Nothing succeeds like success: An approach for evaluating digital preservation efficacy

Stephen Abrams
California Digital Library
University of California
Stephen.Abrams@ucop.edu

Abstract
Digital preservation encompasses the theory and practice ensuring purposeful future use of digital resources. But how can one tell whether it has been effective or not? The evaluation of preservation efficacy has two dimensions: trustworthiness of managerial programs and systems; and successful use of managed resources. While the former has received extensive attention, the latter has been little investigated. This stems from an insufficiently broad conceptualization of the preservation enterprise, which should be viewed expansively as facilitating meaningful human communication across time and concomitant cultural distance. Communicological analysis leads to a semiotic-phenomenological model for preservation-enabled communication cognizant of the elusive nature of use, which is inherently contingent with respect to time, place, person, and purpose. Preservation success is positioned as an individual, rather than universal value, with a benchmark evaluation of situational verisimilitude, rather than absolute fidelity to an illusory canonical state and information experience. The proposed evaluative approach provides new conceptual clarity to preservation theory and practice, a more rigorous basis for illuminating the limits of preservation efficacy, and a more nuanced means of stating, measuring, and evaluating preservation intentions, expectations, and outcomes.

Keywords
Digital preservation, efficacy, trustworthiness, success, communicology, semiotics, phenomenology

1 Introduction
Digital preservation encompasses the theory and practice aimed at ensuring ongoing access to and purposeful use of digital resources into the indefinite future. Preservation strategies can take a variety of forms resulting in a broad range of outcomes. For example, an attempt to satisfy a future request for a previously preserved resource could variously take the form of providing:

- Original physical media holding the resource (say, a magnetic tape);
- Contemporary media (USB drive);
- A file, about which nothing more is known;
- File in original known format (WordPerfect);
- Derivative file of known format (PDF);
- File and rendering software (Acrobat Reader);
- File and provenance (PREMIS metadata);
- File and token of authenticity (PKI signature);
- File and intellectual description (MARC record);
- File and productive context (DataCite methodology statement);
- File and curatorial context (EAD finding aid);
- File and prior consumptive context (published article citing the resource);

and so on. At what point can one say that the preservation outcome was successful? Without knowing, how can one rationally plan for, reasonably expect, effectively measure, or meaningfully be held accountable for that outcome?

In 2006 Lynch observed that digital preservation is “a metric that’s defied measuring” [1], and little progress has been made since then identifying appropriate metrics for quantifying preservation success. This stems from an unduly narrow conceptualization of the preservation field as synonymous with preservation management. The focal point of that management is conformance to the ISO 14721 OAIS reference model [2, 3], and its accepted evaluative metric is trustworthiness relative to ISO 16363 TDR audit and certification criteria [4, 5]. (The terms “management” and “managers” are used throughout to refer to both external curatorial oversight of an OAIS and the performance of its internal functions.) TDR criterion 4.3.4 mandates that a trustworthy repository “provide evidence of the effectiveness of its preservation activities” [6], but the literature provides little direction as to what form that evidence might take. Regardless, a true understanding of preservation efficacy is ultimately a question of evaluating consumer-experienced outcomes, not just managerial trustworthiness. The basis of that evaluation is how well the outcomes align (or misalign) with productive and managerial intentions and consumer expectations and exploitive purposes (see Figure 1).

However, the preservation field has not yet matured to a point of having established metrics for evaluating the success (or failure) of its outcomes [7, 8], nor has it developed sufficiently robust theoretical models [9] to underpin any such evaluation. Progress towards measurable metrics is dependent upon re-defining the preservation domain on a more expansive, nuanced, and conceptually-sound basis [10]. Rather than focusing exclusively on methodological approaches to the question of the trustworthiness of its managerial subdomain, the preservation community also should be asking what
theoretically-informed measures can and should be used to evaluate the success of the digital preservation enterprise.

![Figure 1: Trustworthy vs. successful measures](image)

2 Re-conceptualizing the preservation domain

Within the literature, the focal definition of the digital preservation field revolves around custodial stewardship by archival institutions [3, 11, 12], most often expressed in terms of appropriate data management activities, e.g., [13-15]. At the center of that management is preservation instrumentality provided by OAIS systems, with controlling responsibility resting with OAIS managers. The emphasis on managerial agency and action, e.g., [16, 17], minimizes opportunities for appropriate consideration of pre-acquisition or post-retrieval activities and the concomitant experiences of information producers and consumers [18, 19], even though the ultimate goal of preservation – the future use of preserved resources – is inseparable from their perspectives [20, 21]. Explicit concern for producers and consumers reframes the preservation imperative to a transfer of information from the former to the latter, or in other words, mediated human communication. This is consistent with an alternative formulation of the domain as “communicating with the future” [22-24]. However, while these authors deploy the metaphor of communication, they do not follow through on its consequences to redefine the field in communicological terms.

