Digital Archive and Preservation (DAP) Framework - A structure designed to depict both the distinct and the interdependent activities and responsibilities of the digital archives and digital preservation domains.

The Digital Archives and Preservation (DAP) Framework defines and disambiguates the terms *digital archives* and *digital preservation.* The DAP Framework emerged from the Digital Preservation Management Workshops curriculum to help navigate these terms are often used interchangeably or incorrectly when discussing activities associated with these domains. The DAP Framework evolved from an earlier Digital Preservation Management (DPM) Framework (McGovern, 2016). This description delineates the components and interactions of the DAP Framework.

Background

The digital archives and digital preservation domains have advanced significantly since the 1990s when these terms came into increasingly frequent use. The blurring of these affiliated, but distinct domains results in ambiguity creating a barrier to furthering their advancement as each continues a natural process of specialization. The DAP Framework enables more productive discussions that reference these terms to support more correct and specific citations for these terms and to avoid confusion by comparing and contrasting the domain activities represented by terms. The use of a framework helps illustrate what is distinct about digital archives and digital preservation, what is interdependent between the two domains, and how the timing and duration of their activities define them.

There are two main complicating factors that contribute to the ambiguity surrounding the terms. The first involves the differences how affiliated domains like information technology use them. In many organizations, one person, team, or unit is responsible for both digital archives and digital preservation activities making it easier to use the terms interchangeably and less necessary to differentiate the meaning of the terms. In practice, the term digital archives is used in diverse ways because the term archives takes on different meanings when applied in a digital context. For example, the IT domain may refer to storing content as archiving and refer to any aggregation of stored digital content as an archive. Similarly, the term digital preservation was used in library contexts to refer to protecting print originals by digitizing them, a practice and term that have been deprecated though still sometimes used. Referring to the practitioner, a digital archivist may refer to anyone who works with digital content of any kind or may refer to an archivist working with records in digital form. To understand digital practice that includes both digital archives and digital preservation, it is useful to be aware of these various archival references (McGovern, 2018).

The second complicating factor stems from the central role of technology in digital practice. Both digital archives and digital preservation are intrinsically related to computer-related technologies because digital content is created using

these technologies and long-term access to digital content relies upon an evolving array of technologies throughout its life. This reliance on technology can result in framing digital preservation as a technology-only problem. It is common and problematic for the related term infrastructure to be limited to mean only the technologies needed for digital practice rather than also to the organizational infrastructure including the human effort and wherewithal needed to continually evaluate, select, implement, support, and replace technologies. This technology focus led many digital preservation community discussions to include an inquiry about what "stack" an institution was using. For example, the Information Technology (IT) Stack (Figure 1) is often used to illustrate layers of technology.

Figure 1. Common representation of an IT Stack

The DAP Framework adopts the IT Stack's use of layers to explore the infrastructure for digital archives and digital preservation, so the Framework may also be called the DAP Stack. The DAP Framework responds to the IT Stack by demonstrating what the stack for digital archives and digital preservation might look like. Increasingly, the digital preservation community's literature demonstrates that the biggest challenge is not technological, but human or organizational—getting people to begin then continue to adopt practices to collect and sustain digital content.

The digital archives and preservation community has no definitive source, yet, for authoritative definitions of key terms. As a starting point, this elaboration of the DAP Framework uses these working definitions:

- *digital archives* refers to all the activities organizations or individuals required in real-time to appraise, acquire, process, describe, secure, and make available specified digital content;
- *digital preservation* refers to all the activities organizations or individuals engage in over-time across generations of technology to ensure the long-term readability and usefulness of specified digital content;

- *digital practice* is an inclusive term to cover activities at any point in the lifecycle that are relevant to digital archives and digital preservation; and
- *infrastructure* refers to the combination of human agency, archival and technological expertise, and technologies, e.g., software, hardware, and other tools and equipment, needed for digital archives and digital preservation.

In subsequent sections, this definition of the DAP Framework identifies the layers of activity that both digital archives and digital preservation address in different ways, defines the DAP Framework roles, and explores the relevant human-technology interactions within and between digital archives and digital preservation in practice.

