other hand, is motivated by a complementary line of inquiry: What evaluative norms are indicative that those obligations have been met? The explicit change in emphasis from a predictive to confirmatory evaluative basis is accompanied by an implicit shift from a managerial to communicative perspective of the preservation imperative as well as an expansion in evaluative focus from the outputs of artifactual trustworthiness to the outcomes of experiential success. This insight augments the preservation field's theoretical and pragmatic foundations with greater appreciation for concerns and impacts regarding its proper teleological goal of purposive usability. It also suggests a path for subsequent derivation of operational criteria and metrics characterizing success in terms of the significant semiotic affordances underpinning preservation as a communicative endeavor. A workable framework for evaluation of digital preservation success is a primary component of accountable stewardship of the digital heritage necessary for future engagement with and understanding of the past.

# Table of Contents

Keywords .....	i
Abstract .....	ii
Table of Contents .....	iv
List of Figures .....	vi
List of Tables.....	vii
List of Abbreviations.....	ix
Acknowledgements .....	xii
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Preservation Imperatives.....	3
1.2 Imperative Success.....	6
1.3 Pursuing Norms of Success.....	12
1.4 Impact of Effective Evaluative Norms.....	16
1.5 Thesis Outline .....	17
<b>Chapter 2: Literature Review .....</b>	<b>19</b>
2.1 Evaluative Scope.....	20
2.2 Evaluative Evidence.....	25
2.3 Explanatory Power.....	29
2.4 Research Gap and Question .....	33
<b>Chapter 3: Research Design.....</b>	<b>37</b>
3.1 Inductive Qualitative Content Analysis .....	38
3.2 Abductive Communicological Analysis .....	56
<b>Chapter 4: Predicate Reduction Analysis.....</b>	<b>65</b>
4.1 Data Processing.....	65
4.2 Data Management .....	69
4.3 Quantitative Results .....	73
<b>Chapter 5: Communicological Analysis.....</b>	<b>81</b>
5.1 Accessibility.....	83
5.2 Integrity.....	88
5.3 Authenticity.....	91
5.4 Usability.....	94
5.5 Norms, Criteria, and Metrics.....	96
5.6 Significant Affordances .....	99
5.7 Multivalent Success .....	105

<b>Chapter 6: Conclusion</b> .....	<b>111</b>
6.1 Key Findings.....	114
6.2 Contributions .....	115
6.3 Limitations.....	120
6.4 Future Research Direction .....	122
6.5 Summary.....	125
<b>Appendices</b> .....	<b>129</b>
Appendix A Predicate Reduction Example .....	129
Appendix B Data File Derivation .....	133
<b>Bibliography</b> .....	<b>139</b>

# List of Figures

<i>Figure 1.1.</i> OAIS functional entities and actorial roles .....	7
<i>Figure 1.2.</i> Digital preservation success as a measure of state alignment .....	10
<i>Figure 2.1.</i> Digital preservation as a data management activity .....	23
<i>Figure 2.2.</i> Digital preservation as a communicative enterprise.....	35
<i>Figure 3.1.</i> Semiotic triangle.....	57
<i>Figure 3.2.</i> Grounded semiotic triangle .....	58
<i>Figure 3.3.</i> Semiotic ladder.....	59
<i>Figure 3.4.</i> Preservation-augmented semiotic ladder.....	59
<i>Figure 3.5.</i> Semiotic matrix.....	61
<i>Figure 3.6.</i> Semiotic model of digital preservation .....	62
<i>Figure 4.1.</i> Distribution of evaluative norm tokens .....	78
<i>Figure 5.1.</i> Semiotic levels of access and use.....	85
<i>Figure 5.2.</i> Semiotic alignment of InSPECT categories and normative scope.....	100
<i>Figure 5.3.</i> Significant characteristic, affordance, understanding, and response (CAUR) facets.....	101
<i>Figure 5.4.</i> Normative scope of semiotic applicability .....	104
<i>Figure 5.5.</i> Multi-dimensional evaluative space .....	108
<i>Figure 6.1.</i> Expanded digital preservation states .....	124



# List of Tables

Table 3.1 <i>Selected Digital Preservation Policy Documents</i> .....	40
Table 3.2 <i>Copula Verb Marker</i> .....	47
Table 3.3 <i>Modal Verb Marker</i> .....	48
Table 3.4 <i>Lexical Verb Marker</i> .....	48
Table 3.5 <i>Propositional Expansion</i> .....	49
Table 3.6 <i>Analytic Predicate Reduction</i> .....	50
Table 3.7 <i>Synthetic Predicate Reduction</i> .....	50
Table 3.8 <i>Predicate Canonicalization</i> .....	51
Table 3.9 <i>Sample Thesaurus Entries</i> .....	52
Table 3.10 <i>Synthetic Kernel Templates</i> .....	52
Table 3.11 <i>Kernel Construction</i> .....	53
Table 4.1 <i>Distribution of Propositional Expansion</i> .....	66
Table 4.2 <i>Single Statement Leading to Single Predicate</i> .....	66
Table 4.3 <i>Single Statement Leading to Multiple Predicates</i> .....	67
Table 4.4 <i>Multiple Statements in Single Structural Context Leading to Multiple Predicates</i> .....	68
Table 4.5 <i>Predicate Reduction Dataset Files</i> .....	69
Table 4.6 <i>Policy Contexts, Statements, Propositions, and Predicates by Document</i> .....	72
Table 4.7 <i>Policy Contexts, Statements, Propositions, and Predicates Across Dataset</i> .....	73
Table 4.8 <i>Policy Contexts, Statements, and Evaluative Kernels by Document</i> .....	74
Table 4.9 <i>Policy Contexts, Statements, and Evaluative Kernels Across Documents</i> .....	75
Table 4.10 <i>Frequency Ranking of Evaluative Norms Across and By Document</i> .....	76
Table 4.11 <i>Frequency Ranking of Primary and Total Evaluative Norms</i> .....	77
Table 4.12 <i>Relative Frequency of Primary Evaluative Norms</i> .....	78
Table 4.13 <i>Data Reduction by Evaluative Norm</i> .....	79
Table 4.14 <i>Analytic-to-Canonical Predicate Consolidation</i> .....	79
Table 4.15 <i>Example Consolidation of Analytic and Canonical Predicates</i> .....	80
Table 5.1 <i>Normative Categorization</i> .....	81
Table 5.2 <i>Communicological Accessibility</i> .....	86
Table 5.3 <i>Communicological Integrity</i> .....	91

Table 5.4 <i>Communicological Authenticity</i> .....	93
Table 5.5 <i>Communicological Usability</i> .....	96
Table 5.6 <i>Alignment of Evaluative Levels</i> .....	102
Table A.1 <i>Statement Identification Example</i> .....	129
Table A.2 <i>Propositional Expansion Example</i> .....	129
Table A.3 <i>Predicate Reduction Example</i> .....	130
Table A.4 <i>Predicate Canonicalization Example</i> .....	130
Table A.5 <i>Kernel Construction Example</i> .....	131
Table A.6 <i>Full Predicate Reduction Example</i> .....	131

# List of Abbreviations

AIAU	Accessibility-integrity-authenticity-usability, pronounced “EYE-oh” /ˈaɪ.oo/
ALCTS	Association for Library Collections and Technical Services ( <a href="http://www.ala.org/alcts/">http://www.ala.org/alcts/</a> )
ANZ	Archives New Zealand ( <a href="https://www.archives.govt.nz/">https://www.archives.govt.nz/</a> )
BMA	Baltimore Museum of Art ( <a href="https://artbma.org/">https://artbma.org/</a> )
CR	Critical Realism (Bhaskar, 1998)
CGE	<i>Cambridge Grammar of English</i> (Carter & McCarthy, 2006)
CAUR	Characteristic-affordance-understanding-response, pronounced “core” /ˈkɔr/
CUL	Cambridge University Libraries ( <a href="https://www.lib.cam.ac.uk/">https://www.lib.cam.ac.uk/</a> )
DA	Data archive (Wright et al., 2018)
DACS	Describing Archives: A Content Standard (SAA, 2013)
DIKW	Data-information-knowledge-wisdom (Baskarada & Kononios, 2013)
ETC	Emergent thematic coding (Stemler, 2001)
ECT	Expectation-Confirmation Theory (Hossain & Quaddus, 2012)
IAAU	Integrity-authenticity-accessibility-usability, pronounced “EE-oh” /ˈi:.oo/
ICPSR	Inter-University Consortium for Political and Social Research ( <a href="https://www.icpsr.umich.edu/">https://www.icpsr.umich.edu/</a> )
InSPECT	Investigating Significant Properties of Electronic Content (Knight, 2009)
IR	Institutional repository (Li & Banach, 2011)
ISO	International Organization for Standardization ( <a href="https://www.iso.org/">https://www.iso.org/</a> )
LAM	Libraries, archives, and museums
LIS	Library and information science (Hjørland 2018)
LISA	Library and Information Science Abstracts

(<https://search.proquest.com/libraryscience?accountid=13380>)

- LISTA Library and Information Science and Technology Abstracts  
(<https://www.ebsco.com/products/research-databases/library-information-science-and-technology-abstracts>)
- MARC Machine-Readable Cataloging (Furrie, 2009)
- NA Nationaal Archief [National Archives of the Netherlands]  
(<https://www.nationaalarchief.nl/>)
- nestor<sup>1</sup> Network of Expertise in long-term STORage  
(<https://www.langzeitarchivierung.de/Webs/nestor/>)
- NLA National Library of Australia (<https://www.nla.gov.au/>)
- NLNZ National Library of New Zealand (<https://natlib.govt.nz/>)
- NLP Natural language processing (Friedman et al., 2013)
- NSM Normative-Semiotic-Modal
- OAIS Open Archival Information System (ISO, 2012b)
- OESPSPP Ontics-empirics-syntactic-performics-semantics-plaistics-pragmatics
- PDF Portable Document Format (ISO, 2017)
- PI Philosophical Inquiry (Andow, 2016)
- PKI Public Key Infrastructure (Adams & Lloyd, 2003)
- PO Preservation Objective (CCSDS, 2019)
- PoS Parts-of-speech (Tufiş & Ion, 2017)
- PR Predicate reduction (Abrams, 2021)
- PREMIS Preservation Metadata: Implementation Strategies (LC, 2015)
- QCA Qualitative Content Analysis (Krippendorff, 2019)
- QoS Quality of service (Happe et al., 2011)
- RQ Research question

---

<sup>1</sup> The preferred presentational form of the acronym for the German Network of Expertise in Long-term Storage of Digital Resources is the all-lower-case “nestor”.

SCAPE	Scalable Preservation Environments (Edelstein et al., 2011)
SLA	Service level agreement (Ahmad & Abawajy, 2014)
TDR	Trustworthy digital repository (ISO, 2012a)
TIB	Technische Informationsbibliothek [Leibniz Information Centre for Science and Technology] ( <a href="https://www.tib.eu/en/">https://www.tib.eu/en/</a> )
TRAC	Trustworthy Repository Audit & Certification (Dale & Ambacher, 2007)
UNESCO	United Nations Educational, Scientific, and Cultural Organization ( <a href="https://unesco.org/">https://unesco.org/</a> )
USB	Universal Serial Bus (IEC, 2015)
ZB MED	Informationszentrum Lebenswissenschaften [Information Centre for Life Sciences] ( <a href="https://www.zbmed.de/en/">https://www.zbmed.de/en/</a> )
ZBW	Leibniz-Informationszentrum Wirtschaft [Leibniz Information Centre for Economics] ( <a href="https://www.zbw.eu/en/home">https://www.zbw.eu/en/home</a> )

# Acknowledgements

I cannot fully express my appreciation and gratitude towards my supervisors, Drs. Sylvia Edwards, Peter Bruza, Anthony Bernier, and Kate Devitt. They have provided steadfast support, direction, and encouragement, and that enabled me to approach – and complete – this work with rigor, concentration, and a great deal of enjoyment. I would also like to acknowledge Drs. Marcia Bates, Michael Buckland, and Clifford Lynch, who provided vital enthusiasm for this research at its incipient stage. They provoked me to refine and clarify many core questions, concepts, and approaches quite important to the final work. This work was pursued through the collaborative Queensland University of Technology/San José State University Gateway PhD program. This unique and innovative program provided a marvelously nurturing environment for my studies. It is difficult to imagine being able to have begun, let alone complete, my study in the manner I did anywhere else. In particular, I'd like to thank Confirmation Panel members Drs. Michelle Chen and Bill Fisher for important feedback at a critical juncture; Final Seminar panelist Dr. Guy Gable; my cohort members Lee Yen Han and Pat Sandercock as well as fellow candidates Jennine Knight and Kathleen McDonald for unstinting encouragement and commiseration; and Gateway faculty and alumni Drs. Salvador Barragan, Mary Bolin, Christine Bruce, Walter Butler, Lettie Conrad, Andrew Demasson, Africa Hands, Deborah Hicks, Sandra Hirsch, Darra Hofman, Geoffrey Liu, Kim Morrison, Richard Okumoto, Cheryl Stenström, Ian Stoodley, Virginia Tucker, and Michele Villagran for feedback and comradery. Throughout the research process I was greatly stimulated and encouraged by generous discussions with Leslie Johnston, Tricia Patterson, Raivo Ruusalepp, Barbara Sierman, and Drs. Christoph Becker, Rhiannon Bettivia, José Borbinha, Christine Borgman, Michèle Cloonan, Devan Donaldson, Christopher Lee, Jerome McDonough, Seamus Ross, and Abby Smith Rumsey. I would also like to thank the anonymous *Archival Science*, *ASIS&T*, *DigiPres*, *iPRES*, and *JCDL* reviewers for perceptive comments and suggestions on earlier presentations of this work. I am grateful to the Beta Phi Mu honor society for awarding me a Eugene Garfield Doctoral Dissertation Fellowship midway through my candidature in 2021.

Dedicated to my father, Dr. Joel Abrams, whose life and career embodied his bedrock belief in the transformative power of education. His immediate and unwavering encouragement was critical to my undertaking this journey to begin with and a constant balm throughout. I am sorry he is not here to see its completion, but would like to believe that he would have been proud.

Finally, Paula. My first, last, and best reader, editor, consoler, and cajoler. It was not possible without her.





# Chapter 1: Introduction

---

The successful persistence of personal, organizational, and social memory in the digital age depends upon ready recourse to the diverse digital record documenting past and present time (Abrams, 2021; Hedstrom & King, 2004; Rumsey, 2016; Rydén, 2019; van der Werf & van der Werf, 2019). However, future engagement with that corpus is inherently fragile with respect to the passage of time. The rapid evolution and often-disruptive obsolescence of the technical systems that retain, read, and render digital information objects pose a significant threat to future meaningful use of those objects. Similarly, the ever-widening distance separating the culturally-situated points of creation and consumption exacerbates the potential for loss of critical associational context necessary for legitimate interpretative understanding. Curatorial stewards at libraries, archives, museums, and other memory institutions have long accepted preservation responsibility for addressing challenges to the persistence of the cultural record. However, pursuit of that goal requires significant sustained administrative commitment, technical expertise, and financial support. In view of these consequential calls on finite institutional resources, it is important for institutional stewards to be confident in knowing whether or not their efforts have been, or are likely to be, successful. In the digital realm, however, the characterization of preservation success still remains “a metric that’s defied measuring” (Lynch, 2006; as cited in Lee & Tibbo, 2007). While accepted theory and practice in the digital preservation field promote or imply evaluative criteria for various aspects of activity, these do not currently encompass operational metrics for success. Consequently, this research pursues better understanding of *why* the derivation and use of measurable metrics has remained problematic, and suggests potential approaches responding to those problems. It approaches this question through the identification and analysis of evaluative attitudinal principles pervasive in domain discourse. Greater clarity regarding the scope and limitations of these norms will facilitate future efforts to augment existing evaluative practices with new complementary concepts and methods effectively characterizing digital preservation success.

This pursuit is structured by a research program looking into the parameters of an evaluative framework necessary for effectively and comprehensively determining

the communicative success of digital preservation activity. This defines the three-part research inquiry underlying this dissertation: first identifying existing socially-constituted norms regarding success; then, assessing their suitability for the task; and finally, proposing ways to enhance current theory and practice to address any shortcomings. Current evaluative norms are identified by subjecting policies to Predicate Reduction (PR), a novel variation of Qualitative Content Analysis (QCA) newly developed for this study (Abrams, 2021). This analysis establishes that commonly-accepted evaluative norms are most often expressed in terms of general precepts rather than specific criteria, let alone obtainable measures. Furthermore, these precepts refer to established archival principles reflecting objective *artifactual* rather than intersubjective *experiential* concerns. That is, they characterize various technical aspects of managed information objects themselves in isolation from any subsequent human engagement *with* those objects (Abrams, 2021). Evaluative parameters pertinent to experiential *usability*, recognized as the proper teleological goal of the preservation enterprise (Conway, 2010; Giaretta, 2011; Gladney, 2006; Menne-Haritz, 2001; Strodl et al., 2007; Traczyk, 2017; Walters & Skinner, 2011), remain underdefined in theory and practice (Dearborn & Meister, 2017; Poole, 2015). Consequently, communicative *success*, a primary characterizing quality of use, is similarly not currently susceptible to effective measure.

One contributing factor for this limitation is insufficiently expansive conceptualization of the preservation field, which remains centered on managerial activity and agency (Abrams, 2018a, 2018b, 2021). The primary unit of managerial attention is conventionally termed a *digital object*, the digital encapsulation of a coherent assemblage of abstract intellectual, affective, and behavioral content (Faulkner & Runde, 2019; Kallinikos et al., 2010). The trustworthiness of managerial processes for objects is an important evaluative factor (Donaldson, 2020; Giaretta, 2011), especially as it can provide *predictive* assurance of subsequent preservation efficacy. However, more complete and compelling evidence of preservation success depends upon some means of *confirmatory* characterization. This research establishes the scope of existing evaluative principles and illuminates why they are insufficient for providing that degree of explanatory power necessary for meaningfully characterizing digital preservation success. Extending conceptual perspective of the preservation enterprise from intermediating managerial *means* towards final

communicative *ends* offers a firmer basis for comprehensive consideration and evaluation of the ultimate success of preservation activity.

## 1.1 PRESERVATION IMPERATIVES

The adoption of information technology and electronic resources in commerce, culture, science, education, and entertainment has burgeoned worldwide since the 1980s (Fox, 2002; Kellerman, 2000; Knezek & Christensen, 2001; Ng, 2012). Most, if not all, critical functions of modern life are now thoroughly reliant upon digital content. This ever-growing technical dependence has raised concerns about the need for preservation solutions to counteract the potential for a “digital dark age” in which significant digital heritage content is subject to irretrievable corruption or loss due to technical obsolescence, malicious attack, shifting institutional mission, or insufficient managerial planning, attention, or response (Bollacker, 2010; Brand, 1999; Jeffrey, 2012; Smit et al., 2011; Whitt, 2017). The extent and severity of these threats may be less than imagined (Anderson, 2015; Johnston, 2020a). However, this favorable perspective assumes widespread availability and adoption of a robust and mature set of policies, procedures, and technologies along with a sustained programmatic commitment to address these risks on an ongoing basis.

The scope and range of these ameliorating factors have emerged through significant research and practice over the past quarter century; see for example, (CLIR, 2002; Corrado & Moulaison Sandy, 2017; Owens, 2018; Traczyk, 2017; Waller & Sharpe, 2006; Waters & Garrett, 1996). The primary goals of these efforts include risk management and mitigation (Barateiro et al., 2010; Frank, 2020); increased trustworthiness in managerial programs and systems (Giarretta, 2011); and the resulting integrity, authenticity, accessibility, usability, understandability, and reliability of the digital collections managed in those system by those programs (Burda & Teuteberg, 2013). Pursuit of these goals is complicated by the fact that future purposive use of preserved digital content often occurs in a manner that was not intended or anticipated at the time of its creation or acquisition (Galloway, 2004). In particular, the epistemic experience of, and phenomenological response to, a preserved digital object depends upon the contingent information needs and goals of that object’s human consumer, who is always positioned in a specific cultural as well as technological time and place.

The outcomes of responsible preservation oversight and intervention can take

many forms. For example, a curatorial steward could respond to a consumer's request for a preserved digital object by variously providing:

- The original physical media hosting that object, for example, a magnetic tape
- A piece of contemporary storage media hosting the object, e.g., a USB flash drive
- An individual file manifesting the object, but about which nothing is otherwise known; in other words, an opaque bitstream
- The file in its original known format, e.g., WordPerfect
- A derivative file in another known format, e.g., PDF
- The file accompanied with software capable of rendering it, e.g., Acrobat Reader
- The file and documentation of its provenance and change history, e.g., PREMIS event metadata (LC, 2015)
- The file and an authoritative token of its authenticity, e.g., a verifiable PKI digital signature (Adams & Lloyd, 2003)
- The file and accompanying intellectual description, e.g., a MARC catalog record (Furrie, 2009)
- The file and documentation of the context of its production, e.g., a methodology statement
- The file and documentation of its curatorial context, e.g., an archival DACS finding aid (SAA, 2013)
- The file and documentation of the context of its prior interpretation, e.g., an article citing the object

and so on (Abrams, 2018a). At what point in this spectrum of responses can one plausibly – if not confidently – assert that the result of preservation activity was successful? Without knowing, how can practitioners and stakeholders rationally plan for, reasonably expect, effectively measure, or be held meaningfully accountable for that result?

The question of success cannot be addressed simply by consideration of the formal characteristics of a preserved digital object itself. These need to be accompanied by a sense of the *intent* of the request for that object and the *purpose* towards which it is applied. Every individual use of a preserved digital object is always situated with respect to the potentially unique context of a particular time, place, person, and purpose (Ball, 2010; Bishop & Hank, 2018; Dearborn & Meister, 2017; Morrissey, 2014). Thus, success for one may very well be failure for another. The underlying intellectual meaning or aesthetic import legitimately attributable to an information object is co-constructed by productive and transmissive acts of all participants in the creative process (Boutard, 2016). Similarly, human understanding of that object arises from the complex intersubjective interplay of meanings that inhere *in* the fabric of the object, that adhere *to* it through context, and that ultimately cohere *about* it in the mind of the interpreting consumer (Buckland, 2013; Fornäs, 2017). In the digital realm, these shifting meanings emerge through contingent computational (re)performance of the object (Becker, 2018; Tredinnick, 2008). In other words, the interpretive response to a purposive transmission of meaning is enacted through a fluid situational process. Thus, it is overly reductive to assume that well-managed digital objects, even if possessing critical qualities of artifactual integrity, authenticity and accessibility, *necessarily* ensure ultimate preservation efficacy from the consumer perspective.

Assurances regarding those three archival qualities form the basis for effective digital preservation *management*, that is, custodial oversight and intervention regarding managed *objects*. However, the ultimate goal of that management is not just persistence of digital objects across time, but also persistence of the *usability* of those objects and the legitimate human experience and understanding of them (Day et al., 2018; Duranti & Thibodeau, 2006; Sacchi, 2015). Thus, the proper teleological imperative of digital preservation activity is not only managerial, but also *communicative*. Despite the centrality of technological intermediation in the digital age, preservation-enabled communication ultimately entails a future human encounter with past informative expression leading to a human response (Belkin, 2005; Rogala & Bialowas, 2016). That preservation is successful if the response is meaningful. Meaningfulness arises if something pertinent to the human user's intended – or serendipitous – purpose is satisfied. That is, is something new is intellectually

understood, emotionally felt, or physically acted upon by that user in a manner that would not otherwise have occurred (Ketelaar, 2012; Kuhlthau, 2017; Savolainen, 2019).

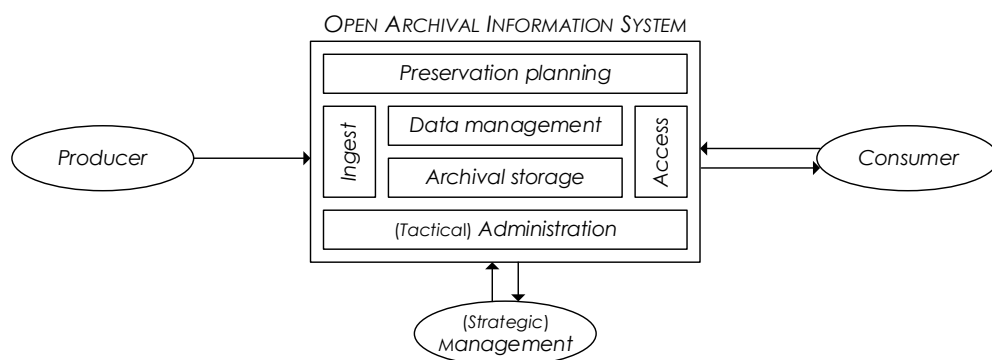
However, the accepted benchmark for evaluating digital preservation activity in contemporary theory and practice remains the managerial trustworthiness of preservation systems and institutional programs (Donaldson, 2016; Maemura et al., 2017). Any assertion of the trustworthiness of those systems and programs follows from a justified belief that they are capable of meeting their obligations (Dryden, 2011). Trustworthiness, especially as codified in the ISO 16363 *Audit and Certification of Trustworthy Digital Repositories* (TDR) standard (Bountouri et al., 2018; ISO, 2012a), is a useful measure for evaluating the efficacy of preservation management (Yoon, 2014). However, it does not provide similar illumination regarding consequent user engagement with managed objects (Ross, 2006). Just as *usability* is the primary characterizing imperative of the preservation enterprise (Conway, 2010; Menne-Haritz, 2001; Strodl et al., 2007; Traczyk, 2017; Walters & Skinner, 2011), *success* is the primary characterizing quality of that use. A determination of success indicates that the encounter with the object satisfied the purposive intent underlying that encounter. Assessment of the technical and institutional characteristics of trustworthy digital object management provides a necessary evaluative foundation. However, it is not sufficient for determining whether preservation's communicative goal has or has not been satisfactorily met. Existing evaluative metrics concerned with quantifying managerially-trustworthy outputs need to be complemented with those qualifying the experiential epistemic and phenomenological outcomes indicative of successful human use of preserved digital material.

## **1.2 IMPERATIVE SUCCESS**

What is meant by digital preservation success? Digital preservation is a highly specialized activity most often performed in the context of a service-provider/stakeholder relationship, whether internal to or across institutional boundaries (Lavoie & Dempsey, 2004; Waters & Garrett, 1996). Many libraries, archives, and museums have established special-purpose digital programs for dealing with their own preservation needs; see for example (Bermès & Fauduet, 2011; Kirchhoff, 2008; Ravenwood et al., 2015). Various non-profit and commercial organizations also offer

membership- or fee-based preservation services for those without the capacity or desire to implement in-house solutions; see for example (*Partners for Preservation*, 2019; Altman et al., 2009). The pertinent characteristic of these organizational arrangements is the explicit division between curatorial and operational responsibility. Stewarding curators provide primary intellectual, strategic, and policy oversight while service-providers contribute technical expertise, capacity, and operational control. The activities of both of these groups, however, are directed towards satisfying the needs and goals of a third: the consuming stakeholders who affirmatively seek out or serendipitously discover preserved content of interest. The level of satisfaction engendered by such a provider/stakeholder relationship is measured by the degree of alignment between actorial aspirations and resulting outcomes (Mason & Simmons, 2012). That is, satisfaction is predicated on the tangible realization of a provider's intentions in a manner that fulfils stakeholder expectations (Liao et al., 2007; Oliver & Burke, 1999). A provider intention refers to an affirmative decision by that provider to perform some future stakeholder-facing behavior (Smith, 2017; Söderlund & Öhman, 2005). A stakeholder expectation is a predictive belief that the provider will in fact perform that behavior (Almsalam, 2014; McKinney et al., 2002).

The provider/stakeholder relationship holds a central position in digital preservation theory and practice. The ISO 14721 *Open Archival Information System (OAIS) Reference Model* (ISO, 2012b) is widely accepted as the controlling framework for theoretical analysis and pragmatic design and operation of preservation activity (Brunsmann et al., 2012; Xie & Matusiak, 2016). The OAIS model, which encompasses institutional programs as well as technical systems, codifies three primary preservation roles: producers, managers, and consumers (see *Figure 1.1*).



*Figure 1.1.* OAIS functional entities and actorial roles

Adapted from (ISO, 2012b)

These groups respectively provide, preserve, and request/retrieve the digital content hosted by an OAIS. The OAIS model draws a distinction between managers proper, concerned with high-level policy, governance, and oversight (i.e., strategic management), and administrators, focused on operational responsibilities (i.e., tactical management). While these roles encompass differential levels of concern and practice, that difference is one of degree rather than kind. Fundamentally, they are both intermediaries, holding and acting upon *delegated* stewardship responsibilities on behalf of other stakeholders. In doing so, they are markedly distinguished from the very different originating intellectual concerns of content producers and exploitative concerns of consumers. Thus, the OAIS managerial/administrative division is not pertinent to this investigation, and both are subsequently subsumed under the single broad concept of actorial “management,” the institutional or programmatic role intermediating between producers and consumers.

Preservation success is dependent upon the alignment of the aspirational positions of its participants. Thus, evaluative determinations of success would be simplified if explicit expressions of managerial intention and consumer expectation were readily available. These could be provided, for example, in the form of the preservation intention statements proposed by the National Library of Australia (Webb et al., 2013). Unfortunately, this documentary form has not received widespread adoption. Search of both domain-specific and general-purpose scholarly abstracting and indexing services – ProQuest LISA,<sup>2</sup> EBSCO LISTA,<sup>3</sup> and Google Scholar<sup>4</sup> – returns no substantive references to intention statements other than citations to the original NLA publication and examples of internal NLA use. Fortunately, other avenues for understanding aspirational positions are available. The intentions and expectations attributable to content producers, managers, and consumers are defined indirectly by policy statements promulgated by preservation service providers (Beagrie et al., 2008; Dressler, 2017; Innocenti et al., 2010; Noonan, 2014). In a provider/stakeholder context, these statements bind the participants together in terms of a governing psychological and social, if not legal, service “contract” (Jeong et al.,

---

<sup>2</sup> Library and Information Science Abstracts, <https://proquest.libguides.com/lisa>

<sup>3</sup> Library, Information Science and Technology Abstracts, <https://www.ebsco.com/products/research-databases/library-information-science-and-technology-abstracts>

<sup>4</sup> <https://scholar.google.com/>



2018). These policies may be articulated in the form of specific service level agreements (SLAs) or general value propositions. An SLA is a formal commitment regarding the parameters of expected service activity between a provider and stakeholder (Happe et al., 2011), often with associated metrics for Quality of Service (QoS) (Ahmad & Abawajy, 2014). A value proposition is a more informal expression of beneficial services, products, and results offered to stakeholders by providers (Kaplan & Norton, 2004). In either case, policy statements express, either explicitly or tacitly, the set of intentional programmatic obligations publicly accepted by service-providers. These can be represented schematically as (Abrams, 2021):

*“Provider P will perform activity A to ensure condition C for stakeholder S.”*

In view of such a published commitment, it is rational for stakeholders to hold realistic complementary assumptions of the form:

*“Stakeholder S expects provider P to perform activity A to ensure condition C.”*

These intentional and expectational positions suggest a natural benchmark metric of digital preservation success. Since the fundamental outcome of digital preservation is future stakeholder use of preserved material, the measure of the success of that use is the degree to which the stakeholder is satisfied with the provider. In other words,

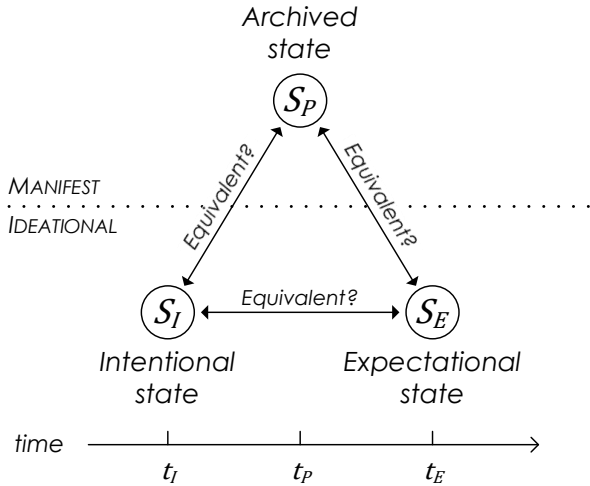
*“Did provider P perform activity A to ensure condition C for stakeholder S?”*

In terms of Expectation-Confirmation Theory (ECT) (Bhattacharjee, 2001; Kim, 2012), service satisfaction derives from confirmation of consumer expectations regarding perceived utility (Hossain & Quaddus, 2012) as well as perceived benefit (Mamun et al., 2020). Utility refers to the degree to which the experience resulting from the provider/stakeholder engagement is pertinent to the consumer’s contextual purpose; and benefit, the degree to which that experience essentially fulfils the needs, goals, and aspirations underlying that purpose. The human experience of digital preservation outcomes is similarly conditioned by assessments regarding utility and benefit. Consequently, this study proposes an ECT-informed metric of digital preservation success defined in terms of the mutual equivalence relations between

discrete artifactual and attitudinal states central to preservation-enabled managerial/consumer interactions (see *Figure 1.2*).

The state of a digital object is a unique point-in-time configuration of the values of its characterizing properties (Weber, 2012). The three states critical to consideration of digital preservation success are:

1. The *intended* state  $S_I$  of a preserved object as committed to by the responsible service-provider at some time  $t_I$ ;
2. The *archival* state  $S_P$  of the object resulting from the provider’s realized intentions at time  $t_P > t_I$ ; and
3. The *expected* state  $S_E$  of that object as anticipated by a stakeholder in light of a particular situated context and affirmative or serendipitous purpose at the time of retrieval request  $t_E > t_P$ .



*Figure 1.2.* Digital preservation success as a measure of state alignment

The measure of satisfaction is characterized by the extent to which the intentional, archival, and expectational states are mutually equivalent. Additional states can be defined to characterize other important aspects of digital object production, curatorial acquisition, and post-retrieval consumption (see *Figure 6.1*). However, incorporating those states into similar analytical consideration is left to future activity (see Section § 6.4.3). This study concentrates on the core intentional, expectational, and archival states whose equivalence relationships lie at the center of the evaluation of digital preservation satisfaction and success.

Any attempt to determine state equivalence presupposes a viable means for representing state values in a manner amenable to comparison. However, because the three core states occupy distinct ontological positions, the nature of those representations and the modality of their recovery will also be distinct. The intentional and expectational states are fundamentally ideational. That is, they exist primarily as hypothesized aspirations within the imaginations of managers and consumers. This is not to say that a description of these states cannot be expressed in concrete form. However, the existence of that description does not affect the ontological status of the actual state being described, which remains an intangible mental position. Archival state, on the other hand, is explicitly manifest. That is, tangibly instantiated in terms of physical bits on a storage medium. Ostensibly, an object's value state can be evaluated in terms of its *significant properties*, the set of attributes that define and characterize the object's essential nature and whose invariance over time constitutes an important preservation imperative (Giaretta et al., 2009; Hedstrom & Lee, 2002; Hockx-Yu & Knight, 2008). For the manifest archival state, these property values can be established directly through an understanding of the object's process of creation or acquisition in conjunction with interrogation of its preserved physical manifestation. The ideational states of managerial intention and consumer expectation, on the other hand, must be approached more indirectly. As discussed previously, a sense of controlling intentions and expectations broadly accepted by the digital preservation community can be identified from obligatory and aspirational attitudes explicitly referenced or tacitly implied in published policy statements.

While the concept of significant properties appears to provide a useful structure for taxonomizing potential state-characterizing values, the concept has proved difficult to put into practice. The idea that significance can be reductively fixed in an objective manner is illusory (Yeo, 2010). It arises from an inappropriate assumption that applicable properties are those attendant to a digital object as a standalone artifact independent of the subjective context of its use (Becker, 2018). A better sense of attributes capable of characterizing the behavioral dimension of usage is captured by the psychological notion of *affordance* (Hedstrom & Lee, 2002). An affordance is a factor within a system or environment that enables the possibility of human action or response (Cheikh-Ammar, 2018; Withagen et al., 2012). Conceptually extending the concept of significant properties to encompass significant affordances emphasizes that

the function of those attributes is applicable on an experiential as well as artifactual level (Abrams, 2018b). That is, affordances provide a lens for understanding not only what a preserved digital object *is*, but also what that object permits one to *do* and subsequently *know*. The human-experienced quality of that doing and knowing properly underlies the measure of digital preservation success.

### 1.3 PURSUING NORMS OF SUCCESS

This study pursues initial progress towards the future development of measurable metrics of digital preservation success through better understanding of *why* such metrics have eluded meaningful definition and operational application to date. The various risks potentially impeding that success arise in an intersubjective as well as a nominally-objective technical context (Frank, 2020). The actions and perceptions leading to a determination of communicative success are socially contingent as well. Thus, putative evaluative norms for success emerge as social constructions in terms of attitudinal positions embedded in the consensual social fabric of domain discourse. In view of these foundational perspectives, this investigation begins by establishing and critiquing evaluative principles and criteria accepted across the preservation community. The results of that critique are then used to suggest meaningful complementary enhancements to current evaluative theory and practice. The relevant discursive sources for this study are digital preservation policy statements that, as explicated by Expectation-Confirmation Theory, tacitly establish the controlling intentional obligations and expectational aspirations underlying service-provider/stakeholder interactions. These, in turn, are determinants of consequent service satisfaction or success.

This investigation proceeds from a metaphysical position of Critical Realism (CR). This perspective assumes a fundamentally realist ontology but interpretivist epistemology (Bhaskar, 1998; Mingers et al., 2013). In other words, it posits that the “real” world exists objectively independent of our sense or thought, yet is knowable to us only through our subjectively-situated perception and cognition (Danermark et al., 2019). That knowledge is therefore contingent and inherently fallible, although we have the capacity to recognize and distinguish between better and poorer explanation (Radulescu & Vessey, 2008). The former arises through critical, theoretically-sound, and well-structured conceptual abstraction and inferential interpretation of the phenomena of which we become aware (Reed, 2009). The intellectual form of that

inferencing is ultimately abductive, rather than deductive or inductive in nature (Overton, 2012). That is, the logical truth-standard underlying its claims tends towards the best-possible, rather than the causally-necessary or probabilistically-most-likely explanation (Reichertz, 2014). This epistemological position is consistent with the overall pragmatic perspective of this study.

The pragmatic research paradigm is an alternative to the extremes of experimental positivist and ethnographic constructivist approaches (Creswell, 2014b; Morgan, 2007). It relies upon a methodological eclecticism similar to that deployed in mixed methods research (Feilzer, 2010). This imparts a freedom to rely upon various investigatory techniques and strategies based on their fitness for research purpose (Teddle & Tashakkori, 2012) as well as their exploratory and confirmatory power (Onwuegbuzie & Leech, 2005). Thus, this study entails both initial inductive Qualitative Content Analysis (QCA) to identify current parameters of evaluative practice regarding preservation activity and subsequent Philosophical Inquiry (PI) to establish the suitability of those parameters as effective norms for preservation success. QCA provides methods for systematically ascertaining the meaning of textual content (White & Marsh, 2006) and is particularly useful for uncovering latent meanings underlying a text's manifest form (Schreier, 2013). The subjective undertones of QCA can raise legitimate concerns regarding the validity of analytic interpretation (Maier, 2018). However, a formal research method relying upon sound reasoning and rigorous adherence to a well-defined analytic process provides confidence in the reliability and replicability of results (Krippendorff, 2019). Predicate Reduction (PR), a novel variant of QCA, was newly developed for this research program (Abrams, 2021). As described in Chapter 3, PR defines a series of iterative textual transformations that systematically reduce narrative policy terms into unitary propositional form, concise predicates expressing core intentional/expectational imperatives, and finally, implied evaluative norms. These norms are then critiqued in terms of an open-ended Philosophical Inquiry.

PI seeks to understand and enhance the conceptual structures that provide meaning to experience (Burbules & Warnick, 2006; Grace & Perry, 2013). That understanding follows from abductive questioning of fundamental domain assumptions and conceptual definitions to derive new, more comprehensive explanatory structures (Andow, 2016; Pesut & Johnson, 2008; Sheffield, 2004). The

dimensions of critical scrutiny underlying this stage of inquiry stem from tripartite semiotic concerns. These encompass investigation of the core processes through which communication occurs, the expressive sign vehicles underlying those processes, and the embodied experience of domain actors engaging with those vehicles through those processes (Eicher-Catt & Catt, 2008; Lanigan, 2010a). The domain in question here is that of preservation-enabled communication of digital information across time, while the explanatory concern of the PI is the efficacy of current evaluative practice with regard to the communicative success of the preservation effort.

These metaphysical and methodological approaches are manifest throughout the research program, particularly regarding core conceptual abstractions and critical methods. The repositioning of digital preservation as a communicative enterprise follows from CR's explication of how fundamental interpretive processes intersubjectively mediate between the world as it is and the world as we can know it. Preserved digital objects are contingent phenomenal representations of some slice of the ontologically-transcendent world, with which we have no otherwise direct access (Danermark et al., 2019; Reed, 2009). Thus, preservation concerns should embrace not only managerial custodianship of those objects as stand-alone representational vehicles, but also the relational – and therefore communicative – processes by which we attempt to exploit those objects to engage with and understand the world. In the context of preservation-enabled communication, the locus of CR meaning-making is the interpretive experience enacted through the service-provider/stakeholder relationship. In consequence, the teleological imperative for the preservation enterprise is assurance of the purposive usability of the digital artifacts underlying that relationship. Evidence of aspirational evaluative attitudes germane to that relationship comes from discursive artifacts – preservation policy statements – that are leavened with socially-constructive traces of pertinent domain norms. Once established through inductive QCA of a representative set of policy documents, those norms are subject to PI-based critique to determine their suitability for evaluative purposes.

Historically, the digital preservation field has been largely preoccupied with practical and methodological concerns rather than theoretical constructs (Flouris & Meghini, 2007; Ross, 2012). There is little inquiry into foundational theory (Flouris & Meghini, 2007; Xie & Matusiak, 2016) and expanded funding support is needed to support new research and promotion of new theoretical models (NDSA, 2014). The

term “theory” is often deployed in the literature in a somewhat restricted sense of a newly posited thesis; see, for example, (Moore, 2008; Owens, 2018; Watry, 2007). “Theory” also carries a more expansive sense of a coherent system of intellectual abstraction, inference, and explanation (Gregor, 2017). However, the contexts in which these more inclusive references occur are generally based on reductive logical and mathematical formalisms (Abrams, 2018b). That is, they rely upon a tacit underlying assumption that preserved digital objects completely encapsulate the knowledge-states and intentions of their creators and that those states are capable of being unambiguously (re)presented to, and (re)experienced by, future consumers, see, for example, (Cheney et al., 2001; Flouris & Meghini, 2007; Giaretta et al., 2011). This position conflicts with the post-modernist tenet regarding the inherently contingent nature of all human exchange of information (Cook, 2001; Hansson, 2005; Tan et al., 2009). That contingency implies that any future use of preserved information will always be contextually-situated with regard to a specific time, place, and purpose of use, and cannot be reductively generalized (Anderson & Colvin, 2003). Given a prevalent view of digital preservation enterprise as enabling digitally-mediated “*communication with the future [emphasis added]*” (Brocks et al., 2010, p. 197; Mois et al., 2009, p. 1; Moore, 2008, p. 64); see also (Bell & Grey, 2001; Caon, 2018; Thibodeau, 2002), this study examines the evaluative success of that enterprise through the lens of Communicology.

Communicology is the study of embodied human discourse (Eicher-Catt & Catt, 2008; Lanigan, 2013). It conceives of that discourse as a semiotic system in which the meaning of expressive signs emerges through contingent interpretation by individuals in the purposive context of their own lived experience as well as institutional and cultural positioning. This semiotic foundation is an appropriate theoretical basis for investigation into the representation, acquisition, and mediated transmission of information (Mingers & Willcocks, 2017; Pai, 2016). It provides an analytic toolbox explicitly cognizant of the inherently contextual and contingent nature of preservation-enabled human communication. The findings resulting from Communicological analysis provide new insight into why effective measurement of digital preservation success has remained problematic to date. It also suggests a promising path forward for the development of a new, more comprehensive procedural framework for characterizing success. Once developed, that framework and its underlying theoretical

and analytical apparatus will provide the digital preservation community with a better means for conceiving, implementing, and assessing the efficacy of its critical activities.

#### **1.4 IMPACT OF EFFECTIVE EVALUATIVE NORMS**

This work illuminates the limited scope and explanatory power of current managerial- and artifactual-centric evaluative practices for characterizing digital preservation efficacy. Those extant practices coalesce around determinations of the trustworthiness of preservation managers and management. While this is an important foundational metric, it is insufficient to encompass the ultimate success of preservation's communicative imperative of enabling future purposive human use of preserved digital objects (Abrams, 2021). The subsequent formalization of experiential success as the degree of relative alignment of intended, expected, and realized object states provides a principled framework for future development of more comprehensive and conceptually-sound principles, criteria, and operational metrics in a rigorous and compelling manner. When available, these should prove beneficial as benchmark measures through which scholars can gain greater insight into foundational imperatives and aspirations of the field. Similarly, practitioners will be able to approach their programmatic mission more responsibly, allocate finite programmatic objects more productively, and be held accountable to stakeholders more effectively.

This research's positioning of service-provider intention and stakeholder expectation at the center of a newly-formalized definition for digital preservation success led to the identification of preservation policy statements as viable sources for establishing those attitudinal positions. This in turn spurred the development of the Predicate Reduction technique for recovering pertinent evaluative attitudes from their often-tacit expression in those policies. The PR technique can be repurposed in future for reliable unobtrusive recovery of attitudinal positions embedded in other discursive forms, genres, and domains. In the digital preservation context, the attitudes and associated principles established through the PR process are found through Communicological critique to be insufficient for evaluating the success of preservation activity. That activity is essentially communicative and experiential, rather than managerial and artifactual, in teleological purpose. Extant evaluative metrics of the preservation enterprise provide insight into, and confidence about, the trustworthiness of institutional processes leading to the persistence of authentic digital information objects. However, they are inadequate to provide complementary characterization of



the successful persistence of communicative opportunities for legitimate information experiences.