Communicology is the science of embodied discourse [25], with theoretical and methodological foci on the semiotic affordances of communicated messages and phenomenological experiences of communicating actors [26]. Communicative processes have been analyzed from a variety of perspectives, including the propagation of signals independent of interpretation, as well as the subjective experience of human participants [27]; the degree to which a common field of experience undermines the interpretation of messages, and the alignment of intent and consequence as reflected in the effect messages have upon their receivers [28]; the psychological and anthropological implications of communication across a hierarchy of intrapersonal, interpersonal, group, and cultural structures [27]; and the context of expressive and interpretive coding/decoding strategies, and the real or conceivable external referents of communicated epistemic meaning [26]. These various aspects can be aligned and compared by reference to a formally-defined semiotic meta-model [29], which provides a compelling framework for such meta-analysis.

Semiotics is the study of signs and systems of signification, that is to say, things that somehow carry communicable meaning or affect, and the ways in which they are expressed, exchanged, and experienced [30]. The semiotic affordances propounded by Peirce are threefold: abstract meaning, expressive form, and interpretive understanding, the latter of which concretizes in relation to some situated context [31]. The antecedents of Peircean semiotics reach back to scholastic and classical philosophy [32, 33], which assumed purely analog sign transmission, e.g., spoken words, ink on paper, paint on canvas, etc. The advent of the digital age necessitated the augmentation of semiotic concerns for greater applicability to digitally-mediated communication. These extensions include distinguishing a sign’s abstract expressive form from its concrete symbolic representation and tangible manifestation [34]. Other aspects of digital communication warranting entry into the semiotic canon include the performative behavior producing humanly perceptible form [35, 36], assessment of archival and informational integrity and authenticity [9], and the manifold ways in which digital information can, should, or must evolve over time to ensure ongoing accessibility and usability [37, 38]. The full complement of these concerns has not previously been integrated into a unified model applicable to the digital preservation domain.

The practice in that domain, moreover, does not rest on firm theoretical foundations [3, 39, 40]. Claims of theoretical advance often deploy “theory” in a narrow sense of a proposed thesis or pragmatic solution, e.g., [23, 41]. Others conform to the more expansive notion of theory as a coherent system of abstraction, explanation, and inference, but rely upon mathematical formalisms tacitly assuming that preserved resources fully encapsulate the knowledge-states of their producers, which can be unambiguously (re)experienced by consumers, e.g., [37, 38, 42]. This position is at odds with the post-modernist view of the essential contingency of human communication and epistemic behavior [43, 44], implying that any use of a preserved resource is inherently situated with respect to time, place, person, and purpose. Thus, the theoretical and methodological constructs of communicology – the cultural semiotics of information expression and reception [45], and the embodied phenomenology of information experience [46] – are an appropriate basis for a re-conceptualization of the digital preservation field.

There are many contemporary forms of digitally-enabled communication, e.g., email, text, telephony, streaming video, social media, etc. How can digital preservation be distinguished meaningfully from these alternatives? The primary differentiating characteristic is preservation’s focal attention to the potentially corrosive impact of time. The communicological literature does not address this temporal concern; instead, communication is tacitly assumed synchronous in time. Conversely, while the preservation
literature is strongly focused on temporal consequences, it
does not incorporate communicological perspectives.
Melding these two philosophical and methodological
traditions provides a richer set of analytic tools applicable to
the question of preservation efficacy.

3 Modeling preservation communication

Communicological analysis of the preservation domain has
three stages: (1) process modeling of preservation-enabled
communication; (2) identification of pertinent semiotic
affordances of modeled resources and processes; and (3)
derivation of appropriate measures of the phenomenological
experience of modeled actors. A proper model for digital
preservation should be grounded in a conceptualization of its
problem domain as human communication rather than data
management; emphasizing the consequences of digital
communication across time and concomitant technical and
and cultural distance; analyzing the entire post-custodial
enterprise, not just the subdomain of custodial management;
promoting the primary value of success alongside
trustworthiness; evaluating success through operational, not
just descriptive evidence; and placing theoretical concerns
on equal par with pragmatic ones, providing opportunity for
consideration of the situated contingency of preservation-
enabled communication [10].

The goal of that communication is to transfer an
intangible but intentional unit of cognitive meaning or
psychological affect, which are internal epistemic states of
mind arising from the situated context of their production
and referencing some external entity or idea, from their
producer to a consumer across temporal, technical, and
cultural distance (see Figure 2). A complementary process
follows reception of the perceptual form of a rendered
resource: new epistemic understanding arising in a
consumptive context relative to an interpretive referent.
Considerations of trustworthiness are pertinent only to
mediating curatorial actions and their consequences
regarding the integrity and authenticity of a resource’s
dynamic state. The measures of success are twofold: a
secondary consideration of equivalence between intentional
and interpretive meanings/understandings and their
underlying referents; and a primary concern for the
substantive fulfillment of consumptive purpose.