DAP Framework Layers

There are six layers in the DAP Framework, each including both real-time digital archives and over-time digital preservation activities.

1. Governance includes:
- decision-making to enact policies that inform and guide digital practice;
- setting priorities to develop and sustain digital practice;
- investing in organizational and technological infrastructure; and
- managing digital archives and digital preservation programs.

2. Collection Scope includes:
- determining what digital content is in scope for the repository;
- establishing and applying criteria to consistently identify what is in scope;
- ensuring in-scope digital content is acquired and preserved; and
- evolving the collection scope as requirements and digital content evolves.

3. Acquisitions includes:
- ingesting selected digital content in accordance with criteria and practice;
- aligning acquisition practices with current policies and regulations;
- defining and following practices that reflect community expectations; and
- ensuring secure, controlled, and documented intake of selected digital content.

4. Workflows includes:
- defining human-based steps to acquire, process, and share content;
- developing tool-enabled actions to ensure consistent completion of steps;
- building workflows with human-tool integration to maximize resources; and
- committing to the responsive evolution of workflows to enhance capacity.

5. Lifecycle Storage includes:
- planning for right-sized lifecycle storage to enable long-term access;
- establishing mechanisms to ensure everyday protection of digital content;
- implementing and maintaining protocols for emergency preparedness; and
- engaging in ongoing review to identify and implement optimal options.

6. Monitoring includes:
- assessing current practice, needs, and options to inform priorities;
- engaging in activities to ensure incremental improvement for digital practice;
- committing to periodic audit to evolve good practice and mitigate risks; and
- monitoring infrastructure to avoid obsolescence and harness new potential.

Governance

Collections Scope

Acquisition

Workflows

Lifecycle Storage

Monitoring

Figure 2. The DAP Framework layers

These six layers illustrate (Figure 2) the purpose and objectives of the core areas of activity to ensure long-term access to digital content that is managed by a repository.

In addition to the six DAP layers, there are two other components. The DAP Framework is built on a sustainable, compliant technological foundation represented by the IT Stack. Collaborative Services represent the combination of services a repository provides to support creators and users and to enable long-term access and re-use that should build on, not duplicate, the DAP Framework and IT Stack to leverage the repository's cumulative strengths. Collaborative Services include access services, research and data services, instruction services, and other service areas that may overlap with or extend beyond digital archives and digital preservation services.

DAP Framework Roles

There are a number of domains and roles that participate in and contribute to digital practice. The primary focus of the DAP Framework is on digital archives and digital preservation roles because of the confusion that results when these terms are used interchangeably. A closer look at digital archives and digital preservation roles illustrates the balance between these paired responsibilities and activities in the DAP Framework layers. In practice, real-time responsibility examples include appraise, acquire, secure, and make available digital content; and over-time responsibility

examples include invest in digital infrastructure, and protect and sustain digital content. There are distinct characteristics of activities for digital archives and for digital preservation in each layer of the DAP Framework (Table 1).

Table 1. Examples and characteristics of functional activities for the DAP Framework layers.

Functional Area	Real-time: Digital Archives	Over-time: Digital Preservation
Governance	managing digital collections	sustaining digital preservation
Collection Scope	what digital content to preserve	how to preserve digital content
Acquisitions	transfer and process digital content	validate acquisition process
Workflows	operations, lifecycle metadata	preservation metadata and plans
Lifecycle Storage	submission and dissemination storage	preservation storage
Monitoring	detect errors and implement response	anticipate and respond to risks

These are examples of the symbiotic connection because digital archives is responsible for completing real-time activities and accountable for ensuring that over-time digital preservation activities occur; and similarly, digital preservation is responsible for over-time activities and accountable for ensuring that real-time digital archives activities are completed.