The conceptual shift of primary evaluative consideration away from the managerial artifact towards the communicative experience of that artifact is consistent with the theoretical position that digital artifacts are not actually susceptible to preservation, but only the computational (re)performance of those artifacts (Becker, 2018; Ross, 2012; Sacchi, 2015; Tredinnick, 2008). Given the inherent situated context integral to any experience of such a performance, effective assessment of that experience cannot be reduced to positivist objectivity, but rather must embrace intersubjective contingency. The human understanding arising from the experience of a preserved digital object is best considered in terms of Peircean pragmatics (Mingers & Willcocks, 2014). This holds that the meaning of a thing is not solely inherent to its fabric, but rather, is encompassed by the totality of the intersubjective perceptual, epistemic, and phenomenological *effects* that the thing provokes in the human actor. The characterizing quality of that experience in the context of digital preservation is communicative success. Success is the relative degree to which the communicative experience leads to satisfactory alignment of intentional, archival, and expectational states of the digital object that is the underlying vehicle for the communicative act. The insights regarding criteria and metrics of success uncovered by this investigation have practical import for preservation practitioners and stakeholders as well as providing firmer conceptual and theoretical foundation for subsequent digital preservation research.

## **1.5 THESIS OUTLINE**

This introductory chapter summarized the problematic state of extant evaluative principles for digital preservation success. It situated that problem within the context of current theory and practice and outlined a research program for attaining better understanding of the critical factors and constraints leading to that problem. Finally, it defined core concepts as well as theoretical, methodological, and analytical structures pertinent to the subsequent investigation. Chapter 2 surveys current thinking in the digital preservation field regarding evaluation of the efficacy of its activities. It identifies pertinent gaps regarding the evaluation of preservation success. This leads to the primary research question pursued in this dissertation to provide insight into how evaluative norms are constructed through relevant domain discourse. Chapter 3

describes the research methodology for the investigation. In particular, it defines the Predicate Reduction technique for Qualitative Content Analysis newly developed and deployed for this purpose. It also derives a semiotic model of digital preservation activity for purposes of subsequent Philosophical Inquiry. Using that methodology, Chapter 4 establishes existing evaluative norms commonly accepted in scholarly and professional practice as tacitly referenced in digital preservation policy statements. Chapter 5 subjects those norms to critical Communicological analysis to determine their applicability to characterize the communicative success of domain activity in a meaningful manner. Finally, Chapter 6 summarizes the overall research findings and their implications. It also proffers a set of recommendations regarding principles for more effective evaluation of digital preservation success and an outline for subsequent inquiry extending this study.

## Chapter 2: Literature Review

---

Greater insight into the problematic state of determining success begins with critical examination of how the field defines its legitimate evaluative dimensions (discussed in Section § 2.1), the factors commonly deployed as evaluative evidence along those dimensions (Section § 2.2), and the explanatory power – and limitations – of resulting evaluative characterizations (Section § 2.3). In other words, examining the literature’s current consensus regarding which aspects of digital preservation activity are susceptible to meaningful appraisal – and by implication, which are not; what pertinent metrical factors underlie such appraisals; the scope of the impact of those appraisals on theory and practice; and what aspects of the preservation enterprise remain unexamined, uncharacterized, or unexplained. This investigation reveals a significant research gap in the literature. This in turn suggests the research question underlying the subsequent research program (Section § 2.4).

The conceptual definition of any field of common concern or practice both establishes the parameters for, and prescribes the boundaries, of legitimate scholarly investigation (Condon, 2014). The digital preservation field is most commonly defined in terms of custodial management; see for example (Becker et al., 2011; Burgi et al., 2019; Chen, 2007; CLIR, 2002; Corrado & Moulaison Sandy, 2017; Gallinger et al., 2017; Gladney, 2006; Traczyk, 2017; Waller & Sharpe, 2006; Waters & Garrett, 1996; Xie & Matusiak, 2016). That custodianship encompasses managerial actors and processes (Moore, 2008; Strodl et al., 2007; Wilson, 2017) with imperatives to provide assurances regarding the integrity (Ross, 2006); accessibility (Burda & Teuteberg, 2013); authenticity (Adam, 2010); intelligibility (Giaretta et al., 2011); understandability (Donaldson, 2016); and usability (Walters & Skinner, 2011) of managed digital objects. Ideally, evaluation of the preservation enterprise would incorporate criteria and metrics capable of characterizing each of these qualitative imperatives. Of these, usability is the teleologically-preeminent goal (Conway, 2010; Traczyk, 2017) and is best evaluated through a benchmark of communicative *success*; that is, a confirmatory measure of the purposive exploitability of past informative expression encapsulated in digital form (Abrams, 2018a, 2018b, 2021). However, operationalizable measures of that success continue to remain elusive in both theory

and practice (Anderson & LeFurgy, 2006; Dearborn & Meister, 2017; Lee & Tibbo, 2007; Poole, 2016). Extant operational evaluation remains focused on the technical and institutional components of preservation management, without corresponding attention to “softer” considerations of subjective user experience (Jääskeläinen, 2015). This study pursues insight into the benefits of, and impediments to, the application of experiential concerns into evaluative practices for the digital preservation field.

Three related terms are commonly used in academic and professional discourse regarding the ongoing stewardship of digital material: digital *preservation*, digital *archiving*, and digital *curation* (Feng & Richards, 2018; Kowalczyk, 2018; Yakel, 2007). While all three carry the imperative of ensuring future accessibility and usability, archiving is most clearly distinguishable from the other two through its programmatic emphasis on records management and evidential integrity (Cunningham, 2008). The preservation/curation distinction hinges of the latter’s focus on enhancing, rather than just conserving, the value of digital objects (Higgins, 2011) and its embrace of concerns across the full information object lifecycle (Feng & Richards, 2018; Walters & Skinner, 2011). Digital curation was originally promoted as a more encompassing term, explicitly subsuming preservation and archiving concerns, and was intended to reduce potential ambiguity and inconsistent usage (Beagrie, 2006; Dallas, 2016; Lord et al., 2004). A parallel terminological label of *data* curation has been applied more narrowly to custodial stewardship of research datasets (Palmer et al., 2013; Weber et al., 2012). This has led to a prevalent assumption that curation is pertinent only to scholarly or scientific information (Giaretta, 2011). Regardless, use of preservation and curation – and to a lesser degree, archiving – as interchangeable cognate concepts is still widespread (Ball, 2010; Dallas, 2016; Nadal, 2017; Palmer et al., 2013). Basing literature searches on all three terms is necessary to achieve broad coverage of the field; see for example (Feng & Richards, 2018; Maemura et al., 2017). Thus, this literature review assumes a conceptual synonymy of digital preservation, curation, and archiving.

## **2.1 EVALUATIVE SCOPE**

The *Encyclopedia of Archival Science* defines digital preservation as “the processes and controls that enable digital information objects to survive over time” (Thibodeau, 2017, p. 160). This object- and process-centric emphasis conceptually positions the preservation enterprise as a managerial activity. That is, a set of things

done *to* objects to ensure the persistence of their characteristics over time, without corresponding attention to what subsequently can be done *with* them. Detail regarding the intent and mechanism of those processes is addressed by the Association for Library Collections and Technical Services (ALCTS), which promotes three parallel definitions of digital preservation – short, medium, and long – purposefully formulated with incrementally increasing levels of detail (ALA, 2009). The short definition, presumably offering the most concise expression of core concern, expresses that core as the “policies, strategies and actions that ensure *access* to digital content over time [emphasis added].” The United Nations Educational, Scientific, and Cultural Organization (UNESCO) similarly promotes digital preservation as the “processes aimed at ensuring the continued *accessibility* of digital materials [emphasis added]” (UNESCO, 2019a). In archival practice, access refers to the ability and permission to find and retrieve information relevant for a specific purpose (SAA, 2020). In this formulation, access is explicitly positioned as an enabling factor for subsequent usage, which remains a distinct phenomenon. In other words, the effectuating agency underlying these definitions is bounded by the procedural effort ensuring access and does not encompass the hypothetical, let alone actual, user who might take purposive advantage of that access. Enforcing a clear separation of preservation and usage issues at the system level is technically appropriate and operationally prudent (Keller, 2009; Moore et al., 2005; Wilson, 2017). However, when considering digital preservation as a service, let alone a conceptual enterprise, the preservation/use distinction can become teleologically problematic. The consensual weight of repeated assertions of the *operational* primacy of accessibility implicitly positions digital preservation *conceptually* as an essentially managerial activity. That is, a set of activities concerned with direct custodial responsibility for the acquisition, documentation, persistence, visibility, and retrievability of digital objects. The ability to retrieve an object, however, is distinct from a subsequent ability to make productive use of it. Thus, an imperative goal of accessibility represents a perspective of the preservation enterprise from the managerial viewpoint. It sets the boundary of managerial responsibility at the point at which the object leaves managerial control. Usability, on the other hand, is concerned with purposive post-managerial experience.

Digital preservation-enabled access and use exist in a symbiotic relationship. Successful re-use *of* preserved digital objects presupposes prior accessibility *to* those

objects (Belkin, 2005; Menne-Haritz, 2001), without which there cannot be any use at all. The Digital Preservation Coalition’s *Digital Preservation Handbook* asserts an explicit synonymy of the two concepts: “access is assumed to mean continued, ongoing *usability* of a digital resource, retaining all qualities of authenticity, accuracy and functionality deemed to be essential for the purposes the digital material was created and/or acquired for [emphasis added]” (DPC, 2015). Efforts addressing these imperatives encompass preservation acts that both maintain and add value to managed digital objects (Beagrie, 2006). This pursuit is enacted through professional stewardship (Lee & Tibbo, 2007); proactive management (Thibodeau, 2017; Yakel, 2007); and socio-technical processes (Harvey et al., 2020). The intent of these efforts is to provide and keep access to managed objects for current and future use (Becker & Rauber, 2011; Traczyk, 2017) as well as mitigate obsolescence and other factors that would otherwise impede that use (Burda & Teuteberg, 2013). The emphasis in these prescriptions on *actions*, *activity*, *management*, *stewardship*, *providing*, *processes*, *keeping*, *retaining*, and *mitigating* implies the prior existence of responsible *actors*, *managers*, *stewards*, *providers*, *processors*, *keepers*, *retainers*, and *mitigators* ensuring the accessibility necessary for the desired use. All of these cognate actorial roles are hereinafter subsumed under the common label of digital preservation “manager.” This actorial emphasis also explicitly elevates the managerial role – and implicitly, managerial evaluation – above that of the future consumer who might reap the benefit of that management.

In addition to reiterating a central concern for accessibility, the ALA medium-length definition articulates a preservation goal of “accurate *rendering* of authenticated content,” to which the long definition also adds an imperative programmatic mission of “preserv[ing] digital content for future *use* [emphasis added]” (ALA, 2009). In terms of definition, these additions complement the centrality of physical accessibility with the opportunity for subsequent behavioral *experience* of accessed material. However, the proper delineating scope of digital preservation as a field of common concern and practice is unsettled (Langley, 2019). Some authors advocate for the subsumption of use as an integral consideration of preservation proper, see for example (DPC, 2015; Traczyk, 2017; Yakel, 2007), while others position preservation and use as independent, albeit mutually supportive, considerations (Kaplan, 2008; Walters & Skinner, 2011; Wilson, 2017). Regardless, the evaluative quality of experiential use is dependent upon the degree to which a preserved object can be exploited “to do something sensible with the information it contains” (Giaretta, 2011, p. 167). In order

to ensure future usability, preservation managers must remain cognizant of the “*responsibilities, functions and characteristics* of comprehensive and reliable digital preservation programmes [emphasis added]” (UNESCO, 2019b). However, that set of managerial concerns does not encompass the means for measuring and assessing resulting programmatic *outcomes*. Such verification is important for purposes of determining whether the programmatic strategies and operational procedures leading towards those outcomes were fit for purpose in the context of a future use of preserved information (Ball, 2010). Much recent work in the field has focused on the development of appropriate social and technical structures ensuring such fitness.

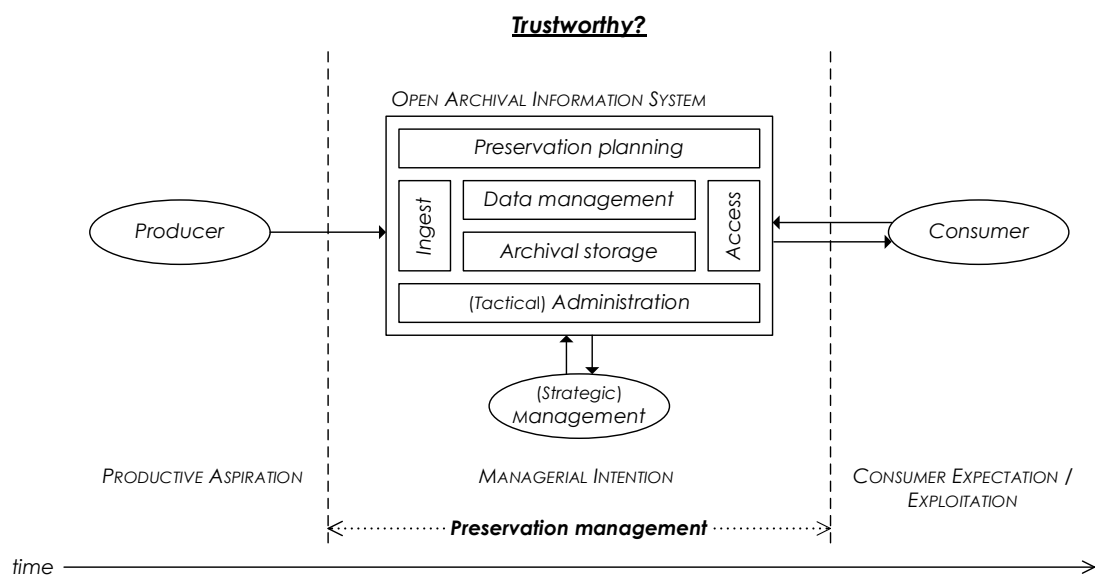


Figure 2.1. Digital preservation as a data management activity  
Adapted from (Abrams, 2018b)

The ISO 14721 *Open Archival Information System (OAIS) Reference Model* (Brunsmann et al., 2012; Giaretta, 2011; Nadal, 2017; Thibodeau, 2002; Xie & Matusiak, 2016) is the primary programmatic framework referenced by the digital preservation community for analysis, design, and operation (see Figure 1.1). Under OAIS, technical instrumentality for preservation is provided by OAIS systems while controlling – albeit delegated – agency is exercised by OAIS managers, rather than producers or consumers (Abrams, 2018b). Preservation itself is defined by OAIS as “The act of *maintaining* information [emphasis added], Independently Understandable by a Designated Community, and with evidence supporting its Authenticity over the

Long Term” (ISO, 2012b, p. 1-13).<sup>5</sup> Within this conceptual framing, digital preservation is implicitly positioned as being synonymous with preservation management (Abrams, 2018b). Similarly, the scope of managerial purview implicitly circumscribes the borders of the preservation act itself (see *Figure 2.1*).

The role of consumers and the activities of actual usage are not directly addressed in the OAIIS model (Gladney, 2006; Nicholson & Dobрева, 2009). While the OAIIS Access functional entity is concerned with consumer-initiated search and retrieval, it does not encompass consideration of the actual phenomenological experience of the subsequent utilization of retrieved material. Since the programmatic mission of preservation is to enable future *use* of preserved objects, the varied perspectives of those objects’ *users* should be incorporated into its evaluation (Caplan, 2008; Chowdhury, 2010; Yakel, 2007). The concept of post-custodial stewardship (Dallas, 2016) acknowledges the agency of all participants involved with preservation concerns, inclusive of information producers and consumers as well as preservation managers (Davis, 2017; Lee & Tibbo, 2007; Moulaison Sandy & Corrado, 2018; Rusbridge et al., 2005). Despite this recognition, there has not been a corresponding expansion of perspectival scope regarding the evaluation of that enterprise, which continues to emphasize assessment only of activities under managerial control (Xie & Matusiak, 2016) and treats discovery, delivery, and use of preserved materials as out of scope (Wilson, 2017). In view of the fact that preservation goals can be articulated as ensuring that preserved objects remain fit for purpose (Dallas, 2007; Ross, 2006), that the primary imperative underlying fitness is to facilitate future use of those objects (Conway, 2010; Traczyk, 2017), and that it is the future user who exercises ultimate discretion regarding the time, place, and manner of that use (Belkin, 2005), the primary focus of preservation evaluation should focus on the successful outcomes of consumer experience (Abrams, 2018b). However, the OAIIS reference model does not provide specific guidance regarding the identification or measurement of that success. Instead, it recommends follow-on effort to develop appropriate evaluative tools and strategies for characterizing the fulfillment of programmatic OAIIS responsibilities. The OAIIS case is reflective of a broader consensus in the field that the primary evaluative

---

<sup>5</sup> Following ISO practice, initial capitalization of key terms in the OAIIS text indicates that they are entities formally defined in the standard document.



dimension of the digital preservation enterprise is managerial, not experiential.

## 2.2 EVALUATIVE EVIDENCE

The evidence deployed in those evaluative determinations includes essential details of how archival programs and systems are designed and implemented (Johnston, 2008; Thibodeau, 2007), the collection scope and range of supported service functions of those programs (Yakel et al., 2009), the ability of stakeholders to find, retrieve, and make use of managed content (Lubell et al., 2008), and trustworthiness (Becker & Rauber, 2011). As mentioned above, the OAIS standard recommends follow-on activity to complement its foundational modelling with evaluative tools. The primary focus of that activity has been inquiry into preservation trustworthiness (Donaldson, 2016; Traczyk, 2017). Trustworthiness is a significant general characteristic of any information system addressing customer concerns over uncertainty, vulnerability, and technological dependence (Corritore et al., 2003; Kelton et al., 2008). In the preservation context, trustworthiness is a justified belief that systems and programs are capable of meeting their preservation obligations (Dryden, 2011). Trustworthiness may be demonstrated through reference to standardized assessment tools such as nestor/DIN 31644 *Criteria for Trusted Digital Repositories* (Maemura et al., 2017; nestor, 2009),<sup>6</sup> *CoreTrustSeal* (CoreTrustSeal, 2019; L'Hours et al., 2019), and ISO 16363 *Audit and Certification of Trustworthy Digital Repositories* (TDR) (ISO, 2012a; Witt et al., 2012). However, these evaluative benchmarks primarily define trustworthiness through *descriptive* programmatic and technical features, rather than *predictive* ones characterizing the outcomes of those programmatic technologies. Descriptive trust is garnered through what has been *said* about an underlying phenomenon. Predictive trust, on the other hand, arises from a review of previous *results* of that phenomenon.

Trust is descriptive if its veracity is dependent upon attributions or testimonials such as stated intentions, contractual assurances, or institutional reputation; and predictive if the presumed state of future events or conditions are extrapolated from past history to new contexts (Dryden, 2011). Descriptive evidence of trustworthiness

---

<sup>6</sup> The preferred lexical form of the acronym for the German Network of Expertise in Long-term Storage of Digital Resources [*Kompetenznetzwerk Langzeitarchivierung*] is the all-lower-case “nestor”.

can be abstracted into the following logical schema

*Because process P has attribute A, presumptively-associated with outcome O, applying P to resource R should result in O.*

This association is non-operational in that it is rooted in intuitive belief rather than experience. In distinction, predictive evidence follows the schema:

*Because process P applied to resource Q resulted in actual outcome O, applying P to resource R, which is similar to Q, should result in O.*

While differing in the degree of engendered confidence, the two schemas are related: the rational basis for deriving A of P follows from analysis of why P led to O relative to R (and Q). The *a posteriori* predictive formulation is inductively stronger in its reliability relative to the *a priori* descriptive assumption, as it has been subject to, and extrapolated from, prior operational scrutiny. However, in practice, trustworthiness continues to be assessed largely in descriptive terms. Thus, as currently constituted, trustworthiness should be viewed as an enabling quality most closely associated with preservation management (Yoon, 2014) and not subsequent use of the information objects preserved through that management. That is to say, trustworthiness is primarily an evaluation of what preservation managers *do* (Xie & Matusiak, 2016), rather than the consumer experience *enabled* by that doing. In this respect, the application of trustworthiness as a benchmark norm is consistent with the general conceptual emphasis within the preservation field on the central position of managerial agency and activity regarding programmatic assessment (Becker et al., 2011; Wilson, 2017).

At best, trustworthiness illuminates the presumptive possibility, but not the substantiated actuality, of preservation activity (Donaldson, 2016). In other words, while it can bolster confidence in what *should* occur, it does not necessarily confirm what *has* occurred. This limitation has been recognized in the OAIS context. The working draft of the proposed 3rd revision of the standard introduces a new concept of *preservation objective* (PO). A PO is a “specific achievable aim which can be carried out using the Information Object” (CCSDS, 2019, p. 1-13). Furthermore, POs “make it possible *to test whether the information actually is* Independently Understandable by members of the Designated Community now and into the future [emphasis added]” (CCSDS, 2019, p. 2-8). In other words, POs are intended to form

the basis for confirmational, rather than descriptive, characterization. However, some initial applications of preservation objectives remain focused narrowly on technical considerations. See for example (Burgi et al., 2019), which defines objectives regarding replication policy, fixity audit, representational formats, and other managerial concerns couched in terms of managerial intentions. While it is possible to infer that these intentions arise from consideration of underlying user goals, needs, or aspirations, that is not explicitly stated. Giaretta and Conway, on the other hand, do derive managerial objectives from assumptions about future patterns of use by a given designated community (Giaretta & Conway, 2011). This underscores their claim that the resulting objectives are specific, actionable, measurable, and realistic. However, their specificity and measurability do not extend beyond a generic statement that the future user “should be able to *correctly* interpret [emphasis added]” (p. 249) the preserved information content, without suggesting accompanying evidentiary measures. The reference to the concept of correctness implicitly ties any subsequent determination to the context and purpose for which the information is being referenced. Regardless, much like the situation regarding preservation intention statements discussed in Section § 1.2, the literature does not provide strong evidence of current adoption of preservation objectives as a routine, public-facing component of preservation activity. Until the expression of aspirational goals becomes widespread, the use of preservation objectives as a benchmark for evaluation will remain problematic. In the absence of explicit documentation of imperative objectives and accompanying measures, visibility of appropriate norms for evaluation is best provided indirectly through examination of reciprocal service-provider intentions and stakeholder expectations as established by publication of programmatic preservation policies.

Accurate and well-defined policies are a critical complement of systems and services for effective preservation (Bountouri et al., 2018). Under the ISO 16363 standard for the audit and certification of Trusted Digital Repositories, policy statements play an important role as documentary evidence (Sanett, 2013) that the “activities of the repository *will be understood by stakeholders and management* [emphasis added]” (ISO, 2012a, p. 3-5). In doing so, TDR explicitly recognizes the contractual relationship of complementary intentions and expectations implicitly established by policies controlling the parameters of manager/stakeholder

engagements. However, TDR continues with an assertion that existence of documentation “*ensures* that repository policies and procedures *are carried out* in approved, consistent ways [emphasis added]” (ISO, 2012a, p. 3-5). This overstates the causal certainty that the existence of controlling policies *necessarily* results in satisfactory fulfillment of policy intentions. Preservation policies play an important role in promoting ultimate success, but it is an enabling role, not a conclusive one. Preservation policies do enumerate important programmatic obligations, but any subsequent demonstration that those obligations have been successfully fulfilled requires verification. Verification of digital preservation success depends upon availability of evidentiary criteria and metrics that remain undefined by TDR and similar assessment frameworks.

Despite its limitations, trustworthiness continues to be the primary means for assessing the digital preservation enterprise. While the parameters of trustworthiness are well defined, success remains a much more elusive concept, let alone a metric (Anderson & LeFurgy, 2006; Lee & Tibbo, 2007). A viable conceptual definition of success has not found scholarly consensus, due in large part to the strongly contingent and contextualized aspects of its inherent nature (Dearborn & Meister, 2017). Trustworthiness does have the advantage of being a leading indicator that can be asserted *before* the fact, albeit provisionally, as a harbinger of anticipated outcomes. Trust in a service-prover is also an important prior consideration in future determinations regarding customer satisfaction with a provided service (Kim, 2012). Success, on the other hand, as a property of the actual outcomes of preservation-enabled communication, is a measure of actual satisfaction and can be asserted unconditionally, although only *after* the fact. Given an option to choose between trustworthy and untrustworthy solutions, a decision to favour the trustworthy alternative may appear obvious. However, if the decision is reframed not as a choice between trustworthy and untrustworthy alternatives, but rather, between trustworthy and *successful* ones, the decisive factors become more nuanced (Abrams, 2018b).

Success *can* occur through untrustworthy as well as trustworthy means. While the former case is less likely, it is nevertheless possible. It would be difficult, however, to associate ultimate trustworthiness with a stewardship system or program that is clearly unsuccessful. Thus, the two qualities of success and trustworthiness share a similar relationship to that of the claimed philosophical priority of states of actuality

over those of potentiality (Cohen & Reeve, 2020). Success is definitionally prior to trustworthiness in that the latter is ultimately formulated in terms of the former. That is, it is rational for a system to be considered trustworthy if it potentially can be, or has been proven to be, successful. Success is also prior to trustworthiness in practice in that while successful outcomes may result from untrustworthy processes, putatively trustworthy processes resulting in unsuccessful outcomes risk losing their designation, as they have failed to achieve their final purpose. This sense of priority bolsters the need for the digital preservation community to develop effective standards and practices for characterizing the ultimate communicative efficacy of its activity. This need represents an extension of the current consensus in the field regarding the primary role of managerial trustworthiness as the benchmark evaluation for digital preservation activity.

### 2.3 EXPLANATORY POWER

Many prevailing expressions of preservation goals and implied evaluative criteria emphasize the imperative persistence of authentic information objects (Becker & Rauber, 2011; Thibodeau, 2002; Traczyk, 2017). For certain classes of digital content, such as interactive games and artworks whose performative behavior is integral to the full information experience, stewardship of the objects themselves must be complemented by preservation of the necessary intermediating software environments (Abbott, 2012; Day et al., 2018; Winget, 2011). But in fact, *all* digital objects rely upon software to render the native digital representation of their underlying information content into analog human-perceptible form (Abrams, 2015; Becker, 2018; Tredinnick, 2008; Zierau, 2012). Without persistent recourse to those – or functionally equivalent – mediating environments, objects that are otherwise “perfectly” preserved as bitstreams will not be susceptible to legitimate understanding (da Silva Júnior & Borges, 2016; DPC, 2015). The OAIS notion of an object’s *understandability* is inherently conditional as the *-ity* suffix indicates that the object has the presumed capacity of being *understandable* (OED, 2009). However, as a measure of consequent communicative success, that is quite different from the quality of having been *understood*. The OAIS goal of independent understandability of preserved digital objects depends upon those objects being directly open to interpretation and use by a designated community without supplementary external information or expert assistance (Austin et al., 2015; Lavoie, 2014). The concept of a

designated community aggregates the experience, expertise, and information-seeking goals of a discrete group of potential users on the basis of their presumed shared knowledge and professional, personal, or institutionally-focused purpose (Donaldson et al., 2020). The more narrowly a designated community can be defined, the better, particularly with regard to devising effective evaluation metrics (Bak, 2016). While the plausible definition of such groups may be problematic (Bettivia, 2016b; McDonough, 2012), a successful preservation outcome occurs when independent understandability is realized on the part of a real, rather than hypothetical, user. That is, an actual consumer who was able to exploit the preserved object in contextually-meaningful pursuit of a purposive goal.

Paradoxically, digital objects are both easily maintainable as opaque *bitstreams* and openly susceptible to damage or irretrievable loss as information-laden *objects* (Rothenberg, 1999). In theory, bitstreams are infinitely and “perfectly” copyable. In practice, however, the design, implementation, and sustenance of policy and procedural regimes ensuring ongoing perfection are technically difficult and financially prohibitive (Rosenthal, 2010a, 2010b). Regardless, without proper attention to the avoidance or mitigation of various technical, operational, or administrative risks, those bitstreams and the information they carry are vulnerable to preservation failure. These risks include incipient format obsolescence (Johnston, 2020b); actions (or inactions) potentially affecting the qualitative integrity of object identity, availability, authenticity, renderability, and understandability (Vermaaten et al., 2012); and generalized vulnerabilities regarding data, infrastructure, and processes as well as threats from natural disasters, malicious attack, and managerial and legal impediments (Barateiro et al., 2010). In view of preservation’s open-ended time horizon, and the continual evolution – and inevitable disruption – of risk-ameliorating strategies, practices, and infrastructure, progress towards successful preservation outcomes depends upon a series of periodic transitions over time to redeployed technical systems and processes (Janée et al., 2009; Owens, 2018). Similar hand-offs of curatorial responsibility and custody may be necessary in cases of institutional closure, financial constraint, or reprioritization of programmatic scope (Caplan et al., 2010; Corrado & Moulaison Sandy, 2017). Thus, digital preservation should be viewed not as a one-time, fully-sufficient activity, but rather, as a series of incrementally necessary activities tailored to meet the needs and respond to the risks

particular to their positioning in time as well as technical and cultural space.

The success of preservation-enabled communication across time is fundamentally constrained by the inherently provisional nature of the enterprise. Because it is not possible to anticipate the full consequences of the immediate – let alone the far – future, it is not possible to assert categorical evaluative positions that are meaningfully applicable beyond the immediate point-in-time of that assertion (Abrams, 2018b). This condition is conceptually-analogous to the idea of scientific *falsification*. This holds that a theory articulated in falsifiable form – that is, with clearly-identified criteria for verification of truth-claims – can be held provisionally true until such time that it is shown to be definitively false (Persson, 2016; Popper, 1959; Tredinnick, 2006). By analogy, one can legitimately assert digital preservation has been successful *so far* if preservation outcomes do not constitute failure to date. The temporal centrality of archival timespans underlying preservation commitments necessarily implies an ever-growing cultural distance separating the past point of initial content acquisition and the future point of consumption (Ricoeur, 1976; Tan et al., 2009). This in turn emphasizes the importance of the cultural-positioning of all actors implicated in preservation activity and the resulting purposive contingency their experiences of operational outcomes (Ball, 2010; Bishop & Hank, 2018; Dearborn & Meister, 2017; Morrissey, 2014). The state of a given preserved object at any point in time may represent both preservation success *and* failure when viewed variously from the perspectives of different users with different purposive intents (Ross, 2012).

Objects resulting from digitization of tangible originals may provide significant function unavailable from the original. For example, the use of multi-spectral imaging to enhance analysis of otherwise indistinct palimpsestic texts (Howell & Snijgers, 2020). However, whatever functional capabilities may be potentially gained through digital reformatting, something is also always lost in the process (Deegan & Tanner, 2006). That loss could encompass specific aspects of an object’s content that were uncapturable or unrepresentable in digital form (Stanford, 2020) or the more ephemeral notion of Benjamin’s “aura” of originality (Benjamin, 1936; Burns, 2017). The significance of that loss is dependent upon purpose and context. In view of this inevitable contextual contingency, all preserved objects – whether reformatted or born-digital – should be viewed as approximate surrogates rather than exact facsimiles of their nominal underlying abstract content and consumable behavior. While the term

“facsimile” is generally used in the context of the relationship between physical originals and copies (SAA, 2020), it can be applied, as in this study, to the relationship between a tangible, if digital, copy and its abstract essence. In this sense, a facsimile entails a one-to-one mapping in all respects between that intangible essence and its tangible digital representation. This implies an associated objective benchmark for evaluation: the facsimile mapping is either complete or incomplete. The purposive adequacy of a surrogate, on the other hand, is a matter of subjective evaluation along a relative scale of fitness for use.

As contingent surrogates, the relative success of the use of preserved objects should be evaluated in terms of situational verisimilitude, given that the notion of absolute fidelity to some canonical object state or information experience is illusory (Ross, 2012; Yeo, 2010). This condition is conceptually-similar to the assertion of scientific *truthlikeness*. This posits that confidence in a theory’s truthfulness is positioned along a continuum ranging from intuitive plausibility to verified actuality, with varying degrees of accompanying explanatory power (Johansson, 2017; Popper, 1976). Modern relativistic physics is more truthlike in an absolute sense than superseded Newtonian physics, especially when applied on a micro- or macro-scale (Gribbin, 1984). However, Newtonian laws of motion are still truthlike enough for adequate prediction of normal human sensory perceptions and interactions with the physical world (Popper, 1999). By analogy, one can legitimately evaluate preservation success as a *relative* measure. Success can indicate the degree to which preserved objects can be meaningful exploited for some particular purpose in a particular context by a particular user (UNESCO, 2003). Alternatively, success is applicable when the evaluated outcome is below some threshold of acceptable loss (Ries & Palkó 2019). To date, however, there has been inadequate critical investigation into ways to quantify digital preservation verisimilitude, let alone the retention of intended and expected levels of verisimilitude across time and iterative preservation interventions (Ross, 2020).

The relativistic basis for evaluation of verisimilitude is mirrored by a similar tiered approach to considerations of institutional and programmatic maturity with respect to preservation capabilities and capacities. A number of assessment instruments are available for determining the position of a preservation institution along a spectrum of maturity; see for example DRAMBORA (Innocenti et al., 2009),



the Digital Preservation Capability Maturity Model (Ashley & Misic, 2019), and NDSA Levels of Digital Preservation (NDSA, 2019; Phillips et al., 2013). These measures are critical for determining, and improving, an institution's capacity and preparedness to achieve its preservation goals (Maemura et al., 2017). It is plausible to extrapolate from a relative scale for maturity of programmatic capability to another scale applicable to anticipated or realized maturity of outcome, or in other words, success. Given the inherently contingent nature of future use of preserved objects and the finite limits to preservation efficacy over archival timescales, the evaluative outcome of digital preservation stewardship is not so much a question of binary success or failure as it is of relative success-likeness. The effective measure of that likeness, however, has not been sufficiently addressed to date.

## 2.4 RESEARCH GAP AND QUESTION

Digital preservation scholarship and practice have focused on intensive investigation of how the preservation enterprise can be meaningfully evaluated as a *managerial* endeavor. Essentially, the synthetic question underlying prior scholarship regarding preservation assessment is: What characteristics of digital preservation agency and systems bolster confidence in their ability to meet their obligations? In answer, the preservation community has developed and continues to promote an evaluative benchmark of the trustworthiness of stewardship institutions and their socio-technical infrastructures. This perspective is managerial in that it is concerned primarily with organizational, curatorial, and operational considerations regarding the persistence and accessibility of authentic digital *objects*. The notion of access implicitly presumes, if not expects, subsequent use and the quality of usability is often referenced in the literature as a core preservation imperative. However, the human context, experience, or measure of that use is not encompassed by current evaluative theory or practice. In other words, even though future purposive use of preserved objects is recognized as the proper teleological goal of the preservation enterprise, it remains a critical aspect of the enterprise not yet susceptible to meaningful characterization of its consequent efficacy.

The communicative success of preservation stewardship is a measure of the satisfactory exploitation of past informative expression in the context of a future purposive goal. While trustworthy preservation management is an important necessary factor for that success, it is not fully sufficient by itself for a complete

measure of the possible attainment of that teleological goal. Thus, the field's previous inquiry regarding evaluative scope is recast in this study to ask instead how the preservation enterprise can be meaningfully evaluated as a *communicative* activity. However, effective evaluation of communicative success depends upon concepts, criteria, and metrics still underdefined in theory and practice. Any attempt to respond to this situation should begin by trying to understand the reason *why*, despite their importance, those evaluative factors they have so far resisted adequate formalization and operationalization. This provides the basis for the primary research question for this dissertation:

RQ 1 *What are the parameters for a conceptually-sound, yet pragmatically-actionable evaluative framework for determining the communicative success of the digital preservation enterprise?*

Without that information, it will be difficult to avoid, respond to, or mitigate past impediments during subsequent development of new measures for preservation success. This study defines that success as a measure of the mutual alignment of managerial intentions, stakeholder expectations, and the realized archival state of preserved digital objects. There is no accepted mechanism currently in use by the preservation community for the explicit articulation of intentions and expectations. Nevertheless, they can be inferred indirectly from relevant domain discourse. That inferential activity is the initial focus of this study, responding to the subordinate research question:

RQ 1.1 *What socially-constituted norms regarding digital preservation success emerge from evaluative attitudes implicit in domain discourse?*

The resulting findings (see Section § 4.3) offer new insights regarding the nature of success as tacitly understood by the preservation community. That insight can be deployed for subsequent research and development of operational criteria and metrics providing more comprehensive assessment of digital preservation efficacy. This is formalized in a second subordinate question:

RQ 1.2 *How suitable for purpose are existing evaluative norms for digital preservation success?*

The suitability of norms is determined through Communicological critique cognizant of preservation's essential communicative function. The results of that critique (see

Sections § 5.1 - 5.5) suggest useful avenues of pursuit regarding a better means to qualify the outcomes as well as quantify the outputs of digital preservation success. This leads to the final subordinate research question:

RQ 1.3 *What complementary enhancements to existing evaluative theory and practice are necessary for more effective and comprehensive characterization of digital preservation success?*

The conjunction of the findings of this last line of inquiry (see Sections § 5.6 - 5.7) with the first two address the fundamental concerns raised by RQ 1 by presenting evidence of the current state-of-the-field regarding the evaluation of success, assessing that state, and proposing ways to mitigate its shortcomings.

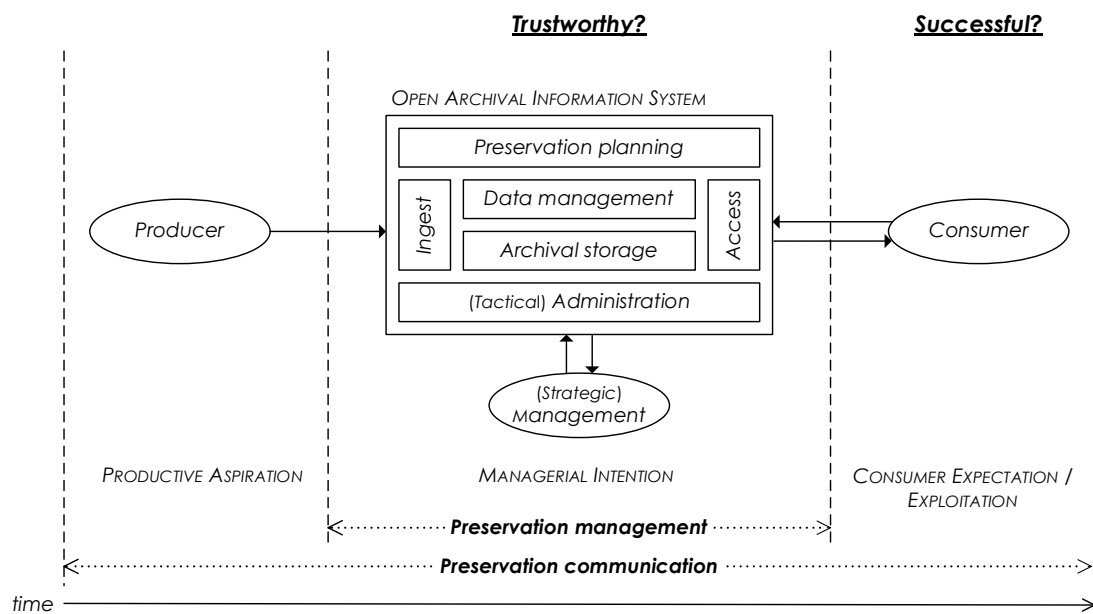


Figure 2.2. Digital preservation as a communicative enterprise  
Adapted from (Abrams, 2018b)

The instigating hypothesis for this investigation was that extant evaluative norms emphasize the programmatic, managerial, artifactual, and predictive aspects of the preservation enterprise at the expense of the actorial, communicative, experiential, and confirmatory (see Figure 2.2). That is, those norms are suitable for characterizing custodial technical and risk mitigation activities as applied to digital objects independent of the circumstances of their use and are therefore suggestive, but not conclusive, regarding eventual stakeholder satisfaction with that use. The validation of this hypothesis proves useful to subsequent attempts to define other norms more

applicable for characterizing preservation efficacy in terms of contextually-situated and purposively-driven stakeholder experience. Thus, the hypothesis implicitly positions criteria for digital preservation success as a matter of intersubjectively-contingent stakeholder assessment. That in turn supports the notion that evaluative norms for success are emergent social constructions. Consequently, success norms are identified through Qualitative Content Analysis of relevant domain discourse using the Predicate Reduction technique newly developed for this study. Once the norms are established, Communicological analysis is deployed to determine their suitability – and limitations – as the basis for comprehensive assessment of the success of the digital preservation enterprise. That information then provides the foundation for a new multivalent definition of digital preservation success and a corresponding multi-dimensional evaluative space in which preservation results can be assessed in terms of pertinent imperative norms, semiotic dimensions, and evaluative modalities.

## Chapter 3: Research Design

---

This research deploys two novel approaches to the question of characterizing digital preservation success. Section § 3.1 introduces the inductive Predicate Reduction (PR) technique of qualitative content analysis, newly developed for this purpose. Section § 3.2 describes the application of abductive Communicological analysis to the data revealed through Predicate Reduction. The chapter concludes with a discussion of potential limitations of this research design (Section § 3.3).

This study provides new understanding of why the derivation and application of effective measures of digital preservation success have remained elusive. The underlying evidence for that understanding comes from establishing and critiquing evaluative attitudes regarding preservation success that permeate relevant domain discourse, if only tacitly. The investigation starts from a conceptual positioning of digital preservation as an act of digitally-mediated human communication unfolding across archival timespans. This position accepts that the preservation enterprise depends upon trustworthy data management ensuring persistent access to integral and authentic information objects. However, it also promotes the importance of complementing that output with the persistence of opportunities for legitimate communicative experiences. As the goal of meaningful communication properly lies at the core of the definition of digital preservation (Abrams, 2018b, 2021), this research is an exercise in Communicology, the critical study of human discourse built upon a foundation of semiotic phenomenology (Catt & Eicher-Catt, 2010; Lanigan, 2008). This theoretical position supports a core insight underlying this research study, namely, that preservation-enabled communication is enacted through the intersubjective experience of its human participants (Lanigan, 2010b). That experience encompasses a range of communicative acts expressing, persisting, transmitting, perceiving, interpreting, and, ultimately, responding to culturally-coded signs. In other words, communicative meaning is an emergent phenomenon and engagement with a preserved digital information object is an inherently constructivist act. However, while this research is studying constructivist *phenomena*, its research *design* relies on pragmatic – and not constructivist – methodological principles.

The pragmatic research paradigm bridges realist and idealist positions regarding

the ontological status of reality, emphasizing the centrality of human experience over metaphysical speculation (Creswell, 2014a). It views that experience as necessarily informed and constrained by the fundamental nature of reality, a borrowing from positivism, as well as the individual contextualized responses to that reality, a hallmark of constructivism (Kankam, 2019; Morgan, 2014). Pragmatic investigation into that experience is characterized by an intersubjective stance. That is, it recognizes that both complete objectivity or subjectivity are implausible standards, and accepts the validity of appropriate researcher intuition and interpretation arising from prior experience, expertise, and deliberate self-reflection as well as consistency with, and reactive refinement of, other relevant research activity (Morgan, 2007; Revez & Borges, 2019). Similarly, the scope of applicability of pragmatic results is not intended to be universal to all possible contexts or narrowly constrained to the specific context of the original investigation. Rather, pragmatic insights strive to be maximally transferable, in whole or in part, to other suitable situations in which they can provide meaningful illumination and explanation of otherwise problematic phenomena (Shannon-Baker, 2016). The findings presented in this study offer such illumination to the long-unaddressed question of what constitutes effective measures of digital preservation success.

### **3.1 INDUCTIVE QUALITATIVE CONTENT ANALYSIS**

This inquiry identifies and critiques existing evaluative attitudes towards digital preservation success through Qualitative Content Analysis (QCA) of institutional preservation policy statements. As argued in Section § 1.2, preservation policies establish the terms of the controlling social contract of reciprocal service-provider intentions and stakeholder expectations whose alignment with the actual preserved state of a preserved resource lies at the core of a determination of success. This study's newly developed QCA technique of Predicate Reduction (PR) mechanistically reduces obligatory policy terms into implied evaluative norms through iterative rule-based textual transformations (Abrams, 2021). Since the identified norms arise from critical examination of preservation policies, those norms can be viewed as emergent thematic codes (ETC). ETC codes are those derived from, rather than imposed upon, underlying data sources (Amundsen & Sohbat, 2008; Gibbs, 2007; Stemler, 2001). Consequently, the PR technique was designed to produce results consistent with criteria appropriate for establishing ETC codes. These include being responsive, exhaustive, mutually

exclusive, sensitizing, and congruent (Merriam & Tisdell, 2015; Schreier, 2012). That is, the resulting norms are directly applicable to the specific research question; they encompass all pertinent concepts implicated by that question; each relevant granule of analyzed data contributes a single norm; terminologically, the norms are allusively-connotative as well as directly-denotative of the described phenomenon; and the norms are defined at equivalent levels of conceptual abstraction. These qualities ensure that the final results are plausible, reliable, and reproduceable.

### 3.1.1 Data Sources

A set of 95 digital preservation policy documents articulating the internal standards and practices of international memory institutions was assembled from existing datasets. These were the results of prior research activity conducted by the Library of Congress (Sheldon, 2013) and the SCAPE project (SCAPE, 2016). These sources were supplemented by a general Internet search with Google (www.google.com) conducted on 21 February 2019 with the query string:

“digital preservation” (policy OR policies)

which expands to two matching criteria: “digital preservation policy” and “digital preservation policies”. The Library of Congress data contributed 29 of the policies, one of them uniquely; the SCAPE results provided 44 documents, five uniquely; and the Google result set, 83 documents, 47 uniquely. Twenty-three of the policy documents were enumerated in two of the lists and 19 in all three. The deduplicated list of documents is managed in a spreadsheet (Abrams, 2020) with descriptive fields for organizational name; parent organization, if relevant; geopolitical jurisdiction; organizational sector based on the Ringgold classification (Ringgold, 2018); Carnegie higher-education classification, for US-based academic organizations (Carnegie, 2018); organizational mission; policy title, version, identifier, date, and URL; and source; i.e., Sheldon, SCAPE, or Google.<sup>7</sup>

Six representative preservation policy documents were chosen from the full set using paradigmatic case sampling (Robinson, 2014). That is, the six were chosen as being prototypically-emblematic of the fundamental characteristics of the larger

---

<sup>7</sup> The policy document dataset is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <<https://doi.org/10.17605/OSF.IO/ZHTQJ>>.

institutional universe cognizant of the diversity of their geopolitical jurisdiction, sectorial role, and mission-orientation (see Table 3.1).