The workings of this model can be illustrated thus
(Figure 2): this paper originated as an idea regarding
measurement of preservation efficacy on the part of the
author (its intentional meaning) with respect to an ur-notion
of preservation success (intentional referent) and reflective
of the author’s accumulated professional experience and
acclimation (productive context). That idea was expressed
as a scholarly paper with abstract, literature review,
citations, etc. (expressive form), encoded in Word Office
Open XML (symbolic representation), saved as
“Abrams.docx” (tangible representation), and accompanied
by CrossRef and PREMIS description of its informational
content and function as a container (content and container
description). Subject to proactive preservation stewardship
(managerial context), the file was migrated (curatorial
intervention) to “Abrams.pdf” (new tangible manifestation)
with PDF/A encoding (new symbolic representation) but
equivalent expression (curatorial authenticity) as a scholarly
paper (expressive form) and underlying idea (intentional
meaning). The PDF can be rendered with Acrobat (revealing
behavior) into a readable document (perceptual form),
instigating a state of newly acquired consumer knowledge
(interpretive understanding), conditioned with respect to that
consumer’s experience, acculturation, understanding of the
author’s reputation, and imprimatur of peer review and
presentation venue (consumptive context). The reading is
successful if it fulfills the consumer’s expectation for
purposeful use: a new understanding of preservation efficacy
informed by recovered productive intent.

Each component of the model supports its own set of
affordances, for which individual metrics and criteria can be
derived. Any instance or category of use entails a
characteristic subset of those affordances and is
evaluatively-susceptible to their associated metrics,
including the equivalence of expressive and perceptual form,
deriving intentional meaning and interpretive
understanding, and intentional and interpretive reference.
Impediments to success can arise at each point of transition
between model components [26], which suggests that
success metrics need to apply not only to the static
configuration of the component entities, but also their
dynamic unfolding through iterative preservation
interventions across archival timespans.

4 Preliminary findings and next steps

Like any formal discipline, digital preservation should be
viewed as a complex of actors, policies, technologies, and
practices [30, 47]; its maturity is dependent on its capacity
for reflective self-evaluation [48]. There are two primary
measures of preservation efficacy: trustworthiness of
managerial systems and programs; and successful use of
preserved resources. Trustworthiness is most properly
associated with processes and actors leading to preservation
outcomes, but not the outcomes themselves [49]. The
attention heretofore given to trustworthiness is not
misplaced, but is insufficient for a true sense of preservation
efficacy. The measure of success complements and validates
prior designations of trustworthiness through consideration
of the operational actuality of outcomes alongside the
descriptive potentiality of trustworthiness.

Because of the open-ended time horizon of preservation
commitments, preservation success should be understood as
a provisional, rather than absolute value. One can’t make
categorical assertions beyond the ever-forward-moving
point of now, since the consequences of the future cannot be
fully anticipated [50]. This bears similarity to the concept of
scientific falsification under which a theory expressed in
falsifiable form is held conditionally true so long as it has
not been shown definitively false [51]; so too it is legitimate
to assert the success of digital preservation so far.

The temporal distance that is the primary impediment to that success necessarily implies concomitant cultural distance [52] and culturally-situated contingency. Thus, success should be evaluated according to a standard of situational verisimilitude, rather than universal fidelity to some illusory notion of canonical state and information experience. This bears similarity to the concept of scientific truthlikeness under which the truth of a theory ranges along a spectrum of plausibility, and superseded but still-explanatory truths may be subsumed as special cases of more general theories [53]; so too it is legitimate to evaluate success as the relative degree to which preserved resources can be purposefully exploited.

The general approach presented here is still under development. Its semiotic-phenomenological modeling of preservation-enabled communication and consequent identification of semiotic affordances will be further refined through Conceptual Framework Analysis [54] for more detailed representation of the complexity and nuance of the preservation domain. In particular, public preservation policy statements [55] will be subject to close content analysis [56, 57] to surface assumptions, ambiguities, and service level obligations that define intentions, constrain actions, and set expectations. Subsequent to that, effort will turn to the last stage of communicological analysis, the derivation of specific criteria and metrics for evaluating digital preservation efficacy and success.

5 Conclusion

Digital information is indispensable to contemporary commerce, culture, science, education, and entertainment. No future understanding of a prior time in the digital age is possible without proactive preservation of our digital heritage. A communicological model of the semiotic and phenomenological affordances of the digital preservation enterprise provides a better means to analyze, explicate, and understand the domain. The model is consistent with, but extends OAIS-defined constructs for richer representation and broader applicability. It can lead to new criteria and metrics for evaluating success in a manner cognizant of post-custodial context and post-modernist contingency. The availability of conceptually-sound yet operationally-feasible measures will offer scholars new insights into the theory, practice, and limits of efficacy of the enterprise. An evaluation rubric based upon those measures will provide practitioners with the means to assert meaningful nuanced distinctions regarding intentions, expectations, and outcomes in a concise yet precise manner. It also will form the basis for rational prioritization of strategic goals, allocation of programmatic resources, and transparent accountability to stakeholder communities.

![Figure 2: Preservation-enabled communication](image)

Acknowledgement

My thanks to Kate Devitt, Peter Bruza, and Anthony Bernier for support and direction; Marcia Bates, Christoph Becker, Michael Buckland, Cliff Lynch, and Jerry McDonough for perceptive comments on earlier presentations of my work; and Raivo Ruusalepp for first raising the question of verisimilitude.

References