There is an observable overlap in the kinds of software and other tools that are used in digital archives and digital preservation work that varies depending on the kinds of digital content a repository manages and how a repository defines and implements its workflows, among other factors. That overlap of common tools by both domains contributes to the tendency to conflate or misuse the terms digital archives and digital preservation. Regardless of the context, the frequency and specifics of use of common tools will differ. For example, fixity checking typically occurs at the point when digital archives staff acquire digital content then periodically until the repository's workflows have created preservation objects for the incoming digital content and deposited the preservation objects into preservation storage for long-term preservation. Once the preservation objects are in preservation storage, fixity checking becomes the responsibility of digital preservation. If the digital content in preservation storage needs to be re-ingested for some reason, responsibility for fixity checking temporarily shifts to digital archives while digital content is repackaged into updated preservation objects. In this scenario, responsibility for fixity checking reverts to digital preservation from digital archives when the new versions go to preservation storage. This checksum tool example illustrates the interchange of real-time digital archives and over-time digital preservation roles during the lifecycle of digital content.

In addition to the digital archives and digital preservation roles, there are IT

roles that build and maintain the IT Stack that enables the DAP Framework, and there are other repositories roles that build on the DAP Framework to extend discovery and access services. Rounding out the roles, creators and stewards of the selected digital content populate the repositories that digital archives and digital preservation manage. When creators hold onto content in a post-custodial context, creators may implicitly take on some of the digital archives roles, possibly in collaboration with an existing repository. Current and future users help to motivate and demonstrate the need for digital archives and digital preservation. Collaboration across these domain roles is key to sustaining access to selected digital content over-time (McGovern, 2018).

Partnering Humans and Technology

Within the DAP Framework, the roles focus on the interchange between digital archives and digital preservation for the horizontal layers. The human-tool continuum (Figure 3) emphasizes the vertical interactions within the DAP Framework. For both digital archives and digital preservation:

- *top-down*: decisions as captured in high-level policies are a top-down activity. These decisions frame and evolve the scope, purpose, and objectives of digital archives and digital preservation programs, including ways to measure progress and success.
- *bottom-up*: recommendations, experiments, and demonstrations that enable solutions and enhancements to implement and improve workflows and sustained activities tend to be bottom-up activities.
- *both directions*: rules that reflect decisions at all levels enable both real-time and over-time activities to be implemented and updated as requirements, technology, capacity, and good practice evolves.

Organizational

High-level organizational policies
reflect the intentions of the organization

Lower-level organizational policies
document the decisions of the organization

Individual policy statements
regulate the actions of the organization

Encoded policy statements
translate organization's policies into actions

Technological

Figure 3. Human-Tool Continuum

To be able to demonstrate good practice, it is a community expectation for both digital archives and digital preservation at all levels of the human-tool continuum that decisions will have current and cumulative documentation and that the results of applying those decisions will be similarly recorded.

The full DAP Framework diagram (Figure 4) builds on the human-tool continuum to illustrate that the Governance, Collection Scope, and Acquisitions layers are primarily human-driven; and Workflows, Lifecycle Storage, and Monitoring are more technology-enabled. The human-tool continuum applies to both digital archives and digital preservation, though in relation to technology, another distinction between these two domains is that digital archives tends to be more engaged with current developments pertaining to tools and implementation that support real-time operational activities; and digital preservation tends to be focused on infrastructural issues for durable digital preservation storage, monitoring to mitigate the impact of technological change and auditing to demonstrate conformance with standards. Within community discussions, digital forensics is an example of an ongoing topic that primarily relates to digital archives because of its relationship to acquisition activities; digital preservation storage criteria is a topic of interest for digital preservation. Individuals may be involved with technologies for both digital archives and digital preservation without thinking about real-time and over-time implications, contributing to a blurring of distinctions.

The IT Stack is depicted as smaller in the DAP Framework (Figure 4) because the focus is on digital archives and digital preservation; and the Stack is offset below the layers in the Framework to emphasize the parts that are human-driven and the parts of the IT Stack that are technology-enabled. The Stack is also offset because it supports other programs in repositories in addition to digital archives and digital preservation. The entire diagram represents the infrastructure needed for sustaining access to selected digital content over time.