Table 3.1

*Selected Digital Preservation Policy Documents*

Organization	Jurisdiction	Sector	Mission
Baltimore Museum of Art (BMA, 2016)	United States	Cultural heritage	Museum
Cambridge University Libraries (CUL, 2018)	United Kingdom	Academic	Library
Inter-University Consortium for Political and Social Research (ICPSR, 2018b)	United States	Academic	Data archive
Leibniz-Informationszentrum Wirtschaft [Leibniz Information Centre for Economics] (ZBW, 2018)	Germany	Research	Institutional repository
Nationaal Archief [National Archive of the Netherlands] (NA, 2015)	Netherlands	Government	Archive
National Library of New Zealand/Archives New Zealand (NLNZ, 2012)	New Zealand	Government	Library/Archive

There are no clear methodological guidelines for determining the minimal or optimal sample size for Qualitative Content Analysis (Elo et al., 2014). Nevertheless, an important principle governing sampling strategy is that the resulting sample set should be adequate for the specific research question (Drisko & Maschi, 2015). That is, the samples are information-rich in a manner explanatory of the phenomenon under study (Vasileiou et al., 2018). Meaningful explanation of domain phenomena can be achieved with small sample sets if they constitute rich and comprehensive information sources and are subject to rigorous analysis (Young & Casey, 2018). A sampling is considered adequate when a threshold of data or thematic saturation is reached (Hennink & Kaiser, 2022). That is, the point at which additional samples do not yield further insight. Young & Casey’s metastudy (2018) reports that over 90% saturation is achievable with as few as four to seven cases. Small sample sizes can be justified in terms of the nature of the research question, the rigor with which samples are subject, and the homogeneity of the sampled population (Boddy, 2016).



As argued in Section § 1.2, policy documents provide primary evidence of attitudinal positions regarding digital preservation success. The Predicate Reduction technique for Qualitative Content Analysis introduced in Section § 3.1.2 defines a rigorous formal structure for identifying those positions in the policies. Sheldon (2013) found high levels of correspondence between the relative frequency of 19 taxonomic categories for policy terms in 31 examined library and archive documents. The SCAPE project found similar consistency regarding 10 guidance policy categories for its corpus of 44 policies (Sierman, 2014). These two corpora include three of the six policies examined in this research. The other three policies were published after the Sheldon and SCAPE studies were completed. Five of the six policies cover 89.5% or more of Sheldon’s taxonomic categories and 90% or more of SCAPE’s guidance policy categories (Abrams, 2023).<sup>8</sup> The NLNZ policy’s coverage of these categories is 73.7% and 70%, respectively. This is explained by the fact that the NLNZ policy explicitly excludes several categories as out of scope and covered by other, external policy statements. As summarized below, the six policy selections are emblematic of commonly-shared policy intentions, themes, and terms as well as spanning institutional types significantly engaged in digital cultural heritage stewardship. In light of this, this research’s sample size of six paradigmatic policies is justified and appropriate.

Digital preservation imperatives are central to the vision and mission of a variety of memory institutions. This is particularly so for libraries, archives, and museums (LAMs), which have long-established stewardship responsibilities for cultural and documentary heritage (Corrado & Moulaison Sandy, 2017; Langley, 2019; Oyelude, 2019). Traditional definitions of LAM institutions assert a primary emphasis on stewardship of published information carriers, records and unpublished information carriers, and dimensional artifacts, respectively. The three LAM types are also distinguishable by imperative missions providing ongoing access to documentary collection, preserving evidential collections necessary for construction of future historiographic narrative, and offering interpretive presentation of artifactual significance, respectively (Robinson, 2012). However, this is an increasingly artificial

---

<sup>8</sup> The policy term consistency dataset is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <<https://doi.org/10.17605/OSF.IO/CSVBM>>.

and strained distinction (Hedstrom & King, 2004; Martin, 2007), which has led to use of “memory institution” as the embracing conceptual term (Dempsey, 2000; Rasmussen, 2019). In practice, many LAM institutions engage to some degree in all of these concerns (Given & McTavish, 2010; Marty, 2010). Distinctions across institutional types regarding long-term digital preservation concerns are more of degree rather than kind.

Newer institutional forms with similar stewardship aspirations include data archives (Kim & Choi, 2016; Pinnick, 2019) and institutional repositories (Asadi et al., 2019; Jones et al., 2006). A data archive (DA) provides centralized management of and access to research data (Wright et al., 2018) collected in accordance with target thematic and format criteria (Borgman et al., 2016). The institutional repository (IR) mission, on the other hand, is to provide organizational commitment for the long-term stewardship (Li & Banach, 2011) of the intellectual output of an institution or community (Hockx-Yu, 2006). In the digital age, all of these institutional stewards – LAMs as well as DAs and IRs – incorporate aspects of digital preservation as mission-critical activity. Consequently, the sample set for this study encompasses all five institutional categories in order to represent policy perspectives, concerns, and practices as they are widely deployed across the digital preservation community. Institutional selection was biased explicitly in favour of larger, well-known, and long-established preservation programs. Their significant history of preservation activity makes it more likely that they encompass the most sophisticated analysis of pertinent concerns and present the most accessible and comprehensive articulation of policy terms (Abrams, 2021). Furthermore, because of their visibility within the community, it is more likely that they will function as exemplars of model digital preservation policy regimes for more recent entrants to the field (Sierman, 2014).

The Baltimore Museum of Art (BMA) is a public cultural heritage institution founded in 1914 with a current mission to connect “art to Baltimore and Baltimore to the world, embodying a commitment to artistic excellence and social equity in every decision from art presentation, interpretation, and collecting” (BMA, 2021). Its policy establishes a “framework for long-term preservation and access to the Museum’s digitized and born-digital assets” and “inform[s] the development of detailed plans and procedures for implementing digital preservation activities” (BMA, 2016, p. 1). The controlling impetus for these obligations arises from the Museum’s Strategic Plan,

Collections Management Policy, Records Retention Schedule, and Records Access Policy.

The policy of the Cambridge University Libraries (CUL) governs both the main research library as well as other affiliated libraries across the University, all supporting its diverse Colleges, Schools, Faculties, and Departments. The Libraries were first established in 1416 and now provide “expertise, partnership, services and collections that underpin the University’s mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence,” which includes “harness[ing] the power and potential of the digital age to transform the cultivation and sharing of knowledge” (CUL, 2019, p. 2). CUL is a legal deposit library entitled to receive copies of all UK publications, whether in tangible or digital form (BL, n.d.).

The Inter-University Consortium for Political and Social Research (ICPSR) is a collaborative of over 750 international academic institutions and research organizations founded in 1962. Its imperative mission “advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations” (ICPSR, n.d.). Its policy “makes explicit ICPSR's commitment to preserving the digital assets in its collections” (ICPSR, 2018b, p. 1) in alignment with the organization’s overall Strategic Plan (ICPSR, 2021).

The Leibniz-Informationszentrum Wirtschaft [Leibniz Information Centre for Economics] (ZBW) is the world’s largest research infrastructure for economic literature, founded in 1919 and now affiliated with Christian-Albrechts-University (ZBW, n.d.). Its mission is to acquire, preserve, and make accessible the literature and subject-area data in the fields of economics and business studies, which is increasingly available only in digital form (ZBW, 2018). The ZBW’s preservation policy builds upon a joint strategic consensus of the three German national subject libraries, the other two of which are the Technische Informationsbibliothek [Leibniz Information Centre for Science and Technology] (TIB) and the Informationszentrum Lebenswissenschaften [Information Centre for Life Sciences (ZB MED)] (ZBW, 2017).

The Nationaal Archief [National Archives of the Netherlands] (NA) is the governmental archive for the Netherlands. Its mission is to facilitate interactions

“between the worlds of history and current affairs, that of the archive creator and the archive user, that of the old and new media and that of the public and private domain” (NA, n.d.) by offering “information and provid[ing] insight into [the Netherland’s] past” (NA, 2015, p. 5). The NA’s strategic imperatives arise from the Netherland’s Public Records Act that explicitly encompasses digital information objects, “including the entire range of interpretations of archive files, records and digital documents” (NA, 2015).

The National Library of New Zealand (NLNZ) was established in 1945, with antecedents stretching back to 1858. Its operational imperative is to “collect, connect, and co-create knowledge to power New Zealand” (NLNZ, n.d.), consistent with a legal deposit mandate and a statutory mission to “preserve, protect, develop and make accessible for all the people of New Zealand the collections of that library in perpetuity” (NLNZ, 2012, p. 1). The NLNZ’s digital preservation policy arises from a strategic obligation to steward digital alongside physical materials (NLNZ, 2016) and is shared by Archives New Zealand (ANZ). The ANZ has a legislative mandate for the “preservation and access of the digital record of [the New Zealand] government” and to “make sure that the digital information is there when today’s and tomorrow’s New Zealanders need it” (ANZ, 2022).

The six selected policies are issued by well-established and long-standing cultural heritage institutions occupying leadership positions in the diverse LAM/DA/IR stewardship landscape. As evident from these contextual summaries, digital preservation concerns and activities play a central role in the strategic and operating principles and priorities of all six institutions. Thus, for purposes of this study, they provide a small, but well-representative, paradigmatic sampling of the policies available for possible analytic consideration. As described below, the novel Predicate Reduction technique for Qualitative Content Analysis developed for this purpose is highly mechanistic in nature. However, while it may be susceptible to future machine automation, for this research project the QCA was carried out manually. In this context, a smaller, highly representative set of policies is both methodologically desirable and appropriate.

### **3.1.2 Analytic Method**

The core activities of Qualitative Content Analysis are data reduction and subsequent abductive inferencing (Krippendorff, 2019). That is, refining original

source data into a more compact representation of pertinent characteristics and then explicating the meaning of those characteristics relative to the underlying research question. The Predicate Reduction technique addresses the data reduction phase of this research study. Its design was informed by specific aspects of the earlier QCA methods of Syntagmatic Analysis (SA) and Evaluative Assertion Analysis (EAA) (Abrams, 2021). The SA method provides tools for examining how informative meaning arises from the associational context of word groupings (Green, 1991; White & Marsh, 2006).<sup>9</sup> That is, it seeks to understand the interpretive implications of a particular sequence of words in evoking intended or serendipitous nuanced connotations of meaning. SA has been deployed successfully for establishing the metaphoric parameters of implicit domain models for the semantic concepts of information (Green, 1991) and libraries (Nitecki, 1993) broadly held across the LIS profession. The central SA technique is the derivation of “atomic syntagmatic combinations” (Green, 1991, p. 133), that is, short unitary phrases distilled from often-complex expressions for subsequent metaphoric analysis. PR relies on a similar process of normalizing its source material in a manner facilitating the synthetic construction of implied evaluative criteria (see Section § 3.1.3, Steps 2 and 3, below). PR also follows SA in incorporating a step of normalizing non-semantically-significant lexical variations, such as inflections for grammatical tense, voice, and aspect, into canonical form to aid clustering of cognate concepts (see Section § 3.1.3, Step 4).

The EAA technique relies upon psycholinguistic principles to determine anticipated attitudinal responses by readers to core concepts cited in texts (Krippendorff, 2019; Osgood, 1959; Osgood et al., 1956). It seeks to establish and rate the intensity of the affective association – positive or negative – regarding the concepts underlying analyzed expressions. Like SA’s fabrication of atomic syntagmatic combinations, EAA manipulates source texts into normalized expressive form to facilitate subsequent analysis. For that purpose, EAA establishes a canonical *object-verb-object* schema to represent the evaluative relationship between individual

---

<sup>9</sup> The QCA technique referred to here as “Syntagmatic Analysis” is left unnamed in the literature, where it is referenced by its developer’s name, i.e., “Green’s methodology” (Nitecki, 1993), rather than descriptive label. The SA label used hereinafter for easy reference is derived from the technique’s reliance on atomic syntagmatic combinations as the unit of analysis.

attitudinal objects. PR borrows from EAA the idea of relying on a representational schema for expressing synthetic units in canonical form (see Section § 3.1.3, Step 5). However, while EAA is useful for exposing attitudes towards conceptual expressions explicitly manifest in the text, it is not intended to uncover latent concepts that may be implied by the text. Since the policy documents examined in this study do not articulate their obligatory terms as explicit criteria or metrics for success, EAA is insufficient for establishing the range and scope of those measures. Instead, PR is used to uncover pertinent metrics from the implicative expression of policy imperatives.

The PR method shares with SA and EAA a reliance on transformative textual manipulation. However, SA relies on an initial lexicographic search of source documents for pre-determined concepts to identify relevant contextual snippets for analysis. In the case of concepts such as “information” or “library” a word stem search for *inform-* and *librar-* provides satisfactory results. In this study, however, the relevant concepts – evaluative criteria for success – are not known *a priori* or identified as such within the source texts. Similarly, EAA relies on the researcher’s intuitive sense of what phrases whose evaluative meanings are susceptible to legitimate variant interpretation by readers. PR, on the other hand, removes the reliance on *a priori* and intuitive assessment by incorporating grammatical, rather than lexicographic or intuitional, criteria for the identification of textual passages relevant as the starting point for further analysis (see Section § 3.1.3, Step 1). The grammatical classification of source texts in PR follows the usage established by the *Cambridge Grammar of English* (Carter & McCarthy, 2006).<sup>10</sup> A fully worked-through example of the Predicate Reduction technique is found in Appendix § A.

### 3.1.3 Predicate Reduction Process

The Predicate Reduction technique systematically identifies pertinent policy obligations and recasts them as synthetic expressions of imperative commitments, presumptions, and criteria appropriate for measuring their alignment. PR encompasses five sequential steps: four initial analytic activities of (1) statement identification; (2) propositional expansion; (3) predicate reduction; and (4) predicate canonicalization; and a final synthetic activity of (5) kernel construction (Abrams, 2021). Since the

---

<sup>10</sup> Hereinafter referenced as CGE. Following internal CGE practice, subsequent citations are given parenthetically with the relevant section number rather than page number, for example, “(CGE § 227)”.

evaluative norms underlying those final implied evaluative kernels were derived throughout the formal data reduction process, they function as emergent – rather than *a priori* – thematic codes (ETC) (Amundsen & Sohbat, 2008; Stemler, 2001).

**Step 1: Statement Identification.** Within a policy document, relevant contextual statements expressing core policy obligations are indicated by specific grammatical markers: a copula verb, a modal auxiliary verb, or a lexical verb of obligation (CGE § 227):

- **Copula Verb Markers.** The copula verb “to be”, generally encountered in its inflected forms “is”, “are”, etc., asserts a semantic equivalence between its grammatical subject and subject complement, nominally of the form *subject-copula-complement* (CGE § 279b). In other words, the subject complement provides a substitutable *definition* of its subject. Thus, in the context of PR-analyzed preservation policies, copula verbs express a state or action of existential necessity on the part of their subjects (see, for example, Table 3.2).

Table 3.2

*Copula Verb Marker*

---

*Statement:* “Monitoring and reporting is an essential aspect of digital preservation activities [emphasis added]” (NLNZ, 2012).

---

In this example, monitoring and reporting are asserted as fundamental obligatory components of digital preservation activity.

Note that the copula verb is distinct from the auxiliary form of “to be” indicating progressive voice (CGE § 224-225), e.g., “the system *is preserving* the object”, or passive voice (CGE § 478), e.g., “the object *was preserved*”. In these cases, the pertinent grammatical marker is not the auxiliary “is/are”, but rather, the augmented lexical verb “preserving/preserved”.

- **Modal Verb Markers.** Modal auxiliary verbs (e.g., “must”, “will”, “shall”, etc.) assert a degree of commitment that the subject brings to a lexical verb action relative to its object, nominally of the form *subject-*

*modal-lexical-object* (CGE § 379). In the context of PR-analyzed preservation policies, modal verbs express agential intentions to fulfil preservation imperatives (see, for example, Table 3.3).

Table 3.3

*Modal Verb Marker*

---

*Statement:* “The BMA [Baltimore Museum of Art] will provide authenticity, discovery, and access to digital assets for current and future generations [emphasis added]” (BMA, 2016).

---

In this example, the policy indicates the strongest possible commitment on behalf of the BMA regarding provision of archival authenticity, discovery, and access.

Table 3.4

*Lexical Verb Marker*

---

*Statement:* “ICPSR [Inter-University Consortium for Political and Social Research] preserves social science digital assets and provides its members with ongoing access to its digital collections [emphasis added]” (ICPSR, 2018b).

---

- ***Lexical Verb Markers.*** Lexical verbs assert an action, event, or state (CGE § 228). In the context of PR-analyzed preservation policies, lexical verbs express affirmative obligations on behalf of their subjects towards their objects (see, for example, Table 3.4). In this example, ICPSR asserts an affirmative obligation regarding the preservation of its digital collections.

Statements identified in Step 1 are considered in scope for subsequent analysis only when they entail an obligation with respect to the preserved state of digital *objects*. Statements relating to operational, financial, or administrative concerns of preservation programs and systems are not considered relevant for this research. All



object-centric statements are recorded along with their structural context within the document; that is, they are associated with the specific named or numbered section under which they are found.

**Step 2: Propositional Expansion.** Many of the obligatory statements identified in Step 1 are compound grammatical constructions. These include *coordinated* statements joined by combining conjunctions (e.g., “and”, “or”) (CGE § 271) or statements expressing an imputed *composition* of independent concepts. Every such compound statement is expanded into a set of singular propositional clauses (CGE § 539), each with a nominal form of *subject-verb-object* (see, for example, Table 3.5).

Table 3.5

*Propositional Expansion*

---

*Statement:* “The NA [Nationaal Archief / National Archive of the Netherlands] ensures that users are able to understand *and* use the information that it has made available [emphasis added]” (NA, 2015).

→ *Proposition:* “The NA ensures that users are able to understand the information”

→ *Proposition:* “The NA ensures that users are able to use the information”

→ *Proposition:* “The NA ensures that information has [been] made available”

---

In this example, the original compound statement contains two main clauses linked by the inclusive conjunction “and” indicating that *both* clauses are subject to the NA’s obligatory assurance. These two main clauses are factored into three singular propositions. The first two are derived from the *grammatical* expansion of the coordinated phrases linked by “and”. The final proposition results from an implied *semantic* expansion justified by recognizing that the concept of availability must be asserted implicitly before considering the implications arising from explicit references to understanding or use. It is not possible to understand or use a preserved digital object that is not readily available for that understanding or use.

PR’s propositional expansion is equivalent to the second step in Green’s Syntagmatic Analysis of deriving atomic syntagms, or propositions, from coordinated or otherwise complex narrative statements (Green, 1991).

**Step 3: Predicate Reduction.** Because the specific propositional subjects are not relevant to subsequent analysis, every expanded proposition is reduced to its corresponding analytic predicate. These *verb-object* formulations (CGE § 539) capture the central intentional obligations underlying the full propositions (see, for example, Table 3.6).

Table 3.6

*Analytic Predicate Reduction*

---

*Statement:* “Archives New Zealand ... and the National Library of New Zealand ... have agreed to give access to digital objects [emphasis added]” (NLNZ, 2012).

→ *Proposition:* “Archives New Zealand ... [has] agreed to give access to digital objects [emphasis added]”

→ *Analytic Predicate:* “give access”

---

In many cases, the terms of the imperative predicate are analytically extracted directly from the propositional text, as shown in the example above. In other instances, the predicate must be constructed synthetically from an interpretive sense of the central obligation underlying the proposition. In the example in Table 3.7, the assertion of the fundamental goal of being able to access preserved objects presupposes a complementary agential responsibility for the affirmative *assurance* of that access.

Table 3.7

*Synthetic Predicate Reduction*

---

*Proposition:* “The primary objective of digital preservation activities is the ability to meaningfully access digital content over time [emphasis added]” (BMA, 2016).

→ *Synthetic Predicate:* “[ensure] access”

---

**Step 4: Predicate Canonicalization.** Given that policy documents are expressed in free expository form, many cognate variations are found of common obligatory concepts. All of the verbs and objects in the reduced predicates are passed through a

thesaurus (Abrams, 2022b) for canonicalization in terms of standardized vocabulary (see, for example, Table 3.8). This procedure facilitates the clustering of conceptually-related predicates, and their derivative kernels, through a simple lexicographic sort.

Table 3.8

*Predicate Canonicalization*

---

*Analytic Predicate:* “give access”

→ *Canonical Predicate:* “ensure accessibility”

---

In this example, “give” is replaced by “ensure” to capture a more proactive sense of service-provider obligation. “Access” is replaced by “accessibility” to emphasize a sense of affirmatively-provisioned agential capacity via the *-ability* suffix.

The thesaurus was constructed during the Predicate Reduction process. The preferred terms were selected for denotative as well as connotative clarity of meaning and to enforce consistent inflection (see Table 3.9).<sup>11</sup> In some instances, the thesaurus mapping is dependent on a source term’s functional context. For example, the predicate verb “archive” is normalized to “ensure” when it is applied to intangible qualities such as accessibility or integrity. On the other hand, it is normalized to “preserve” in the context of tangible items such as objects or metadata. New entries were added to the thesaurus as they were encountered until full saturation was achieved. In terms of general analytic coding, saturation refers to the state when no new meaningful data emerges from the underlying data (Saldaña, 2016). The thesaurus also classifies the mapping function between source entries and their preferred terms to distinguish between semantic synonymy or syntactic variation and conceptual sub/superordination. The relational tags “USE” and “BT” (broader term), as defined by ISO 2788 and ANSI/NISO Z39.19 (Aitchison et al., 2000), are used for these purposes, respectively.

PR’s predicate canonicalization is an implementation of the third step in Green’s Syntagmatic Analysis that aggregates variant morphological and syntactic expressions into groups on the basis of their underlying abstract concepts (Green, 1991; Nitecki,

---

<sup>11</sup> The thesaurus is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <<https://doi.org/10.17605/OSF.IO/X4SDN>>.

1993). Green and Nitecki perform this activity by determining cognate meaning of the atomic syntagms in their native form. In PR, the core predicates are explicitly re-expressed in canonical form to facilitate mechanistic lexicographic manipulation.

Table 3.9

*Sample Thesaurus Entries*

*Adapted from (Abrams, 2021)*

Entry	Relation	Preferred term
access	USE	accessibility
access conditions	USE	security
accuracy	USE	authenticity
acquisition decision	BT	provenance
adhere to	USE	ensure
administrative metadata	BT	metadata
archive [intangible quality]	USE	ensure
archive [tangible entity]	USE	preserve
assets	USE	objects
assure	USE	ensure
availability	USE	accessibility
⋮	⋮	⋮

**Step 5: Kernel Construction.** Analytically-derived canonical predicates are the source for the construction of synthetic kernel phrases expressing underlying assertions of preservation service-provider obligation, reciprocal stakeholder expectation, and the basis for the relational evaluation between the two. These kernels are formed using a set of three templates (see Table 3.10). The properly-inflected forms of predicate verbs and objects are inserted into the placeholder slots indicated by *underlined italics*.

Table 3.10

*Synthetic Kernel Templates*

Kernel role	Template
Service-provider intentional obligation	“P intends / to <u><i>verb object</i></u> / for S”
Stakeholder expectational result	“S expects / P / to <u><i>verb object</i></u> ”

References to preservation service-providers and stakeholders are common to all kernels. To streamline kernel structure, they are represented as generic agential classes by the symbols “P” and “S”, respectively.

The first two kernels express the core intentional and expectational positions. The third expresses a general evaluative metric for success in terms of intentional/expectational alignment (see, for example, Table 3.11). PR kernel construction is similar in function to EEA’s reliance of templated forms of analyzed expressions (Krippendorff, 2019; Osgood, 1959).

Table 3.11

*Kernel Construction*


---

*Canonical Predicate:* “ensure usability”

→ *Kernels:*                    “P intends / to ensure usability / for S”  
     “S expects / P / to ensure usability”  
     “Did P / ensure usability / for S?”

---

Because of the standardized vocabulary enforced by the predicate canonicalization in Step 4, a simple alphabetical sort of the kernels automatically clusters cognate instances of intentional and expectational imperatives as well as resulting evaluative norms.

### 3.1.4 Frequency Analysis

The constructed kernels are first examined through quantitative word-count-based analysis (Guest et al., 2014). They are subsequently subjected to a qualitative Communicological critique (see Chapters 4 and 5, respectively). In the former analysis, relative frequency of appearance is assumed to be a reliable indicator of conceptual significance (Gaur & Kumar, 2018; Krippendorff, 2019; White & Marsh, 2006). It is important to recognize that this assumption is reliable only when well-justified with regard to the particular context of a research study (Schreier, 2013). The vagaries introduced by rhetorical style, decontextualized analysis, and misalignment between denotative and connotative semantics may invalidate the use of frequency

metrics as a sole proxy for significance (Kracauer, 1952/2022; Mayring, 2014; Stemler, 2001). In the case of this research, however, the methodological reliance on frequency counts is justified by the nature of the source texts, contextually-sensitive selection of countable units, and confirmatory consistency of the final results with other pertinent data.

Digital preservation policy statements are intentionally created to provide unambiguous guidance regarding programmatic responsibilities and obligations (Sierman et al., 2013). Rhetorically, they should be expressed “in such a way that they will actually be used and referred to, actively enabling the work of preservation” (Madsen & Hurst, 2019, pp. 37-38). In other words, as formal technical documents, policy statements aspirationally represent a factual enumeration of controlling principles, definitions, standards, and rules. This objective intent for policy language minimizes concerns for potential semantic allusiveness and elusiveness. Thus, frequency analysis of the obligatory norms found in a representative sampling of digital preservation policy documents is an appropriate benchmark for the community-accepted evaluative significance of those attitudes.

As illustrated in Section § 4.1, any number of duplicative canonical predicates can result from Predicate Reduction of a single identified policy statement or from multiple statements found in a single named and/or numbered structural context. These duplications may result from rhetorical convention, stylistic lapses, or the thematic coherence reasonably expected within a given expository context. For this reason, frequency analysis relies on counts of predicate expressions that are unique to a given context. For example, multiple references to a single norm within a given context would increment that norm’s count only by one. This minimizes the potential for inappropriate inflation of frequency metrics due to the presence of non-conceptually-significant instances. References to norms found across contexts, on the other hand, are assumed to reflect independent articulations of the evaluative importance of those norms.

The Predicate Reduction results presented in Section § 4.3 represent empiric recovery of the primary evaluative norms tacitly accepted by the digital preservation community from the policies establishing the controlling intentions and expectations of that community. These norms also are consistent with core preservation imperatives expressed in other vehicles of domain discourse as discussed in the Literature Review

in Chapter 2; see, for example, (Adam, 2010; Burda & Teuteberg, 2013; Ross, 2006; Walters & Skinner, 2011). This is an example of both methodological and data triangulation (Flick, 2018), providing confidence in results through parallel derivation via distinct approaches and distinct data sources: PR vs. literature review, and policy documents vs. the scholarly and professional literature, respectively. Taken together, the objective nature of those policies, the contextually-sensitive selection of countable units, and the triangulatory confirmation of the final findings all justify the use of frequency counts as the basis for post-PR analysis.

### **3.1.5 Predicate Reduction for Qualitative Content Analysis**

The intentionally-overt rhetorical expression of digital preservation policy documents, as discussed in Section § 3.1.4, suggested the potential for a mechanistic textually-transformative approach to QCA for identification of common evaluative attitudes tacitly underlying those policies. The resulting Predicate Reduction technique conforms to the definitional goal of content analysis of “making replicable and valid inferences from texts ... to the contexts of their use” (Krippendorff, 2019, p. 24). The central procedural core of content analysis encompasses the recording and coding of meaningful data points (Schreier, 2013; White & Marsh, 2006). In traditional QCA, these steps are performed by human observers/analysts guided by specific instructions (Saldaña, 2016) and interpreted in light of pertinent experience and intuition (Krippendorff, 2019). Predicate Reduction similarly encompasses recording and coding, but does so in a manner providing explicit stepwise external visibility of otherwise internal analytic decision-making.

The recording of coding units takes place through policy statement identification with relevancy based on well-defined grammatical markers of intentional obligation, as presented in Step 1 of Section § 3.1.3. The coding process subsequently unfolds through the iterative stages of propositional expansion (Step 2), predicate reduction (Steps 3 and 4), and kernel construction (Step 5), all of which are mechanistic manipulations of pertinent grammatical components, i.e., the subjects, verbs, and objects of the identified statements. This textually-transformative technique gains its validity in view of the fact that preservation policy documents are explicitly concerned with expressing unambiguous programmatic obligations and the PR technique is designed specifically to recover those expressed obligations. The technique is a direct translation of the accepted functional goals and requirements of QCA into a new

operational, and potentially automatable, framework.

### 3.2 ABDUCTIVE COMMUNICOLOGICAL ANALYSIS

The applicability and effectiveness of evaluative norms for the robust characterization of digital preservation success depends upon first positioning those norms within the full range of activity and actors encompassed by the preservation enterprise. Consistent with the communicative – as opposed to managerial – conceptualization of digital preservation activity described in Section § 1.1, the analytical framework for this study is Communicological. Communicology studies embodied human discourse as the contingent interplay of meaning-laden expressive signs and individual sign consumers (Lanigan, 2015). Its central focus on individual human actors distinguishes Communicology from the concerns of disembodied information-theoretic machine-to-machine communication (Lanigan, 2008) and socially-embodied mass communication (Catt, 2014). Communicological explication of discourse begins with the development of a comprehensive model of the communication environment and its constituent processes to provide a framework for subsequent analysis. That analysis pertains to the information object underlying a communicative act as well as the perceptual experience of that object and the resulting interpretive effect it has cognitively, affectively, and conatively on the human consumer (Eicher-Catt & Catt, 2008; Lanigan, 2010b). This twin emphasis on the structure and functioning of communicative vehicles alongside the intersubjective human response to those vehicles leads to Communicology’s description as a method of “semiotic phenomenology” (Mancino, 2020, p. 17).

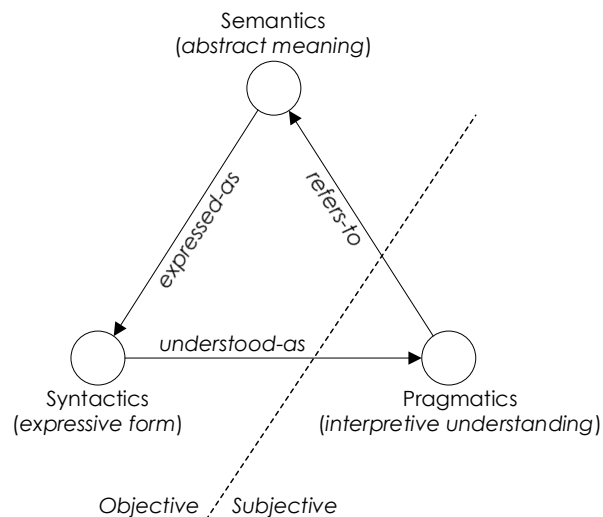
Semiotics is the science of signs and their encompassing systems of signification (Pelc, 2000). A sign is something imbued with communicable cognitive meaning or psychological affect, while signification is the process by which a sign is established, transmitted, experienced, and understood (Eco, 1976; Nöth, 1990). Two distinct theoretical schools exist in semiotic scholarship following either Saussure’s dyadic *signifier-signified* distinction (Harris, 1987; Saussure, 1983) or Peirce’s triadic *signifier-signified-referent* formulation (Peirce, 1932, 1991). The Peircean triad provides a more complete foundation for understanding preservation-enabled communication given its explicit cognizance of the roles played by external informative referents, actors, contexts, and consequences of communicative acts (Mingers & Willcocks, 2014). The semiotic affordances of sign-based communicable



information in the Peircean tradition, particularly as expounded by Morris, are threefold (Mingers & Willcocks, 2017; Morris, 1964; Peirce, 1932):

1. Semantics, encompassing a semiotic information object's abstract intellectual meaning or emotional affect;
2. Syntactics, encompassing the concrete form expressing that object's underlying information context; and
3. Pragmatics, encompassing that object's epistemic interpretation and phenomenological understanding by a human agent.

The cyclic relationships adhering between these aspects constitute the so-called *triangle of reference* or *semiotic triangle* (Eco, 1976; Nöth, 1990) (see *Figure 3.1*).



*Figure 3.1.* Semiotic triangle

Adapted from (Nöth, 1990)

The literature deploys a wide variety of descriptive labels for the vertices of the triangle; see, for example, the manifold usages documented in (Eco, 1976; Nöth, 1990). The labels used in this research – semantics, syntactics, pragmatics – were selected for interpretive clarity by a non-specialist audience. Pragmatics, concerned with explicating the genesis and consequence of internal mental states of human sign consumers, is inherently a subjective affordance. Semantics and syntactics, on the other hand, whose referents exist external to the consumer, are objective in nature.

The origins of the Peircean triangle are found in classical and scholastic philosophy (Deely, 1982; Nöth, 1990). At that time, the semiotic medium was

conceived as solely analog: words, sung or spoken; marks carved into stone; ink written or printed on paper; paint brushed on canvas; etc. The advent of the digital age introduced new opportunities for potential transmissive media, necessitating an extension of semiotic concerns to explicate the nuanced characteristics of technology-dependent channels of communication. In response, Stamper segments the traditional conceptualization of syntactics into three distinct affordances (Beynon-Davies, 2010; Mingers & Willcocks, 2017; Stamper, 1993):

1. Syntactics proper, encompassing the aspects of rhetorical expressive abstraction;
2. Empirics, encompassing the aspects by which that expressive form is represented through symbolic encodings; and
3. Physics, encompassing the tangible manifestation of empiric form in computational infrastructure; in other words, actual bits in memory, on storage media, or across networks.

Stamper also proposes a new *social* affordance concerned with the broader intersubjective context underlying pragmatic understanding. This is equivalent to Peirce’s notion of a semiotic *ground*, the allusive network of intuitions and associations within which interpretative understanding concretizes (Peirce, 1991) (see *Figure 3.2*).

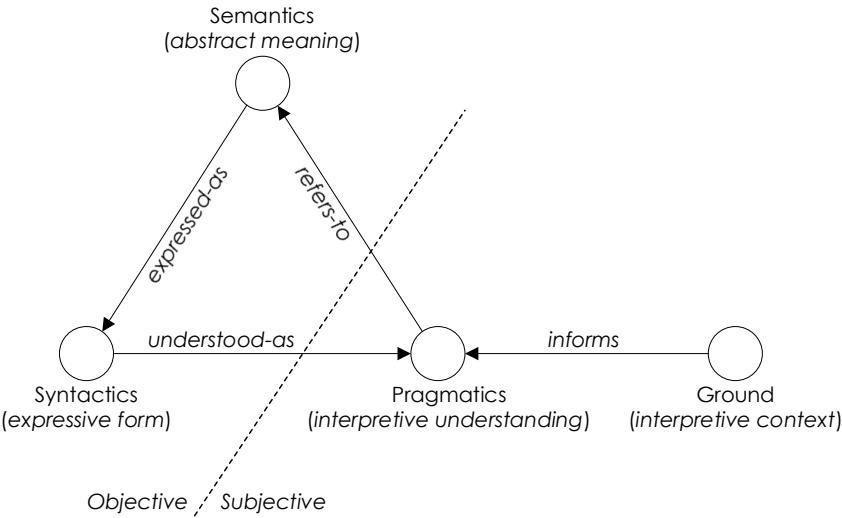
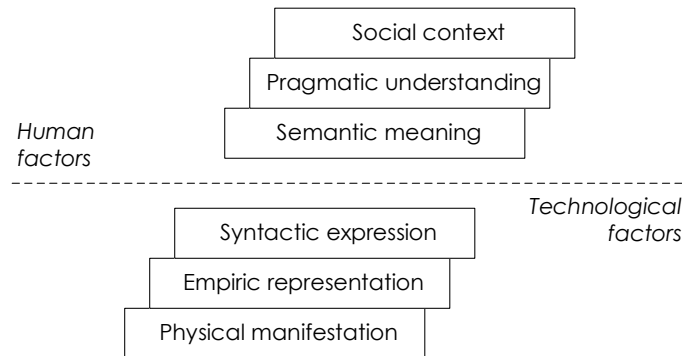


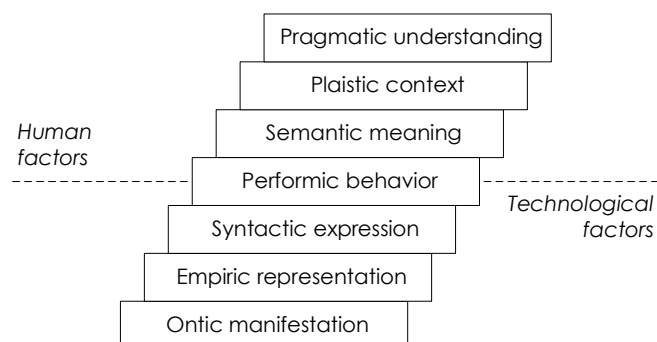
Figure 3.2. Grounded semiotic triangle

Stamper’s enhanced formulation, referred to as the *semiotic ladder*, classifies its component “rungs” as functioning primarily in either a human or technological sphere (see *Figure 3.3*).



*Figure 3.3.* Semiotic ladder  
Adapted from (Stamper, 1993)

Digital resources are inherently dependent upon mediating technological behaviour to render their digital representations into analog form perceptible by human sensory modalities (Becker, 2018; Flouris & Meghini, 2007; Heslop et al., 2002; Morrissey, 2014). This behavioral mechanism is an inherent component of the semiotic process for digital objects. Consequently, subsequent analysis is based upon a preservation-augmented extension of the Stamper ladder to include a new *performics* rung as a liminal affordance mediating between the technological and human realms of the opaque digital and sensate analog (see *Figure 3.4*).



*Figure 3.4.* Preservation-augmented semiotic ladder  
Cf. *Figure 3.3*

For greater terminological consistency, Stamper’s foundational *physical* rung is re-labeled as *ontics*, from the Greek *ὄντος* [*ontos*], “of that which is”, while the *social*

rung is re-labeled as *plaistics*, from *πλαίσιο* [*plaisio*], “frame” or “context”. The consistent application of the *-ics* suffix to all semiotic affordances, denoting them as branches of knowledge or fields of activity (OED, 2009), has rhetorical appeal. Plaistics is also repositioned between semantics and pragmatics to reflect its critical role in conditioning the pragmatic response to semantic consumption. The scope and function of all seven affordances in the augmented ladder – ontics, empirics, syntactics, performics, semantics, plaistics, pragmatics (OESPSPP) – span the significant concerns of a comprehensive semiotic model of preservation-enabled communication.

The contours of communicative processes have been subject to a variety of formulations reflecting various conceptual perspectives. Shannon’s information theory is concerned with modelling a narrow subset of the communication problem: that of the technical transmission of physical signals independent of subsequent human interpretation (Shannon, 1948; Tzafestas, 2018). While propagation of preserved materials from producer to manager to consumer is a foundational component of digital preservation activities, Shannon’s narrow perspective is insufficient to explicate the teleologically-imperative use of preserved resources by human actors. The subjective human participation absent from Shannon’s formulation is explicitly incorporated in Berlo’s sender-message-channel-receiver (SMCR) model (Berlo, 1960; Tzafestas, 2018), but without reference to actorial context. Schramm’s extension to SMCR posits that the success of the communicative act depends upon a common field of experience shared by the participants of that act and underlying their individual interpretations of a message (Rogala & Bialowas, 2016; Schramm, 1954), but does not inquire into the nature of the consumer experience. In Laswell’s persuasive communication model, success is also dependent upon the alignment of the productive intent of a communicative act and the consequent effect the communicated message has on its consumer (Lasswell, 1948; Rogala & Bialowas, 2016). Jakobson’s linguistic perspective gives greater attention to context and the expressive coding and interpretive decoding strategies underlying a communicated message (Jakobson, 1960; Lanigan, 2013). These strategies are central to an understanding of the communicative functions afforded by preserved resources to preservation actors. The philosophical concerns of Alexander focus on explicating the nuanced distinctions between a message’s underlying meaning and the actual – or conceivable possible – referents of

that meaning, as well as identifying the cusp points in the modelled communication process where failures can occur (Alexander, 1988; Lanigan, 2013). Those potential points of failure represent specific risks that preservation planning and intervention attempts to ameliorate or remediate. None of these prior modelling efforts, however, includes any explicit, or even implicit, reference to the effect of temporal distance on the efficacy of the modelled communication.

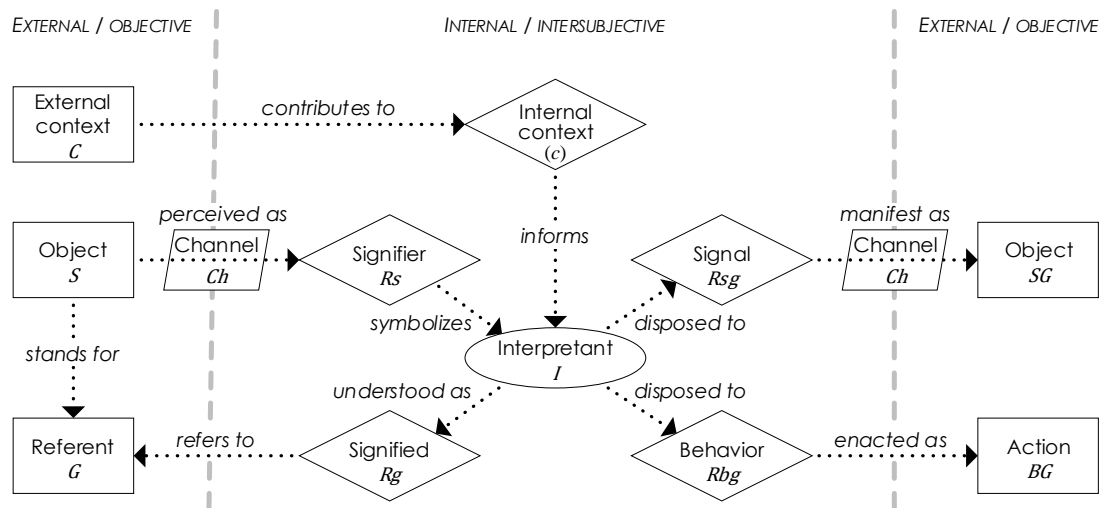


Figure 3.5. Semiotic matrix

Adapted from (Krampen, 1997a)

Krampen’s semiotic matrix<sup>12</sup> is a meta-model defining a set of descriptive abstractions for documenting arbitrary semiotic activity (Krampen, 1997a) (see Figure 3.5). For purposes of presentational clarity, several of the original descriptive names in Krampen’s matrix have been replaced with more accessible terms. For example, Krampen’s use of “interpretandum” and “interpretatum” in reference to the tangible and intangible sources of semiotic representation are hereinafter referred to as “object” and “referent”, respectively. However, Krampen’s shorthand labels, e.g., “S” and “G”, are retained throughout as an aid for mapping between the nomenclatures.

<sup>12</sup> In the technical literature, “semiosics” refers to the relational properties of signs, while “semiosis” refers to processes through which sign-based activities unfold, and “semiotics”, the general science and study of sign-making, interpreting, and understanding (Pelc, 2000). For clarity of exposition, “semiotics” is used throughout this dissertation as an encompassing term for all three aspects.

Any semiotic process entails a set of tangible and intangible entities external and internal to the semiotic actor. (External entities are represented in *Figure 3.5* by rectangles, internal entities by rhomboids, and actors by ovals.) A semiotic object  $S$ , defined by Peirce as “something which stands to somebody for something [its referent  $G$ ] in some respect or capacity” (Mingers & Willcocks, 2017; Peirce, 1932), is perceived by an interpretational agency  $I$  through some techno-physiological channel  $Ch$  as an abstract signifier  $Rs$ . Under the intersubjective influence of external and internal contexts  $C$  and  $(c)$ ,  $I$  interprets  $Rs$  as a set of signified cognitive or affective consequences  $Rg$ . These leave  $I$  disposed to perform subsequent semiotic signalling  $Rsg$  or conative behavior  $Rbg$ , effectuated through physio-technological channels as an external semiotic object  $SG$  or physical action  $BG$ .

Krampen’s matrix is defined from the perspective of a single semiotic actor. Given preservation’s position as a communicative activity, it implicates three actorial categories: information producers, managers, and consumers (see *Figure 2.2*). The application of the matrix in such dialogic situations requires the concatenation of multiple matrix instantiations, where the resulting object  $SG$  of one forms the initial  $S$  of another (Krampen, 1997b). In the context of preservation-enabled communication, matrix modelling must capture the full spectrum of activities of the productive, managerial, and consuming actors implicated in preservation activity. The pertinent components of the enhanced model can be associated with the primary semiotic function as defined by the augmented semiotic ladder (see *Figure 3.6*).

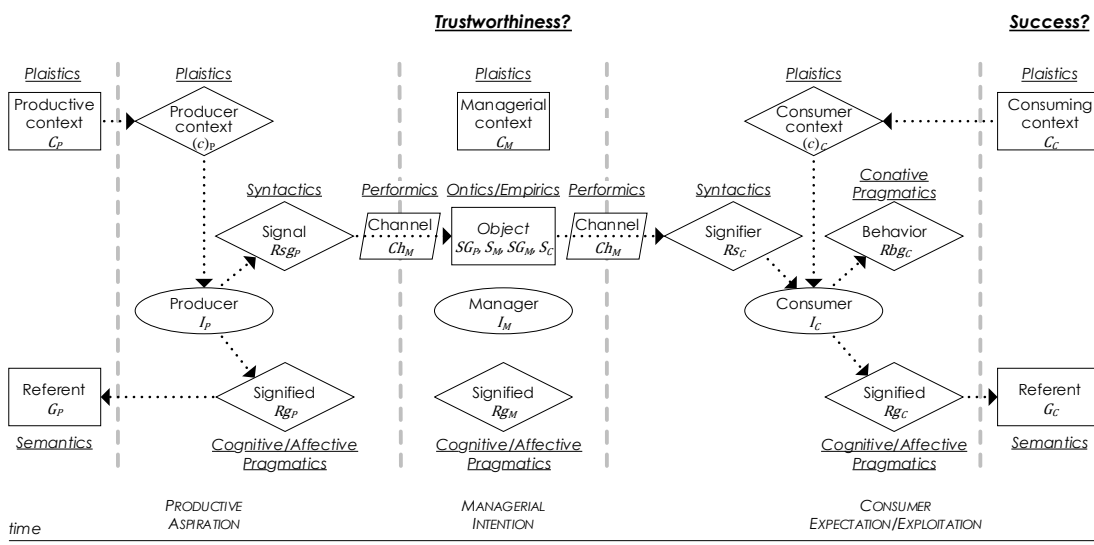


Figure 3.6. Semiotic model of digital preservation

Cf. *Figure 2.2* regarding the OAIS-based model of preservation communication

Consistent with the ECT-informed explication of digital preservation success proposed in Section § 1.2, that success is dependent upon the alignment of productive aspirations, managerial intentions, and consumer expectations. That alignment is determined with regard to the various preserved states of the underlying digital object as it passes through the semiotic process. Full alignment could be ensured if, naively, the final consumed object was identical to the intermediate managed object and the originally produced object, i.e.,  $S_C = S_{GM} = S_M = S_{GP}$ . However, that situation is unlikely to predominate over archival timespans given persistent incremental technological innovation and periodic disruptive transformation. In response, it is prudent to assume some form of migration or emulation intervention (Strodl et al., 2007).

In the first case, there will not be a singular managerial object, but rather, a multiplicity of objects over time, each derived from its predecessor in a manner avoiding or ameliorating potential risk of contemporaneous damage or loss. In the second, there is a multiplicity of behavioral platforms for performing a singular object over time. In both cases, however, consideration over ever-increasing archival time horizons increases the introduction of accumulating subtle or overt differences in expressive representation (Day, 2002), behavioral experience (Hedstrom et al., 2006), and eventual pragmatic response. Therefore, the focus of digital preservation effort is more properly aimed at ensuring the approximate but appropriate pragmatic *equivalence* – but not necessarily the exact ontic, empiric, and syntactic *equality* – of the three associated signified states, i.e.,  $R_{gc} \cong R_{gm} \cong R_{gp}$ . “Equivalence” is used hereinafter in the Fregean sense of being freely interchangeable without loss of conceptual integrity (May, 2001).<sup>13</sup> That is, an equivalence relation holds when “The sign A and the sign B have the *same conceptual content*, so that everywhere we can put B for A and conversely [emphasis added]” (Frege, 1879, §8, p. 15; Weiner, 2004). This formulation, with its emphasis on epistemological and phenomenological equivalence rather than ontological equality, provides the primary basis for the

---

<sup>13</sup> Frege’s symbology for what he called an identity, rather than equivalence, relation used the triple bar symbol, “≡”. Since the common meaning of “identity” doesn’t capture Frege’s nuance of *conceptual* substitutability, the symbol “≅” is used instead. As defined as codepoint U+2245 in the Unicode standard (Unicode, 2021), this symbol indicates an APPROXIMATELY EQUAL TO relation, which comports better with the conceptual dimension of measure.

assessment of the suitability of identified tacit evaluative norms to meaningfully characterize digital preservation success.



# Chapter 4: Predicate Reduction Analysis

---

Four primary evaluative norms – accessibility, integrity, authenticity, usability (AIAU)<sup>14</sup> – emerge from subjecting representative digital preservation policy statements to Predicate Reduction analysis and synthesis. These implicate service-provider assurances regarding the archival qualities of digital materials under proactive preservation stewardship. This chapter documents the application of the PR method, the procedural derivation of those four norms, and the data management framework representing the quantitative results. Subsequent qualitative Communicological analysis regarding the suitability of the emergent norms as the basis for effective and operationalizable metrics of digital preservation success is provided in Chapter 5.

## 4.1 DATA PROCESSING

As described in Section § 3.1.3, the PR technique systematically transforms obligatory policy statements into kernel expressions of core evaluative intentions, expectations, and criteria. Applying PR against the six policy documents paradigmatically selected in Section § 3.1.1 results in the identification of 266 statements expressing relevant obligatory service-provider intentions. These statements are found within the specific structural contexts of 104 topically named or numbered sections. The statements are often complex or coordinated in nature (see Section § 3.1.2, Step 2), and a single statement may expand into multiple propositions. Overall, the 266 statements contribute 543 individual propositions for subsequent PR processing (see Table 4.1). The propositional counts tally the original – or only – proposition derived from statements as well as any expanded propositions. Thus, a statement with *two* expansion increments the tally by *three* – the original plus the two expanded propositions.

---

<sup>14</sup> For declamatory purposes, the AIAU acronym can be pronounced “EYE-oh” /'ai.ou/, as in the common English usage for the name of the moon of Jupiter, Io. Cf. note 22, p. 100.

Table 4.1

*Distribution of Propositional Expansion*

Statements		Expansions	Propositions	
122	45.0%	0	122	22.5%
70	26.3%	1	140	25.8%
45	16.9%	2	135	24.9%
15	5.6%	3	60	11.0%
4	1.5%	4	20	3.7%
6	2.3%	5	36	6.6%
2	1.5%	6	14	2.6%
2	1.5%	7	16	2.9%
266	100%		543	100%

In just under half the statement cases (122 of 266, or 45.0%), there is a one-to-one relationship between the statement and proposition (see, for example, Table 4.2).

Table 4.2

*Single Statement Leading to Single Predicate*


---

*Statement:* “These digital assets are an essential component of the overall institutional strategy and the BMA is dedicated to their preservation” (BMA, 2016).

→ *Proposition:* “BMA is dedicated to preservation of digital assets”

→ *Analytic Predicate:* “preserve digital assets”

→ *Canonical Predicate:* “preserve objects”

---

For the remaining 144 statements, the maximum degree of statement-to-proposition expansion is 7, the mean degree of expansion is 1.92, the median is 2, and the standard deviation is 1.28. This expansion contributes 421 (77.5%) of the total 543 propositional instances. In general, the counts of propositional instances *decrease* as the degree of expansion *increases*. The majority of expanded propositions (275 of 421, or 65.3%) result from one or two degrees of expansion. Thus, the policy documents appear to conform with established compositional best-practice guidance for expository writing that deprecates the use of extensive coordinated, run-on, and fused sentences (Butler, 2021).

Statement-level predicate duplication (see, for example, Table 4.3) contributes 73 of the total 543 predicate instances (13.4%). The maximum number of predicates added through statement-level duplication is 2, the average degree of statement-level duplication is 1.17, the median is 1, and the standard deviation is 0.37. This indicates that the majority of cases of statement-level predicate duplication (20 or 24, or 83.3%) result from a single degree of duplication.

Table 4.3

*Single Statement Leading to Multiple Predicates*

---

*Statement:* “The program will strive to care for both born-digital and digitized material throughout the lifecycle of the digital asset, *maintaining the intellectual property rights of creators and copyright holders* [emphasis added]” (BMA, 2016)

→ *Proposition:* “program will maintain intellectual property rights of creators”

→ *Analytic Predicate:* “maintain intellectual property rights”

→ *Canonical Predicate:* “ensure IPR”

→ *Proposition:* “program will maintain intellectual property rights of copyright holders”

→ *Analytic Predicate:* “maintain intellectual property rights”

→ *Canonical Predicate:* “ensure IPR”

---

Similar predicate duplication occurs at the level of structural contexts, that is, named and/or numbered policy document sections. For most contexts (59 of 104, or 70.2%), *multiple* instances of the *same* canonical predicate are derived from *multiple* independent statements embedded within a single context (see, for example, Table 4.4). Contextual-level predicate duplication contributes 116 of the total 543 predicate instances (21.3%). The maximum number of duplicated predicates in any given context is 6, the population mean is 1.24, the median is 1, and the standard deviation is 0.74. This indicates that the majority of cases of context-level predicate duplication (50 of 59, or 84.7%) result from a single degree of duplication.

It is reasonable to expect some degree of local predicate duplication in view of a number of factors. For example, the presumed topical coherence of any given

statement or structural context; intentional or tacit conformance to rhetorical convention; or simple inadequate copyediting. Thus, any consequently-inflated predicate counts would not provide compelling evidence regarding the relative significance of the duplicated predicate obligations. On the other hand, duplication *across* structural contexts can be assumed to be more topically uncorrelated. In view of this, the counts of these duplicative instances can be assumed to be reasonably indicative of the broader policy-wide importance of the underlying evaluative norms. Thus, subsequent analysis is based only on the counts of unique-to-context canonical predicates.

Table 4.4

*Multiple Statements in Single Structural Context Leading to Multiple Predicates*

---

*Statement:* “Metadata is created and/or transformed to meet relevant standards”  
(CUL, 2018, § 3.2.7).

→ *Proposition:* “Metadata is transformed to meet standards”

→ *Analytic Predicate:* “[preserve] metadata”

→ *Canonical Predicate:* “preserve metadata”

*Statement:* “Digital content created by CUL always has accompanying standards-based metadata created” (CUL, 2018, § 3.2.7)

→ *Proposition:* “Digital content has accompanying standards-based metadata”

→ *Analytic Predicate:* “[preserve] metadata”

→ *Canonical Predicate:* “preserve metadata”

*Statement:* “In order for digital content to be acquired by CUL, it must be accompanied by a minimum amount of metadata” (CUL, 2018, § 3.2.7)

→ *Proposition:* “Digital content must be accompanied by metadata”

→ *Analytic Predicate:* “[preserve] metadata”

→ *Canonical Predicate:* “preserve metadata”

---

For example, seven instances of the predicate “preserve metadata” are derived from Section § 3.3.2, *Ingest*, of the Nationaal Archief’s policy document (NA, 2015, § 3.3.2). For tallying and analytical purposes, however, this counts as a *single* unique-to-context predicate. All six policy documents show consistent relative proportions of

the number of predicates accounted for in each of the three tallying categories: total count of predicates, the count of unique-to-statement predicates, and the count of unique-to-context predicates (see Tables 4.8 and 4.9). Thus, the choice of the unique-to-context count does not negatively bias analytic integrity.

## 4.2 DATA MANAGEMENT

The PR analysis dataset (Abrams, 2022a) is represented by 11 tabular data files (see Table 4.5). These document the initial, intermediate, and final processing stages producing relative frequency rankings of evaluative norms for digital preservation success tacitly unpinning representative policy documents.<sup>15</sup> The files' organization is described in Section § 4.2.1 and their derivation in Appendix § B. File information content is a combination of literal values for obligatory policy statements, propositions, and analytic and canonical predicates resulting from application of the Predicate Reduction technique to the six policy documents as well as summary statistics, including token and type counts, expansion and duplication metrics, and kernel frequency rankings calculated automatically through formulas.

Table 4.5

*Predicate Reduction Dataset Files*

Data File	Data
1. Analysis_1-raw	Data as encountered the natural reading order
2. Analysis_3-propositions_d	Data sorted by propositions per-document
3. Analysis_3-propositions_t	Data sorted by propositions across documents
4. Analysis_4-a-predicates_d	Data sorted by analytic predicates per-document
5. Analysis_4-a-predicates_t	Data sorted by analytic predicates across documents
6. Analysis_5-c-predicates_d	Data sorted by canonical predicates per-document
7. Analysis_5-c-predicates_t	Data sorted by canonical predicates across documents
8. Analysis_6-kernels_d	Data sorted by evaluative kernels per-document
9. Analysis_6-kernels_t	Data sorted by evaluative kernels across documents
10. Analysis_7-rankings_d	Data sorted by frequency rankings per-document
11. Analysis_7-rankings_s	Data sorted by frequency rankings across documents

<sup>15</sup> The Predicate Reduction dataset is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <<https://doi.org/10.17605/OSF.IO/75Q29>>.

By sorting on various key fields, the proposition, predicate, kernel, and ranking data files present specific summary statistics about the full dataset on a per-document and document-spanning basis (the “\_d” and “\_t” suffixed data file names, respectively). For example, the file “Analysis\_4-a-predicates\_d” calculates counts of lexicographically-sorted analytic predicates individually for each policy document, while “Analysis\_5-c-predicates\_t” calculates analogous canonical predicate counts across all six policy documents.

#### 4.2.1 Data File Structure

The data files all share the same internal structure. Their 72 columnar fields are organized into nine thematic groups aligned with the various Predicate Reduction processing steps:

1. **Document** group, containing institutional names and associated sampling unit values. The various managerial unit values used in subsequent analysis – sampling as well as context, coding, and reporting – are defined in Section § 4.2.2. All units are assigned on both a per-document and document-spanning basis.
2. **Context** group, containing contextual section titles and page numbers and associated context unit values for identified obligatory statements.
3. **Statement** group, containing obligatory statements (resulting from Predicate Reduction Step 1) and associated coding unit values and propositional expansion metrics.
4. **Propositions** group, containing expanded propositions (PR Step 2) and associated reporting units and propositional token and type counts. The distinction between token and type is described in Section § 4.2.3.
5. **Analytic Predicates** group, containing reduced analytic predicates (PR Step 3) and associated counts and frequency rankings.
6. **Canonical Predicates** group, containing canonicalized predicates (PR Step 4) and associated counts, unique-to-statement and unique-to-context duplication metrics, and frequency rankings.
7. **Synthetic Kernels** group, containing synthetic intentional, expectational, and evaluative kernels (PR Step 5) and associated token

and type counts and frequency rankings.

8. **Evaluative Kernels Unique-to-Statement** group, containing unique-to-statement evaluative kernels and associated token and type counts and frequency rankings.
9. **Evaluative Kernels Unique-to-Context** group, containing unique-to-context evaluative kernels and associated frequency token and type counts and rankings.

More detailed field-level definitions are found in the PR dataset's accompanying codebook.<sup>16</sup>

The evaluative kernel frequency rankings in Group 7 are calculated in terms of all kernels synthetically-derived from the six policy documents. The unique-to-statement frequency rankings in Group 8 are calculated in terms of only those kernels unique to the statements from which they are derived. This documents the cases where propositional expansion and predicate canonicalization (Steps 2 and 4 of PR) lead to multiple identical evaluative kernels from a given policy statement. The unique-to-context frequency rankings in Group 9 are calculated in terms of only those kernels unique to the structural context in which they are found, that is, a named and/or numbered document section. The contextually-unique rankings are the basis for subsequent analysis.

#### **4.2.2 Managerial Units**

In Content Analysis, pertinent data elements are assigned managerial unit numbers to provide unambiguous identification and reference. Sampling units are those items selected for review, coding units are those more granular items significant for analytic purposes, context units are those providing the structural setting in which the coding units are found, and reporting units are those more granular elements fully described as analytic outputs (Krippendorff, 2019; Schreier, 2013). For PR analysis, the sampling units are the six selected policy documents. Coding units are the individual policy statements identified in Step 1 of PR. Context units are the named and/or numbered document sections in which, and page numbers on which, those

---

<sup>16</sup> The codebook is available in Word (.docx) and PDF (.pdf) format at <<https://doi.org/10.17605/OSF.IO/75Q29>>.





### 4.3 QUANTITATIVE RESULTS

The six institutional policy documents contribute 266 statements of intentional digital preservation obligation in 104 named and/or numbered structural contexts (see Table 4.6). These statements are often complex or coordinated, so they expand into 543 individual propositional tokens. (There are also 543 analytic predicate and 543 canonical predicate tokens. For brevity, these token counts are not repeated in Table 4.6; only their type counts are reported.) Of the 543 propositions, 432 are unique propositional types with respect to the document in which they are found. That is, each of these 465 propositional instances is found only once in its document. As expected, predicate reduction and canonicalization produces a significant consolidation of analytic predicates and a consequent lowering of analytic type counts relative to those for expanded propositions. Analytic type counts are lower by -23.8% (i. e.,  $\frac{329-432}{432} \cdot 100$ ) relative to propositional counts. Similarly, canonical type counts are lower by -69.9% ( $\frac{99-329}{329} \cdot 100$ ) relative to analytic types, and -77.1% ( $\frac{99-432}{432} \cdot 100$ ) relative to propositional types.

Table 4.7

*Policy Contexts, Statements, Propositions, and Predicates Across Dataset*

*Cf. Table 4.6*

Scope	Contexts	Statements	Propositions		Predicates		
			Tokens	Types	Analytic Types	Canonical Types	
Dataset	104	266	543	432	241	28	
					$\Delta$ relative to Proposition type count =	-44.2%	
					$\Delta$ relative to Proposition type count =		-93.5%
					$\Delta$ relative to Analytic Predicate type count =		-88.4%
					$\Delta$ relative to per-document counts in Table 4.6 =	-26.7%	-71.7%

In the context of the counts across all six policy documents in the full dataset (see Table 4.7), there is again consolidation of analytic predicate type counts relative to propositions and canonical predicate type counts relative to analytic predicates and propositions, with the tallies lowered by -44.2% (i. e.,  $\frac{241-432}{432} \cdot 100$ ), -88.4% ( $\frac{28-241}{241} \cdot 100$ ), and -93.5% ( $\frac{28-432}{432} \cdot 100$ ), respectively. Proposition token and type

counts are identical when summed from the individual per-document totals (as shown in Table 4.6) as opposed to being directly summed across all documents. However, because the same analytic and canonical predicate types appear in multiple documents, the global counts for these predicate metrics ungrouped by document are less than the sum of the local per-document counts. The analytic predicate type count across the full dataset is lower by -26.7% (i. e.,  $\frac{241-330}{330} \cdot 100$ ) relative to the per-dataset sum. The canonical type count is similarly, though more significantly, lower by -71.7% ( $\frac{28-99}{99} \cdot 100$ ).

Table 4.8

*Policy Contexts, Statements, and Evaluative Kernels by Document*

*Cf. Table 4.9*

Scope	Contexts		Statements		Evaluative Kernels							
					All Kernels				Unique-to-Statement		Unique-to-Context	
					Tokens	Types	Tokens	Types	Tokens	Types		
BMA	5	4.8%	27	10.2%	77	14.2%	17	17.2%	61	13.9%	57	12.8%
CUL	22	21.2%	51	19.2%	135	24.9%	17	17.2%	107	23.7%	106	24.2%
ICPSR	12	11.5%	30	11.3%	54	9.9%	11	11.1%	49	10.0%	44	10.2%
NA	23	22.1%	77	28.9%	135	24.9%	20	20.2%	122	25.4%	111	24.6%
NLNZ	27	26.0%	54	20.3%	91	16.8%	16	16.2%	87	17.4%	82	17.9%
ZBW	15	14.4%	27	10.2%	51	9.4%	18	18.2%	48	9.6%	44	10.0%
<i>N =</i>	104	100%	266	100%	543	100%	99	100%	474	100%	444	100%

$$\Delta N \text{ relative to All-Kernel token count} = -12.7\%$$

$$\Delta N \text{ relative to All-Kernel token count} = -40.0\%$$

$$\Delta N \text{ relative to Unique-to-Statement token count} = -31.2\%$$

In Content Analysis, the frequency of occurrence of a reporting unit can be assumed to be a reliable proxy for conceptual significance (Krippendorff, 2019; Stemler, 2001). (Discussion of the basis for this assumption is found in Section § 3.1.4.) Thus, frequency rankings of the synthetic evaluative kernels derived from policy obligations are used as a proxy for broad, if tacit, community understanding of the relative importance of those kernel’s underlying evaluative norms. The counts of these kernels are calculated across all those derived from the six policy documents, those uniquely derived from their underlying statements, and those uniquely derived from their underlying structural contexts (see Table 4.8). The kernel type counts are

identical for All-Kernel, Unique-to-Statement, and Unique-to-Context tallying categories. Consequently, these values are presented in Table 4.8 once for the All-Kernel category and then not repeated.

Given Predicate Reduction’s one-to-one mapping of canonical predicates and synthetic evaluative kernels, token and type counts of those two tallying categories are identical. As expected, the unique-to-statement kernel token counts are lower by -12.7% (i.e.,  $\frac{474-543}{474} \cdot 100$ ) relative to the count of all kernels. Similarly, unique-to-context kernel type counts are lower by -31.2% ( $\frac{444-474}{474} \cdot 100$ ) relative to unique-to-statement types, and -40.0% ( $\frac{444-543}{543} \cdot 100$ ) relative to all kernel types. However, the relative proportions of the per-document token counts are consistent across the three token tallying categories. Thus, the choice of the unique-to-context kernel token counts for subsequent calculation of frequency rankings does not negatively bias analytic integrity.

Table 4.9

*Policy Contexts, Statements, and Evaluative Kernels Across Documents*

*Cf. Table 4.8*

Scope	Contexts	Statements	Evaluative Kernels			
			All Kernels		Unique-to-Statement	Unique-to-Context
			Tokens	Types	Tokens	Types
Dataset	104	266	543	28	474	326
					$\Delta$ relative to All-Kernel type count =	-12.7%
					$\Delta$ relative to All-Kernel type count =	-40.0%
					$\Delta$ relative to Unique-to-Statement type count =	-31.2%
					$\Delta$ relative to per-document counts in Table 4.8 =	-71.7%
						0.0%
						0.0%

The global token count for all evaluative kernels is identical when summed from the individual totals per-document (as shown in Table 4.8) as opposed to being summed directed across all documents (see Table 4.9). However, because most kernel types are shared by the six documents, the total all-kernel type count directly tallied across all documents is significantly lower, -71.7% (i.e.,  $\frac{28-99}{99} \cdot 100$ ), than the sum of the individual per-document totals.

The 28 evaluative kernel types correspond to evaluative norms implicitly defined by the policy document sample set.<sup>17</sup> Nine of these 28 norms (32.1%) are referenced in at least five of the six policies, and 15 (53.5%) in at least four (see Table 4.10).

Table 4.10

*Frequency Ranking of Evaluative Norms Across and By Document*

Evaluative Norm	Token Counts and Frequencies							
	Dataset	BMA	CUL	ICPSR	NA	NLNZ	ZBW	
Preserve objects	57 17.5%	4 11.1%	14 19.2%	10 31.3%	10 12.2%	14 21.5%	5 13.2%	
Ensure accessibility	56 17.2%	5 13.9%	12 16.4%	8 25.0%	16 19.5%	9 13.8%	6 15.8%	
Preserve metadata	31 9.5%	3 8.3%	9 12.3%	2 6.3%	9 11.0%	5 7.7%	3 7.9%	
Ensure integrity	26 8.0%	3 8.3%	6 8.2%	1 3.1%	6 7.3%	9 13.8%	1 2.6%	
Ensure authenticity	25 7.7%	2 5.6%	2 2.7%	2 6.3%	10 12.2%	8 12.3%	1 2.6%	
Ensure usability	23 7.1%	2 5.6%	3 4.1%	2 6.3%	8 9.8%	3 4.6%	5 13.2%	
Ensure IPR	15 4.6%	4 11.1%	5 6.8%		3 3.7%	1 1.5%	2 5.3%	
Ensure security	12 3.7%	1 2.8%	2 2.7%	3 9.4%	2 2.4%	3 4.6%	1 2.6%	
Ensure provenance	9 2.8%	2 5.6%	3 4.1%	—	1 1.2%	2 3.1%	1 2.6%	
Preserve bitstreams	7 2.1%	1 2.8%	—	—	2 2.4%	3 4.6%	1 2.6%	
Preserve orig. objects	6 1.8%	—	—	1 3.1%	1 1.2%	2 3.1%	2 5.3%	
Preserve descr. metadata	6 1.8%	1 2.8%	2 2.7%	—	1 1.2%	—	2 5.3%	
Preserve derivatives	5 1.5%	—	—	1 3.1%	1 1.2%	1 1.5%	2 5.3%	
Preserve orig. bitstreams	5 1.5%	1 2.8%	—	—	1 1.2%	2 3.1%	1 2.6%	
Preserve PIDs	4 1.2%	—	1 1.4%	1 3.1%	1 1.2%	—	1 2.6%	
Other (13)	39 12.0%	7 19.4%	14 19.2%	1 3.1%	10 12.2%	3 4.6%	4 10.5%	
<i>N</i> =	326 100%	36 100%	73 100%	32 100%	82 100%	65 100%	38 100%	

<sup>17</sup> The Predicate Reduction dataset is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <<https://doi.org/10.17605/OSF.IO/75Q29>>.

Those 15 norms are manifest in those documents as 287 (88%) of the 326 instances of unique-to-context evaluative norms. The other 13 norms with 39 instances constitute the remaining 12%. When ranked in decreasing order of frequency, only the topmost six of the 28 norms (21.4%) are referenced in all six documents with a global dataset count greater than 5% of the total. The “Ensure security” norm also is found in all six documents, but its global count is only 3.7%. Since none of the remaining 22 norms are uniformly referenced across the sample set in significant number, they are not considered broadly reflective of evaluative positions in the community. As such, they are excluded from subsequent analysis.

Table 4.11

*Frequency Ranking of Primary and Total Evaluative Norms*

Evaluative Norm	Tokens	Per $N_P$	Per $N_T$
Ensure accessibility	56	43.1%	17.2%
Ensure integrity	26	20.0%	8.0%
Ensure authenticity	25	19.2%	7.7%
Ensure usability	23	17.7%	7.1%
<hr/>			
$N_P =$	130	100.0%	39.9%
Other	196		60.1%
<hr/>			
$N_T =$	326		100.0%

Of the six universal norms, two (or 33.3% of the remaining set) refer to assurances with respect to high-level generic entities: “preserve objects” and “preserve metadata”. In both instances, the predicating verb “preserve” leads to a tautological statement regarding a generic preservation obligation to preserve. As this does not offer practical detail regarding evaluation of the preservation task, these norms also are excluded from further analytic consideration. The four remaining predicates (66.7%) are defined with respect to more specific archival characteristics, i.e., accessibility, integrity, authenticity, and usability (AIAU). The distinction between the quality-based norms and entity-based norms is significant. The former are useful to provide more granular definitional detail as to the meaning of the latter. In essence, the metrics illuminate specific constitutive aspects of the preservation of objects and metadata central to the evaluation of the outcome of the preservation act. Consequently, these four quality-based norms are considered representative of the

primary evaluative positions held by the digital preservation community regarding the evaluation of preservation success.

The total count of evaluative norms  $N_T$  is 326. Instances of the four primary norms constitute less than half of that total (130 of 326, or 39.9%) The “Ensure accessibility” norm contributes just over one-sixth of all instances (56 of 326, or 17.2%). The other three primary norms together contribute 22.8% (26 + 25 + 23 = 74 of 326), while the remainder (196 or 326, or 60.1%) are non-primary (“Other”). Considering the norm counts in the context of only the four primary norms with total count  $N_P$  of 130, the accessibility norm contributes less than half of the instances (56 of 130, or 43.1%), while the other three contribute the remaining 74 (56.9%). The accessibility norm’s instances are 2.15 to 2.43 times more prevalent than those of the other three norms, suggesting a significant degree of relative evaluative importance (see Table 4.12).

Table 4.12  
*Relative Frequency of Primary Evaluative Norms*

Evaluative Norm	Tokens	Relative to accessibility
Ensure accessibility	56	
Ensure integrity	26	46.4% = 26/56 = 1 / 2.15
Ensure authenticity	25	44.6% = 25/56 = 1 / 2.24
Ensure usability	23	41.1% = 23/56 = 1 / 2.43
$N_P =$		130

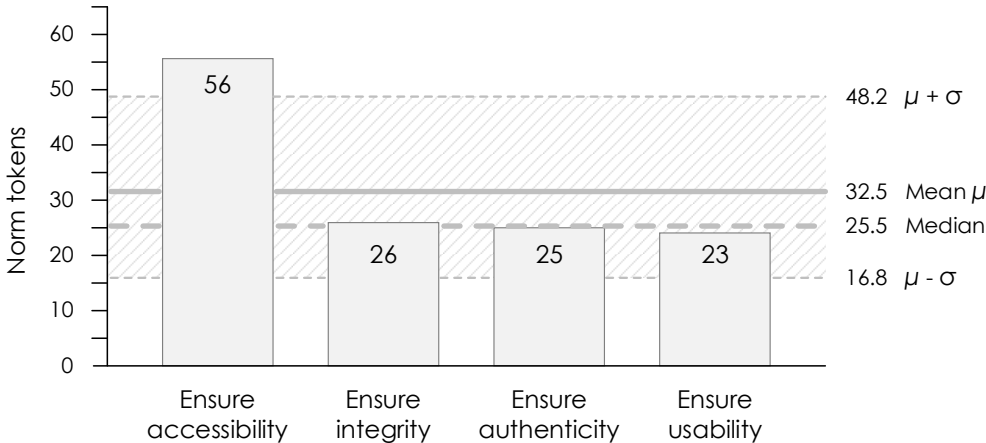


Figure 4.1. Distribution of evaluative norm tokens

The Predicate Reduction technique results in non-trivial data reduction reflected

in the tallies of evaluative norm token counts while progressing from consideration of all evaluative kernels to those unique-to-context, which correspond to the established norms (see Table 4.13). There is a 12.7% reduction in the number of unique-to-statement instances relative to the full set (i.e.,  $\frac{474-543}{543} \cdot 100$ ), a 31.2% reduction in the number of unique-to-context instances relative to unique-to-statement ( $\frac{326-474}{474} \cdot 100$ ), and 40.0% reduction of unique-to-context relative to the full set ( $\frac{326-543}{543} \cdot 100$ ).

Table 4.13

*Data Reduction by Evaluative Norm*

Evaluative Norm	Evaluative Kernel Token Counts					
	All		Unique-to-Statement		Unique-to-Context	
Ensure accessibility	104	19.2%	91	19.2%	56	17.2%
Ensure integrity	62	11.4%	50	10.5%	26	8.0%
Ensure authenticity	38	6.4%	34	7.2%	25	7.7%
Ensure usability	40	7.4%	30	6.3%	23	6.1%
Other	302	56.6%	269	56.8%	196	56.1%
<i>N =</i>	543	100%	474	100%	326	100%
	$\Delta$ relative to All-Kernel token count =		-12.7%		-40.0%	
	$\Delta$ relative to Unique-to-Statement Kernel token count =				-31.2%	

The advantage of this data reduction is illustrated by the fact that the six policy documents articulate 32 distinct analytic predicate type variations that eventually coalesce into the single canonical predicate – and eventual unique-to-context kernel and evaluative norm – of “ensure accessibility”. The PR reduction and canonicalization steps simplify data management and analysis by aggregating these variations into a smaller set of standardized normative categories (see Table 4.14).

Table 4.14

*Analytic-to-Canonical Predicate Consolidation*

Evaluative Norm	Predicate Types		Data Reduction
	Canonical	Analytic	
Ensure accessibility	1	32	96.9%
Ensure integrity	1	32	96.9%
Ensure authenticity	1	19	94.7%
Ensure usability	1	22	95.5%
	4	105	96.2%

For example, it permits analysis of the single norm “ensure accessibility” in place of independent analysis of all of the variant analytic predicate forms (see Table 4.15).

Table 4.15  
*Example Consolidation of Analytic and Canonical Predicates*

Analytic Predicates		Canonical Predicate
check access	check availability	} ensure accessibility
deliver content	deliver digital content	
distribute assets	enable access	
enable accessibility	enable location	
enable retrieval	enable search	
ensure access	ensure accessibility	
ensure availability	ensure delivery	
ensure disbursement	ensure identification	
ensure release	ensure retrievability	
exchange content	facilitate accessibility	
give access	keep accessible	
maintain access	make available	
manage access	prevent disappearance	
provide access	provide content	
provide discovery	provision access	
support access	support accessibility	

The four primary qualitative norms – accessibility, integrity, authenticity, usability – represent consensus evaluative attitudes tacitly underpinning digital preservation service-provider/stakeholder relationships, as articulated indirectly through expectation- and intention-setting obligatory policy statements. While they should form the basis for a viable evaluative framework and operationalizable metrics characterizing the experiential success of human engagement with preserved digital objects, as discussed in Section § 2.4, this is not the case in practice. Chapter 5 investigates the possible reasons for the lack of accepted evaluative criteria as well as the suitability of these norms to characterize digital preservation success effectively.



# Chapter 5: Communicological Analysis

---

This chapter submits each of the four primary evaluative norms identified in Chapter 4 to Communicological analysis (Sections § 5.1 – 5.4). Section § 5.5 then discusses their suitability to function variously as norms, criteria, and metrics of digital preservation success. Section § 5.6 promotes extending current evaluative practice’s emphasis on artifactual properties to embrace communicative affordances. Finally, Section § 5.7 describes a new evaluative regime based upon a multi-valent measure of success.

Critical examination of the four identified evaluative norms for digital preservation success – assurances regarding the accessibility, integrity, authenticity, and usability (AIAU) of preserved objects – relies upon a Communicological perspective that views digital preservation as an intersubjective act of technically-mediated human communication unfolding against the passage of time and ever-accumulating technical and cultural distance. While accessibility, integrity, and authenticity are well-established archival concerns, they are most applicable to characterizing programmatic digital object management and individual managed digital objects. These managerial qualities are necessary normative components for evaluating preservation-enabled communication, particularly regarding assessment of its programmatic and artifactual trustworthiness. However, they are not fully sufficient for the meaningful teleological characterization of the successful purposive use of programmatically-preserved digital objects.

Table 5.1  
*Normative Categorization*

Norm	Conceptual framing	Focus	Measure	Benchmark	Applicability
Accessibility					
Integrity	Managerial	Artifactual	Objectively quantifiable	Definitive	Universal
Authenticity					
Usability	Communicative	Experiential	Intersubjectively qualifiable	Relative	Situational

The three managerial qualities – accessibility, integrity, authenticity – are descriptive of objects as independent ontological entities, rather than intersubjective constituents of relational epistemic and responsive phenomenological processes. That is, they quantify what an object *is* rather than qualifying what it enables its consumer to *do* and *know* (see Table 5.1). Thus, they are pertinent to evaluative questions such as: Is this object susceptible to retrieval? Is this object whole and uncorrupted from its accepted form? Is this object what it purports to be? Being inherent to the fabric of the *object*, these qualities, and their underlying evidential facts and implications, are essentially *objective* in nature. As such, they are definitive in determination and universal in applicability. That is, for any reasonable preservation stakeholder, an object either is or is not accessible, integral, or authentic. By definition, an object that cannot be fully retrieved is not accessible, one that is not entirely whole is not integral, and one that is not fully true regarding its claimed substance is not authentic. What these norms don't address, however, are the implications of these artifactual determinations within a broader social environment of relational agency, intention, expectation, and action. Those communicative considerations fall under the purview of usability. However, usability of managed objects remains an under-defined concept in domain discourse and practice (Abrams, 2021; Hirtle, 2008; Ross, 2012)

The normative managerial and communicative qualities are distinct in nature and descriptive power but complementary in result. Plausible assertions of integrity and authenticity increase the degree of confidence that an artifactual vehicle is acceptable for subsequent consumer use (Ross, 2006). Accessibility also empowers effective user agency regarding the conditions and contexts of that use (Menne-Haritz, 2001), providing users with autonomy of information-seeking and exploitive meaning-making. However, none of the three managerial norms address the experiential conditions of that consuming activity. Instead, they are limited in characterization to the monadic qualities of a preserved object-in-itself. The quality of usability, on the other hand, is descriptive of the triadic relation between objects, human users of those objects, and the intersubjective contexts of those uses. Communicatively, the fundamental evaluative question is: Is this *object* meaningful to the purpose of this *user* in this contingent *situation*?

To encompass beneficial serendipity, the parameters of that usage should be characterized in terms of purposiveness rather than purposefulness. The latter asserts

an affirmative prospective *intent*, while the former indicates retrospective recognition that the use *contributed*, whether by design or accident, to fulfilment of a meaningful contextual purpose. Thus, an evaluative basis of purposiveness supports the widest range of information-seeking and consuming behaviors. It also emphasizes the contingent nature of those behaviors as being specific to an individual user and inherently positioned with regard to the context and modality of behavioral use. Similarly, it accepts the potential informational and experiential gap that may separate productive intention from consuming expectation, and that expectation from possibly variant actualization.

Sections § 5.1 – 5.4 examine each of the four evaluative norms regarding their strengths and weaknesses for characterizing digital preservation success from a Communicological perspective. Next, Section § 5.5 looks at the suitability of these norms as the basis for actionable evaluative criteria and metrics and proposes an explanation of why benchmarks for preservation success have not yet been widely accepted or operationalized within the preservation community. Section § 5.6 discusses a more effective approach based on refactoring the current concept of significant properties as significant affordances. Finally, Section § 5.7 presents the implications of this analysis and a set of recommendations for how meaningful assessment of success can be incorporated into digital preservation theory and practice.

## **5.1 ACCESSIBILITY**

Although the four identified evaluative norms are referenced in varying degrees throughout the digital preservation literature, they are generally not accompanied with formal definitions. However, various reference works for the domain do indicate the range of their meanings. For the past 20 years, the InterPARES project has investigated the challenges of “reliable, accurate, and authentic digital records” (Duranti, 2007, p. 113). One of the project deliverables is a Glossary of digital archival terminology (InterPARES, 2008). This Glossary defines accessibility in terms of the twin qualities of availability and usability of information. However, the Glossary does not define either of those subordinate concepts. The cognate concept of *access* is defined, but with a strong instrumental emphasis as the “right, opportunity, or means of finding, using or approaching documents and/or information.” In this case, access implies subsequent usage, although the meaning of “use” is itself not formally propounded. On the other hand, access *privileges* are defined in terms of authority to

“compile, classify, register, retrieve, annotate, read, transfer or destroy” an information record. While the Glossary elsewhere distinguishes between human and machine readability, it is not clear in which sense “read” is intended in the definitional context of access or use.

The complementary InterPARES Dictionary (InterPARES, 2022) provides more detailed delineation of the meaning of “access”, again focusing on its instrumental characteristics. For example, access encompasses “permission to locate and retrieve information for use” and the “ability to locate, gain entry to, and use something, such as a building or a database”.<sup>18</sup> Only the latter entry expresses a sense of actual use; the former merely asserts access as a prerequisite for subsequent access. Of the 11 variant definitions of access in the Dictionary, only three – a restatement of the Glossary entry as well as the two definitions quoted above – mention use as a definitional component or imply exploitive use as synonymous with or a possible consequence of access. The other eight relegate access to concerns of discovery and retrieval as *enabling* managerial conditions distinct from considerations of actual consummating use. As such, accessibility is primarily conceptualized as a characteristic of a preserved digital object as an artifactual vehicle rather than its consuming experience.

A Communicological perspective of the preservation domain leads to explicit consideration of both the preserved digital object as an expressive semiotic carrier and the subsequent phenomenological reception of and response to that object by its user. This artifactual/experiential distinction is explicated by reference to the preservation-augmented semiotic ladder and preservation model introduced in Section § 3.2 (see *Figures 3.4* and *3.6*). Access in its purest instrumental sense of simple physical custody of a retrieved object corresponds to possession of the ontic, i.e., physical, manifestation of that object. At minimum, this entails an internal bitstream, as made accessible, for example, through a bitstream reader such as HexDump (FileFormat.Info, 2022) (see *Figure 5.1(a)*). Dependent upon the technical environment of access, accessibility may also encompass accompanying file-level properties such as name, location, and size. However, a fully accessible, but otherwise opaque bitstream is unlikely to be sufficient for fulfilling all possible consumable

---

<sup>18</sup> [http://www.interpares.org/ip2/display\\_file.cfm?doc=ip2\\_dictionary.pdf&CFID=28189515](http://www.interpares.org/ip2/display_file.cfm?doc=ip2_dictionary.pdf&CFID=28189515)

purposes.


11111111	SOI	799 x 1086	Half-length,		Documentary evidence of late 19th-century social and artistic feminism
11011000	APP0 JFIF 1.1	1 x 1 density	three-quarter		
11111111	DQT Luma	YCbCr	view of		
11100000	DQT Chroma	color	seated		
00000000	SOF0	28:1	woman with		
00010000	DHT DC	compression	head resting		
01001010	DHT AC		on upraised		
01000110	DHT DC		hand of right		
01001001	DHT AC		arm draped		
01000110	SOS 3		across		
000000...	EOI		chairback		
Ontic manifestation	Empiric representation	Syntactic expression	Performic behavior	Semantic meaning	Pragmatic understanding
File (physical)	File (symbolic)	Image	Painting	Portrait	Cognitive/affective/conative reaction
Opaque bitstream	Internal JPEG data structure	Expressive image elements	Epistemic perception of color, line, mass, and texture	Ellen Day Hale portrait of Charlotte Perkins Stetson Gilman	Intersubjective phenomenological response to work of pioneering feminist artists
(a)	(b)	(c)	(d)	(e)	(f)

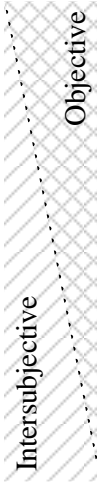
Figure 5.1. Semiotic levels of access and use

Ontic access may be sufficient for a systems administrator whose responsibility does not extend beyond concern for the existence of physical files on storage media. For example, confirming that the correct number of files, with the correct names and sizes, are found at the correct storage locations. Engagement with higher-level function depends upon access at higher semiotic levels. For example, empiric knowledge of the bitstream's digital encoding format (*Figure 5.1(b)*), as made accessible through an appropriate format-aware editor such as JPEGsnoop (Hass, 2017). This in turn enables syntactic accessibility to the abstract expressive components of the JPEG image such as size, sampling resolution, color space, and compression (*Figure 5.1(c)*). These elements are embodied as a perceptual image of a painting of a seated woman (*Figure 5.1(d)*) through performic accessibility by rendering software such as Irfanview (Skiljan, 2022). That reveals the semantic content recognizable as the ca. 1880 portrait of the pioneering feminist author and social reformer Charlotte Perkins Gilman by the artist Ellen Day Hale (*Figure 5.1(e)*) (Hale, ca. 1880). In conjunction with all prior semiotic affordances, plaistic accessibility situating artist and sitter in their contextual matrix of late 19th-century social and artistic feminism, e.g., (Allen, 2009; Fitzpatrick, 2010), conditions the consumer's intersubjective pragmatic response to the portrait (*Figure 5.1(f)*).

The range of meanings for accessibility in domain discourse aligns with that of common usage, which similarly span considerations of instrumental availability and performative enablement, including the potential of being “readily reached or got hold of,” “received, acquired, or made use of,” and “(readily) understood or appreciated [parentheses in original]” (OED, 2011a).<sup>19</sup> A more contemporary alternative meaning of accessibility in both common and specialized senses concerns suitability to the needs of consumers with various physical, sensory, or cognitive limitations (Abascal et al., 2016; Mack et al., 2021). This sense is often referenced in the context of assistive technologies intended to alleviate impediments to the fullest possible parameters of usage raised by those limitations (Botelho, 2021). From this perspective, determinations of accessibility should not be made under an assumption that access is a singular universal quality. Instead, accessibility should be evaluated in terms of the diverse needs and inherent cognitive and physical experiential capabilities and constraints of individual consumers, which are potentially multifarious in view of inherent contingencies particular to the consumer actor and the time, place, and purpose of the consuming act.

Table 5.2

*Communicological Accessibility*

Semiotic Dimension		Evaluative Factor
Ontic		Access to manifest bitstream and external file-level properties
Empiric		Access to symbolic representation and internal encoded properties
Syntactic		Access to abstract rhetorical structure and expressive properties
Performic		Access to technically-mediating behavior and perceptual properties
Semantic		Access to underlying meaning and affect and ontological properties
Plaistic		Access to contextual relationships and environmental properties
Pragmatic		Access to purposive intellectual and psychological understanding and epistemic properties

In the digital preservation context, this multivalent view of accessibility can be

<sup>19</sup> <https://www.oed.com/view/Entry/1034>

formalized in terms of the augmented semiotic ladder (see *Figure 3.4*). Each ladder rung or dimension corresponds to a set of informational and experiential affordances, the availability of which enables successful epistemic engagement at each dimension (see *Table 5.2*). For purposes of evaluating successful accessibility, metrics can be devised specific to each of the individual dimensions on the basis of appropriate and purposive access to their individual semiotic entailments. In current digital preservation practice, however, accessibility focuses on the ability to retrieve ontic-level representations into the technical context of the human consumer. In doing so, it is insufficient for characterizing an expansive sense of digital preservation success that encompasses an imperative for experiential engagement. Additional evaluation criteria contributing to a reliable determination of success need to address higher-level concerns. Beyond an object's manifest bitstream, file-level properties encompass those reported by POSIX functions (IEEE, 2016; Lewine, 1991) such as `stat()` and `chmod()`, including pathname, size, ownership, modification/access timestamps, access permissions, etc, for ontic manifestations conforming to a filesystem storage abstraction. Similar properties are also retrievable from storage systems implementing an object store abstraction. The availability of these characteristics is important for establishing higher-level archival qualities such as integrity.

Symbolic encoding properties are those specific to the internal representation of the ontic manifestation. The JPEG file referenced in *Figure 5.1(b)* defines data structures for indicating the image sampling densities and height and width dimensions as well as colorimetric information (ISO, 2013; ISO/IEC, 1994). Access to these properties is dependent on understanding that the JPEG format encoding has been used, so that it can be appropriately decoded. The encoding format can be known through either direct extrinsic knowledge or inferentially by file-level format characterization (Abrams et al., 2009) that interrogates the bitstream for JPEG-conforming structures. The availability of these characteristics is important for supporting format-specific performances of preserved objects to provide access at a perceptual level, itself a determinant of higher-level semantic recovery and pragmatic interpretation.

Rhetorical properties are those particular to the mode of expressing an object's underlying cognitive and affective content. In the case of JPEG file, the image can be described in abstract terms according to the conventions of artistic portraiture, as

shown in *Figure 5.1(c)*. These features are dependent on successful availability of lower-level ontic and empiric qualities permitting behavioral performance amenable to human visual perception and recognition. To a certain extent, accessibility at the ontic, empiric, and syntactic levels are tightly coupled. Any given syntactic expression can be symbolically represented in a variety of encoding standards. Each specific encoding corresponds to a fixed manifest bitstream. Thus, explicit ontic accessibility to a bitstream implicitly assures availability of one specific associated empiric encoding and abstract syntactic expression. Note, however, that the same syntactic expression could be faithfully represented by alternative encodings each again with its own unique bitstream. This close ontic/empiric/syntactic affinity is indicated graphically in *Figure 5.1* by the relative proportions attributed to the objective- and subjective-ness of the seven semiotic dimensions. This varies from objectively fixed at the ontic level to subjectively contingent at the syntactic level. That is, once an expression is subjectively selected, the choices for its representation are somewhat constrained by the requirements of that expression but are nevertheless themselves still subjectively contingent. However, after the expression and representation are both chosen, the resulting manifestation is fully determined. The ontic/empiric/syntactic coupling may explain why these dimensions are not considered in isolation in common digital preservation practice. Instead, ontic accessibility, that is, access to a physical bitstream, is the primary concern that is explicitly recognized. Given their coupling, ontic accessibility can function as a reliable proxy for empiric and syntactic accessibility. However, this does not provide useful characterization of higher-level concerns that are important fully meaningful evaluation of Communicological success.

## **5.2 INTEGRITY**

In archival discourse, integrity is the quality of an artifact being complete and unaltered in its essential nature relative to an accepted state (ICA, 2016; InterPARES, 2008; SAA, 2020). Evaluation of integrity needs to draw a critical distinction between an artifact's abstract information content and the tangible manifestation of that information (Hamid, 1998; Harvey et al., 2020). Controlled or monitored modification of the latter, whether intentional or natural, does not necessarily invalidate the former. For physical items, material degradation is an inevitable entropic consequence of the passage of time (DeSilvey, 2006; Domínguez Rubio, 2014), as for example, the fading of ink or discoloration of paper (Daniels, 1996). Other forms of material damage may



result from inadvertently-inappropriate handling, intentionally-malicious acts, or simple overuse. In response, affirmative conservation may be called for to stabilize an artifact in its current state or restore it to its original or some known prior state (Cloonan, 2015). In all cases, these actions represent some degree of change in material condition that constitutes a violation of physical integrity, i.e., the thing is – in some way, large or small – no longer what it once was. However, this may not affect the unity and cohesion of higher-level integrity of some curatorially-designated “essence” (Adam, 2010). For example, faded ink still may be readable, a page’s background discoloration may not affect the legibility of its foreground text, a page tear may not intrude into the text block. In this respect, the manifestation/information dichotomy underlying the InterPARES, ICA, and SAA definitions corresponds to contrasting considerations of the syntactic and semantic components of the classic semiotic triangle (see *Figure 3.1*). As discussed in Section § 3.2, classical syntactics subdivides into distinct ontic, empiric, and syntactic rungs of the augmented semiotic ladder (see *Figure 3.4*). A more comprehensive definition of integrity should expand beyond the view of a singular quality assessed against the totality of an archival artifact. It should additionally encompass integrity as a multivalent quality independently-considered relative to each semiotic dimension.