Applying the DAP Framework

The DAP Framework illustrates how the domains of digital archives and digital preservation are distinct, how they can and should work together, and how they are supported. Ideally, the technological foundation for the DAP Framework and the collaborative services that build on the DAP Framework contribute to achieving a common objective that it shares with digital archives and digital preservation. If the DAP Framework is not in place or not known to the other roles, the foundational technological layer and the comprehensive services layer try to fill that void by taking on work that is needed to extend the life of digital content. The DAP Framework reflects a balance of real-time and over-time layers of activities with designated roles that cumulatively provide the requisite infrastructure to ensure that selected digital content is available to meet current and future user needs.

The DAP Framework can be used by repositories and within digital practice contexts to:
- Specify digital archives and digital preservation roles and responsibilities regardless of where in the organization the activities take place;

- Update policies and workflows for digital archives and digital preservation;
- Raise awareness about digital archives and digital preservation; and
- Explain core concepts for training and instruction purposes.
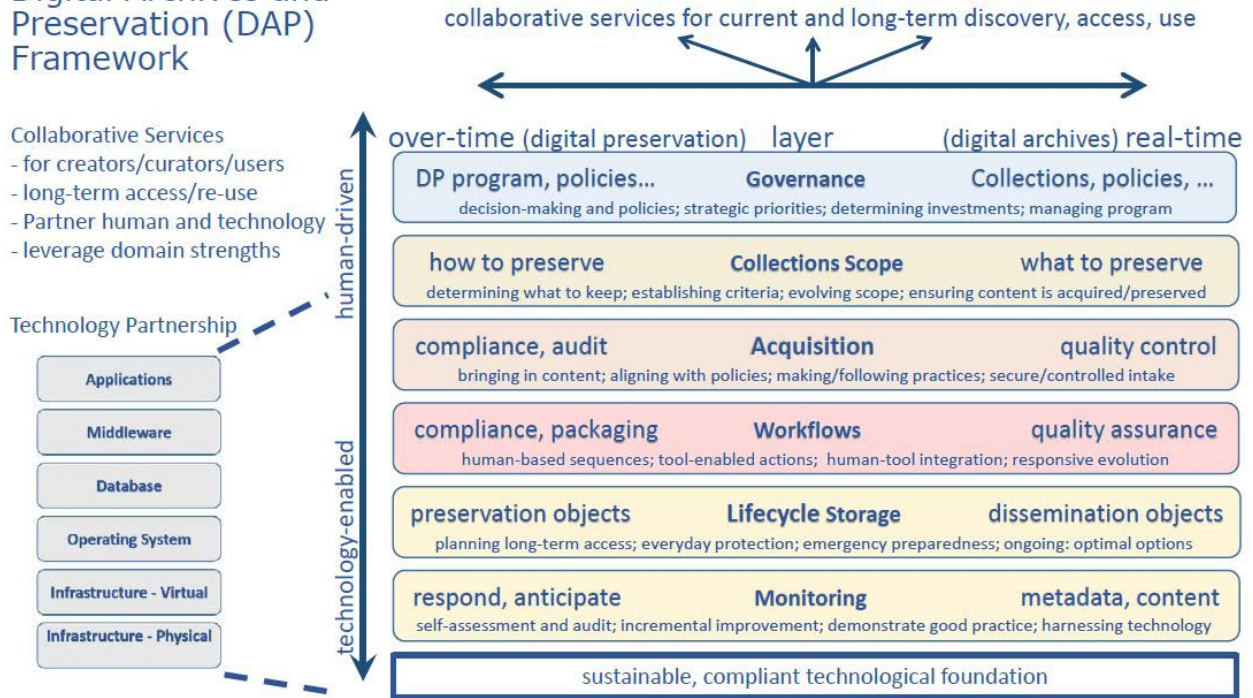
Figure 4. The Digital Archives and Preservation (DAP) Framework.

The author appreciates the contributions of Kari Smith, Institute Archivist at MIT.

References

McGovern, Nancy Y. "DPM 'Stack': A Management Infrastructure Frame for Digital Preservation that Parallels Technical Infrastructure." *DataOne Webinar Series*, December 2016.

McGovern, Nancy Y. "Radical Collaboration: An Archival View," Research Library Issues, no. 296, December 2018. https://doi.org/10.29242/rli.296

Keywords: digital archives, digital preservation, IT Stack