In the digital preservation realm, however, integrity often carries a narrower meaning synonymous with the concept of bit-level fixity, that is, “that we have in hand the same set of sequences of bits that came into existence when the object was created” (Lynch, 2000, p. 38). The NDSA Levels of Digital Preservation rubric (Phillips et al., 2013) conflates data integrity and file fixity into a single functional category. The tiered set of recommendations in that category, however, refer solely to fixity, implicitly asserting synonymy of the concepts of integrity and fixity. Similar conceptual synonymy is found throughout the literature; see for example (Baucom, 2021, p. 31; Bountouri et al., 2018, pp. 369-370; Tallman, 2021, pp. 2-3). In these cases, the general term “integrity” is used tacitly as short-hand for bit-level or *ontic* integrity.

For digital preservation purposes, cryptographic hashing is used to indicate bit-level integrity or fixity. A cryptographic hash algorithm uses a one-way mathematical function to map the arbitrary bit sequence of a source message into a smaller, fixed-size numeric value, or message digest, that provides an essentially unique and invariant

“signature” of the message (Chi & Zhu, 2017). The comparison of a stored message digest value against one freshly calculated on a digital file can detect the smallest bit-level variation (Spencer, 2019). For example, two states of a digitized raster still image could differ by a single bit, say, a value of 123 rather than 124 for the 8-bit red channel of a single pixel. This constitutes an absolute violation of ontic integrity. However, it is probably significantly below the noise threshold of the digitization process. For example, two digitized images captured in immediate succession with the same camera setup inevitably will differ by more than a single bit due to random variations in the physical functioning of the camera sensor. This degree of minor variation is likely perceptually unnoticeable by a human consumer (Chanod et al., 2010). In cases like this, the determination of integrity can be recast away from a reliance on absolute fidelity to one of relative similarity (Hao et al., 2021). Perceptual hashing provides an alternative to cryptographic techniques that accommodates the subjective nature of human perception.

Perceptual hash algorithms are designed to provide unique compact signatures of semantic content (Du et al., 2020) that are invariant with respect to “content preserving modifications” (Samanta & Jain, 2021, p. 204). In other words, invariance is no longer solely a strictly objective ontic property of a preserved digital object. Instead, it is also a property pertinent to a subjective performance of the object. For example, a given semantic proposition is susceptible to multiple cognate expressions, each corresponding to a unique empiric and ontic form. The textual propositions “I painted the house” and “The house was painted by me” both express an equivalent primary assertion of house-painted-ness causality and agential responsibility. If the former value was the one originally subject to preservation stewardship, during the course of which the syntactic/empiric/ontic forms were shifted to those of the latter, it could be legitimate, subject to pragmatic interpretive context, to claim that the perceptual integrity of the semantics is maintained. In *Figure 5.1* perceptual integrity operates at the level of the perceived *image* distinct from the cryptographic integrity of the manifest image *file* (Tiknonov, 2019). In other words, it definitively is – or isn’t – in its proper and accepted physical *form* in a mathematically rigorous manner. That is, it approximates – more or less – its proper and accepted intellectual/aesthetic/emotive *essence* for an interpreting human consumer.

Of the two hashing types – cryptographic and perceptual – only the first is well

integrated into digital preservation practice. Thus, preservation community concerns for archival integrity emphasize ontic fixity and minimize considerations of integrity at higher semiotic levels. As with accessibility, each rung of the augmented preservation ladder (*Figure 3.4*) corresponds to a set of properties, each of which may or may not possess the quality of integrity; that is, being whole and uncorrupted from their accepted state. Fully successful epistemic engagement with a preserved digital object depends upon integrity across all these semiotic dimensions (see *Table 5.3*). Similar to the case of contingent accessibility (*Table 5.2*), the notion of perceptual integrity introduces a sliding scale of subjectiveness associated with the various dimensions.

Table 5.3  
*Communicological Integrity*

Semiotic Dimension		Evaluative Factor
Ontic		Integrity of bit-level fixity and file-level properties
Empiric		Integrity of symbolic representation and internal encoded properties
Syntactic		Integrity of rhetorical structure and expressive properties
Performic		Integrity of technically-mediating behavior and perceptual properties
Semantic		Integrity of underlying meaning and affect and ontological properties
Plaistic		Integrity of contextual relationships and environmental properties
Pragmatic		Integrity of purposive intellectual and psychological understanding and epistemic properties

While identified through Predicate Reduction analysis of the source digital preservation policy documents as an evaluative norm in its own right, integrity has been viewed conceptually as a subcomponent of authenticity (Duranti, 2005).

### 5.3 AUTHENTICITY

Authenticity is defined by the SAA Dictionary as “The quality of being genuine” (SAA, 2020) and therefore trustworthy as evidence. However, according to the InterPARES Glossary, that trustworthiness applies to “a *record as a record* [emphasis added]” (InterPARES, 2008). This formulation draws upon the two traditional

complementary aspects of a record's function: as a set of abstract *information* externally-documenting a thing, event, or condition, and at the same time the manifest *carrier* of that information (Lester, 2018). In light of these fluid roles, the definition of archival trust can be restated more explicitly as applying to “a [physical/informational] record as a [purposive] record,” the purpose of which is archival evidence. Thus, the quality of authenticity inheres to an archival artifact *qua* artifact rather than the relational experience of its use. In this regard, archival authenticity conforms can be characterized as an objective evaluative standard, rather than a socially constructive or intersubjective one (Mochocki, 2021). An archival record is considered authentic through attestation to that effect by a responsible archival agency (Duranti & Blanchette, 2004). The primary consideration for such attestation is continuous care and a well-documented chain of associated provenance under managerial custody.

Authenticity is a distinct archival quality from reliability (Duranti, 1995; Kastenhofer, 2015). Authenticity asserts the evidentiary trustworthiness of a tangible information *carrier* while reliability asserts the intellectual trustworthiness of the abstract carried *information* itself (MacNeil, 1998, 2000). Because that information content is implicitly situated within a specific domain of practice and concern, the determination of its reliability necessitates contextual pragmatic interpretation dependent upon specific domain knowledge (Greene, 2002; Kastenhofer, 2015). Such specialist knowledge is primarily a curatorial or consumer responsibility and falls outside the normal purview and capacity of digital preservation managerial agency. Nevertheless, managerial responsible does encompass preservation of authentic plaistic context providing non-managerial actors with a basis for reasoned determinations of reliability.

Reference to reliability in the examined policy document set correlates with institutional missions. Only two of the six documents refer to reliability as a preservation goal, and then only with low relative frequency. In both cases, less than 4% of articulated policy imperatives concern reliability, compared to over 12% for authenticity. Those two issuing institutions are archival in mission: the Nationaal Archief of the Netherlands (NA, 2015) and Archives New Zealand, which shares a joint policy with the National Library of New Zealand (NLNZ, 2012). Assurance regarding the evidentiary reliability of records is one of the two central foci of the

institutional archival mission (Thomassen, 2001), the other relating to the continuity of social meaning and memory (Cook, 2013). The four institutions omitting any reference to reliability include a museum (BMA, 2016), an academic research library (CUL, 2018), a datacenter (ICPSR, 2018b), and an institutional repository (ZBW, 2018). Historically, libraries have placed a lower emphasis of the evaluation of the truthfulness of their collections (Jatkevicius, 2005; Lor et al., 2021), so this absence is not unexpected.

In archival practice, authenticity is considered an objective determination – “a record is either authentic or not” (Duranti, 1995, p. 215) – while reliability is inherently subjective to the contingent context of the consumer (Rogers, 2015b). Thus, it is appropriate to distinguish measures of the *authenticity* of preserved digital information *objects* and the *legitimacy* of digital information *experiences*. Whereas a determination of authenticity carries the connotation of singular objective universality, legitimacy is pertinent to situated intersubjective plurality. This is considered most appropriately at the pragmatic level of contingent individual response (see Table 5.4).

Table 5.4  
*Communicological Authenticity*

Semiotic Dimension	Evaluative Factor
Ontic	Authenticity of bit-level fixity and file-level properties
Empiric	Authenticity of symbolic representation and internal encoded properties
Syntactic	Authenticity of rhetorical structure and expressive properties
Performic	Authenticity of technically-mediating behavior and perceptual properties
Semantic	Authenticity of underlying meaning and affect and ontological properties
Plaistic	Authenticity of contextual relationships and environmental properties
Pragmatic	<i>Legitimacy</i> of purposive intellectual and psychological understanding and epistemic properties

Contemporary archival practice has extended traditional approaches for evaluating authenticity into a new domain of digital diplomatics (Rogers, 2015a). However, the legitimacy of pragmatic experience falls outside the purview of existing diplomatic procedures. As that experience is central to the notion of teleological

preservation success, new criteria and measures will be necessary to support future determinations of that success. Object-level authenticity, along with accessibility and integrity, are enabling constituents that provide an opportunity for success. However, they do not reductively determine that success. By definition, a consumer cannot engage with an inaccessible object, while a non-integral or unauthentic object, however accessible, may not fulfil all consumer needs. The contingent circumstances underlying those needs establish the purposive parameters of need-fulfilling usage of accessible, integral, and authentic objects. The degree of that fulfilment is a measure of consumer satisfaction or preservation-enabled communicative success. Thus, the three managerial or artifactually-centric evaluative qualities of accessibility, integrity, and authenticity must be supplemented with that of experiential *usability* in order to meaningfully characterize digital preservation success.

#### **5.4 USABILITY**

The concept of usability is not given formal definition in the SAA Dictionary or InterPARES Glossary (InterPARES, 2008; SAA, 2020), both prominent points of reference in the preservation community. The SAA Dictionary does provide a definition of the related concept of “access”, but this emphasizes access as an enabling function, that is, a quality facilitating retrieval of a preserved object *for use*. However, there is no commensurate definition detail regarding use itself. ISO 15489 is the international standard for concepts and principles of archival records management (ISO, 2016). While it does not include usability in its formal glossary, the concept is defined in the narrative text as the quality permitting a record to be “located, retrieved, presented and interpreted” (p. 5). The first three characteristics more properly fall under the enabling umbrella of accessibility as defined in this study (see Section § 5.1). Given the inherent communicative nature of the digital preservation enterprise, any act of engagement with preserved digital materials is an act of intersubjective interpretation. Thus, the fourth ISO characteristic, interpretability, informs an important constituent aspect of usability. Within the scholarly literature and professional best practice guidance, usability is often referenced as a central preservation imperative. However, these references do not generally provide specific detail regarding the constitution of “use”, let alone successful use; see for example (Caplan, 2008; Heslop et al., 2002; Traczyk, 2017; Waters & Garrett, 1996; Yakel, 2007). In the absence of explicit definition, a general common sense must be assumed,

for example, “The act of putting something to work, or employing or applying a thing, for any (esp. a beneficial or productive) purpose” (OED, 2011b).<sup>20</sup> While this generically captures the inherent purposive nature of use in the Communicological context, it is necessarily silent with regard to the implications for the evaluation of digital preservation success.

The Digital Preservation Consortium is a leading international membership organization dedicated to promoting the world’s digital legacy in the face of strategic, cultural, and technological challenges (DPC, 2022a). DPC membership of national and academic research libraries, archives, museums, institutional repositories, and data archives (DPC, 2022b) reflects the same range of mission orientation as the institutions publishing digital preservation policy documents described in Section § 3.1.1, and from which the six specific policies examined in this study were drawn. As part of its mission to encourage and support preservation activity, the DPC publishes a Glossary of key preservation concepts. In it, usability is defined indirectly in terms of the persistence of artifactual characteristics that a user would reasonably deem indicative of productive or managerial intention (DPC, 2015). However, the scope of this definition does not give due consideration to the purposive aspiration on the part of a consuming user or the subsequent communicative response of meaningful intellectual, emotional, or physical consequence *to* that user (Abrams, 2021). The impediments to articulating a theoretically and pragmatically sound definition for usability arise from the fact that the purposive needs and experiential contexts of the user are inherently intersubjective (Bishop & Hank, 2018). Thus, the evaluation of those preservation-enabled experiences cannot rely on the assumption of singular canonical *use*. Instead, it must acknowledge the potential of a diversity of individual *uses* (Abrams, 2018b). Given their intersubjective context, these various uses cannot be fully anticipated, especially considering their inevitable evolution across archival timespans.

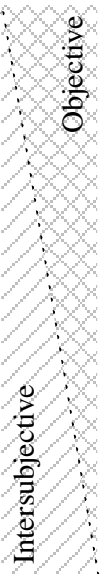
Thus, an important consideration at the center of any effective framework for evaluating the success of preservation-enabled usability is cognizance of, and response to, the nuanced contingent contexts of users and uses. This suggests the necessity of recasting the prevalent singular conceptualization of usability into a multivalent set of semiotic concerns as was previously done in Sections § 5.1 – 5.3 for the other three

---

<sup>20</sup> <https://www.oed.com/view/Entry/220635>

preservation norms. Table 5.5 emphasizes the teleological dependence of usability on the prior accessibility, integrity, authenticity/legitimacy of each of the semiotic dimensions. These can range from technical ontic concerns to teleologically-fulfilling pragmatic understanding of and response to a preserved object. This granular notion of usability enables derivation of more appropriate evaluative norms, formal criteria, and operational metrics applicable to the digital preservation needs, goals, and aspirations of diverse varieties of user and use.

Table 5.5  
*Communicological Usability*

Semiotic Dimension		Evaluative Factor
Ontic		Usability of accessible, integral, and authentic bit-level fixity and file-level properties
Empiric		Usability of accessible, integral, and authentic symbolic representation and internal encoded properties
Syntactic		Usability of accessible, integral, and authentic rhetorical structure and expressive properties
Performic		Usability of accessible, integral, and authentic technically-mediated behavior and perceptual properties
Semantic		Usability of accessible, integral, and authentic underlying meaning and affect and ontological properties
Plaistic		Usability of accessible, integral, and authentic contextual relationships and environmental properties
Pragmatic		Usability of accessible, integral, and legitimate purposive intellectual and psychological understanding and epistemic and phenomenological properties

**5.5 NORMS, CRITERIA, AND METRICS**

The three managerial norms of accessibility, integrity, authenticity identified in this research are consistent with earlier expressions of non-functional digital preservation requirements found in the literature (Burda & Teuteberg, 2013). Now, however, they have been established empirically through Predicate Analysis of preservation policy statements determinative of reciprocal contractual service-provider intentions and stakeholder expectations. In order to be operationalized in practice, these norms must be translated into high-level evaluation *criteria* as well as actionable *metrics*. A criterion is a generic evaluative quality, while a metric is a specific standard by which one can obtain a measurement of a relevant quality (Black



et al., 2008; Seffah et al., 2006). The construction of an effective measurement system necessitates sufficiently granular and detailed conceptual understanding of a domain in order to establish appropriate evaluative categories, the scope of evaluative factors within those categories, and procedures for interrogating those factors (BIPM, 2012).

The concepts of accessibility, integrity, and authenticity are widely deployed in archival theory and practice as important qualities of archived objects (Abrams, 2021). An object is accessible if its existence is known and it can be requested and retrieved subject to legal, technical, and policy considerations; it is integral if it is whole and uncorrupted in form and structure; and it is authentic if it is what it purports to be (Duranti, 2005; SAA, 2020). Since these three qualities are well-formalized, they can act as the normative basis for assessing the efficacy of digital preservation activities, outputs, and outcomes; see for example (Korenkova & Hägerfors, 2011). The pertinent level of detail provided by their definitions also facilitates the derivation of evaluative criteria and associated metrics, such as bit-level cryptographic fixity for validating ontic integrity (Bountouri et al., 2018), descriptive standards and discovery platforms for support of performic accessibility (Bak & Armstrong, 2008; Whitelaw, 2012), and digital diplomacy for characterizing semantic authenticity (Rogers, 2015a). These semiotic characteristics are important evaluative considerations for preservation efficacy. However, they function primarily as ontological rather than epistemological or phenomenological characterizations. That is, they provide important information about the existential fabric of managerially-preserved digital *artifacts*, but not the consequent behavioral *experience* and communicative *understanding* and *response* on the part of the artifactual consumer. Thus, accessibility, integrity, and authenticity are necessary enabling factors for preservation activity. However, they are not fully sufficient to ensure the teleological imperative of that activity: the purposive use of preserved digital objects.

Usability has not been formalized in community discourse to the same extent as accessibility, integrity, and authenticity. Use of the term in that discourse relies upon vague definition or tacit assumption. For example, “by usable we mean that someone is able to do something sensible with the information it [a preserved digital object] contains” (Giaretta, 2011, p. 167). While the basic tenor of this definition aligns generally with preservation’s teleologically-communicative goal, it does not specifically explicate the range or context or possible “somethings” or measures of

“sensible-ness”. Similarly, “[a] useable record is one that can be located, retrieved, presented and interpreted within a time period deemed reasonable by stakeholders” (ISO, 2016, p. 5). In addition to emphasizing the instrumental aspects of access, this definition also places consumer interpretation at the center of preservation focus. However, while this indicates a communicative goal, it does not explicate that goal in terms of granular Communicological functions. Thus, while usability can function as a high-level evaluative norm, its informal conceptualization makes programmatic comparison of evaluations problematic. Furthermore, the non-rigorous fluidity of the concept’s definitional deployment makes it unsuitable for translation into specific measurable criteria and implementable metrics. The Communicological segmentation of the broad concept of usability into seven more-specific analytic dimensions presented in Section § 5.4 provides a new viable structure for greater definitional specificity based on granular semiotic concerns and evaluative norms. This should enable easier identification of relevant assessable criteria and associated metrics. These metrics are necessary to ascertain degrees of alignment and equivalence of the intentional, archived, and expectational states of a preserved digital object (see *Figure 1.2*).

As introduced in Section § 1.2, establishment of pertinent significant properties of preserved objects is widely posited as an appropriate basis for preservation assessment (Giaretta et al., 2009; Hedstrom & Lee, 2002; Hockx-Yu & Knight, 2008). Existing frameworks for deriving workable properties, such as InSPECT (Knight, 2009), can be insufficient for appropriate characterization of complex digital objects or behaviors (Sacchi & McDonough, 2012). Subsequent extension of InSPECT focuses on parallel object and stakeholder analyses (Stepanyan et al., 2012). The latter introduces dynamic epistemological and situated phenomenological concerns of relational, behavioral, and experiential nature. These concerns supplement the ontological consideration of the static properties of isolated objects in a purely managerial context. This reemphasis is consistent with a view of digital objects not as fixed, but rather, fluid carriers of technically-mediated but socially-negotiated meaning (Rozenberg, 2021). This in turn accords well with the Pragmatic theory of meaning as arising from the conditions and practical effects that engagement with a meaning-laden artifact has upon its consumer (Mingers & Willcocks, 2014). Explication of object-consumer interaction can be couched in terms of affordance

rather than property to accentuate the critical sense of purposively-instigated human action. An affordance is the nexus of factors intrinsic and extrinsic to object and environment that enables opportunities for those actions (Cheikh-Ammar, 2018; Withagen et al., 2012). Communicological application of affordances to the evaluation of digital preservation success necessitates extension of prior processes for deriving evaluative norms, criteria, and metrics.

## 5.6 SIGNIFICANT AFFORDANCES

The preservation concept of significant properties provides information characterizing what an object *is* in its managerial context. Recasting these evaluative norms as functional affordances shifts the conceptual emphasis to what those properties enable the object's consumer to do, understand, and act upon in the context of a communicative process. Prior research has established proposals for the significant properties of various content genres (van Veenendaal et al., 2018), including journalism (Heravi et al., 2021), relational databases (Freitas & Ramalho, 2010), research data (Knight & Pennock, 2009), software (Matthews et al., 2008), spreadsheets (van Veenendaal et al., 2019), and video games (Bettivia, 2016a). Most of these efforts rely on some form of the InSPECT framework (Knight, 2009), which groups properties into five high-level categories for purposes of analysis, characterization, and application:

1. *Structure*, concerned with characterization of internal encoded form and external relational associations;
2. *Rendering*, concerned with internal expressive form and external instrumental dependencies of subsequent perceptual form;
3. *Behavior*, concerned with experiential interaction;
4. *Content*, concerned with abstract intellectual essence; and
5. *Context*, concerned with environmental factors of production and intentional meaning.

These concerns align with the rungs of the extended semiotic ladder (see *Figure 5.2*). However, the InSPECT framework is defined at coarser granularity: the Structural group conflates characterization at both the ontic and empiric dimensions and there is no category corresponding to the pragmatic dimension's concerns for

epistemic understanding and cognitive, affective, and conative response. The four primary evaluative norms also can be placed in alignment with the ladder and InSPECT categorization, but again, without strict one-to-one correspondence. Usability spans both the Content and Context categories just as authenticity spans Rendering and Structure, while integrity applies most closely to the ontic manifestation subset of the Structure category. Defining a future set of norms, criteria, and measures scoped more tightly to each of the semiotic dimensions will provide greater confidence that the evaluative process appropriately incorporates considerations of the full set of significant Communicological concerns. Prior criticism of the concept of significance emphasizes it as an inherently indeterminant factor (Yeo, 2010) due to the subjectivity of human-centered affordances (Hedstrom & Lee, 2002) contingent to time, place, person, and purpose. Recasting significance in terms of a semiotic framework implicitly cognizant of the full range of human communicative engagement offers the potential for the targeted derivation of appropriate characterizing elements.

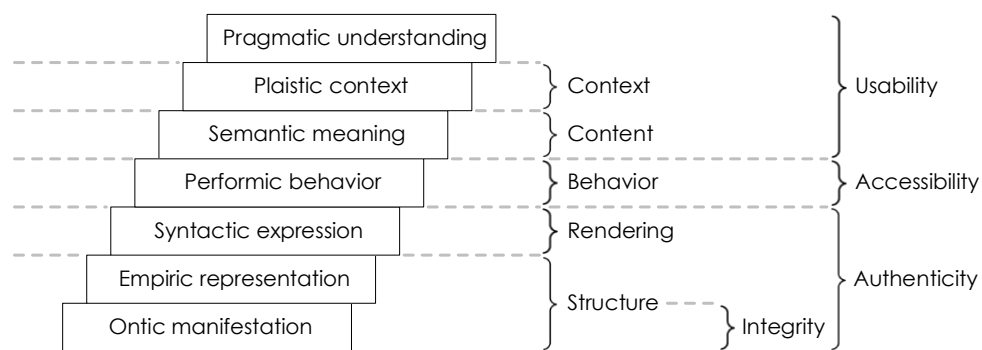
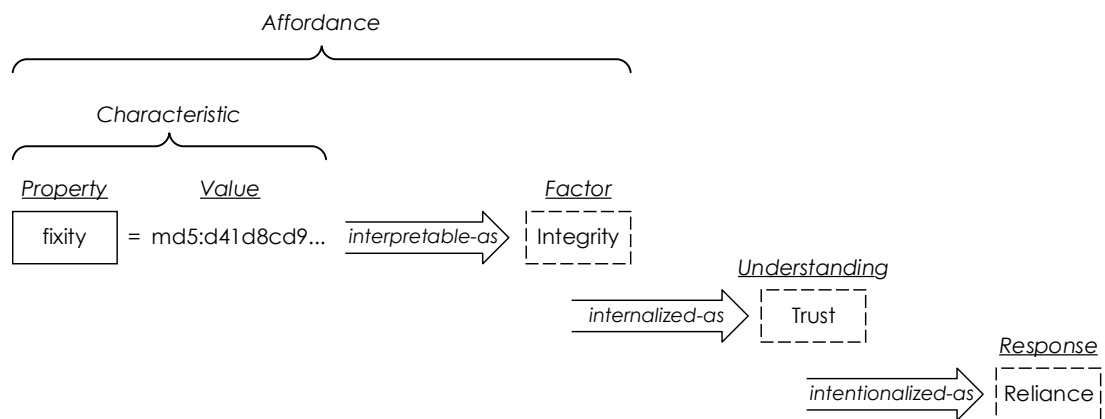


Figure 5.2. Semiotic alignment of InSPECT categories and normative scope

The InSPECT framework defines an analytic procedure for identifying relevant sets of significant properties of preserved digital objects. For example, the concept of cryptographic fixity is a Structural property of ontic manifestation. However, the static *property* of fixity inhering to a preserved object supports an associated relational *affordance* of object integrity on the part of the object’s consumer. That is, fixity enables the rational determination of integrity. That determination in turn bolsters a dynamic *response* of consumer confidence that the object is whole and uncorrupted from its accepted form. While “significant property” is the accepted term-of-art within digital preservation community discourse, the underlying concept has been usefully extended to that of significant *characteristic* by distinguishing between a property *per se* as an abstract metrical standard capable of measurement and its value as an actual

instantiated measure of that metric (Dappert & Farquhar, 2009) (see *Figure 5.3*). An affordance is a further extension of a characteristic expressed from the human perspective of desirable, and ultimately actionable, information. This in turn enables an intentional response to informational understanding in light of contingent purposive considerations.



*Figure 5.3.* Significant characteristic, affordance, understanding, and response (CAUR) facets  
Property-value-characteristic structure adapted from (Dappert & Farquhar, 2009)

The characteristic-affordance-understanding-response (CAUR)<sup>21</sup> progression shares an affinity with the tiers of the data-information-knowledge-wisdom hierarchy (DIKW) (Rowley, 2007) often used to model and explicate human cognitive processes. The DIKW pyramid is subject to criticism regarding the theoretical imprecision of its internal definitional boundaries and transformational processes (Frické, 2019). However, it is susceptible to a semiotic formulation (Baskarada & Kononios, 2013) pragmatically useful for purposes of a Communicological framework for evaluating digital preservation success. The individual CAUR facets align conceptually with semiotically-based DIKW tiers (see Table 5.6). The ontological Characteristic facet as a knowable-attribute of a preserved digital object corresponds to the DIKW Data tier of objective facts. Similarly, the epistemological Affordance facet as the means-of-knowing objective attributes corresponds to the Information tier of emergent interpretations on the part of an informed human actor; just as Understanding as the nexus of pragmatic cognitive and affective responses by that actor corresponds to Knowledge as individually-justified belief; and Response as an actor's subsequent

<sup>21</sup> For declamatory purposes, the acronym CAUR can be pronounced "core" /'kòr/.

pragmatic conative action corresponds to Wisdom as socially-acceptable reliance belief. The latter two correspondences – Understanding/Knowledge and Response/Wisdom – both function at the phenomenological level of internalized and socialized experience.

Table 5.6

*Alignment of Evaluative Levels*

Primary Concern	CAUR Facet		DIKW Tier	
Ontological	Characteristic	Knowable attributes of preserved object	Data	Objective fact
Epistemological	Affordance	Means-of-knowing object attributes	Information	Emergent interpretation
Phenomenological	Understanding	Known cognitive/affective pragmatics	Knowledge	Individual justified belief
	Response	Consequential conative pragmatics	Wisdom	Socially-acceptable reliance

The four CAUR facets and associated DIKW tiers also correspond to primary components of digital preservation semiosis as modelled in Section § 3.2 (see *Figure 3.6*). Characteristic/Data and Affordance/Information encompass the knowable and means-of-knowing aspects of a preserved digital object  $S_C$ , the technically-mediating performic channel  $Ch_M$  through which an object is engaged with by its human consumer  $I_C$ , and its consumer-perceivable form as signifier  $R_{S_C}$ . Understanding/Knowledge encompasses the known, or signified, cognitive/affective pragmatics of  $R_{G_C}$  relative to semantic referent  $G_C$ , while Response/Wisdom encompasses the consequential conative pragmatics of behavior  $R_{b_{G_C}}$ . Referring to the example digital object in *Figure 5.1*, the physical file ( $S_C$ ) possesses data characteristics of ontic filename, bitstream, size, and cryptographic fixity (*Figure 5.1(a)*), empiric symbolic encoding in terms of the JPEG format standard (*Figure 5.1(b)*), and syntactic rhetorical expression in terms of representational painterly convention (*Figure 5.1(c)*), all supporting informational affordances of the qualities of accessibility, integrity, authenticity, and usability. These affordances become actionable through the mediation of a behavioral rendering process ( $Ch_M$ ) that transforms the ineffable digital into a tangible visual representation ( $R_{S_C}$ ) perceptible to human sensory capabilities (*Figure 5.1(d)*), interpretive agency ( $I_C$ ), and subsequent semantic interpretation ( $G_C$ ) as the Hale portrait of Gilman (*Figure 5.1(e)*). This in turn instigates individual but

intersubjectively-contingent purposive cognitive and affective pragmatic understanding of the portrait (*Rgc*) and subsequent conative behavior (*Rbgc*) within a larger environmental domain of common concern and practice (*Figure 5.1(f)*). The close conceptual synonymy between the elements of the semiotic model, CAUR, and DIKW validates the use of the semiotic toolset underlying Communicological analysis as the basis for this critique of preservation success evaluation factors.

Within the semiotic/CAUR/DIKW formulation, data are objective ontological facts embodying what is knowable about a domain of interest. However, they are inherently meaningless in isolation from consuming agency. The meaning of data emerges only through an epistemological process of intersubjective human interpretation affording opportunities for a response of informative knowing. Interpretation is cognitively and affectively internalized as phenomenological understanding of knowledge adjudged by individual agency. Intentional conative action in response to newly acquired knowledge may have broader social visibility where it will be judged acceptable or not in light of established social norms and conventions. For example, the fixity characteristic property value of “md5:d41d8c...” is by itself a textual string of no inherent evaluative meaning. It accrues meaning when interpreted as an affordance for integral completeness and absence of corruption of a specific preserved digital object. That meaning forms a rational basis for an object’s consumer to adjudge a quality of trust regarding the object. That trust in turn permits subsequent reliable purposive use of, and response to, the object in a manner socially judged as warranted.

Reconceptualizing evaluative measures as dynamic epistemological affordances with responsive phenomenological consequences, rather than static ontological properties, suggests an alternative definitional scheme for evaluative norms. Usability is the normative umbrella for the consequential response that is the teleological goal of communicative digital preservation activity. Rather than a distinct norm on an equivalent conceptual plane as accessibility, integrity, and authenticity, usability can be viewed as a primary evaluative quality conceptually encompassing the other three, which represent subordinate concerns. From this perspective, normative qualities are meaningful only with respect to the use to which they can be put. Thus, the traditional sense of integrity as bit-level correctness is better considered as a particular significant affordance of ontic usability. That is, in order to make productive purposive use of a

preserved object’s ontic content, it is necessary for that object to exhibit a degree of integrity appropriate for that purpose. Similarly, the quality of accessibility can be repositioned in terms of performic usability, that is, the mediating behavioral instrumentalities that afford use of an object’s ontological characteristics by interpreting epistemic agency. However, within the structure of this normative reframing, authenticity is problematic, as it spans concerns of ontic, empiric, and syntactic usability (refer to *Figure 5.2*). This many-to-one mapping suggests ambiguity regarding normative derivation, since norms within the same authenticity group could apply more narrowly to some or only one of the associated semiotic rungs. A more satisfactory refinement is to consider *all* four norms – accessibility, integrity, authenticity, as well as usability – as being applicable at *each* of the rungs of the ladder (see *Figure 5.4*).

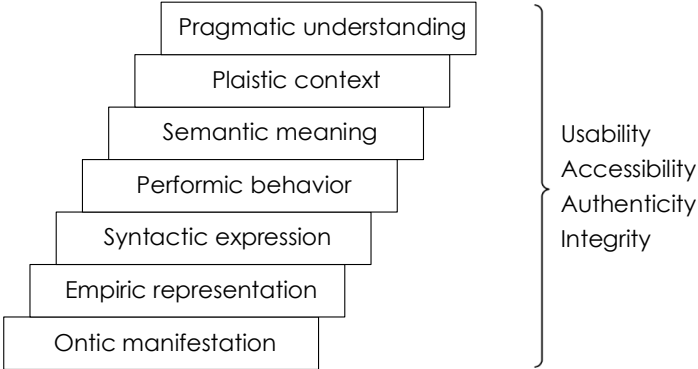


Figure 5.4. Normative scope of semiotic applicability

For example, the ultimate usability of ontic characteristics is dependent upon the accessibility, integrity, and authenticity of those characteristics. Just like an entire preserved digital object, that object’s characteristic fixity value is subject to damage or loss. Knowing that a particular fixity value is complete and uncorrupt, that it is accessible for retrieval, and that it is the authentic value are all preconditions for successful exploitation of that value as the basis of an epistemic affordance and subsequent phenomenological response regarding individual consumer trust and socially domain-acceptable reliance of that object. The imperative of ensuring that reliance lies at the heart of the digital preservation enterprise, and at the heart of effective measures of the success of that enterprise. In order to function as an effective composite measure of that success, a more broadly conceived norm of usability should encompass traditional notions of archival accessibility, integrity, and authenticity



complemented with concerns for intersubjective contingency, contextually-purposive relevance, instrumental affordance, and, ultimately, experiential satisfaction.

## 5.7 MULTIVALENT SUCCESS

The digital preservation enterprise exists within a larger domain of digital information seeking and exploiting activities. Within that domain, usability represents the quality of enabling opportunities for “effective, efficient and satisfactory task accomplishment” in the interaction of human actors, technically-mediating systems, and digitally-manifest informative content (Tsakonas & Papatheodorou, 2008, p. 1238). The centrality of stakeholder satisfaction in any determination of usability prioritizes success as the characterizing benchmark of usability. In the digital preservation context, success is a measure of the satisfactory-for-purpose alignment of service-provider intention, stakeholder expectation, and preserved actuality. Given that usability is the teleological aim of digital preservation attention and activity, preservation success is the preeminent evaluative factor for that activity.

In the absence of explicit articulation of preservation intentions and expectations, the implicit definition of those two actorial conditions are established through preservation policy statements, as discussed in Section § 3.1.1. Policy terms define a social, if not legal, contract of reciprocal obligations and assumptions delineating the parameters of service-provider/stakeholder interaction. Predicate Reduction analysis of representative policy statements identifies four evaluative qualities – accessibility, integrity, authenticity, usability – broadly accepted as normative imperatives across the digital preservation domain theory and practice, as presented in Section § 4.3. In terms of their traditional definition and application as discussed in Sections § 5.1 – 5.4, the four norms group naturally into two distinct categories (see Table 5.1). The first three are essentially managerial in scope and artifactual in focus. That is, they primarily characterize the outputs of managerial oversight and intervention in terms of the static properties of preserved digital objects isolated from direct consideration of the circumstances or consequences of their use. Usability, on the other hand, is inherently communicative in scope and experiential in focus. It is concerned with the relational outcome of an object in its role as a communicative vehicle for past productive expression of informative content consumed in a contemporaneous context. A successful act of consumption results in newly emergent cognitive and affective mental states and responsive conative actions on the part of the object’s human

consumer in satisfactory furtherance of some meaningful purpose. That purpose is contingently positioned in relation to time, place, person, and impetus.

The norm of accessibility corresponds to a high-level evaluative question, “Is the object of interest retrievable *for* use?” Similarly, integrity corresponds to the question, “Is the object complete *for* use?”; authenticity, “Is the object trustworthy *for* use?”; and usability, “Is the object helpful *in* use?” The *for/in* distinction in the formulation of these questions emphasizes that the three managerial/artifactual norms support a predictive assessment modality, providing a basis for justified supposition regarding what *should* result from preservation attention and, as necessary, affirmative intervention. The more successful the managerial retention of the accessibility, integrity, and authenticity of significant artifactual affordances of a preserved object, the stronger the possibility of its subsequent purposive usability. In this regard, usability is an aggregate quality of affordances available across the full semiotic semiotic spectrum of communicative concerns (see *Figure 5.4*). More than that, however, usability implies an assessment concern not only with what should result, but also with what experientially *did* result from preservation action. The *should/did* dichotomy corresponds to the distinction between outputs and outcomes in LIS assessment theory.

An output is a quantifiably-measurable result of an activity, such as counts or enumerations of the generated states or productions of a system or process, while an outcome is a qualitatively-assessable benefit of an output (Bertot, 2004; Dugan & Hernon, 2002; Kyriellidou, 2002). Thus, an outcome focuses on the experiential impact or difference an output has on the part of its recipient (Tsakonas & Papatheodorou, 2011). Traditional measures of digital preservation success focus on outputs as represented by the managerially-preserved state of artifactually-significant characteristics. The recasting of significance in terms of enabling affordances promotes a concomitant shift of evaluative attention from outputs to outcomes. Conceptualizing evaluative success in terms of affordances rather than characteristics shifts the basis of normative benchmarks from a primary concern for the *existence* of quantifiable ontological properties to that of the qualitative *effect* epistemologically afforded by a property with respect to purposive experiential use and consequent phenomenological response.

Heretofore, the four identified primary evaluative norms have been uniformly

presented in order of their inverse frequency as unique-to-context tokens across all six policies in the Predicate Reduction dataset: accessibility (43.1%), integrity (20.0%), authenticity (19.2%), usability (17.7%) (see Table 4.11). In terms of preservation imperatives, however, they are more properly ordered as a progression from initial necessity to final sufficiency: integrity, authenticity, accessibility, usability (IAAU).<sup>22</sup> The integrity of a preserved digital object can be adjudged independent of its authenticity or higher-order qualities. Similarly, authenticity is independent of accessibility, and accessibility, independent of usability. While an integral and authentic but inaccessible object may not present meaningful opportunities for consumer exploitation, that lack does not affect the object's possession of underlying integrity and authenticity. This suggests that the success of the digital preservation enterprise should be understood as an inherently multivalent quality.

Preservation success – that is, the degree of alignment of service-provider intention, consumer expectation, and preservation actuality and the corresponding level of stakeholder satisfaction resulting from that alignment – can, and should, be evaluated independently for each of the four normative elements. Furthermore, each of those elements can, and should, be evaluated independently in terms of the seven semiotic dimensions of the augmented ladder. Finally, each of those dimensions is susceptible to assessment in terms of its consequent outputs and outcomes. The resulting multi-dimensional evaluative space defines 56 distinct combinations of evaluative attention on the basis of consequent evaluative Norm, Semiotic dimension, and Modality (NSM) (see *Figure 5.5*).

For example, the point  $(n,s,m)$  highlighted in *Figure 5.5* represents the evaluative concerns of the outcomes of semantic authenticity. With that established as a defining principle, associated evaluative criteria could include the individual trustworthiness and community-warranted reliance confirmed with regard to the intellectual meaning of, and emotional response to, the behavioral performance of a preserved digital object. For example, in terms of the painting in *Figure 5.1*, is the perceived visual image actually Ellen Day Hale's portrait of Charlotte Perkins Gilman? Can it be used reliably for purposes of legitimate understanding of late 19th-century feminist culture and

---

<sup>22</sup> Following the precedent for AIAU (see note 14, p. 59), the acronym IAAU can be pronounced “EE-oh” /i:.oʊ/, a more historically and phonetically-correct pronunciation of Io.

cultural production? Metrics for trust could include verifiable statements of curatorial provenance for the underlying tangible artwork and the methodology of its colorimetrically-accurate capture in digital form at a sampling resolution consistent with thresholds of human optical acuity. All of this bolsters confidence that the resulting digital object authentically represents the authentic museum artifact at the point of object production. Once transferred to a responsible preservation program for ongoing stewardship, continual tracking of auditable change history provides further confidence in the authenticity of the object as being what it purports to be. The preservation of this auxiliary informative corpus alongside, or embedded within, the object itself ensures the successful persistence of an Authentic Semantic Outcome ( $n, s, m$ ) in future consuming contexts. A similar Communicologically-grounded process can be used to derive criteria and metrics for the other evaluative norms.

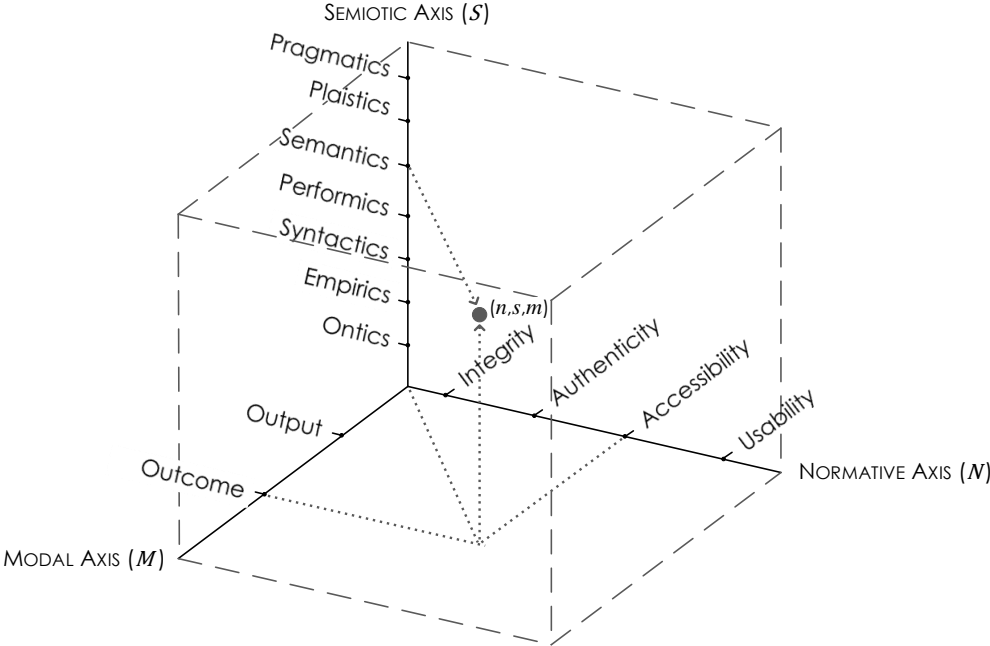


Figure 5.5. Multi-dimensional evaluative space

The success at each discrete normative position in the evaluative space can be determined independently of the others. As discussed in Sections § 5.1 – 5.4, each norm is situated along a scale of objectivity and intersubjectivity, suggesting that, in the general case, success is a quality of relative degree rather than absolute kind. This evaluative principle is particularly pertinent to the norms defined on the outcome plane of the evaluative space ( $n, s, m_o$ ), where  $n$  and  $s$  are free variables while  $m_o$  is a constant bound to the outcome modality. In order to provide a summary indication of

overall digital preservation success, the successes of individual norms can be combined into an aggregate scoring function:

$$\zeta = \frac{1}{N_N \cdot N_S \cdot N_M} \sum_{n=1}^{N_N} \sum_{s=1}^{N_S} \sum_{m=1}^{N_M} \omega_{n,s,m} \cdot \sigma_{n,s,m}$$

where  $\zeta$  is the final composite success score;  $N_N$ ,  $N_S$ , and  $N_M$  are the number of discrete evaluative points along the respective Normative ( $N$ ), Semantic ( $S$ ), and Modal ( $M$ ) axes of the evaluative space (see *Figure 5.5*);  $\omega_{n,s,m} \in [0.0,1.0]$  is a real-valued weighting function for the specific evaluative factor ( $n, s, m$ ), for  $n \in N, s \in S, m \in M$ , contingent on time, place, person, and purpose; and  $\sigma_{n,s,m} \in [0.0,1.0]$  is a real-valued success score specific to evaluative factor ( $n, s, m$ ) contingent on that same time, place, person, and purpose. The weighting function  $\omega$  is necessary to account for the unique determination of relative importance placed on the various individual norms in any given stakeholder context. The leading scaling factor  $1/N_N \cdot N_S \cdot N_M$  normalizes the composite score to the inclusive real-valued range  $[0.0,1.0]$ . Along this continuum, 1.0 represents complete digital preservation success and 0.0, total preservation failure. In practice, it is likely that both of those terminal points are theoretical conditions approached asymptotically but never fully realized.

Dependent on a particular weighting function  $\omega$  – particularly if it is skewed towards the origin point of the evaluative space, in other words, prioritizing managerial/artifactual semiotic concerns, lower-order normative factors, and resulting outputs – it is possible for  $\zeta = 1.0$ . This would be an indication that the digital object at play in the preservation-enabled communicative act is a complete contemporaneous *facsimile* of *all* significant ontological characteristics, epistemological affordances, and phenomenological experiences canonically, if situationally, understood to constitute the essence of that object. In the more general – and possibly realistic – case, success will fall somewhere within the exclusive range  $(0.0,1.0)$ , aspirationally-skewed towards the upper end of the range,  $0.0 \ll \zeta < 1.0$ , where  $\lim_{e \rightarrow 0.0} \zeta + e = 1.0$ . In this scenario, preserved objects should be fundamentally conceptualized as *surrogates* approximating *some* – presumably the most significant – characteristics, affordances, and experiences.

The intersubjectivity contingency and time-boundedness of an evaluation of digital preservation success positions it as a situational quality continually approached, but never definitively achieved. A digital object successfully preserved as of today could be at risk tomorrow as the actual state of the object's condition as well as the service-provider intentions and stakeholder expectations surrounding its use fluctuate. Nevertheless, a meaningful characterization of success provides service-providers with data central to prudent, responsible, and accountable stewardship. Measures of success similarly provide stakeholders with understanding critical to the formulation of rational information seeking plans, availability of plausible information engagement opportunities, and likelihood of beneficial information experiences regarding preserved digital collections.

The norms, criteria, and metrics underlying assessment of digital preservation success inhabit a multidimensional evaluative space. That space encompasses all existing evaluative factors tacitly endorsed by the digital preservation community as evidenced by the four primary imperative evaluative qualities underlying representative preservation policy statements. However, Communicological critique of these factors reveals the limitations of their traditional conceptualization emphasizing a too-narrow focus on the ontological characterization of artifactual properties. That characterization is of primarily managerial relevance within the wider contours of the preservation enterprise. The alternative evaluative framework presented above leverages the foundational evaluative power of the traditional approach. However, it also complements it with new capability to incorporate critical consideration of contingent epistemological and phenomenological concerns that underlie digital preservation activity, which is a fundamentally communicative endeavor. At its core, that endeavor is expansively intersubjective and humanistic rather than objectively technical. This framework supports digital preservation service-providers and stakeholders in the important work of communicating with the future, ensuring the persistence and continuity of personal, organizational, and social memory in the digital age.

## Chapter 6: Conclusion

---

This chapter highlights the findings (Section § 6.1); the intellectual, methodological, and practical contributions (Section § 6.2); and the potential limitations (Section § 6.3) of this doctoral study. It then outlines the direction of future opportunities to extend the scope and impact of the research program (Section § 6.4). Finally, it concludes with a summary of the full research activity (Section § 6.5).

Digital preservation is an information age enterprise grounded in an intentional act of memory, providing the future with critical illumination of the near or far past. However, despite the clear cultural importance of this function, the digital preservation community currently does not have at its disposal an actionable evaluative framework – let alone specific criteria and metrics – for characterizing the *success* of its activities. Instead, existing theory and practice focus on evaluative determinations of the *trustworthiness* of preservation programs and systems. While this is an important foundational component of assessment, it deals with the operational *means* of preservation activity rather than its teleological *ends*. Trustworthy programs and systems may be more prone towards final success, but success could nevertheless arise from seemingly, or even patently, untrustworthy conditions. On the other hand, it would be difficult to ascribe long-term confidence to purportedly trustworthy archival environments that consistently fail to yield successful outcomes. Thus, trustworthiness and success are complementary but independent measures of the preservation enterprise, with the former being an aspirationally-desirable but not fully-sufficient determinant of the latter. However, while trustworthiness is well-explicated in the scholarly and professional literature as well as being prominently incorporated into operational practice, the evaluative concept of success remains largely unexamined and unformalized.

The contemporary state of understanding of digital preservation assessment is contextualized within a tacit *managerial* conceptualization of the domain. That is, the legitimate scope of evaluative attention is assumed to be circumscribed by the programmatic environments, intentions, and practices of preservation service-providers (see *Figure 2.1*). The underlying motivation for prior research efforts can

be synthesized as an imperative to define the characteristics of digital preservation agency and systems bolstering confidence in their ability to perform their obligations. Trustworthiness emerges from this line of inquiry as the predominant evaluative factor for preservation management. However, as argued in this dissertation, a managerial perspective is insufficient for characterizing the stakeholder-perceived success of the fundamental preservation imperative of enabling future purposive *use* of preserved material, however trustworthily managed. Essentially, trustworthiness is a predicative metric for what is *expected* to happen, while success is a confirmatory determination of what *actually* did happen. In order to progress towards an actionable framework for elucidating final efficacy, this study investigates the pervasive conceptual and operational conditions contributing to the elusiveness of the effective measure of success.

Given that success has not received significant prior scrutiny, there is little explicit consideration of its measure found in the scholarly literature or professional practice. Consequently, this research relies upon indirect methods for revealing and critiquing tacit contemporary positions regarding the contours of success. The overall approach of this research is set by the primary research question: What are the parameters for a conceptually-sound, yet pragmatically-actionable evaluative framework for determining the communicative success of the digital preservation enterprise? Subsequent research activity is structured by three subordinate questions. First: What socially-constituted norms regarding digital preservation success emerge from evaluative attitudes implicit in domain discourse? The norms identified through this line of inquiry are then subject to the second question: How suitable for purpose are existing evaluative norms for digital preservation success? Under Communicological analysis, the norms are determined to be pertinent to a communicative as well as managerial perspective of the digital preservation enterprise. However, while the managerial norms are well-incorporated into current theory and practice, the communicative norms have been largely unexplored to date. This leads to the final subordinate question: What complementary enhancements to existing evaluative theory and practice are necessary for more effective and comprehensive characterization of digital preservation success? The evaluative model proposed here embraces a multivalent definition of success and a multi-dimension evaluative space in which to benchmark the results of digital preservation activity in a manner providing



a complementary sense of the experiential outcomes of that activity.

This dissertation research offers fresh insights into the nature of digital preservation activity, its conceptual foundations, and the operational assessment of its practice. These findings provide a path forward for significant improvements in the evaluative power and comprehensiveness of assessment of programmatic digital preservation activity. With more targeted feedback on the outcomes of those various activities, they can be planned and implemented in a more responsive, effective, and sustainable manner. Greater efficacy in managerial action will enhance the consequent communicative success as experienced by all stakeholders involved in and benefiting from the digital preservation enterprise.

The primary sources of evidence for this inquiry are representative digital preservation policy statements promulgated by a range of memory institutions whose missions encompass long-term stewardship of digital heritage. The terms of these policies implicitly establish the social contract of reciprocal intentions and expectations that underlie engagement between preservation service-providers and stakeholders, particularly regarding the preserved digital objects that are the informative vehicles for those interactions. In terms of Expectation-Confirmation Theory (ECT), stakeholder satisfaction regarding a product or service depends upon the degree of alignment between expectations, intentions, and the actual delivered product or service. In the digital preservation context, ECT suggests that policy terms function as tangible embodiments of underlying attitudinal positions pertinent to actorial and institutional satisfaction, and thus, preservation success. The evaluative norms implied by those attitudes can be recovered through Predicate Reduction, a novel form of Qualitative Content Analysis developed specifically for this study. Subsequently, the recovered norms are subject to Communicological analysis to identify their relative strengths and limitations as the basis for evaluation of preservation success. Under this critique, the recovered norms, as broadly constituted and deployed in contemporary theory and practice, are shown to be less than sufficient for a fully comprehensive measure of preservation efficacy in enabling future purposive use of past informative expression. However, the Communicological theory underlying this critique suggests a path forward towards a more effective evaluative framework. Repositioning digital preservation as an inherently semiotic act of meaningful signification of digitally-encoded information across time, and the ever-

growing technical and cultural distance consequent to the passage of time, provides a new, more comprehensive and evaluatively-powerful basis for assessing the ultimate success of the preservation enterprise.

## **6.1 KEY FINDINGS**

The research program for this dissertation unfolded in two distinct methodological phases. First, the newly-developed Predicate Reduction technique for inductive Qualitative Content Analysis was applied to a representative set of preservation policy statements. This established a set of primary evaluative norms tacit to controlling service-provider intentions and stakeholder expectations broadly accepted across the preservation domain (see Chapter 4 for more detail). Second, a Communicologically-grounded abductive philosophical inquiry was performed regarding the suitability of those norms as the basis for comprehensive and actionable assessment of the success of preservation outputs and outcomes (see Chapter 5). The complementary use of these two approaches provides new clarity regarding the historical elusiveness in deploying success as an operational measure of the digital preservation enterprise. That understanding, in turn, is critical to the future design and implementation of more robust principles and systems for characterizing preservation efficacy.

### **6.1.1 Tacit Evaluative Norms**

Twenty-eight unique normative attitudes were revealed through Predicate Reduction-based QCA of six paradigmatically-selected policies. However, only six norms are found in all six policy documents (see Table 4.10). Of these, two represent high-level statements of obligation – “preserve objects” and “preserve metadata” – that are too broad and unspecified for use as actionable evaluative metrics. The four remaining norms express more targeted evaluative obligations regarding assurances for the ongoing archival accessibility, integrity, authenticity, and usability (AIAU) of preserved digital objects (see *Figure 3.6*). In essence, these norms begin to delineate the more detailed, and measurable, obligations underlying the generic imperative to “preserve” objects and metadata.

### **6.1.2 Evaluative Suitability of Norms**

Relative to the ultimate communicative goal of digital preservation activity, three of the evaluative norms – integrity, authenticity, and accessibility – are

subordinate managerially-enabling qualities, while the fourth – usability – is a superordinate communicatively-enabled quality. That is, the first three characterize fundamental aspects of the ongoing management of preserved digital information objects in their role as information-bearing artifacts. They are functional assertions about the ontological state of a preserved digital object at a particular point of time in its managed history, namely, that the characterized object is, at the time of its characterization, susceptible to appropriate request and retrieval, that it is whole and uncorrupted relative to an accepted state, and that it is what it purports to be and can be relied upon as an informative artifact. Usability, on the other hand, addresses the efficacy of the epistemological process of human information *experience*. That is, the degree to which contingent use of a managerially-preserved object results in contextually-satisfactory purposive phenomenological effect. These two approaches toward evaluative assessment – the managerial/artifactual and the communicative/experiential – are complementary and mutually necessary for fully sufficient determination of the teleological success of the digital preservation enterprise.

## 6.2 CONTRIBUTIONS

The research presented in this dissertation makes several important contributions. In general, they promote an alternative controlling conceptual metaphor for the field, viewing digital preservation activity as fundamentally a communicative, and not only a managerial activity. More particularly, the contributions advance the theory and practice of assessment of the digital preservation domain in a more rigorous, comprehensive, and teleologically-relevant direction. These results firmly ground digital preservation activity as an ultimately humanistic – and not solely a technical – endeavor. That is, while those activities are embodied and enacted *somehow*, they are performed *by someone* on behalf of *someone*. The human productive and consuming agencies exist in co-equal partnership with, if not in a more teleologically-fundamental position regarding managerial instrumentality (see *Figure 2.2*). Digital preservation does remain an endeavor foundationally-concerned with the objective artifactual persistence and ontological integrity, authenticity, and accessibility of digital information objects. Beyond that, however, the preservation enterprise also should embrace as its final imperative the persistence of opportunities for legitimate communicative experiences of managerially-preserved objects. Those intersubjective, and thus situationally-contingent, experiences are adjudged successful

if they confer epistemologically-cognitive and -affective as well as phenomenologically-conative purposive outcomes. The determination of that success is critical for transparent managerial value-propositions, accountability, self-reflection, and improvement. This work provides new understanding of longstanding conceptual and practical impediments to the effective derivation of actionable norms, criteria, and metrics that should prove helpful in the future evaluation of the communicative success of preservation endeavors. All data collected for and produced during this study are available for future investigatory research into the evaluative or other aspects of the digital preservation enterprise.<sup>23</sup>

### **6.2.1 Communicological Perspective**

The expansion of digital preservation's conceptual basis from data management to the enablement of human communication across time (see *Figure 2.2*) provides a more rigorous and comprehensive foundation for scholarly investigation and professional practice in the field. Communicological theory (see the discussion in Section § 3.2) better explicates the full range of human productive, managerial, and consuming engagement with and response to preserved digital objects at each of the semiotic levels inherent to the ontological, epistemological, and phenomenological parameters of those engagements. Future application of a Communicological approach to the derivation of actionable metrics for evaluating the results of those engagements will provide scholars and practitioners with better means to express and assess preservations intentions, expectations, and outcomes with formal precision and rhetorical concision.

### **6.2.2 Formalization of Success**

Heretofore, the concept of evaluative success has not been applied to the digital preservation enterprise. Promoting success as the primary characterizing norm for preservation activity, particularly when repositioned as a communicative and not just managerial endeavor, enhances critical understanding of the field's theoretical and conceptual foundations. Given the service-provider/stakeholder relationship that is inherent to the effective practice of the enterprise, Expectation-Confirmation Theory

---

<sup>23</sup> All datasets and accompanying codebooks are freely available under the CC-By Attribution 4.0 International license at <<https://doi.org/10.17605/OSF.IO/ZHTQJ>>, <<https://doi.org/10.17605/OSF.IO/CSVBM>>, <<https://doi.org/10.17605/OSF.IO/X4SDN>>, and <<https://doi.org/10.17605/OSF.IO/75Q29>>.

(see the discussion in Section § 1.2) provides a useful guide for formalizing the concept of digital preservation success. In ECT terms, preservation activity is successful to the extent that service-provider intentions align with stakeholder expectations with regard to the preserved state of the digital object that is the communicative vehicle of the provider/stakeholder interaction. This formulation reduces the core evaluative problem to the availability and measurable equivalence of operative norms, criteria, and metrics capable of expressing the three pertinent states of intention, expectation, and actuality (see *Figure 1.2*). A formal definition of success is critical to future development of metrical systems to measure that success.

Prior research and practice in the field regarding the significant properties of digital objects, such as the InSPECT framework, offers a good foundation for characterizing object state (see the discussion in Section § 5.6). Similarly, the NLA's proposal for preservation intention statements and the new concept of explicit preservation objectives introduced in the proposed revision to ISO 14721 both hold out promise for expressing provider intentions, although neither has been implemented in widespread practice (see Sections § 1.2 and 2.2). However, any such proposed measures depends upon more complete understanding of the normative attitudes and imperatives emerging from underlying intentional and expectational positions.

### **6.2.3 Predicate Reduction Methodology**

Given that expressions of evaluative norms regarding digital preservation success are not explicit in domain discourse, alternative means are necessary to uncover tacit traces of pertinent evaluative attitudes. Under the ECT-derived definition of success, the contours and parameters of service-provider/stakeholder interactions are circumscribed by controlling preservation policy terms. The obligatory, conditional, and aspirational terms found in policy statements – i.e., “*P* will *X*”, “*P* may *X*”, “*P* should *X*” – outline imperative provider commitments while also instigating corresponding stakeholder assumptions. For example, if service-provider *P* affirmatively asserts a realistic intention *X*, then it is rational for stakeholder *S* to expect that *P* will make good on *X*. In other words, published policy documents are a manifest form of contract – generally more social than legal, but ideally no less professionally binding – between *P* and *S*.

Predicate Reduction is a new technique for Qualitative Content Analysis developed specifically for this inquiry. It defines a mechanistic process (see Section

§ 3.1.3) for identifying relevant policy statements by their grammatical function, expands those statements into singular propositions, reduces those propositions into imperative predicates, and finally uses the normalized forms of those predicates to construct synthetic evaluative kernels expressing core, if tacit, evaluative norms. The PR method relies on unobtrusive data collection of physical documents rather than human engagement modalities such as surveys or focus group sessions that can be problematic to arrange and moderate. Furthermore, the algorithmic nature of PR suggests future potential for automation that could permit its application to a much more widespread and comprehensive sampling of pertinent data sources. PR also may be found useful in other fields of inquiry in which important domain attitudes are recoverable from implicit expression in discursive artifacts.

#### **6.2.4 Empiric Normative Recovery**

Four primary evaluative norms regarding assurances of archival accessibility, integrity, authenticity, and usability (AIAU) of preserved digital objects emerge clearly through Predicate Reduction of representative digital preservation policy documents (see Chapter 4). These norms are broadly consistent with prior expressions of imperative preservation qualities in the scholarly and professional literatures, although these are often presented via anecdotal argument or as the result of *a priori* analysis. Furthermore, it has been unclear to what extent their literary expression has been incorporated into actual evaluative practice. Now, a fuller understanding of contemporary normative attitudes towards preservation success has been established empirically. PR analysis of the social “contracts” of policy terms that control preservation activities – albeit often implicitly – reveals the primary evaluative norms implicit to the discursive domain instruments that establish the working obligatory service-provider intentions and reciprocal stakeholder expectations underlying actual provider/stakeholder interactions.

Frequency of unique-to-context occurrence of the four primary evaluative norms in the policy sample set is used as a proxy for evaluative importance (see Section § 4.3). The norms are lexically-structured in decreasing order of normative significance; that is, as a normative value, accessibility is accorded twice the importance of usability (see *Figure 4.1*). However, for purposes of programmatic assessment of the digital preservation enterprise, these four norms are better placed in integrity-authenticity-accessibility-usability order (IAAU). This places them along an axis of successively

added-value function and assurance building from imperative necessity towards evaluative sufficiency (see Section § 6.1.1). Regardless, under Communicological critique (see Chapter 5) the four recovered evaluative norms are found to be insufficient for a comprehensive assessment of digital preservation success. A Communicological perspective offers a promising alternative approach addressing this limitation.

### 6.2.5 Significant Semiotic Affordances

Communicology views informative human-to-human interactions as inherently semiotic activity. That is, it unfolds through the creation, transmission, reception, and response to tangible information-bearing signs intersubjectively-signifying epistemic meaning and purposive phenomenological import (see the discussion in Section § 3.2). Human agency thus instigates and completes the semiotic process (see *Figure 2.2*) even if, in the digital preservation context, the primary vehicle underlying that process is a preserved digital preservation object. It is important, therefore, for evaluative characterizations of preserved objects to embrace the full range of agential commitments to and modalities of engagement with those objects. Thus, this research promotes extension of the concept of significant properties of objects to encompass the more evaluatively-relevant idea of significant affordances (see the discussion in Section § 5.6). Whereas *property* connotes a static ontological characteristic of an object in isolation, *affordance* positions the object within a dynamic multivalent relationship with controlling productive, managerial, and consuming agencies. Organizing evaluative assessment in terms of significant affordances conceptually shifts the perspective of evaluative attention from the simple characteristic property/value pair of properties to include interpretable factors important to actorial consumption and response. For example, a specific objective value of the object property of fixity supports the agentially-intersubjective factor of integrity (see *Figure 5.3*). This provides affordances with an evaluative range beyond the ontological characteristics of what a preserved digital object *is*, to also include the epistemological and phenomenological aspects of what an object permits its human consumer to *do* and *know*.

The evaluative embrace of significant affordances emphasizes the proper position of a preserved object within the preserved-enabled communicative process. The intersubjective positioning of the last two components of the characteristic-

affordance-understanding-response sequence (see *Figure 5.3*) illuminates the semiotic agent-centricity of that process. Whereas significant characteristics are externalities of a tangible preserved object, the understanding of and response to that object are internal to the consuming human agent. Affordances occupy the important liminal boundary between the two. Thus, they enable the critical evaluative transition from the foundationally-objective to the teleologically-intersubjective.

For comprehensive assessment of the relative success or failure of digital preservation activity, relevant evaluative affordances need to span the full range of semiotic concerns entailed by digital objects being proactively managed across archival timespans and accompanying technical and cultural distance. The preservation-augmented semiotic ladder – ontics, empirics, syntactics, performics, semantics, plaistics, pragmatics – divides the semiotic perspective into granular segments of object significance and agential concern (see *Figure 5.2*). These segments form one axis of a proposed multidimensional evaluative space encompassing orthogonal concerns for the four primary evaluative norms and critical distinction between quantitative preservation outputs and qualitative outcomes (see *Figure 5.5*). These structures form the basis for future research activity to begin deriving actionable evaluative norms, criteria, and metrics capable of measuring the success of the digital preservation-enabled communication.

### **6.3 LIMITATIONS**

The six policy documents examined in this study form a small sample (~6%) of the full document corpus assembled in the initial research phase. These six were selected purposefully through paradigmatic case sampling to act as exemplars fairly representing the scope and range of the policy obligations found throughout the corpus. In Qualitative Content Analysis, sample size is often the determinant as to whether a study's results should be considered suggestive but meaningfully-transferable, rather than indicative and fully generalizable (Jenson, 2008). The current results of this research fall into the first category. Future research effort should focus on extending the scope of analysis. The Predicate Reduction method appears susceptible to automation through natural language processing (NLP) (Friedman et al., 2013). NLP can be applied, for example, to the critical first PR step of statement identification through parts-of-speech (PoS) tagging (Tufiş & Ion, 2017) to determine relevant statements marked by copula, modal, or lexical verbs of obligation. If proven



effective, NLP, along with other automated workflows for semantic expansion of coordinated statement components, thesaural canonicalization, and kernel construction could simplify the application of the PR technique to a larger, if not the full set of collected policy documents for greater confidence in generalizable results.

As discussed in Section § 3.1.4, the use of evaluative norm frequency counts as proxies for conceptual significance is justified by the rhetorical function of the source policy documents, the contextually-sensitive selection of countable units, and the consistency of analytic findings relative to other pertinent expressions of preservation imperatives in domain discourse. Nevertheless, for purposes of bolstering greater confidence in this methodological assumption, future research activity should consider independent re-analysis of the sample policy set with an alternative approach emphasizing human assessment, coding, and determination of significance (Kracauer, 1952/2022; Mayring, 2014; Stemler, 2001). If, as expected, these new results confirm those presented here, that would strengthen credence in the current methodological design.

While the Predicate Reduction process is highly mechanistic in terms of its iterative grammatically-based textual transformations (see Section § 3.1.5), it does rely of interpretive analysis to identify the preferred terms and mapping rules for the predicate-canonicalizing thesaurus (Step 4 in Section § 3.1.2). Future research should investigate the potential for a more algorithmic approach, again potentially leveraging NLP concepts and technologies, such as concept extraction (Fu et al., 2020; Gul et al., 2022) and topic detection (Wartena & Brussee, 2008; Yang & Tang, 2022).

Programmatic digital preservation commitments can be expressed in terms of *guidance* or *control* policies (Becker et al., 2014; Sierman et al., 2013). Guidance policies define high-level strategic obligations or express general aspirational principles. Control policies, on the other hand, operate at a more tactical level of specific actions or conditions that are required, recommended, permissible, or prohibited (Madsen & Hurst, 2019). While the former outline the overall contours of evaluative norms, they do not immediately translate into actionable criteria or metrics. The detailed definitions of control policies make them more susceptible to direct evaluative application. This study did not attempt to position examined policy documents, or their obligatory terms, to a specific location along the guidance/control spectrum. Future investigations should classify policy terms by their intentional role

to ensure greater heterogeneity of examined policy documents and facilitate a more nuanced analytic examination of derived evaluative norms. This analysis should accept that those norms derived from guidance policies should not be expected to function as actionable metrics, whether due to lack of conceptual formalization or inadequate levels of implementable detail. This is the case, for example, with the norm of usability. Given the lack of definitional detail accompanying its references, it should be classified as an imperative guidance principle rather than an actionable control policy. As such, the analytic conclusion presented in Section § 5.4 that usability is not directly deployable as an operational evaluative metric is not surprising. Ideally, the policy obligations underlying each derived norm would be expressed twice within a policy document: first, as a guidance principle establishing the programmatic context, justification, and intended outcome of the obligation; and second, as one or more control policies providing the detail necessary for deployment of the obligatory norm as an actionable metric.

The institutional scope of preservation policies may be construed narrowly, but then be complemented with additional policies focusing on related concerns. For example, ICPSR publishes a policy regarding access to managed datasets that is distinct from its policy on the preservation of those datasets (ICPSR, 2018a). Thus, it is possible for a preservation policy to include no references to usability as the result of an intentional decision regarding policy scope. (This is not actually the case with ICPSR, as their preservation and access policies both incorporate references to usability issues.) For this study, the search for relevant policy documents was limited to those explicitly branded as “digital preservation” policies (see Section § 3.1.1). Future research efforts should expand the scope of collection to include all relevant policy documents. Sets of institutionally-interlocking policies should not be analyzed independently at the individual document level, but rather, with respect to the overall policy regime established by the document-spanning aggregation of pertinent policy terms. This will help to ensure the inclusion of all appropriate preservation obligations in future determinations of normative metrics for digital preservation success.

#### **6.4 FUTURE RESEARCH DIRECTION**

This dissertation represents one stage within a larger research program concerned with development and implementation of conceptually-rigorous and operationally-actionable evaluative measures of digital preservation activities,

especially when conceptually positioned as a managerially-mediated, but fundamentally communicative endeavor. Several subsequent research possibilities furthering that program emerge from the results and conclusions presented here.

#### **6.4.1 Expanded Sample Set**

While the six documents used in this study are paradigmatically-representative of preservation policy obligations broadly asserted by national, academic, public, and private libraries, archives, museums, datacenters, and institutional repositories, the sample set is relatively small (6 of 95 known published policies, or 6.1%). Performing additional PR analysis on a larger sampling of policy documents drawn from across the LAM, DC, and IR categories should provide further confirmation and higher levels of confidence in the results and conclusions presented here. Large-scale automation of the PR process in whole or part would facilitate PR analysis of the largest possible set of policies.

#### **6.4.2 Predicate Reduction Automation**

The mechanistic nature of the Predicate Reduction process suggests the potential for automated implementation. For example, Natural Language Processing (NLP) tools for Parts-of-Speech (PoS) analysis could be used to distinguish the characteristic copula, modal, and lexical verb markers indicative of obligatory statements of intention in policy documents (see Section § 3.1.3). If reliable, this would streamline the initial Statement Identification step of the PR process. Subsequent steps of Propositional Expansion and Predicate Reduction also could benefit from PoS analysis by demarcating propositional subjects from predicate verbs and objects. Automated Predicate Canonicalization is dependent on the availability of the canonicalizing thesaurus. Construction of the thesaurus, however, will probably require some form of human effort and review, although it is unclear what support could be provided by NLP Deep Learning and Concept Mapping algorithms. The final PR step of Kernel Construction and subsequent statistical manipulation, such as the manually-processed quantitative results presented in Section § 4.3, also appear to be easily scriptable.

#### **6.4.3 Expanded Evaluative Model**

As discussed in Section § 1.2, the intentional states of preservation service-providers, the expectational states of stakeholders, and the archived states of preserved digital object are central to the determination of preservation success (see *Figure 1.2*).

However, these three are a subset – albeit the critical subset for evaluative purposes as investigated in this study – of the full range of manifest artifactual and ideational actorial state-characterizations implicated in acts of preservation-enabled communication (see *Figure 6.1*).

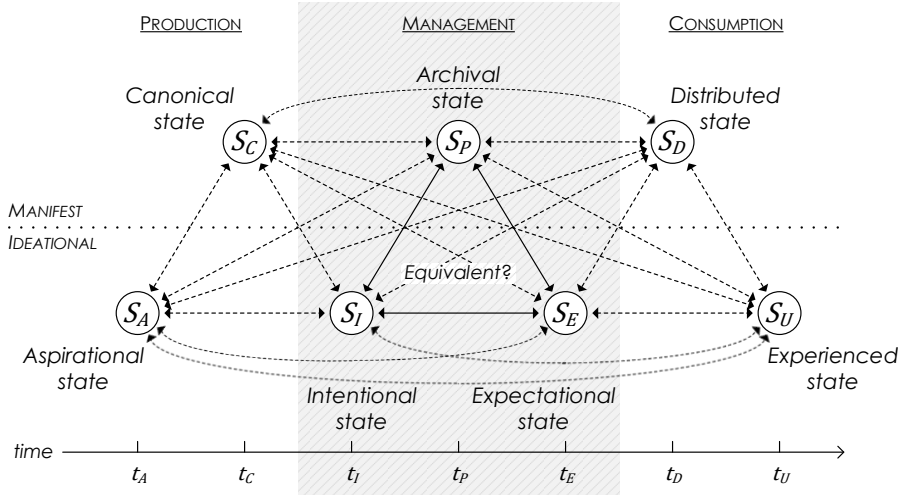


Figure 6.1. Expanded digital preservation states  
Cf. Figure 1.2

In particular these three are concerned with the mediating managerial component of preservation activity, characterizing what a service-provider *intends* to do ( $S_I$ ), what a stakeholder *expects* that provider to do ( $S_E$ ), and what the provider *actually* does ( $S_P$ ). However, additional states are necessary to encompass pre-managerial concerns of production, both the aspirational state of the producing agent ( $S_A$ ) and the resulting canonical state of the produced digital object ( $S_C$ ). Similarly, the consuming side of the preservation process contributes additional states to represent the actuality of the distributed preserved object ( $S_D$ ), which, as a uniquely-situated (re)performance, may differ in critical ways from its archival state  $S_P$  (see Section §1.4), and the final, and evaluatively-preminent, experiential state of epistemological interpretation and phenomenological response ( $S_U$ ). Expanding the scope of significant affordances to represent the evaluatively-critical aspects of all seven Communicologically-valid states would provide a more comprehensive evaluation of preservation outcomes.

**6.4.4 Actionable Criteria and Metrics**

The most significant follow-on research activity is the application of the findings presented here to the derivation of actionable evaluative norms, criteria, and metrics.

The multivalent evaluative space expressing the inherent interdependencies of evaluative norms, semiotic affordances, and result modalities (see *Figure 5.5*) provides a structure principle for the derivation task. All 56 normative-semiotic-modal combinations, from (Integrity, Ontics, Output) to (Usability, Pragmatics, Outcome), specify a bounded locus of evaluative attention within the full volumetric evaluative space, e.g., assessment of the integrity of ontic outputs such as fixity property values or usability of pragmatic outcomes such as legitimate purposive result. The formal definitions for the axial scales – the evaluative norms (Section § 6.1.1), the preservation-augmented semiotic ladder (see Section § 3.2), and the modal output/outcome distinction (Section § 5.7) – provide helpful guidance for Communicological derivation of concomitant metrics measuring the relative success of the pertinent preservation imperatives.

## 6.5 SUMMARY

Current evaluative theory and practice in the digital preservation domain focus on the trustworthiness of managed digital information objects, and the institutional programs and processes of their management. They do not incorporate concomitant consideration of the communicative success of the contingent information experience when engaging with those programmatically-preserved objects. Managerial trustworthiness remains an important evaluative factor for preservation outputs, but primarily as a suggestive measure that is *predictive* of ultimate preservation efficacy. A more complete and compelling assertion of the teleological success of preservation activity depends upon complementary *confirmatory* metrics capable of qualifying as well as quantifying the outcomes of actual purposive use enabled by prior trustworthy management. Existing criteria and metrics for assessing preservation trustworthiness are sufficient for addressing the objective *artifactual* persistence of integral, authentic, accessible, and usable digital objects. They are, however, less effective in characterizing the intersubjective *experiential* persistence of opportunities for legitimate human exploitation of those objects. Reconceptualizing digital preservation as a communicative enterprise, and not just a managerial one, helps to support a more comprehensive and rigorous foundation for evaluating success. It shifts the basis of evaluative focus from narrow consideration of intermediating managerial *means* to more expansive consideration of final communicative *ends*. The various conceptual distinctions explored in this study – managerial/communicative, artifactual/

experiential, objective/intersubjective, output/outcome – highlight and clarify important issues pertinent for new research in the digital preservation field. The new Communicological framework for the multi-dimensional evaluation of digital preservation success offers a promising path forward towards the derivation of actionable multivalent success metrics. The availability of such metrics will offer preservation practitioners better tools for responsibly fulfilling institutional imperatives, productively allocating finite programmatic resources towards that task, and effectively increasing relevancy, transparency, and accountability to stakeholders.

Even assuming the eventual availability of actionable evaluative metrics, it is important to emphasize that digital preservation success still will not be able to be asserted in a final and definitive manner. The factors complicating such an assessment include the open-ended time horizon of the preservation commitment; the impossibility of forecasting, and thus forestalling, all possible programmatic and technological risks, innovations, and disruptions; and the inexorable evolution and dislocation of cultural context and memory through the passage of time. Thus, the determination of success is an inherently relative rather than absolute process, as well as one that is ever-ongoing. The nature of digital information objects is fluid with respect to time and conditions of stewardship, users, and use. In McKemmish's formulation, "The [archival] record is always in a process of becoming" (Reed, 2005, p. 128). It follows that the experiential reception of a record – or more generally, a digital information object – also exists in a state of perpetual becoming, with actorial persona, context, and purpose unique to each act of communicative experience. Consequently, the measure of success for that experience is necessarily intersubjective and provisional. Given this, the most realistic aspirational goal for preservation activity is that preservation outcomes continually approach success asymptotically.

Progress towards this goal requires an effective framework for assessing the success of those outcomes; essentially, determining the – hopefully – vanishingly-small asymptotic gap in alignment of preservation service-provider intention, stakeholder expectation, and the semiotic affordances of the preserved digital object central to the provider-stakeholder interaction. This study has shown that the existing evaluative norms broadly accepted, if only tacitly, in scholarly and professional practice are insufficient for this purpose. However, the study also proposes a pathway towards sufficient norms, criteria, and actionable metrics through the application of

Communicological principles to the evaluative exercise. These principles refocus preservation attention to assessment of the enterprise's teleological goal, complementing existing measures of the accessibility and integrity of authentic digital information objects with those that effectively characterize consequent legitimate information experiences. The Communicological critique of evaluative norms for digital preservation success reveals significant impediments to the comprehensive measure of digital preservation efficacy due to fundamental constraints inherent in contemporary theory and practice. However, that critique's Communicological terms of reference also suggest an alternative approach to the evaluative problem. This both leverages the strengths of the artifactual perspective of current practice and augments that practice with necessary concern for the experiential aspects of the digital preservation enterprise. The resulting multivalent framework of evaluative norms, semiotic affordances, and consequential modalities offers the promise of fully effective and meaningful measures of satisfaction regarding the long-term stewardship of digital heritage through successful communicative acts of digital preservation.





# Appendices

---

## Appendix A

### Predicate Reduction Example

The following example illustrates the complete analytic/synthetic processing of the Predicate Reduction technique applied to a single policy statement, as described in Section § 3.1.3.

**Step 1: Statement Identification.** A coordinated statement is identified by a passive auxiliary lexical verb phrase expressing the fundamental mission of the preservation institution (see Table A.1). In this example, the statement is coordinated via “and” conjunctions regarding the component imperatives of the BMA’s mission, i.e., providing both care *and* access, and the objects of those imperatives, i.e., both collections *and* records.

Table A.1

*Statement Identification Example*

---

*Statement:* “As a public museum, the BMA is charged with caring for and providing access to its art collection and the records that support it, including a growing number of items in digital formats [emphasis added]” (BMA, 2016).

---

**Step 2: Propositional Expansion.** The coordinated statement is expanded into four singular propositions by applying each of the two imperatives to each of the two objects (see Table A.2).

Table A.2

*Propositional Expansion Example*

---

→ *Proposition:* “BMA is charged with caring for its art collection”  
→ *Proposition:* “BMA is charged with providing access to its art collection”  
→ *Proposition:* “BMA is charged with caring for [supporting] records”  
→ *Proposition:* “BMA is charged with providing access to [supporting] records”

---

**Step 3: Predicate Reduction.** Each proposition is reduced to its core *verb-object* predicate (see Table A.3). Given that all policy terms are concerned with the preservation of digital material, the specific objects of the second and fourth predicates are subject to reasonable implication and do not need to be represented explicitly, i.e., “provide access [to the collection]” and “provide access [to supporting records]”, respectively.

Table A.3

*Predicate Reduction Example*

---

→ <i>Analytic Predicate:</i>	“care for collection”
→ <i>Analytic Predicate:</i>	“provide access”
→ <i>Analytic Predicate:</i>	“care for [supporting] records”
→ <i>Analytic Predicate:</i>	“provide access”

---

**Step 4: Predicate Canonicalization.** Each predicate is passed through the thesaurus (Abrams, 2022b) for expression in canonical form (see Table A.4). In this case, the concepts of the predicative objects “collection” and “[supporting] records” are generalized to “objects” and “metadata”, respectively, as legitimate foci of preservation attention. Similarly, “care” and “provide” are mapped to the established terms “preserve” and “ensure”. Note that this canonicalization results in two instances of the predicate “ensure accessibility”.

Table A.4

*Predicate Canonicalization Example*

---

→ <i>Canonical Predicate:</i>	“preserve objects”
→ <i>Canonical Predicate:</i>	“ensure accessibility”
→ <i>Canonical Predicate:</i>	“preserve metadata”
→ <i>Canonical Predicate:</i>	“ensure accessibility”

---

**Step 5: Kernel Construction.** Each unique predicate is the basis for construction of three synthetic kernels expressing the service-provider obligation, stakeholder expectation, and implied evaluative criteria for their alignment (see Table A.5).

Table A.5

*Kernel Construction Example*


---

→ <i>Kernels:</i>	“P intends / to preserve objects / for S” “S expects / P / to preserve objects” “Did P / preserve objects / for S?”
→ <i>Kernels:</i>	“P intends / to ensure accessibility / for S” “S expects / P / to ensure accessibility” “Did P / ensure accessibility / for S?”
→ <i>Kernels:</i>	“P intends / to preserve metadata / for S” “S expects / P / to preserve metadata” “Did P / preserve metadata / for S?”

---

These kernels implicitly define three evaluative norms for digital preservation success regarding imperatives to preserve objects and metadata as well as ensure accessibility. The evaluative power of all PR-derived norms is subsequently assessed in the context of preservation’s communicative function in facilitating the transmission of past informative expression across time and accompanying technical and cultural distance for future consumption and understanding.

The complete set of mechanistic PR manipulations is shown in Table A.6.

Table A.6

*Full Predicate Reduction Example*


---

<i>Statement:</i>	“As a public museum, the BMA <u>is charged with</u> caring for <u>and</u> providing access to its art collection <u>and</u> the records that support it, including a growing number of items in digital formats [emphasis added]” (BMA, 2016).
→ <i>Proposition:</i>	“BMA is charged with caring for its art collection”
→ <i>Analytic Predicate:</i>	“care for collection”
→ <i>Canonical Predicate:</i>	“preserve objects”
→ <i>Kernels:</i>	“P intends / to preserve objects / for S” “S expects / P / to preserve objects” “Did P / preserve objects / for S?”

---

---

→ *Proposition*: “BMA is charged with providing access to its art collection”

→ *Analytic Predicate*: “provide access”

→ *Canonical Predicate*: “ensure accessibility”

→ *Kernels*: “P intends / to ensure accessibility / for S”

“S expects / P / to ensure accessibility”

“Did P / ensure accessibility / for S?”

→ *Proposition*: “BMA is charged with caring for [supporting] records”

→ *Analytic Predicate*: “care for [supporting] records”

→ *Canonical Predicate*: “preserve metadata”

→ *Kernels*: “P intends / to preserve metadata / for S”

“S expects / P / to preserve metadata”

“Did P / preserve metadata / for S?”

→ *Proposition*: “BMA is charged with providing access to [supporting] records”

→ *Analytic Predicate*: “provide access”

→ *Canonical Predicate*: “ensure accessibility”

→ *Kernels*: “P intends / to ensure accessibility / for S”

“S expects / P / to ensure accessibility”

“Did P / ensure accessibility / for S?”

---

## Appendix B

### Data File Derivation

Predicate Reduction results are managed in 11 data files (Abrams, 2022a), as described in Section § 4.1.<sup>24</sup> All files share the same structural organization of 72 fields or columns, conventionally labeled A through BT. The files are produced iteratively through the processing steps described below. Except where explicitly noted, all sorting is performed in ascending order, i.e., numerically smallest-to-largest and lexicographically by alphabetical order. The files for propositions, predicates, kernels, and rankings are paired, with one each for per-document and per-dataset statistics. The per-document files are sorted first by policy document and then by the reported element, i.e., propositions, predicates, etc. The per-data files are sorted first by the reported element and then by document.

#### B.1 RAW

The Raw data file (“Analysis\_1-raw”) presents the relevant policy statements and their derived expanded propositions, reduced analytic predicates, canonical predicates, and synthetic kernels in the natural reading order in which the statements were encountered in the six source documents. All other data fields are automatically calculated by formulas. At this stage of the analysis, the values for proposition and analytic and canonical predicate token and type counts, and expansion and duplication metrics are placeholders; the actual values are calculated in subsequently derived data files.

#### B.2 PROPOSITIONS

The per-document Proposition data file (“Analysis\_2-propositions\_d”) presents the propositions grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the Raw data file (“Analysis\_1-raw”);

---

<sup>24</sup> The Predicate Reduction dataset is available in Excel (.xlsx) and CSV (.csv) format, and its accompanying codebook in Word (.docx) and PDF (.pdf) format, at <https://doi.org/10.17605/OSF.IO/75Q29>.

2. Delete all instances of square brackets “[“ and “]”;
3. Select the Document, Context, Statement, and Proposition Reporting Unit columns A:N and paste back in place as literal values;
4. Select the Analytic Predicate frequency columns Y:AB and paste back in place as literal values;
5. Select the Canonical Predicate frequency columns AN:AQ and paste back in place as literal values;
6. Select the Synthetic Kernel frequency columns AY:BB and paste back in place as literal values; and
7. Select rows 4:546 and sort by Sampling unit (A), Proposition (O), per-document Coding Unit (H), per-document Context Unit (D), and Pg (F).

This automatically clusters Propositional instances and, since the formulas for their per-document token and type counts in R:S were left in place, recalculates those counts.

The per-dataset Proposition data file (“Analysis\_2-propositions\_t”) presents the propositions grouped across documents in globally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Proposition data file (“Analysis\_2-propositions\_d”);
2. Select the per-document Proposition Count columns (R:S) and paste back in place as literal value; and
3. Sort rows 4:546 by Proposition (O), per-dataset Coding Unit (G), per-dataset Context Unit (C), Pg (F), and Sampling Unit (A).

This re-clusters Propositional instances and recalculates their per-dataset token and type counts in P:Q.

### **B.3 ANALYTIC PREDICATES**

The per-document Analytic Predicate data file (“Analysis\_3-a\_predicates\_d”) presents the reduced predicates grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Proposition data file (“Analysis\_2-propositions\_t”);
2. Select the per-dataset Proposition Count columns (P:Q) and paste them back in place as literal values;
3. Select rows 4:546 and sort by Sampling Unit (A), Analytic Predicate (T), Proposition (O), per-document Coding Unit (H), per-document Context Unit (D), and Pg (F).

This automatically clusters Analytic Predicates instances and, since the formulas for their per-document token and type counts in W:X were left in place, recalculates those counts.

The per-document Analytic Predicate data file (“Analysis\_3-a\_predicates\_d”) presents the reduced predicates grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Analytic Predicate file (“Analysis\_3-a-predicates\_d”);
2. Select the per-dataset Analytic Predicate Count columns (W:X) and paste them back in place as literal values; and
3. Select rows 4:546 and sort by Analytic Predicate (T), Proposition (N), per-dataset Coding Unit (G), per-dataset Context Unit (C), and Pg (F), and Sampling Unit (A).

This re-clusters Analytic Predicate instances and recalculates their per-dataset token and type counts in U:V.

#### **B.4 CANONICAL PREDICATES**

The per-document Canonical Predicate data file (“Analysis\_4-c-predicates\_d”) presents the canonicalized predicates grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Analytic Predicate data file (“Analysis\_3-a-predicates\_t”);
2. Select the per-dataset Analytic Predicate Count columns (U:V) and paste back in place as literal values; and

3. Select rows 4:546 and sort by Sampling Unit (A), Canonical Predicate (AC), Analytic Predicate (T), Proposition (O) , per-document Coding Unit (H), per-document Context Unit (D), and Pg (F).

This automatically clusters Canonical Predicate instances and, since the formulas for their per-document token and type counts in AF:AG were left in place, recalculates those counts.

The per-dataset Canonical Predicate data file (“Analysis\_4-c-predicates\_t”) presents the predicates in globally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Canonical Predicate file (“Analysis\_4-c-predicates\_d”);
2. Select the per-dataset Canonical Predicate Count columns (AF:AG) and paste back in place as literal values; and
3. Sort rows 4:546 by Canonical Predicate (AC), Analytic Predicate (T), Proposition (N), per-dataset Coding Unit (G), per-dataset Context Unit (C), and Pg (F), and Sampling Unit (A).

This re-clusters Canonical Predicate instances and recalculates their per-dataset token and type counts in AD:AE.

## **B.5 KERNELS**

The per-document Kernels data file (“Analysis\_5-kernels\_d”) presents the evaluative kernels grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Canonical Predicate data file (“Analysis\_4-c-predicates\_t”);
2. Select the per-dataset Canonical Predicate Count columns (AD:AE) and paste them back in place as literal values; and
3. Sort rows 4:546 and Sort by Sampling Unit (A), Evaluative Kernel (AT), per-document Coding Unit (H), per-document Context Unit (D), and Pg (F).



This automatically clusters Evaluative Kernel instances and, since the formulas for their per-document token and type counts in AW:AX were left in place, recalculates those counts.

The per-dataset Kernels data file (“Analysis\_5-kernels\_t”) presents the evaluative kernels in globally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Kernel file (“Analysis\_5-kernel\_d”);
2. Select the per-document Kernel Count columns (AW:AX) and paste them back in place as literal values;
3. Select the per-document Unique-to-Statement Kernel Count columns (BF:BG) and paste them back in place as literal values;
4. Select the per-document Unique-to-Context Kernel Count columns (BO:BP) and paste them back in place as literal values; and
5. Sort rows 4:546 by Evaluative Kernel (AT), per-dataset Coding Unit (G), per-dataset Context Unit (C), and Pg (F), and Sampling Unit (A).

This re-clusters Evaluative Kernel instances and recalculates their per-dataset token and type counts in AU:AV.

## **B.6 RANKINGS**

The per-document Results data file (“Analysis\_6-rankings\_d”) presents evaluative norm frequency rankings grouped by document in locally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document data file (“Analysis\_5-kernels\_t”);
2. Select the per-dataset Kernel Count columns AU:AV and paste them back in place as literal values;
3. Select the per-dataset Unique-to-Statement Kernel Count columns BD:BE and paste them back in place as literal values;
4. Select the per-dataset Unique-to-Context Kernel Count columns BM:BN and paste them back in place as literal values; and
5. Sort rows 4:546 and Sort by Sampling unit (A), unique-to-context document Evaluative Kernel Frequency (BT) *in descending order*,

Evaluative Kernel (AT), per-document Coding Unit (H), per-Context Unit (D), and Pg (F).

By sorting first by sampling unit, this file reports the relative rankings of the evaluative norms grouped by policy document, with the most frequently referenced norm at the top and the least frequently referenced at the bottom of each document section.

The per-dataset Results data file (“Analysis\_6-rankings\_t”) present evaluative norm frequency rankings in globally-sorted lexicographic order. It is derived as follows:

1. Make a copy of the per-document Ranking file (“Analysis\_6-rankings\_d”);
2. Select the per-document Unique-to-Statement Frequency columns (BI:BJ) and paste back in place as literal values;
3. Select the per-document Unique-to-Context Frequency columns (BS:BT) and past back in place as literal values; and
4. Sort rows 4:546 by unique-to-context dataset Evaluative Kernel Frequency (BR) *in descending order*, Evaluative Kernel (AT), per-dataset Coding Unit (G), per-dataset Context Unit (C), Pg (F), and Sampling Unit (A).

By not first sorting by sampling unit, this file reports the relative rankings of the evaluative norms globally across all six documents, with the most frequently referenced norm at the top and the least frequently referenced at the bottom. These rankings are used to establish the common evaluative norms for digital preservation success tacitly underlying domain policy documents.

# Bibliography

---

- Abascal, J., Barbosa, S. D. J., Nicolle, C., & Zaphiris, P. (2016). Rethinking universal accessibility: A broader approach considering the digital gap. *Universal Access in the Information Society*, 15(2), 179-182. <https://doi.org/10.1007/s10209-015-0416-1>
- Abbott, D. (2012). Preserving interaction. In L. Konstantelos, J. Delve, D. Anderson, C. Billenness, D. Baker, & M. Dobрева (Eds.), *Preservation of Complex Objects, Volume 2: Software Art* (pp. 61-70). JISC. [http://radar.gsa.ac.uk/2806/1/pocos\\_vol\\_2\\_final\\_release\[1\].pdf](http://radar.gsa.ac.uk/2806/1/pocos_vol_2_final_release[1].pdf)
- Abrams, S. (2015, November). *A foundational framework for digital curation: The Sept domain model* 12th International Conference on Digital Preservation, Chapel Hill. <https://phaidra.univie.ac.at/o:429533>
- Abrams, S. (2018a). *Nothing succeeds like success: An approach for evaluating digital preservation efficacy* iPRES 2018, 15th International Conference on Digital Preservation, Boston. <https://escholarship.org/uc/item/8cp160c9>
- Abrams, S. (2018b). *Theorizing success: Measures for evaluating digital preservation efficacy* JCDL '18, 18th ACM/IEEE Joint Conference on Digital Libraries, Fort Worth. <https://escholarship.org/uc/item/7xt368b2>
- Abrams, S. (2020). *Digital Preservation Policy Documents* [Dataset]. <https://doi.org/10.17605/OSF.IO/ZHTQJ>
- Abrams, S. (2021). Tacit attitudinal principles for evaluating digital preservation success. *Archival Science*, 21(3), 295-315. <https://doi.org/10.1007/s10502-021-09360-5>
- Abrams, S. (2022a). *Predicate Reduction Results* [Dataset]. Open Science Framework. <https://doi.org/10.17605/OSF.IO/75Q29>
- Abrams, S. (2022b). *Predicate Reduction Thesaurus* [Dataset]. Open Science Framework. <https://doi.org/10.17605/OSF.IO/X4SDN>
- Abrams, S. (2023). *Policy Term Consistency* [Dataset]. <https://doi.org/10.17605/OSF.IO/CSVBM>
- Abrams, S., Morrissey, S. M., & Cramer, T. (2009). 'What! So what': The next-generation JHOVE2 architecture for format-aware characterization. *International Journal of Digital Curation*, 4(3), 123-136. <https://doi.org/10.2218/ijdc.v4i3.122>
- Adam, S. (2010). Preserving authenticity in the digital age. *Library Hi-Tech*, 28(4), 596-604. <https://doi.org/10.1108/07378831011096259>
- Adams, C., & Lloyd, S. (2003). *Understanding PKI: Concepts, Standards, and Deployment Considerations* (2nd ed.). Addison-Wesley.
- Ahmad, M., & Abawajy, J. H. (2014). Service level agreements for the digital library. *Procedia - Social and Behavioral Sciences*, 147, 237-243. <https://doi.org/10.1016/j.sbspro.2014.07.166>
- Aitchison, J., Gilchrist, A., & Bawden, D. (2000). *Thesaurus Construction and Use: A Practical Manual* (4th ed.). Aslib IMI.
- ALA. (2009). *Definitions of Digital Preservation*. American Library Association. Retrieved 2016-08-20 from <http://www.ala.org/alcts/resources/preserv/defdigpres0408>

- Alexander, H. G. (1988). *The Logic and Language of Philosophy*. University Press of America. (Revised and enlarged edition, (c) 1972 University of New Mexico Press)
- Allen, J. A. (2009). *The Feminism of Charlotte Perkins Gilman: Sexualities, Histories, Progressivism*. University of Chicago Press.
- Almsalam, S. (2014). The effects of customer expectation and perceived service quality on customer satisfaction. *International Journal of Business and Management Invention*, 3(8), 79-84.
- Altman, M., Adams, M. O., Crabtree, J., Donakowski, D., Maynard, M., Pienta, A., & Young, C. H. (2009). Digital preservation through archival collaboration: The Data Preservation Alliance for the Social Sciences. *American Archivist*, 72(1), 170-184. <https://doi.org/10.17723/aarc.72.1.eu7252lhnrp7h188>
- Amundsen, C., & Sohbat, E. (2008). The untapped learning potential of CMC in higher education. In S. Kelsey & K. St. Amant (Eds.), *Handbook of Research on Computer Mediated Communication* (Vol. 1, pp. 1-14). IGI Global. <https://doi.org/10.4018/978-1-59904-863-5.ch001>
- Anderson, D. (2015). The digital dark age. *Communications of the ACM*, 58(12), 20-23. <https://doi.org/10.1145/2835856>
- Anderson, J. A., & Colvin, J. W. (2003). An ecology of communication: An acknowledgement of Gregory Bateson. *American Journal of Semiotics*, 19(1-4), 35-67. <https://doi.org/10.5840/ajs2003191/42>
- Anderson, M., & LeFurgy, B. (2006). *The archive ingest and handling test: Implications of diverse content and diverse repository practices* Digital Curation and Trusted Repositories: Seeking Success Workshop, Chapel Hill. [https://ils.unc.edu/tibbo/JCDL2006/Lefurgy\\_jcdl\\_paper\\_3\\_wgl%202.pdf](https://ils.unc.edu/tibbo/JCDL2006/Lefurgy_jcdl_paper_3_wgl%202.pdf)
- Andow, J. (2016). Abduction by philosophers: Reorienting philosophical methodology. *Metaphilosophy*, 47(3), 353-370. <https://doi.org/10.1111/meta.12191>
- ANZ. (2022, February 11, 2022). *Digital Preservation*. Archives New Zealand. Retrieved January 16, 2023 from <https://www.archives.govt.nz/about-us/our-work/preservation/digital-preservation>
- Asadi, S., Abdullah, R., Yah, Y., & Nazir, S. (2019). Understanding institutional repository in higher learning institutions: A systematic literature review and directions for future research. *IEEE Access*, 7, 35242-25263. <https://doi.org/10.1109/ACCESS.2019.2897729>
- Ashley, L., & Mistic, M. (2019). Digital Preservation Capability Maturity Model (DPCMM): Genesis and practical uses. In S. Katuu (Ed.), *Diverse Applications and Transferability of Maturity Models* (pp. 152-167). IGI Global. <https://doi.org/10.4018/978-1-5225-7080-6.ch006>
- Austin, C., Brown, S., Humphrey, C., Leahey, A., & Webster, P. (2015). *Guidelines for the Deposit and Preservation of Research Data in Canada*. R. D. Canada. [https://ecccdocs.techno-science.ca/documents/ECCC\\_STB\\_ASTD\\_GuidelinesforDepositofResearchDatainCanada.pdf](https://ecccdocs.techno-science.ca/documents/ECCC_STB_ASTD_GuidelinesforDepositofResearchDatainCanada.pdf)
- Baggini, J., & Fosl, P. S. (2003). Types and tokens. In *The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods* (pp. 173-175). Blackwell.
- Bak, G. (2016). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, 16(4), 373-402. <https://doi.org/10.1007/s10502-015-9257-1>

- Bak, G., & Armstrong, P. (2008). Points of convergence: Seamless long-term access to digital publications and archival records at Library and Archives Canada. *Archival Science*, 8(4), 279-293. <https://doi.org/10.1007/s10502-009-9091-4>
- Ball, A. (2010). *Review of the State of the Art of the Digital Curation of Research Data* (ERIM Project Document erim1rep091103ab12). <http://opus.bath.ac.uk/19022/2/erim1rep091103ab12.pdf>
- Barateiro, J., Antunes, G., Freitas, F., & Borbinha, J. (2010). Designing digital preservation solutions: A risk management-based approach. *The International Journal of Digital Curation*, 5(1), 4-17. <https://doi.org/10.2218/ijdc.v5i1.140>
- Baskarada, S., & Kononios, A. (2013). Data, information, knowledge, wisdom (DIKW): A semiotic theoretical and empiric exploration of the hierarchy and its quality dimension. *Australasian Journal of Information Systems*, 18(1), 5-24. <https://doi.org/10.3127/ajis.v18i1.748>
- Baucom, E. (2021). Chapter 7. Maintaining digital materials over time. *Library Technology Reports*, 57(4), 30-33. <https://proquest.com/scholarly-journals/chapter-7-maintaining-digital-materials-over-time/docview/2535488229/se-2>
- Beagrie, N. (2006). Digital curation for science, digital libraries, and individuals. *The International Journal of Digital Curation*, 1(1), 14. <https://doi.org/10.2218/ijdc.v1i1.2>
- Beagrie, N., Semple, N., Williams, P., & Wright, R. (2008). *Digital Preservation Policies Study*. C. B. Limited.
- Becker, C. (2018). Metaphors we work by: Reframing digital objects, significant properties, and the design of digital preservation systems. *Archivaria*, 85, 6-36. <http://muse.jhu.edu/article/694332>
- Becker, C., Antunes, G., Barateiro, J., Vieira, R., & Borbinha, J. (2011). Control objectives for DP: Digital preservation as an integrated part of IT governance. *Proceedings of the American Society for Information Science and Technology*, 48(1), 1-10. <https://doi.org/10.1002/meet.2011.14504801124>
- Becker, C., Faria, L., & Duretec, K. (2014). Scalable decision support for digital preservation. *OCLC Systems & Services*, 30(4), 249-284. <https://doi.org/10.1108/OCLC-06-2014-0025>
- Becker, C., & Rauber, A. (2011). Decision criteria in digital preservation: What to measure and how. *Journal of the American Society for Information Science and Technology*, 62(6), 20. <https://doi.org/10.1002/asi.21527>
- Belkin, N. J. (2005). Anomalous state of knowledge. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of Information Behavior* (pp. 44-48). Information Today.
- Bell, G., & Grey, J. (2001). Digital immortality. *Communications of the ACM*, 44(3), 28-31. <https://doi.org/10.1145/365181.365182>
- Benjamin, W. (1936). The work of art in the age of mechanical reproduction (H. Zohn, Trans.). In H. Arendt (Ed.), *Illuminations*. Schocken Books. <https://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm>
- Berlo, D. (1960). *The Process of Communication: An Introduction to Theory and Practice*. Holt, Rinehart & Winston.
- Bermès, E., & Fauduet, L. (2011). The human face of digital preservation: Organizational and staff challenges, and initiatives at the Bibliothèque nationale de France. *International Journal of Digital Curation*, 6(1), 226-237. <https://doi.org/10.2218/ijdc.v6i1.184>

- Bertot, J. C. (2004). Libraries and networked information services: Issues and consideration in measurement. *Performance Measurement and Metrics*, 5(1), 11-19. <https://doi.org/10.1108/14678040410530340>
- Bettavia, R. S. (2016a). *Mapping significance of video games in OAIS* iPRES 2016, 13th International Conference on Digital Preservation, Bern. <https://phaidra.univie.ac.at/o:503173>
- Bettavia, R. S. (2016b). The power of imaginary users: Designated communities in the OAIS reference model. *Proceedings of the Association for Information Science and Technology*, 53(1), 1-9. <https://doi.org/10.1002/pra2.2016.14505301038>
- Beynon-Davies, P. (2010). The enactment of significance: A unified conception of information, systems, and technology. *European Journal of Information Systems*, 19, 389-408. <https://doi.org/10.1057/ejis.2010.34>
- Bhaskar, R. (1998). Philosophy and scientific realism. In M. Archer, R. Bhaskar, A. Collier, T. Lawson, & A. Norrie (Eds.), *Critical Realism: Essential Readings* (pp. 16-47). Routledge. (1997)
- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351-370. <https://doi.org/10.2307%2F3250921>
- BIPM. (2012). *International vocabulary of metrology – Basic and general concepts and associated terms (VIM) (JCGM 200:2012)*. B. I. d. P. e. Mesures. [https://www.bipm.org/utils/common/documents/jcgm/JCGM\\_200\\_2012.pdf](https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf)
- Bishop, B. W., & Hank, C. (2018). Measuring FAIR principles to inform fitness for use. *International Journal of Digital Curation*, 13(1), 35-46. <https://doi.org/10.2218/ijdc.v13i1.630>
- BL. *About Legal Deposit*. British Library. Retrieved January 14, 2023 from <https://www.bl.uk/legal-deposit/about-legal-deposit>
- Black, P. E., Scarfone, K. A., & Souppaya, M. P. (2008). Cyber security metrics and measures. In J. G. Voeller (Ed.), *Wiley Handbook of Science and Technology for Homeland Security* (pp. 1-8). Wiley. <https://doi.org/10.1002/9780470087923.hhs440>
- BMA. (2016). *Digital Preservation Policy*. B. M. o. Art. [http://files.archivists.org/groups/museum/standards//3.%20Records%20Management/BMA\\_DigitalPreservationPolicy\\_2016.pdf](http://files.archivists.org/groups/museum/standards//3.%20Records%20Management/BMA_DigitalPreservationPolicy_2016.pdf)
- BMA. (2021). *2021 Annual Report, July 1, 2020 - June 30, 2021*. B. M. o. Art. [https://cdn.artbma.org/2022/03/28121806/BMA\\_FY21\\_AnnualReport-2.pdf](https://cdn.artbma.org/2022/03/28121806/BMA_FY21_AnnualReport-2.pdf)
- Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research*, 19(4), 436-432. <https://doi.org/10.1108/QMR-06-2016-0053>
- Bollacker, K. D. (2010). Avoiding a digital dark age. *American Scientist*, 98(2), 106-110. <https://doi.org/10.1511/2010.83.106>
- Borgman, C., Van de Sompel, H., Scharnhorst, A., van den Berg, H., & Treloar, A. (2016). Who uses the digital data archive? An exploratory study of DANS. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4. <https://doi.org/10.1002/pra2.2015.145052010096>
- Botelho, F. H. F. (2021). Accessibility to digital technology: Virtual barriers, real opportunities. *Assistive Technology*, 33(sup1), 27-34. <https://doi.org/10.1080/10400435.2021.1945705>
- Bountouri, L., Gratz, P., & Sanmartin, F. (2018). Digital preservation: How to be trustworthy. In M. Ioannides (Ed.), *Digital Cultural Heritage* (pp. 364-374). Springer. [https://doi.org/10.1007/978-3-319-75826-8\\_29](https://doi.org/10.1007/978-3-319-75826-8_29)



- Boutard, G. (2016). Co-construction of meaning, creative processes and digital curation: The transmission of music with live electronics. *Journal of Documentation*, 72(4), 755-780. <https://doi.org/10.1108/JDOC-09-2015-0121>
- Brand, S. (1999). Escaping the digital dark age. *Library Journal*, 124(2), 46-48.
- Brocks, H., Kranstedt, A., Jaschke, G., & Hemmje, M. (2010). Modeling context for digital preservation. In E. Szczerbicki & N. T. Nguyen (Eds.), *Smart Information and Knowledge Management: Advances, Challenges, and Critical Issues* (pp. 197-226). Springer. <https://doi.org/10.1007/978-3-642-04584-4>
- Brunsmann, J., Wilkes, W., Schlageter, G., & Hemmje, M. (2012). State-of-the-art of long-term preservation in product lifecycle management. *International Journal on Digital Libraries*, 12(1), 27-39. <https://doi.org/10.1007/s00799-012-0081-4>
- Buckland, M. (2013). *Document theory: An introduction* Records, Archives and Memory Studies, Zadar, Croatia. <http://people.ischool.berkeley.edu/~buckland/zadardoctheory.pdf>
- Burbules, N. C., & Warnick, B. R. (2006). Philosophical inquiry. In J. L. Green, G. Camilli, P. B. Elmore, A. Skukauskaiti, & E. Grace (Eds.), *Handbook of Complementary Methods in Education Research*. Routledge. <https://doi.org/10.4324/9780203874769.ch29>
- Burda, D., & Teuteberg, F. (2013). Sustaining accessibility of information through digital preservation: A literature review. *Journal of Information Science*, 39(4), 442-458. <https://doi.org/10.1177/0165551513480107>
- Burgi, P.-Y., Cazeaux, H., & Eschernier, L. (2019). *A versatile solution for long-term preservation of research data: Data life-cycle management: The Swiss way* 16th International Conference on Digital Preservation, Amsterdam. <https://doi.org/10.17605/OSF.IO/YG9JT>
- Burns, J. E. (2017). The aura of materiality: Digital surrogacy and the preservation of photographic archives. *Art Documentation*, 36(1), 1-8. <https://doi.org/10.1086/691368>
- Butler, D. B. (2021). Language awareness to improve teaching, learning, and assessment matters. *Academia Letters*, 233, 1-6. <https://doi.org/10.20935/AL233>
- Caon, M. (2018). Towards cross-generational system design. In M. Kurosu (Ed.), *Human-Computer Interaction: Interaction in Context* (Vol. 10902, pp. 258-267). Springer. [https://doi.org/10.1007/978-3-319-91244-8\\_21](https://doi.org/10.1007/978-3-319-91244-8_21)
- Caplan, P. (2008). What is digital preservation? *Library Technology Reports*, 2, 7-9. <https://doi.org/10.5860/ltr.44n2>
- Caplan, P., Kehoe, W., & Pawletko. (2010, March). *Towards interoperable preservation repositories (TIPR)* US-DPIF '10: 2010 Roadmap for Digital Preservation Interoperability Framework Workshop, Gaithersburg. <https://doi.org/10.1145/2039274.2039290>
- Carnegie. (2018). *The Carnegie Classification of Institutions of Higher Education*. Indiana University Center for Postsecondary Research. Retrieved April 23 from <http://carnegieclassifications.iu.edu/>
- Carter, R., & McCarthy, M. (2006). *Cambridge Grammar of English*. Cambridge University Press.
- Catt, I. E. (2014). The two sciences of communication in philosophical context. *Review of Communication*, 14(3-4), 201-228. <https://doi.org/10.1080/15358593.2014.986876>

- Catt, I. E., & Eicher-Catt, D. (2010). Communicology: A reflexive human science. In *Communicology: The New Science of Embodied Discourse*. Fairleigh Dickinson University Press.
- CCSDS. (2019). Reference Model for an Open Archival Information System (OAIS). In (Draft issue 3 ed., Vol. CCSDS 650.0-P-3). Washington, DC: Consultative Committee for Space Data Systems.
- Chanod, J.-P., Dobрева, M., Rauber, A., Ross, S., & Casarosa, V. (2010, July). *Issues in digital preservation: Towards a new research agenda* Automation in Digital Preservation, Dagstuhl. <http://drops.dagstuhl.de/opus/volltexte/2010/2899/pdf/10291.Report.2899.pdf>
- Cheikh-Ammar, M. (2018). The IT artifact and its spirit: A nexus of human values, affordances, symbolic expressions, and IT features [Theory development]. *European Journal of Information Systems*, 27(3), 278-294. <https://doi.org/https://doi.org/10.1080/0960085X.2018.1436025>
- Chen, S.-S. (2007). Digital preservation: Organizational commitment, archival stability, and technological continuity. *Journal of Organizational Computing and Electronic Commerce*, 17(3), 205-2115. <https://doi.org/10.1080/10919390701294012>
- Cheney, J., Lagoze, C., & Botticelli, P. (2001, 2001). Towards a theory of information preservation. In P. Constantopoulou & I. T. Sølvsberg, ECDL 2001, 5th European Conference, Darmstadt.
- Chi, L., & Zhu, Z. (2017). Hashing techniques: A survey and taxonomy. *ACM Computing Surveys*, 50(1), 11.11-11.36. <https://doi.org/10.1145/3047307>
- Chowdhury, G. (2010). From digital libraries to digital preservation research: The importance of users and context. *Journal of Documentation*, 66(2), 207-223. <https://doi.org/10.1108/00220411011023625>
- CLIR. (2002). *The State of Digital Preservation: An International Perspective*. Council on Library and Information Resources. <https://www.clir.org/pubs/reports/pub107/pub107.pdf>
- Cloonan, M. I. V. (2015). Conservation and preservation of library and archival materials. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences* (3rd ed., pp. 1250-1268). CRC. <https://doi.org/10.1201/9780203757635>
- Cohen, S. M., & Reeve, C. D. C. (2020). Aristotle's metaphysics. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford University. <https://plato.stanford.edu/archives/win2020/entries/aristotle-metaphysics/>
- Condon, P. B. (2014). *Digital Curation Through the Lens of Disciplinarity in the Development of an Emerging Field* [Dissertation, Simmons College]. Boston. <https://beatleyweb.simmons.edu/scholar/files/original/c79af2ebd58dfb5cdf3a26d95f03419.pdf>
- Conway, P. (2010). Preservation in the age of Google: Digitization, digital preservation, and dilemmas. *Library Quarterly*, 80(1), 61-79. <https://doi.org/10.1086/648463>
- Cook, T. (2001). Archival science and postmodernism: New formulations for old concepts. *Archival Science*, 1(1), 22. <https://doi.org/10.1007/BF02435636>
- Cook, T. (2013). Evidence, memory, identity, and community: Four shifting archival paradigms. *Archival Science*, 13(2-3), 95-120. <https://doi.org/10.1007/s10502-012-9180-7>
- CoreTrustSeal. (2019). *CoreTrustSeal Trustworthy Data Repositories Requirements, 2020–2022*. CoreTrustSeal. <https://doi.org/10.5281/zenodo.3638211>



- Corrado, E. M., & Moulaison Sandy, H. (2017). *Digital Preservation for Libraries, Archives, and Museums*. Rowman & Littlefield.
- Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58(6), 22. [https://doi.org/10.1016/S1071-5819\(03\)00041-7](https://doi.org/10.1016/S1071-5819(03)00041-7)
- Creswell, J. W. (2014a). *Research design : qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Creswell, J. W. (2014b). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). Sage.
- CUL. (2018). *Cambridge University Libraries Digital Preservation Policy*. C. U. Libraries. <https://doi.org/10.17863/CAM.32927>
- Cunningham, A. (2008). Digital curation/digital archiving: A view from the National Archives of Australia. *The American Archivist*, 71(2), 530-543. <http://www.jstor.org/stable/40294529>
- da Silva Júnior, L. P., & Borges, M. M. (2016, November). *Fundamental concepts in digital preservation* TEEM '16, The Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca. <http://dx.doi.org/10.1145/3012430.3012530>
- Dale, R. L., & Ambacher, B. (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist*. C. f. R. Libraries. [https://www.crl.edu/sites/default/files/d6/attachments/pages/trac\\_0.pdf](https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf)
- Dallas, C. (2007). *An agency-oriented approach to digital curation theory and practice*. International Cultural Heritage Informatics Meeting, Toronto. <http://www.archimuse.com/ichim07/papers/dallas/dallas.html>
- Dallas, C. (2016). Digital curation beyond the 'wild frontier': A pragmatic approach. *Archival Science*, 16(4), 421-457. <https://doi.org/10.1007/s10502-015-9252-6>
- Danermark, B., Ekström, M., & Karlsson, J. C. (2019). *Explaining Society: Critical Realism in the Social Sciences* (2nd ed.). Routledge.
- Daniels, V. (1996). The chemistry of paper conservation. *Chemical Society Reviews*, 25(3), 179-186. <https://doi.org/10.1039/CS9962500179>
- Dappert, A., & Farquhar, A. (2009). *Significance is in the eye of the stakeholder* 13th European Conference on Research and Advanced Technology for Digital Libraries, [http://www.planets-project.eu/docs/papers/Dappert\\_Significant\\_Characteristics\\_ECDL2009.pdf](http://www.planets-project.eu/docs/papers/Dappert_Significant_Characteristics_ECDL2009.pdf)
- Davis, J. L. (2017). Curation: A theoretical treatment. *Information, Communication, & Society*, 20(5), 770-783. <https://doi.org/10.1080/1369118X.2016.1203972>
- Day, M. (2002). Digital preservation and long-term access to the content of electronic serials. In W. Jones (Ed.), *E-Serials: Publishers, Libraries, Users, Standards* (2nd ed., pp. 167-195). Haworth Information Press. <http://www.ukoln.ac.uk/metadata/publications/e-serials/>
- Day, M., Pennock, M., Smith, C., Jenkins, J., & Cooke, I. (2018). *Preservation planning for emerging formats at the British Library* 15th International Conference on Digital Preservation, Boston. <https://doi.org/10.17605/OSF.IO/58PJV>
- Dearborn, C., & Meister, S. (2017). Failure as process: Interrogating disaster, loss, and recovery in digital preservation [Research article]. *Alexandria*, 27(2), 83-93. <https://doi.org/10.1177/0955749017722076>
- Deegan, M., & Tanner, S. (2006). Key issues in digital preservation. In M. Deegan & S. Tanner (Eds.), *Digital Preservation* (pp. 1-31). Facet.

- Deely, J. (1982). *Introducing semiotic: Its history and doctrine*. Indiana University Press.
- Dempsey, L. (2000). Scientific, industrial, and cultural heritage: A shared approach. *Ariadne*, 22. <http://www.ariadne.ac.uk/issue/22/dempsey/>
- DeSilvey, C. (2006). Observed decay: Telling stories with mutable things. *Journal of Material Culture*, 11(3), 318-338. <https://doi.org/10.1177/1359183506068808>
- Domínguez Rubio, F. (2014). Preserving the unpreservable: Docile and unruly objects at MoMA. *Theory and Society*, 43(6), 617-645. <https://doi.org/10.1007/s11186-014-9233-4>
- Donaldson, D. R. (2016). The digitized archival document trustworthiness scale. *International Journal of Digital Curation*, 11(1), 252-270. <https://doi.org/10.2218/ijdc.v11i1.387>
- Donaldson, D. R. (2020). Certification information on trustworthy digital repository websites: A content analysis. *PLOS ONE*, 15(12), e0242525. <https://doi.org/10.1371/journal.pone.0242525>
- Donaldson, D. R., Zegler-Poleska, E., & Yarmey, L. (2020). Data managers' perspectives on OAIS designated communities and the FAIR principles: mediation, tools and conceptual models. *Journal of Documentation*, 76(6), 1261-1277. <https://doi.org/10.1108/JD-10-2019-0204>
- DPC. (2015). *Digital Preservation Handbook* (N. Beagrie, Ed. 2nd ed.). Digital Preservation Coalition. <https://www.dpconline.org/handbook>
- DPC. (2022a). *About the Digital Preservation Consortium*. DPC. Retrieved April 10, 2022 from <https://www.dpconline.org/about>
- DPC. (2022b). *Member List*. DPC. Retrieved April 16, 2022 from <https://www.dpconline.org/about/members>
- Dressler, V. A. (2017). The state of affairs with digital preservation at ARL member libraries: A survey and analysis of policy. *Digital Library Perspectives*, 33(2), 137-155. <https://doi.org/10.1108/DLP-08-2016-0030>
- Drisko, J. W., & Maschi, T. (2015). *Content Analysis*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190215491.001.0001>
- Dryden, J. (2011). Measuring trust: Standards for trusted digital repositories. *Journal of Archival Organization*, 9(2), 127-130. <https://doi.org/10.1080/15332748.2011.590744>
- Du, L., Ho, A. T. S., & Cong, R. (2020). Perceptual hashing for image authentication: A survey. *Signal Processing: Image Communication*, 81(115713), 1-23. <https://doi.org/10.1016/j.image.2019.115713>
- Dugan, R. E., & Hernon, P. (2002). Outcomes assessment: Not synonymous with inputs and outputs. *Journal of Academic Librarianship*, 28(6), 376-380. [https://doi.org/10.1016/S0099-1333\(02\)00339-7](https://doi.org/10.1016/S0099-1333(02)00339-7)
- Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria*, 39, 5-10. <https://archivaria.ca/index.php/archivaria/article/view/12063>
- Duranti, L. (2005). The long-term preservation of accurate and authentic digital data: The InterPARES project. *Data Science Journal*, 4, 106-118. <https://doi.org/doi.org/10.2481/dsj.4.106>
- Duranti, L. (2007). Reflections on InterPARES: The InterPARES 2 Project (2002-2007): An overview. *Archivaria*, 64(1), 113-121. <https://archivaria.ca/index.php/archivaria/article/view/13155>
- Duranti, L., & Blanchette, J.-F. (2004, April 20-23). *The authenticity of electronic records: The InterPARES approach* Archiving 2004, San Antonio.

<https://www.ingentaconnect.com/content/ist/ac/2004/00002004/00000001/art00046>

- Duranti, L., & Thibodeau, K. (2006). The concept of record in interactive, experiential and dynamic environments: The view of InterPARES. *Archival Science*, 6(1), 13-68. <https://doi.org/10.1007/s10502-006-9021-7>
- Eco, U. (1976). *A theory of semiotics* (First Midland Book edition ed.). Indiana University Press.
- Edelstein, O., Factor, M., King, R., Risse, T., Salant, E., & Taylor, P. (2011, November 1-4). *Evolving domains, problems and solutions for long term digital preservation* iPRES 2011, 8th International Conference on Preservation of Digital objects, Singapore. <http://www.l3s.de/~risse/pub/ipres2011.pdf>
- Eicher-Catt, D., & Catt, I. E. (2008). What can it mean to say that communication is 'effective' (and for whom) in postmodernity? *Atlantic Journal of Communication*, 16, 119-121. <https://doi.org/10.1080/15456870802086903>
- Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open*, 4(1). <https://doi.org/10.1177/2158244014522633>
- Faulkner, P., & Runde, J. (2019). Theorizing the digital object. *MIS Quarterly*, 43(4), 1279-1302. <https://doi.org/0.25300/MISQ/2019/13136>
- Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1), 6-16. <https://doi.org/10.1177/1558689809349691>
- Feng, Y., & Richards, L. (2018). A review of digital curation professional competencies: Theory and current practices. *Records Management Journal*, 28(1), 62-78. <https://doi.org/10.1108/RMJ-09-2016-0022>
- FileFormat.Info. (2022). *Online HexDump Utility*. In FileFormat.Info. <https://www.fileformat.info/tool/hexdump.htm>
- Fitzpatrick, T. (2010). Ellen Day Hale: Painting the self, fashioning identity. *Woman's Art Journal*, 31(1), 28-34. [https://link.gale.com/apps/doc/A271318528/AONE?u=gmilcs\\_amherst&sid=googleScholar&xid=0ef1dac4](https://link.gale.com/apps/doc/A271318528/AONE?u=gmilcs_amherst&sid=googleScholar&xid=0ef1dac4)
- Flick, U. (2018). *Doing Triangulation and Mixed Methods*. Sage. <https://doi.org/10.4135/9781529716634>
- Flouris, G., & Meghini, C. (2007). *Some preliminary ideas towards a theory of digital preservation* 1st International Workshop on Digital Libraries Foundations, Vancouver, BC. <https://www.ics.forth.gr/isl/publications/paperlink/DLF107.pdf>
- Fornäs, J. (2017). The hermeneutic concept of culture as meaning. In *Defending Culture: Conceptual Foundations and Contemporary Debate* (pp. 49-88). Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-57810-1>
- Fox, G. (2002). E-science meets computational science and information technology. *Computing in Science & Engineering*, 4(4), 84-85. <https://doi.org/10.1109/MCISE.2002.1014985>
- Frank, R. D. (2020). The social construction of risk in digital preservation. *Journal of the Association for Information Science and Technology*, 71(4), 474-484. <https://doi.org/10.1002/asi.24247>
- Frege, G. (1879). *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Verlag von Louis Nebert.

- Freitas, R. A. P., & Ramalho, J. C. (2010, September). *Significant properties in the preservation of relational databases* 14th European Conference on Digital Libraries, Glasgow. <http://hdl.handle.net/1822/13702>
- Frické, M. (2019). The knowledge pyramid: The DIKW hierarchy. *Knowledge Organization*, 46(1), 33-46. <https://doi.org/10.5771/0943-7444-2019-1-33>
- Friedman, C., Rindflesch, T. C., & Corn, M. (2013). Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the National Library of Medicine. *Journal of Biomedical Informatics*, 46(5), 765-773. <https://doi.org/10.1016/j.jbi.2013.06.004>
- Fu, S., Chen, D., He, H., Liu, S., Moon, S., Peterson, K. J., Shen, F., Want, L., Wang, Y., Wen, A., Zhao, Y., Sohn, S., & Liu, H. (2020). Clinical concept extraction: A methodology review. *Journal of Biomedical Informatics*, 109. <https://doi.org/10.1016/j.jbi.2020.103526>
- Furrie, B. (2009). *Understanding MARC Bibliographic Machine-Readable Cataloging*. Library of Congress. <http://www.loc.gov/marc/umb/>
- Gallinger, M., Bailey, J., Cariani, K., Owens, T., & Altman, M. (2017). Trends in digital preservation capacity and practice: Results from the 2nd bi-annual National Digital Stewardship Alliance storage survey. *D-Lib Magazine*, 23(7/8). <https://doi.org/10.1045/july2017-gallinger>
- Galloway, P. (2004). Preservation of digital objects. *Annual Review of Information Science and Technology*, 38(1), 549-590. <https://doi.org/10.1002/aris.1440380112>
- Gaur, A., & Kumar, M. (2018). A systematic approach to conducting review studies: An assessment of content analysis in 25 years of IB research. *Journal of World Business*, 53(2), 280-289. <https://doi.org/10.1016/j.jwb.2017.11.003>
- Giaretta, D. (2011). *Advanced Digital Preservation*. Springer. <https://doi.org/10.1007/978-3-642-16809-3>
- Giaretta, D., & Conway, E. (2011). Advanced preservation analysis. In *Advanced Digital Preservation* (pp. 233-264). Springer. [https://doi.org/10.1007/978-3-642-16809-3\\_14](https://doi.org/10.1007/978-3-642-16809-3_14)
- Giaretta, D., Matthews, B., Bicarregui, J., Lambert, S., Guericio, M., Michell, G., & Sawyer, D. (2009, October). *Significant properties, authenticity, provenance, representation information and OAIS* iPRES 2009, The Sixth International Conference on the Preservation of Digital Objects, San Francisco. <https://escholarship.org/content/qt0wf3j9cw/qt0wf3j9cw.pdf>
- Giaretta, D., Tzitzikas, Y., Marketakis, Y., & Christophides, V. (2011). Preservation of intelligibility of digital objects. In *Advanced Digital Preservation* (pp. 139-166). Springer. <https://doi.org/10.1007/978-3-642-16809-3>
- Gibbs, G. R. (2007). Thematic coding and categorizing. In *Analyzing Qualitative Data* (pp. 38-55). Sage. <https://doi.org/10.4135/9781849208574.n4>
- Given, L. M., & McTavish, L. (2010). What's old is new again: The reconvergence of libraries, archives, and museums in the digital age. *Library Quarterly*, 80(1), 7-32. <https://doi.org/10.1086/648461>
- Gladney, H. M. (2006). Principles for digital preservation. *Communications of the ACM*, 49(2), 111-116. <https://doi.org/10.1145/1113034.1113038>
- Grace, P. J., & Perry, D. J. (2013). Philosophical inquiry and the goals of nursing: A critical approach for disciplinary knowledge development and action. *Advances in Nursing Science*, 36(2), 64-79. <https://doi.org/10.1097/ANS.0b013e3182901921>



- Green, R. (1991). The profession's models of information: A cognitive linguistic analysis. *Journal of Documentation*, 47(2), 130-140. <https://doi.org/10.1108/eb026874>
- Greene, M. (2002). The power of meaning: The archival mission in the postmodern age. *American Archivist*, 65(1), 42-55. <https://doi.org/10.17723/aarc.65.1.1914668v881wv19n>
- Gribbin, J. (1984). *In Search of Schrodinger's Cat: Quantum Physics and Reality*. Bantam.
- Guest, G., MacQueen, K. M., & Namey, E. E. (2014). Introduction to applied thematic analysis. In *Applied Thematic Analysis* (pp. 3-20). Sage. <https://doi.org/10.4135/9781483384436>
- Gul, S., Rübiger, S., & Saygin, Y. (2022). Context-based extraction of concepts from unstructured textual documents. *INFORMATION SCIENCES*, 588, 248-264. <https://doi.org/10.1016/j.ins.2021.12.056>
- Hale, E. D. (ca. 1880). *Charlotte Perkins Stetson Gilman* [Oil on panel]. National Portrait Gallery, Washington. [https://npg.si.edu/object/npg\\_NPG.83.162](https://npg.si.edu/object/npg_NPG.83.162)
- Hamid, S. (1998). Constructing a global time capsule: Challenges in the digital preservation of society's cultural memory archive. *Information Technology and Libraries*, 17(4), 207-211. <http://search.proquest.com/scholarly-journals/constructing-global-time-capsule-challenges/docview/215830719/se-2>
- Hansson, J. (2005). Hermeneutics as a bridge between the modern and postmodern in library and information science [Research paper]. *Journal of Documentation*, 61(1), 12. <https://doi.org/10.1108/00220410510578032>
- Hao, Q., Luo, L., Jan, S. T. K., & Wang, G. (2021, November 15-19). *It's not what it looks like: Manipulating perceptual hashing based applications* ACM SIGSAG Conference on Computer and Communications Security (CCS '21), Korea. <https://doi.org/10.1145/3460120.3484559>
- Happe, J., Theilmann, W., Edmons, A., & Kearney, K. T. (2011). A reference architecture for multi-level SLA management. In P. Wieder, J. M. Butler, W. Theilmann, & R. Yahyapour (Eds.), *Service Level Agreements for Cloud Computing* (pp. 13-26). Springer. <https://doi.org/10.1007/978-1-4614-1614-2>
- Harris, R. (1987). *Reading Saussure: A Critical Commentary on the Cours de linguistique générale*. Open Court.
- Harvey, R., Mahard, M., & Conn, D. (2020). *Preservation Management Handbook: A 21st Century Guide for Libraries, Archives, and Museums* (2nd ed.). Rowman & Littlefield.
- Hass, C. (2017). *JPEGsnoop 1.8.0 - JPEG File Decoding Utility*. In (Version 1.8.0) <https://www.impulseadventure.com/photo/jpeg-snoop.html>
- Hedstrom, M., & King, J. L. (2004). *On the LAM: Library, archive, and museum collections in the creation and maintenance of knowledge communities* (Innovation in the Knowledge Economy: Implications for Education and Learning, Issue. O. f. E. C. a. Development. <https://www.oecd.org/education/innovation-education/32126054.pdf>
- Hedstrom, M., & Lee, C. (2002). *Significant properties of digital objects: Definitions, applications, implications* DLM Forum, Barcelona. <https://pdfs.semanticscholar.org/8735/6572cc184bb86a9a27315dc6dec41e8cd0cd.pdf>

- Hedstrom, M., Lee, C., Olson, J., & Lampe, C. (2006). 'The old version flickers more': Digital Preservation from the User's Perspective. *American Archivist*, 69(1), 159-187. <https://doi.org/10.17723/aarc.69.1.1765364485n41800>
- Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292(114523). <https://doi.org/10.1016/j.socscimed.2021.114523>
- Heravi, B., Cassidy, K., Davis, E., & Harrower, N. (2021). Preserving data journalism: A systematic literature review. *Journalism Practice*. <https://doi.org/10.1080/17512786.2021.1903972>
- Heslop, H., Davis, S., & Wilson, A. (2002). *An Approach to the Preservation of Digital Records*. [http://www.naa.gov.au/Images/An-approach-Green-Paper\\_tcm16-47161.pdf](http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf)
- Higgins, S. (2011). Digital curation: The emergence of a new discipline. *The International Journal of Digital Curation*, 6(2), 78-88. <https://doi.org/10.2218/ijdc.v6i2.191>
- Hirtle, P. B. (2008). The history and current state of digital preservation. In E. Westbrook & K. Jenkins (Eds.), *Metadata and Digital Collections: A Festschrift in Honor of Tom Turner* (pp. 121-140). Cornell University Library. <https://hdl.handle.net/1813/45862>
- Hjørland, B. (2018). Library and information science (LIS), Part 1. *Knowledge Organization*, 45(3), 232-254. <https://doi.org/10.5771/0943-7444-2018-3-232>
- Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. *Program: Electronic library and information systems*, 40(3), 232-243. <https://doi.org/10.1108/00330330610681312>
- Hockx-Yu, H., & Knight, G. (2008). What to preserve? Significant properties of digital objects. *International Journal of Digital Curation*, 3(1), 141-153. <https://doi.org/10.2218/ijdc.v3i1.49>
- Hossain, M. A., & Quaddus, M. (2012). Expectation–confirmation theory in information system research: A review and analysis. In Y. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information Systems Theory: Explaining Our Digital Society, Vol. 1*. Springer. <https://doi.org/10.1007/978-1-4419-6108-2>
- Howell, D., & Snijgers, L. (2020). Spectral imaging. In *Conservation Research in Libraries* (pp. 172-196). De Gruyter. <https://doi.org/https://doi.org/10.1515/9783110375374-008>
- ICA. (2016). *Multilingual Archival Terminology*. I. C. o. Archives. <http://www.ciscra.org/mat/>
- ICPSR. *About ICPSR*. ICPSR. Retrieved January 15, 2023 from <https://www.icpsr.umich.edu/web/pages/about/>
- ICPSR. (2018a). *ICPSR Access Policy Framework*. I-U. C. f. P. a. S. Research. <https://www.icpsr.umich.edu/web/pages/datamanagement/policies/access-policy-framework.html>
- ICPSR. (2018b). *ICPSR Digital Preservation Policy Framework*. I-U. C. f. P. a. S. Research. <https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dpp-framework.html>
- ICPSR. (2021). *ICPSR, The Data Consortium: Strategic Plan, 2017 - 2021*. ICPSR. [https://www.icpsr.umich.edu/files/videos/ExecutiveSummary05\\_2018\\_v5.pdf](https://www.icpsr.umich.edu/files/videos/ExecutiveSummary05_2018_v5.pdf)

- IEC. (2015). Universal serial bus interfaces for data and power -- Part 2-1: Universal Serial Bus Specification. In (2.0 ed., Vol. 62680-2-1). Geneva: IEC.
- IEEE. (2016). Standard for Information Technology — Portable Operating System Interface (POSIX): Base Specifications, Issue 7. In (Vol. Std 1003.1). New York: IEEE.
- Innocenti, P., Ross, S., Maceviciute, E., Wilson, T., Ludwig, J., & Pempe, W. (2009, October). *Assessing digital preservation frameworks: The approach of the SHAMAN project* MEDES '09: International Conference on Management of Emergent Digital EcoSystems, France. <https://doi.org/10.1145/1643823.1643899>
- Innocenti, P., Vullo, G., & Ross, S. (2010). Towards a digital library policy and quality interoperability framework: The DL.org project. *New review of Information Networking*, 15(1), 29-53. <https://doi.org/10.1080/13614571003751071>
- InterPARES. (2008). *InterPARES 2 Project Glossary* (L. Duranti & R. Preston, Eds.). Associazione Nazionale Archivistica Italiana. [http://www.interpares.org/ip2/display\\_file.cfm?doc=ip2\\_glossary.pdf](http://www.interpares.org/ip2/display_file.cfm?doc=ip2_glossary.pdf)
- InterPARES. (2022). InterPARES 2 Project Dictionary. In *InterPARES 2 Project Dictionary*. Retrieved 2022-02-05, from [http://www.interpares.org/ip2/display\\_file.cfm?doc=ip2\\_dictionary.pdf](http://www.interpares.org/ip2/display_file.cfm?doc=ip2_dictionary.pdf)
- ISO. (2012a). Space data and information transfer systems — Audit and certification of trustworthy digital repositories. In (Vol. 16363). Geneva: ISO.
- ISO. (2012b). Space data and information transfer systems — Open archival information system (OAIS) — Reference model. In (Vol. ISO 14721). Geneva: ISO.
- ISO. (2013). Information technology – Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF) – Part 5. In (Vol. 10918-5). Geneva: ISO/IEC.
- ISO. (2016). Information and documentation — Records management — Part 1: Concepts and principles. In (Vol. ISO 15489-1). Geneva: ISO.
- ISO. (2017). Document management — Portable document format — Part 2: PDF 2.0. In (Vol. ISO 32000). Geneva: ISO.
- ISO/IEC. (1994). Information technology — Digital compression and coding of continuous-tone still images: Requirements and guidelines. In (Vol. 10918-1). Geneva: IEEE/IEC.
- Jääskeläinen, A. (2015, May 19-25). *Realizing the potentiality of metrics in digital preservation* IS&T Archiving Conference 2015, Los Angeles. <https://www.ingentaconnect.com/contentone/ist/ac/2015/00002015/00000001/art00017>
- Jakobson, R. (1960). Linguistics and poetics. In K. Pomorska & S. Rudy (Eds.), *Language in Literature* (pp. 62-94). Harvard University Press.
- Janée, G., Frew, J., & Moore, T. (2009). Relay-supporting archives: Requirements and progress. *International Journal of Digital Curation*, 4(1), 57-70. <https://doi.org/10.2218/ijdc.v4i1.78>
- Jatkevicius, J. (2005). Truth vs. authenticity: Dealing with 'good' books going 'bad' and seeking opportunities for catalog record enhancement. *Public Library Quarterly*, 24(5), 35-43. [https://doi.org/10.1300/J118v24n04\\_02](https://doi.org/10.1300/J118v24n04_02)
- Jeffrey, S. (2012). A new digital dark age? Collaborative web tools, social media and long-term preservation. *World Archaeology*, 44(4), 553-570. <https://doi.org/10.1080/00438243.2012.737579>

- Jenson, D. (2008). Transferability. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 888-887). Sage. <https://doi.org/10.4135/9781412963909>
- Jeong, J., Kurnia, S., Samson, D., & Cullen, S. (2018). *Psychological contract in IT outsourcing: A systematic literature review* 51st Hawaii International Conference on System Sciences, Waikoloa Village. <http://hdl.handle.net/10125/49906>
- Johansson, I. (2017). In defense of the notion of truthlikeness. *Journal for General Philosophy of Science*, 48(1), 59-69. <https://doi.org/10.1007/s10838-016-9343-y>
- Johnston, L. (2008). Principles and activities of digital curation for developing successful and sustainable repositories. In K. Skinner, M. Halbert, & M. Battle (Eds.), *Strategies for Sustaining Digital Libraries* (pp. 84-96). Emory University. <https://educopia.org/strategies-for-sustaining-digital-libraries/>
- Johnston, L. (2020a). Challenges in preservation and archiving digital materials. *Information Services & Use*, 1-7. <https://doi.org/10.3233/ISU-200090>
- Johnston, L. (2020b). Implementing a framework for digital preservation risk assessment and mitigation at the US National Archives. *Journal of Digital Media Management*, 8(4), 351-360. <https://www.ingentaconnect.com/content/hsp/jdmm/2020/00000008/00000004/art00007>
- Jones, R., Andrew, T., & MacColl. (2006). *The Institutional Repository*. Chandos.
- Kallinikos, J., Aaltonen, A., & Marton, A. (2010). A theory of digital objects. *First Monday*, 15(6-7). <https://doi.org/10.5210/fm.v15i6.3033>
- Kankam, P. K. (2019). The use of paradigms in information research. *Library & Information Science Research*, 41(2), 85-92. <https://doi.org/10.1016/j.lisr.2019.04.003>
- Kaplan, P. (2008). What is digital preservation? *Library Technology Reports*, 43(2), 7-9. <https://doi.org/10.5860/ltr.44n2>
- Kaplan, R. S., & Norton, D. P. (2004). *Strategy Maps: Converting Intangible Assets Into Tangible Outcomes*. Harvard Business School Press.
- Kastenhofer, J. (2015). The logic of archival authenticity: ISO 15489 and the varieties of forgeries in archives. *Archives and Manuscripts*, 43(3), 166-180. <https://doi.org/10.1080/01576895.2015.1074085>
- Keller, M. (2009). *Establishing a Digital Library* [White paper]. S. Microsystems. [https://www.academia.edu/download/53231066/establishing\\_digital\\_library\\_MAKFeb2009.pdf](https://www.academia.edu/download/53231066/establishing_digital_library_MAKFeb2009.pdf)
- Kellerman, A. (2000). Phases in the rise of the information society. *Info*, 2(6), 537-541. <https://www.ingentaconnect.com/content/mcb/272/2000/00000002/00000006/art00002>
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the Association for Information Science and Technology*, 59(3), 363-374. <https://doi.org/10.1002/asi.20722>
- Ketelaar, E. (2012). Cultivating archives: Meanings and identities. *Archival Science*, 12, 15. <https://doi.org/10.1007/s10502-011-9142-5>
- Kim, D. J. (2012). An investigation of the effect of online consumer trust on expectation, satisfaction, and post-expectation. *Information Systems and E-Business Management*, 10(2), 219-240. <https://doi.org/10.1007/s10257-010-0136-2>



- Kim, S., & Choi, M.-S. (2016, February 3-6). *Study on data center and data librarian role for reuse of research data* KST 2016, 8th International Conference on Knowledge and Smart Technology, <https://doi.org/10.1109/KST.2016.7440517>
- Kirchhoff, A. J. (2008). Digital preservation: Challenges and implementation. *Learned Publishing*, 21(4), 285-294. <https://doi.org/10.1087/095315108X356716>
- Knezek, G., & Christensen, R. (2001). Impact of new information technologies on teachers and students. In D. Watson & J. Andersen (Eds.), *Networking the Learner* (pp. 169-178). Springer. [https://doi.org/10.1007/978-0-387-35596-2\\_17](https://doi.org/10.1007/978-0-387-35596-2_17)
- Knight, G. (2009). *InSPECT Framework Report* (Work package 3.3). <http://www.significantproperties.org.uk/inspect-framework.pdf>
- Knight, G., & Pennock, M. (2009). Data without meaning: Establishing the significant properties of digital research. *The International Journal of Digital Curation*, 4(1), 16. <https://doi.org/10.2218/ijdc.v4i1.86>
- Korenkova, M., & Hägerfors, A. (2011, May 16-19). *Quality criteria for digital information in long-term digital preservation* Archiving 2011 Conference, Salt Lake City. <https://www.ingentaconnect.com/contentone/ist/ac/2011/00002011/00000001/art00010>
- Kowalczyk, S. (2018). *Digital Curation for Libraries and Archives*. Libraries Unlimited.
- Kracauer, S. (1952/2022). The challenge of qualitative content analysis. In J. Kang, G. Gilloch, & J. Abromeit (Eds.), *Selected Writings on Media, Propoganda, and Political Communication* (pp. 322-332). Columbia University Press. <http://www.jstor.org/stable/10.7312/krac15896.27> (Public Opinion Quarterly 16(4): 631-642)
- Krampen, M. (1997a). Models of semiosis. In R. Posner, K. Robering, & T. A. Sebeok (Eds.), *Semiotik: Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur [Semiotics: A Handbook on the Sign-Theoretic Foundation of Nature and Culture]* (Vol. 1). de Gruyter. <https://doi.org/10.1515/9783110095845.1.2.247>
- Krampen, M. (1997b). Semiosis of the mass media: Modeling a complex process. In W. Nöth (Ed.), *Semiotics of the Media: A State of the Art, Projects, and Perspectives* (pp. 87-98). de Gruyter. <https://doi.org/10.1515/9783110803617-006> (1997)
- Krippendorff, K. (2019). *Content Analysis: An Introduction to Its Methods* (4th ed.). Sage.
- Kuhlthau, C. C. (2017). Information search process (ISP) model. In J. D. McDonald & M. Levine-Clark (Eds.), *Encyclopedia of Library and Information Sciences* (4th ed., pp. 2232-2238). CRC Press. <https://doi.org/10.1081/E-ELIS4>
- Kyrillidou, M. (2002). From input and output measures to quality and outcome measures, or, from the user in the life of the library to the library in the life of the user. *Journal of Academic Librarianship*, 28(1-2), 42-26. [https://doi.org/10.1016/S0099-1333\(01\)00299-3](https://doi.org/10.1016/S0099-1333(01)00299-3)
- L'Hours, H., Kleemola, M., & de Leeuw, L. (2019). CoreTrustSeal: From academic collaboration to sustainable services. *IASSIST Quarterly*, 43(1), 1-17. <https://doi.org/10.29173/iq936>

- Langley, S. (2019). Digital preservation should be more holistic: A digital stewardship approach. In J. Myntti & J. Zoom (Eds.), *Digital Preservation in Libraries: Preparing for a Sustainable Future* (pp. 93-128). ALA.
- Lanigan, R. L. (2008). Communicology. In W. Donsbach (Ed.), *The International Encyclopedia of Communication*. Blackwell.  
<https://doi.org/10.1111/b.9781405131995.2008.x>
- Lanigan, R. L. (2010a). Verbal and nonverbal codes of communicology: The foundation of interpersonal agency and efficacy. In D. Eicher-Catt & I. E. Catt (Eds.), *Communicology: The New Science of Embodied Discourse* (Vol. Fairleigh Dickinson University Press, pp. 102-128). Madison & Teaneck.
- Lanigan, R. L. (2010b). Verbal and nonverbal codes of communicology: The foundation of interpersonal agency and efficacy. In D. Eicher-Catt & I. E. Catt (Eds.), *Communicology: The New Science of Embodied Discourse* (pp. 102-128). Fairleigh Dickinson University Press.  
[https://www.academia.edu/attachments/42946370/download\\_file?st=MTQ5MjY1OTQ0NSw2Ny4xODguNTkuMTQ3LDUxNDkyMTQ0&s=swp-toolbar&ct=MTQ5MjY1OTQ1MCwyMiw1MTQ5Mje0NA==](https://www.academia.edu/attachments/42946370/download_file?st=MTQ5MjY1OTQ0NSw2Ny4xODguNTkuMTQ3LDUxNDkyMTQ0&s=swp-toolbar&ct=MTQ5MjY1OTQ1MCwyMiw1MTQ5Mje0NA==)
- Lanigan, R. L. (2013). Information theories. In P. Cobby & P. Schulz (Eds.), *Theories and Models of Communication* (Vol. 1, pp. 58-83). De Gruyter Mouton.
- Lanigan, R. L. (2015). Philosophy of communicology: "Discourse which expresses itself is communication". *Review of Communication*, 15(4), 349-358.  
<https://doi.org/10.1080/15358593.2015.1102408>
- Lasswell, H. D. (1948). The structure and function of communication in society. In B. Lyman (Ed.), *The Communication of Ideas: A Series of Addresses* (pp. 37-51). Jewish Theological Seminary of America.
- Lavoie, B. (2014). *The Open Archival Information System (OAIS) Reference Model: Introductory Guide* (14-02). (DPC Technology Watch Report, Issue. D. P. Coalition. <http://dx.doi.org/10.7207/twr14-02>
- Lavoie, B., & Dempsey, L. (2004). Thirteen ways of looking at ... digital preservation. *D-Lib Magazine*, 10(7/8). <https://doi.org/10.1045/july2004-lavoie>
- LC. (2015). *PREMIS Data Dictionary for Preservation Metadata*. L. o. Congress.  
<http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
- Lee, C. A., & Tibbo, H. (2007). Digital curation and trusted repositories: Steps towards success. *Journal of Digital Information*, 8(2).  
<https://journals.tdl.org/jodi/index.php/jodi/article/view/229/183>
- Lester, P. (2018). Of mind and matter: The archives as object. *Archives and Records*, 39(1), 73-87. <https://doi.org/10.1080/23257962.2017.1407748>
- Lewine, D. (1991). *POSIX Programmer's Guide: Writing Portable UNIX Programs with the POSIX.1 Standard* (D. Dougherty, Ed.). O'Reilly.
- Li, Y., & Banach, M. (2011). Institutional repositories and digital preservation: Assessing current practice at research libraries. *D-Lib Magazine*, 17(5/6).  
<https://doi.org/10.1045/may2011yuanli>
- Liao, C., Chen, J.-L., & Yen, D. C. (2007). Theory of planning behavior (TPB) and customer satisfaction in the continued use of e-service: An integrated model. *Computers in Human Behavior*, 23(6), 2804-2822.  
<https://doi.org/10.1016/j.chb.2006.05.006>
- Lor, P., Wiles, B., & Britz, J. (2021). Re-thinking information ethics: Truth, conspiracy theories, and librarians in the COVID-19 era. *Libri*, 71(1), 1-14.  
<https://doi.org/10.1515/libri-2020-0158>

- Lord, P., Macdonald, A., Lyon, L., & Giaretta, D. (2004). *From data deluge to data curation* e-Science All Hands Meeting, Nottingham. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.7425&rep=rep1&type=pdf>
- Lubell, J., Rachuri, S., & Mani, M. (2008). Sustaining engineering informatics: Towards methods and metrics for digital curation. *The International Journal of Digital Curation*, 2(3), 59-73. <https://doi.org/10.2218/ijdc.v3i2.58>
- Lynch, C. (2000). Authenticity and integrity in the digital environment: An exploratory analysis of the central role of trust. In R. Parry (Ed.), *Authenticity in a Digital Environment* (pp. 32-50). CLIR. <https://www.clir.org/pubs/reports/pub92/>
- Lynch, C. (2006). *Overview of the state of institutional repositories* [Workshop]. JCDL '06, 6th ACM/IEEE-CS Joint Conference on Digital Libraries, Chapel Hill.
- Mack, K., McDonnell, E., Jain, D., Wang, L. L., Froehlich, J. E., & Findlater, L. (2021, May 8-13). *What do we mean by 'accessibility research'? A literature survey of accessibility papers in CHI and ASSETS from 1994 to 2019* CHI Conference on Human Factors in Computing Systems (CHI '21), Yokohama. <https://doi.org/10.1145/3411764.3445412>
- MacNeil, H. (1998). *Trusting records: The Evolution of Legal, Historical, and Diplomatic Methods of Assessing the Trustworthiness of Records, from Antiquity to the Digital Age* [Dissertation, University of British Columbia]. Vancouver.
- MacNeil, H. (2000). *Trusting Records: Legal, Historical, and Diplomatic Perspectives*. Kluwer. <https://doi.org/10.1007/978-94-015-9375-5>
- Madsen, C., & Hurst, M. (2019). Digital preservation policy and strategy: Where do I start? In J. Myntti & J. Zoom (Eds.), *Digital Preservation in Libraries: Preparing for a Sustainable Future* (pp. 37-48). ALA.
- Maemura, E., Moles, N., & Becker, C. (2017). Organizational assessment frameworks for digital preservation: A literature review and mapping. *Journal of the Association for Information Science and Technology*, 68(7), 1619-1637. <https://doi.org/10.1002/asi.23807>
- Maier, M. (2018). Content analysis: Advantages and disadvantages. In M. Allen (Ed.), *The Sage Encyclopedia of Communication Research Methods* (pp. 240-242). Sage. <https://doi.org/10.4135/9781483381411>
- Mamun, M. R. A., Senn, W. D., Peak, D. A., Prybutok, V. R., & Torres, R. A. (2020). Emotional satisfaction and IS continuance behavior: Reshaping the expectation confirmation model. *International Journal of Human-Computer Interaction*. <https://doi.org/10.1080/10447318.2020.1752478>
- Mancino, S. (2020). Doing philosophy and rhetoric historically: A review of the philosophy of communication. *Communication Research Trends*, 39(2), 5-26. <http://search.proquest.com/scholarly-journals/doing-philosophy-rhetoric-historically-review/docview/2430113670/se-2?accountid=11311>
- Martin, R. S. (2007). Intersecting missions, converging practice. *RBM: Journal of Rare Books, Manuscripts, and Cultural Heritage*, 8(1), 80-88. <https://doi.org/10.5860/rbm.8.1.281>
- Marty, P. F. (2010). An introduction to digital convergence: Library, archives, and museums in the information age. *Library Quarterly*, 80(1), 1-5. <https://doi.org/10.1086/648549>

- Mason, C., & Simmons, J. (2012). Are they being served? Linking consumer expectation, evaluation and commitment. *Journal of Services Marketing*, 26(4), 227-237. <https://doi.org/10.1108/08876041211237532>
- Matthews, B., McIlwrath, B., Giaretta, D., & Conway, E. (2008). *The Significant Properties of Software: A Study*. <https://epubs.stfc.ac.uk/manifestation/9506/SignificantPropertiesofSoftware.pdf>
- May, R. (2001). Frege on identity statements. In C. Cecchetto, G. Chierchia, & M. T. Guasti (Eds.), *Semantic Interfaces: Reference, Anaphora, and Aspect* (pp. 1-62). CSLI Publications. <https://www-philosophy.ucdavis.edu/people/rcmay/robert-mays-home-page/Frege.pdf>
- Mayring, P. (2014). *Qualitative Content Analysis: Theoretical Foundations, Basic Procedures and Software Solution*. Leibniz Institute for Social Science. [https://www.ssoar.info/ssoar/bitstream/handle/document/39517/ssoar-2014-mayring-Qualitative\\_content\\_analysis\\_theoretical\\_foundation.pdf](https://www.ssoar.info/ssoar/bitstream/handle/document/39517/ssoar-2014-mayring-Qualitative_content_analysis_theoretical_foundation.pdf)
- McDonough, J. (2012). 'Knee-deep in the data': Practical problems in applying the OAIS reference model to the preservation of video games 2012 45th Hawaii International Conference on System Sciences, Maui. <https://doi.org/10.1109/HICSS.2012.1>
- McKinney, V., Yoon, K., & Zahedi, F. (2002). The measurement of web-customer satisfaction: An expectation and disconfirmation approach. *Information Systems Research*, 13(3), 296-315. <https://doi.org/10.1287/isre.13.3.296.76>
- Menne-Haritz, A. (2001). Access - The reformulation of an archival paradigm. *Archival Science*, 1(1), 57-82. <https://doi.org/10.1007/BF02435639>
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative Research: A Guide to Design and Implementation* (4th ed.). Wiley.
- Mingers, J., Mutch, A., & Willcocks, L. (2013). Critical realism in information systems research. *MIS Quarterly*, 37(3), 795-802. <https://doi.org/2162-9730>
- Mingers, J., & Willcocks, L. (2014). An integrative semiotic framework for information systems: The social, personal and material worlds. *Information and Organization*, 24(1), 48-70. <https://doi.org/10.1016/j.infoandorg.2014.01.002>
- Mingers, J., & Willcocks, L. (2017). An integrative semiotic methodology for IS research. *Information and Organization*, 27(1), 7-36. <https://doi.org/10.1016/j.infoandorg.2016.12.001>
- Mitchell, D. (2015). Type-token models: A comparative study. *Journal of Quantitative Linguistics*, 22(1), 1-21. <https://doi.org/10.1080/09296174.2014.974456>
- Mochocki, M. (2021). Heritage sites and video games: Questions of authenticity and immersion. *Games and Culture*, 16(8), 951-977. <https://doi.org/10.1177/15554120211005369>
- Mois, M., Klas, C.-P., & Hemmje, M. L. (2009). *Digital preservation as communication with the future* DSP '09, 16th International Conference on Digital Signal Processing, <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5201104>
- Moore, R. (2008). Towards a theory of digital preservation. *International Journal of Digital Curation*, 3(1), 63-75. <https://doi.org/10.2218/ijdc.v3i1.42>
- Moore, R., JaJa, J., & Rajasekar, A. (2005). *Storage Resource Broker Data Grid: Preservation Assessment* [Technical report](TR-2006.3). S. D. S. Center. <https://www.academia.edu/download/31215808/SDSC-TR-2006-3-srb.pdf>



- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48-76. <https://doi.org/10.1177/2345678906292462>
- Morgan, D. L. (2014). Pragmatism as a paradigm for social research [Research article]. *Qualitative Inquiry*, 20(8), 1034-1053. <https://doi.org/10.1177/1077800413513733>
- Morris, C. (1964). *Signification and Significance: A Study on the Relations of Signs and Values*. MIT Press.
- Morrissey, S. M. (2014). 'How can we know the dancer from the dance?' Intention and the preservation of digital objects. *New review of Information Networking*, 19(1), 1-15. <https://doi.org/10.1080/13614576.2014.883935>
- Moulaison Sandy, H., & Corrado, E. M. (2018). Bringing context into the picture: Proposing a tri-partite model for digital preservation. *Journal of Library Administration*, 58(1), 1-17. <https://doi.org/10.1080/01930826.2017.1385988>
- NA. *Mission*. National Archief. Retrieved January 16, 2023 from <https://web.archive.org/web/20160512041359/http://www.nationaalarchief.nl/organisatie/missie>
- NA. (2015). *Preservation Policy*. N. A. o. t. Netherlands. <https://www.nationaalarchief.nl/sites/default/files/field-file/National%20Archives%20of%20the%20Netherlands%20preservation%20policy.pdf>
- Nadal, J. (2017). Digital preservation. In J. D. McDonald & M. Levine-Clark (Eds.), *Encyclopedia of Library and Information Sciences* (4th ed., Vol. 1, pp. 1332-1337). CRC Press. <https://doi.org/10.1081/E-ELIS4-120050527>
- NDSA. (2014). *2015 National Agenda for Digital Stewardship: A Report on the Challenges, Opportunities, Gaps, Emerging Trends, and Key Areas for Research and Development That Support the National Capacity for Digital Stewardship*. N. D. S. Alliance. <http://hdl.loc.gov/loc.gdc/lcpub.2013655119.1>
- NDSA. (2019). *Levels of Digital Preservation, Version 2.0*. NDSA. Retrieved April 10 from <https://ndsa.org/publications/levels-of-digital-preservation/>
- nestor. (2009). *Catalogue of Criteria for Trusted Digital Repositories* (nestor materials 8). nestor. [https://files.dnb.de/nestor/materialien/nestor\\_mat\\_08\\_eng.pdf](https://files.dnb.de/nestor/materialien/nestor_mat_08_eng.pdf)
- Ng, S. (2012). A brief history of entertainment technologies. *Proceedings of the IEEE*, 1000, 1386-1390. <https://doi.org/10.1109/JPROC.2012.2189805>
- Nicholson, D., & Dobreva, M. (2009, 5-7 July). *Beyond OAIS: Towards a reliable and consistent digital preservation implementation framework* 16th International Conference on Digital Signal Processing, Santorini-Hellas. <https://doi.org/10.1109/ICDSP.2009.5201126>
- Nitecki, D. (1993). Conceptual models of libraries held by faculty, administrators, and librarians: An exploration of communications in the Chronicle of Higher Education. *Journal of Documentation*, 49(3), 255-277. <https://doi.org/10.1108/eb026915>
- NLNZ. *About the Library*. National Library of New Zealand. Retrieved January 16, 2023 from <https://natlib.govt.nz/about-us>
- NLNZ. (2012). *Digital Preservation Policy Manual*. A. N. Zealand & N. L. o. N. Zealand. <https://digitalpreservation.natlib.govt.nz/assets/NDHA/About-Us/Strategic-Partnerships/Digital-Preservation-Policy-Manual.pdf>
- NLNZ. (2016). *Turning Knowledge Into Value - Strategic Directions to 2030*. N. L. o. N. Zealand. <https://natlib.govt.nz/files/strategy/strategicdirection-2030.pdf>

- Noonan, D. (2014). Digital preservation policy framework: A case study. *EDUCAUSE Review*, 49(4). <http://www.educause.edu/ero/article/digital-preservation-policy-framework-case-study>
- Nöth, W. (1990). *Handbook of semiotics* (First paperback edition ed.). Indiana University Press.
- OED. (2009). -ability, *suffix*. In *OED Online*. Retrieved 2021-03-28, from <https://www.oed.com/view/Entry/251524>
- OED. (2011a). accessible, *adj*. In *OED Online*. Retrieved 2022-02-06, from <https://www.oed.com/view/Entry/1034>
- OED. (2011b). use, *n*. In *OED Online*. Retrieved 2022-04-10, from <https://www.oed.com/view/Entry/220635>
- Oliver, R. L., & Burke, R. R. (1999). Expectation processes in satisfaction formation: A field study. *Journal of Service Research*, 1(3), 196-214. <https://doi.org/10.1177/109467059913002>
- Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology*, 8(5), 375-387. <https://doi.org/10.1080/13645570500402447>
- Osgood, C. E. (1959). The representational model and relevant research methods. In I. De Sola Pool (Ed.), *Trends in Content Analysis* (pp. 33-88). University of Illinois Press.
- Osgood, C. E., Saporta, S., & Nannally, J. C. (1956). Evaluative assertion analysis. *Litera*, 3, 47-102. <http://dergipark.gov.tr/download/article-file/13623>
- Overton, W. F. (2012). Evolving scientific paradigms: Retrospective and prospective. In L. L'Abate (Ed.), *Paradigms in Theory Construction* (pp. 31-66). Springer. <https://doi.org/10.1007/978-1-4614-0914-4>
- Owens, T. (2018). *The Theory and Craft of Digital Preservation*. Johns Hopkins University Press.
- Oyelude, A. A. (2019). Galleries, libraries, archives and museums (GLAMs) and digital preservation activities. *Library Hi Tech News*, 36(6), 8-9. <https://doi.org/10.1108/LHTN-06-2019-0039>
- Pai, K. (2016). Library and library space. In C. J. R. Higuera & T. J. Bennett (Eds.), *Concepts for Semiotics* (pp. 105-122). University of Tartu Press. <https://doi.org/978-9949-77-265-0>
- Palmer, C., Weber, N., Renear, A., & Muñoz, T. (2013). Foundations of data curation: The pedagogy and practice of "purposeful work" with research data. *Archive Journal*, 3. <http://hdl.handle.net/2142/78099>
- Partners for Preservation: Advancing Digital Preservation Through Cross-Community Collaboration*. (2019). (J. Kramer-Smyth, Ed.). Facet.
- Peirce, C. S. (1906). Prolegomena to an apology for Pragmatism. *The Monist*, 16(4), 492-546. <https://www.jstor.org/stable/27899680>
- Peirce, C. S. (1932). *Collected Papers of Charles Sanders Peirce. Volumes I and II: Principles of Philosophy and Elements of Logic*. Harvard University Press.
- Peirce, C. S. (1991). *Peirce on Signs: Writings on Semiotic by Charles Sanders Peirce* (J. Hoopes, Ed.). University of North Carolina Press.
- Pelc, J. (2000). Semiosis and semiosics vs. semiotics. *Semiotica*, 128(3-4), 425-434. <https://doi.org/10.1515/semi.2000.128.3-4.425>
- Persson, U. (2016). Is falsification falsifiable? *Foundations of Science*, 21(3), 461-475. <https://doi.org/10.1007/s10699-015-9420-4>

- Pesut, B., & Johnson, J. (2008). Reinstating the 'Queen': Understanding philosophical inquiry in nursing [Theoretical paper]. *Journal of Advanced Nursing*, 61(1), 115-121. <https://doi.org/10.1111/j.1365-2648.2007.04493.x>
- Phillips, M., Bailey, J., Goethals, N., & Owens, T. (2013). *The NDSA Levels of Digital Preservation: An Explanation and Uses*. [http://www.digitalpreservation.gov/ndsaworking\\_groups/documents/NDSA\\_Levels\\_Archiving\\_2013.pdf](http://www.digitalpreservation.gov/ndsaworking_groups/documents/NDSA_Levels_Archiving_2013.pdf)
- Pinnick, J. (2019). *Building resilience at the National Geoscience Data Center: Enhancing digital data continuity through research data management training* iPRES 2019, 16th International Conference on Digital Preservation, Amsterdam. [https://ipres2019.org/static/pdf/iPres2019\\_paper\\_12.pdf](https://ipres2019.org/static/pdf/iPres2019_paper_12.pdf)
- Poole, A. H. (2015). How has your science grown? Digital curation and the human factor: A critical literature review. *Archival Science*, 15(2), 39. <https://doi.org/10.1007/s10502-014-9236-y>
- Poole, A. H. (2016). The conceptual landscape of digital curation. *Journal of Documentation*, 72(5), 961-986. <https://doi.org/10.1108/JD-10-2015-0123>
- Popper, K. (1959). *The Logic of Scientific Discovery*. Routledge.
- Popper, K. (1976). A note on verisimilitude. *The British Journal for the Philosophy of Science*, 27(2), 147-159. <https://www.jstor.org/stable/686164>
- Popper, K. (1999). *All Life Is Problem Solving* (P. Camiller, Trans.). Routledge. <https://doi.org/10.4324/9780203431900> (Alles Leben ist Problemlösen (Munich: Pipe Verlag, 1994))
- Radulescu, C., & Vessey, I. (2008, July 11-13). *Causality in critical research research: An analysis of three explanatory frameworks* International Association for Critical Realism, London.
- Rasmussen, C. H. (2019). Is digitalization the only driver of convergence? Theorizing relations between libraries, archives, and museums. *Journal of Documentation*, 75(6), 1258-1273. <https://doi.org/10.1108/JD-02-2019-0025>
- Ravenwood, C., Muir, A., & Matthews, G. (2015). Stakeholders in the selection of digital material for preservation: Relationships, responsibilities, and influence. *Collection Management*, 40(2), 83-110. <https://doi.org/10.1080/01462679.2015.1011816>
- Reed, B. (2005). Records. In S. McKemmish, M. Piggott, B. Reed, & F. Upward (Eds.), *Archives: Recordkeeping in Society* (pp. 101-130). Charles Sturt University.
- Reed, M. I. (2009). Critical realism: Philosophy, method, or philosophy in search of a method? In D. A. Buchanan & A. Bryman (Eds.), *The Sage Handbook of Organizational Research Methods* (pp. 430-448). Sage.
- Reichertz, J. (2014). Induction, deduction, abduction. In U. Flick (Ed.), *The Sage Handbook of Qualitative Data Analysis* (pp. 123-135). <https://doi.org/dx.doi.org/10.4135/9781446282243.n9>
- Revez, J., & Borges, L. C. (2019). Pragmatic paradigm in information science research: A literature review. *Qualitative and Quantitative Methods in Libraries*, 8(2), 179-188. <http://www.qqml-journal.net/index.php/qqml/article/view/485>
- Ricoeur, P. (1976). *Interpretation Theory: Discourse and Surplus of Meaning*. Texas Christian University Press.
- Ries, T., & Palkó, G. (2019). Born-digital archives. *International Journal of Digital Humanities*, 1(1), 1-11. <https://doi.org/10.1007/s42803-019-00011-x>

- Ringgold. (2018). *Ringgold Types*. Retrieved March 4 from <https://support.ringgold.com/wp-content/uploads/2018/07/Ringgold-Types-20180719.xlsx>
- Robinson, H. (2012). Remembering things differently: Museums, libraries and archives as memory institutions and the implications for convergence. *Museum Management and Curatorship*, 27(4), 413-429. <https://doi.org/10.1080/09647775.2012.720188>
- Rogala, a., & Bialowas, S. (2016). Theory of communication: Evolution, approaches, models. In *Communication in Organizational Environments: Functions, Determinants and Areas of Influence* (pp. 1-27). Palgrave.
- Rogers, C. (2015a). Diplomatics of born digital documents – considering documentary form in a digital environment. *Records Management Journal*, 25(1), 6-20. <https://doi.org/10.1108/RMJ-03-2014-0021>
- Rogers, C. (2015b). *Virtual Authenticity: Authenticity of Digital Records from Theory to Practice* [Dissertation, University of British Columbia]. Vancouver. <https://dx.doi.org/10.14288/1.0166169>
- Rosenthal, D. S. H. (2010a). Bit preservation: A solved problem? *International Journal of Digital Curation*, 5(1), 134-148. <https://doi.org/10.2218/ijdc.v5i1.148>
- Rosenthal, D. S. H. (2010b). Keeping bits safe: How hard can it be? *Communication of the ACM*, 53(11), 47-55. <https://doi.org/10.1145/1839676.1839692>
- Ross, S. (2006). Approaching digital preservation holistically. In A. Tough & M. Moss (Eds.), *Record Keeping in a Hybrid Environment: Managing the Creation, Use, Preservation, and Disposal of Unpublished Information Objects in Context* (pp. 115-153). Chandos. <https://doi.org/10.1016/B978-1-84334-142-0.50006-2>
- Ross, S. (2012). Digital preservation, archival science and methodological foundations for digital libraries. *New review of Information Networking*, 17(1), 43-68. <https://doi.org/10.1080/13614576.2012.679446>
- Ross, S. (2020). Beyond place: Data curation possibilities for post-custodial archives and libraries. In P. Demers & T. Samek (Eds.), *Minds Alive: Libraries and Archives Now* (pp. 173-183). University of Toronto Press.
- Rothenberg, J. (1999). *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. Council on Library and Information Resources. <http://files.eric.ed.gov/fulltext/ED426715.pdf>
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180. <https://doi.org/10.1177/0165551506070706>
- Rozenberg, S. (2021). Digital records as relational objects–Yuk Hui’s concept of digital objects applied to archival science. *Archival Science*, 21, 193-218. <https://doi.org/10.1007/s10502-021-09357-0>
- Rumsey, A. S. (2016). *When We Are No More: How Digital Memory is Shaping Our Future*. Bloomsbury.
- Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D., Lyon, L., & Atkinson, M. (2005). *The Digital Curation Centre: A vision for digital curation 2005* IEEE International Symposium on Mass Storage Systems and Technology, Sardinia. <https://doi.org/10.1109/LGDI.2005.1612448>



- Rydén, R. (2019). Archivist and time: Conceptions of time and long-term information preservation among archivists. *Journal of Contemporary Archival Studies*, 6(6), 1-18. <https://elischolar.library.yale.edu/jcas/vol6/iss1/6/>
- SAA. (2013). *Describing Archives: A Content Standard*. S. o. A. Archivists.
- SAA. (2020). Dictionary of Archives Terminology. In Retrieved 2021-02-12, from <https://dictionary.archivists.org/>
- Sacchi, S. (2015). *What Do We Mean By 'Preserving Digital Information'? Towards Sound Conceptual Foundations for Digital Stewardship* [PhD Thesis, University of Illinois at Urbana-Champaign]. Urbana, IL. <http://hdl.handle.net/2142/78440>
- Sacchi, S., & McDonough, J. (2012, February 7019). *Significant properties of complex digital artifacts: Open issues from a video game case study* iConference '12, Toronto. <https://doi.org/10.1145/2132176.2132293>
- Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers* (3rd ed.). Sage.
- Samanta, P., & Jain, S. (2021). Analysis of perceptual hashing algorithms in image manipulation detection. *Procedia Computer Science*, 185, 203-212. <https://doi.org/10.1016/j.procs.2021.05.021>
- Sanett, S. (2013). Archival digital preservation programs: Staffing, costs, and policy. *Preservation, Digital Technology & Culture*, 42(3), 137-149. <https://doi.org/10.1515/pdte-2013-0019>
- Saussure, F. d. (1983). *Course in General Linguistics* [Cours de linguistique générale] (R. Harris, Trans.; C. Bally, A. Sechehaye, & A. Riedlinger, Eds.). Open Court. (1916)
- Savolainen, R. (2019). Elaborating the sensory and cognitive-affective aspects of information experience. *Journal of Librarianship and Information Science*, 1-14. <https://doi.org/10.1177/0961000619871595>
- SCAPE. (2016). *Published Preservation Policies*. O. P. Federation. <http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies>
- Schramm, W. (1954). How communication works. In W. Schramm (Ed.), *The Process and Effects of Mass Communication* (pp. 3-26). University of Illinois Press.
- Schreier, M. (2012). *Qualitative Content Analysis in Practice*. Sage.
- Schreier, M. (2013). Qualitative content analysis. In U. Flick (Ed.), *Sage Handbook of Qualitative Data Analysis* (pp. 170-193). Sage. <https://doi.org/10.4135/9781446282243>
- Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14(2), 159-178. <https://doi.org/10.1007/s11219-006-7600-8>
- Shannon-Baker, P. (2016). Making paradigms meaningful in mixed methods research. *Journal of Mixed Methods Research*, 10(4), 319-334. <https://doi.org/10.1177/1558689815575861>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423, 623-656. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sheffield, E. (2004). Beyond abstraction: Philosophy as a practical qualitative research method. *The Qualitative Report*, 9(4), 760-769. <http://nsuworks.nova.edu/tqr/vol9/iss4/9>
- Sheldon, M. (2013). *Analysis of Current Digital Preservation Policies: Archives, Libraries, and Museums*. L. o. Congress. <http://www.digitalpreservation.gov/documents/Analysis%20of%20Current%20Digital%20Preservation%20Policies.pdf>

- Sierman, B. (2014, October 6-10). *The SCAPE Policy Framework, maturity levels and the need for realistic preservation policies* iPRES 2014, the 11th International Conference on Digital Preservation, Melbourne. <https://ipres-conference.org/ipres14/sites/default/files/upload/iPres-Proceedings-final.pdf>
- Sierman, B., Jones, C., Bechhofer, S., & Elstrøm, G. (2013). *Digital policy levels in SCAPE* 10th International Conference on Preservation of Digital Objects, Lisbon. [https://purl.pt/24107/1/iPres2013\\_PDF/Preservation%20Policy%20Levels%20in%20SCAPE.pdf](https://purl.pt/24107/1/iPres2013_PDF/Preservation%20Policy%20Levels%20in%20SCAPE.pdf)
- Skiljan, I. (2022). *Irfanview*. <https://www.irfanview.com/>
- Smit, E., Van Hoeven, J., & Giaretta, D. (2011). Avoiding a Digital Dark Age for data: Why publishers should care about digital preservation. *Learned Publishing*, 24(1), 35-49. <https://doi.org/10.1087/20110107>
- Smith, M. N. (2017). Intentions: past, present, future. *Philosophical Explorations*, 20(Supplement II), 1-12. <https://doi.org/10.1080/13869795.2017.1356360>
- Söderlund, M., & Öhman, N. (2005). Assessing behavior before it becomes behavior: An examination of the role of intentions as a link between satisfaction and repatronizing behavior. *International Journal of Service Industry Management*, 16(2), 169-185. <https://doi.org/10.1108/09564230510592298>
- Spencer, R. (2019). Digital preservation as a thought experiment. In J. Myntti & J. Zoom (Eds.), *Digital Preservation in Libraries: Preparing for a Sustainable Future* (pp. 21-33). ALA.
- Stamper, R. K. (1993, September 6-10, 1993). *A semiotic theory of information and information systems* ICL/University of Newcastle Seminar on 'Information', <https://assets.cs.ncl.ac.uk/seminars/101.pdf>
- Stanford, E. (2020). A field guide to digital surrogates: Evaluating and contextualizing a rapidly changing resource. In K. Brown (Ed.), *The Routledge Companion to Digital Humanities and Art History*. Routledge. <https://doi.org/10.4324/9780429505188-19>
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, 7(17). <https://scholarworks.umass.edu/pare/vol7/iss1/17>
- Stepanyan, K., Gkotsis, G., Kalb, H., Kim, Y., Cristea, A. I., Joy, M., Trier, M., & Ross, S. (2012, October 1-5). *Blogs as objects of preservation: Advancing the discussion on significant properties* 9th International Conference on Preservation of Digital Objects, Toronto. <https://hdl.handle.net/11353/10.293844>
- Strodl, S., Becker, C., Neumayer, R., & Rauber, A. (2007). *How to choose a digital preservation strategy: Evaluating a preservation planning procedure* 7th ACM/IEEE-CS Joint Conference on Digital Libraries, Vancouver. <https://doi.org/10.1145/1255175.1255181>
- Tallman, N. (2021). A 21st century technical infrastructure for digital preservation. *Information Technology and Libraries*, 40(4). <https://doi.org/10.6017/ital.v40i4.13355>
- Tan, H., Wilson, A., & Olver, I. (2009). Ricoeur's theory of interpretation: An instrument for data interpretation in hermeneutic phenomenology [Research article]. *International Journal of Qualitative Methods*, 8(4), 1-15. <https://doi.org/10.1177/160940690900800401>
- Teddle, C., & Tashakkori, A. (2012). Common 'core' characteristics of mixed methods research: A review of critical issues and call for greater convergence. *American*

- Thibodeau, K. (2002, July). Overview of technological approaches to digital preservation and challenges in coming years. *The State of Digital Preservation: An International Perspective*,
- Thibodeau, K. (2007). If you build it, will it fly? *Journal of Digital Information*, 8(2).  
<https://journals.tdl.org/jodi/index.php/jodi/article/view/197/174>
- Thibodeau, K. (2017). Digital preservation. In L. Duranti & P. Franks (Eds.), *Encyclopedia of Archival Science* (pp. 160-163). Rowman & Littlefield.  
<https://ebookcentral.proquest.com/lib/qut/detail.action?docID=2076364>
- Thomassen, T. (2001). A first introduction to archival science. *Archival Science*, 1(4), 373-385. <https://doi.org/10.1007/BF02438903>
- Tikonov, A. (2019). Preservation of digital images: Question of fixity. *Heritage*, 2(2), 1160-1165. <https://doi.org/10.3390/heritage2020075>
- Traczyk, T. (2017). Requirements for digital preservation. In T. Traczyk, W. Ogryczak, P. Pałka, & T. Śliwiński (Eds.), *Digital Preservation: Putting It to Work* (pp. 3-14). Springer.
- Tredinnick, L. (2006). *Digital Information Contexts: Theoretical Approaches to Understanding Digital Information*. Chandos.
- Tredinnick, L. (2008). *Digital Information Culture: The Individual and Society in the Digital Age*. Chandos.
- Tsakonas, G., & Papatheodorou, C. (2008). Exploring usefulness and usability in the evaluation of open access digital libraries. *Information Processing and Management*, 44, 17. <https://doi.org/10.1016/j.ipm.2007.07.008>
- Tsakonas, G., & Papatheodorou, C. (2011). An ontological representation of the digital library evaluation domain. *Journal of the American Society for Information Science*, 62(8), 1577-1593. <https://doi.org/10.1002/asi.21559>
- Tufiş, D., & Ion, R. (2017). Part-of-speech tagging. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics* (2nd ed.). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199573691.013.51>
- Tzafestas, S. G. (2018). Information I: Communication, transmission, and information theory. In *Energy, Information, Feedback, Adaptation, and Self-Organization: The Fundamental Elements of Life and Society* (pp. 157-218). Springer.  
<https://doi.org/10.1007/978-3-319-66999-1>
- UNESCO. (2003). *Guidelines for the Preservation of Digital Heritage*.  
<http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>
- UNESCO. (2019a). *Concept of Digital Preservation*. UNESCO. Retrieved 2019-10-26 from <https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-preservation>
- UNESCO. (2019b). *Digital Preservation Programmes*. UNESCO. Retrieved 2020-12-22 from <https://en.unesco.org/themes/information-preservation/digital-heritage/digital-preservation-programmes>
- van der Werf, T., & van der Werf, B. (2019). Documentary heritage in the digital age: Born digital, being digital, dying digital. In R. Edmondson, L. Jordan, & A. C. Prodan (Eds.), *The UNESCO Memory of the World Programme* (pp. 165-189). Springer. [https://doi.org/10.1007/978-3-030-18441-4\\_12](https://doi.org/10.1007/978-3-030-18441-4_12)
- van Veenendaal, R., Kjærskov, F. H., Sein, K., O'Sullivan, J., Nielsen, A. B., Tømmerholt, P. M., & Takema, J. (2019, September 16-20). *Significant properties of spreadsheets: An update on the work of the Open Preservation Foundation's Archives Interest Group* 16th International Conference on Digital

- Preservation, Amsterdam.  
[https://ipres2019.org/static/pdf/iPres2019\\_paper\\_48.pdf](https://ipres2019.org/static/pdf/iPres2019_paper_48.pdf)
- van Veenendaal, R., Lucker, P. C. M., & Sijtsma, C. D. (2018, September 24-27). *Significant significant properties* 15th International Conference on Digital Preservation, Boston. <https://openpreservation.org/wp-content/uploads/2018/10/Significant-Significant-Properties.pdf>
- Vasileiou, K., Barnett, J., Thorpe, S., & Youn, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, 18(148). <https://doi.org/10.1186/s12874-018-0594-7>
- Vermaaten, S., Lavoie, B., & Caplan, P. (2012). Identifying threats to successful digital preservation: The SPOT model for risk assessment. *D-Lib Magazine*, 18(9/10). <https://doi.org/doi:10.1045/september2012-vermaaten>
- Waller, M., & Sharpe, R. (2006). *Mind the Gap: Assessing Digital Preservation Needs in the UK*. [http://www.dpconline.org/component/docman/doc\\_download/340-mind-the-gap-assessing-digital-preservation-needs-in-the-uk](http://www.dpconline.org/component/docman/doc_download/340-mind-the-gap-assessing-digital-preservation-needs-in-the-uk)
- Walters, T., & Skinner, K. (2011). *Digital Curation for Preservation* (New Roles for New Times, Issue. A. o. R. Libraries. [https://www.arl.org/wp-content/uploads/2011/03/nrnt\\_digital\\_curation17mar11.pdf](https://www.arl.org/wp-content/uploads/2011/03/nrnt_digital_curation17mar11.pdf)
- Wartena, C., & Brussee, R. (2008, September 1-5). *Topic detection by clustering keywords* 19th International Workshop on Database and Expert Systems Application, Turin.
- Waters, D., & Garrett, J. (1996). *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* (CLIR pub 63). C. o. L. a. I. Resources. <http://www.clir.org/pubs/reports/reports/pub63watersgarrett.pdf>
- Watry, P. (2007). Digital preservation theory and application: Transcontinental persistent archives testbed activity. *International Journal of Digital Curation*, 2(2), 41-68. <https://doi.org/10.2218/ijdc.v2i2.28>
- Webb, C., Pearson, D., & Koerbin, P. (2013). 'Oh, you wanted us to preserve that?!' Statements of preservation intent for the National Library of Australia's digital collections. *D-Lib Magazine*, 19(1/2). <https://doi.org/10.1045/january2013-webb>
- Weber, N., Palmer, C., & Chao, T. (2012). Current trends and future directions in data curation research and education. *Journal of Web Librarianship*, 6(4), 305-320. <https://doi.org/10.1080/19322909.2012.730358>
- Weber, R. (2012). Evaluating and developing theories in the information systems discipline. *Journal of the Association for Information Systems*, 13(1), 1-30. <https://doi.org/10.17705/1jais.00284>
- Weiner, J. (2004). *Frege Explained: From Arithmetic to Analytic Philosophy*. Open Court.
- White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library trends*, 55(1), 22-45. <https://doi.org/10.4135/9781446282243>
- Whitelaw, M. (2012, August 20-24). *Towards generous interfaces for archival collections* International Council on Archives Congress, Brisbane. <http://ica2012.ica.org/files/pdf/Full%20papers%20upload/ica12Final00423.pdf>
- Whitt, R. S. (2017). 'Through a glass, darkly': Technical, policy, and financial actions to avert the coming digital dark ages. *Santa Clara High Technology Law Journal*, 33(2), 117-229. <http://digitalcommons.law.scu.edu/chtlj/vol33/iss2/1>



- Wilson, T. C. (2017). Rethinking digital preservation: Definitions, models, and requirements [Conceptual paper]. *Digital Library Perspectives*, 33(2), 128-136. <https://doi.org/10.1108/DLP-08-2016-0029>
- Winget, M. (2011). Videogame preservation and massively multiplayer online role-playing games: A review of the literature. *Journal of the American Society for Information Science and Technology*, 62(10), 1869–1883. <https://doi.org/10.1002/asi.21530>
- Withagen, R., de Poel, H. J., Araújo, D., & Pepping, G.-J. (2012). Affordances can invite behavior: Reconsidering the relationship between affordances and agency. *New Ideas in Psychology*, 30(2), 250-258. <https://doi.org/10.1016/j.newideapsych.2011.12.003>
- Witt, M., Kroll, M., Minor, D., & Reilly, B. (2012, July 12). *ISO 16363: Trustworthy digital certification in practice* 7th International Conference on Open Repositories, Edinburgh. [https://docs.lib.purdue.edu/lib\\_fspres/4](https://docs.lib.purdue.edu/lib_fspres/4)
- Wright, B., Brunner, O., & Nebel, B. (2018). On the importance of a research data archive. *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, 32(1). <https://ojs.aaai.org/index.php/AAAI/article/view/11395>
- Xie, I., & Matusiak, K. (2016). Digital preservation. In *Discover Digital Libraries: Theory and Practice* (pp. 255-279). Elsevier.
- Yakel, E. (2007). Digital curation. *OCLC Systems & Services*, 23(4), 6. <https://doi.org/10.1108/10650750710831466>
- Yakel, E., Rieh, S. Y., Markey, K., St. Jean, B., & Yao, X. (2009). *Secrets of success: Identifying success factors in institutional repositories* 4th International Conference on Open Repositories, Atlanta. <http://hdl.handle.net/1853/28419>
- Yang, S., & Tang, Y. (2022). News topic detection based on capsule semantic graph. *Big Data Mining and Analytics*, 5(2), 98-109. <https://doi.org/10.26599/BDMA.2021.9020023>
- Yeo, G. (2010). 'Nothing is the same as something else': Significant properties and notions of identify and originality. *Archival Science*, 10(2), 85-116. <https://doi.org/10.1007/s10502-010-9119-9>
- Yoon, A. (2014). End users' trust in data repositories: Definition and influences on trust development [Original paper]. *Archival Science*, 14(1), 17-34. <https://doi.org/10.1007/s10502-013-9207-8>
- Young, D. S., & Casey, E. A. (2018). An examination of the sufficiency of small qualitative samples. *Social Work Research*, 43(1), 53-58. <https://doi.org/10.1093/swr/svy026>
- ZBW. *History of the ZBW*. Leibniz-Informationszentrum Wirtschaft. Retrieved January 16, 2023 from <https://web.archive.org/web/20170318090240/https://www.zbw.eu/en/about-us/profile/history/>
- ZBW. (2017). *Preservation Policy of the Three German National Libraries*. <https://www.zbw.eu/en/about-us/key-activities/digital-preservation/preservation-policy-national-libraries>
- ZBW. (2018). *Preservation Policy: Guidelines for Digital Preservation at the ZBW*. L. I. C. f. Economics. <https://www.zbw.eu/en/about-us/key-activities/digital-preservation/preservation-policy/>
- Zierau, E. M.-B. O. (2012). A holistic approach to bit preservation. *Library Hi Tech*, 30(3), 18. <https://doi.org/10.1108/07378831211266618>