

Organizational Assessment Frameworks for Digital Preservation: A literature review and mapping

Authors:

Emily Maemura*
Faculty of Information
140 St. George St. #705
University of Toronto
Toronto, ON, Canada M5S 3G5
Phone: +1 416 978 3234
Fax: +1 416 978 5762
email: e.maemura@mail.utoronto.ca

Nathan Moles
Faculty of Information
140 St. George St. #705
University of Toronto
Toronto, ON, Canada M5S 3G5
Phone: +1 416 978 3234
Fax: +1 416 978 5762
email: n.moles@mail.utoronto.ca

Christoph Becker
Faculty of Information
140 St. George St. #651
University of Toronto
Toronto, ON, Canada M5S 3G5
Phone: +1 416 946 5367
Fax: +1 416 978 5762
email: christoph.becker@utoronto.ca

*Corresponding author

This is a preprint of an article accepted for publication in Journal of the Association for Information Science and Technology copyright ©2016 (Association for Information Science and Technology)

Abstract

As the field of digital preservation matures, there is an increasing need to systematically assess an organization's abilities to achieve its digital preservation goals, and a wide variety of assessment tools have been created for this purpose.

To map the landscape of research in this area, evaluate the current maturity of knowledge on this central question in DP and provide direction for future research, this paper reviews assessment frameworks in digital preservation through a systematic literature search and categorizes the literature by type of research.

The analysis shows that publication output around assessment in digital preservation has increased markedly over time, but most existing work focuses on developing new models rather than rigorous evaluation and validation of existing frameworks.

Significant gaps are present in the application of robust conceptual foundations and design methods, and in the level of empirical evidence available to enable the evaluation and validation of assessment models. The analysis and comparison with other fields suggest that the design of assessment models in DP should be studied rigorously in both theory and practice, and that the development of future models will benefit from applying existing methods, processes, and principles for model design.

Keywords: digital preservation, capability model, maturity model, literature review, assessment

Introduction

Mandates and business requirements are forcing many organizations to develop digital preservation capabilities to retain digital information over the long-term. Digital preservation (DP) aims to ensure that authentic digital objects remain accessible and usable over time, through a combination of people, technology, and processes. A fundamentally interdisciplinary field, DP has seen considerable contributions from information science, archival science, computer science, information systems, and other fields. As DP standards, methods, and concepts developed, organizations began to invest resources to meet their preservation needs. Interest and focus in recent years has subsequently extended from technical DP research and prototypes for individual tools, techniques, or system components to the deployment and operation of full-fledged archival information systems (Atkins, Goethals, Kussmann, Phillips, & Vardigan, 2013; McKinney, Benson, & Knight, 2012; Sinclair et al., 2011). To evaluate how these solutions work in aggregate, an organizational view is needed. This is the focus of organizational assessment, defined broadly here as approaches to systematically evaluate an organization's DP abilities. Organizational assessment is a key component of effective DP since it is necessary to demonstrate efficacy, a prerequisite to effective improvement, and a mechanism to help stakeholders establish trust in digital objects, systems, and repositories.

Existing assessment frameworks range from basic checklists to rigorous audits, and can be applied in different ways, including self-assessments, peer-review, and assessments by third parties and external auditors. Assessment approaches can be broadly grouped in three scenarios aligning with different organizational goals (Maemura, Moles, & Becker, 2015). First, organizations starting a new DP initiative can benefit from frameworks that facilitate **initial planning**. Second, organizations with existing processes and workflows can benefit from assessments that support **improvement** by identifying specific areas for organizational investment and development. Third, experienced organizations with documented and formalized processes can undertake assessment for external **certification** to recognize compliance with standards.

Despite significant attention to the latter, it is important to note that certification itself may not be necessary or even appropriate for every organization. This point is reinforced in the approach of maturity models that facilitate ongoing improvement or growth. Prevalent in related domains such as business process management (Rosemann, De Bruin, & Power, 2006) and IT management (Curley, Kenneally, & Carcary, 2015), maturity models are a type of assessment framework used to address how well organizational needs are identified and how predictably goals are met. Since these concepts are central to assessment, and many of the frameworks developed in DP reference concepts of maturity and capability, our work draws on this existing body of knowledge on maturity models.

This study aims to map the landscape of research in this area, identify limitations to current approaches, provide direction for future research in order to advance organizational assessment, and indicate the current maturity of knowledge on this central question in DP. In addition to reviewing available assessment options, it is equally important to understand the research that supports these approaches, in order to understand the state of organizational assessment in DP. Therefore, we aim to answer the following research questions:

- **(RQ1) How has organizational assessment developed in the field of DP?**
- **(RQ2) What types of research exist on assessment frameworks in DP?**
- **(RQ3) What gaps exist in the current landscape of research on assessment in DP, as compared to related fields, and what opportunities do these present for further research?**

The first sections of this paper provide a brief background of key concepts and developments in organizational assessment, before describing the literature search process and mapping the results based on distinctions from the literature on maturity models in information systems. The concluding discussion highlights gaps in the current spectrum of solutions and identifies opportunities for further research.

Background

DP involves managed activities directed towards ensuring ongoing access to digital objects. To achieve this goal, DP activities must address the challenges and threats to access through changing technological environments, shifting semantics, and evolving user groups. Time frames for DP are sufficiently long to necessitate solutions for media failure and technological obsolescence, while also addressing the need to keep digital objects meaningful to future users (ISO, 2012). DP's objective is to preserve digital objects while maintaining their authenticity, integrity, and usability.

Trust is paramount in this process, and requires a systematic, traceable, and accountable approach to managing digital objects (Research Libraries Group, 2002). From an external perspective, stakeholders can rely on certification to be assured a repository meets its responsibilities. This has long been recognized as a goal for many in DP (Garrett & Waters, 1996). From an internal perspective, certification is a means to other ends, useful where required or valuable for outreach and trust. More broadly however, assessment is a mechanism to facilitate diagnosis of where an organization stands in terms of capabilities and stages of growth and maturity. Potentially, this diagnosis can provide invaluable insights for future improvements. Approaches to determine the state of organizational capabilities for successful and effective DP in any given organization often correspond to concepts of maturity models, and the body of knowledge on maturity models provides examples of rigorous assessment and certification.

The concepts of capability and maturity stem from research on Quality Management by Crosby (1979), as applied in the influential Capability Maturity Model (CMM) for software engineering processes developed at the Software Engineering Institute (SEI) (Paulk, Curtis, Chrissis, & Weber, 1993). This original CMM divides capabilities into 18 key process areas with specific activities and maps each capability to five maturity levels: Initial (Level 1), Repeatable (2), Defined (3), Managed (4), and Optimizing (5) (Paulk et al., 1993). CMMs subsequently developed in different fields leading to many definitions for the concepts of capability and maturity, with varying agreement or consistency in the use of the terms (Maier, Moultrie, & Clarkson, 2012). For our purposes, *capability* is understood in an organizational context and defined as the ability to achieve a particular goal, typically supported by a combination of people, technology, and processes (Antunes, Barateiro, Becker, Borbinha, & Vieira, 2011). *Maturity* is the state of development an entity has reached. The maturity of a capability therefore describes aspects such as the consistency and reliability with which a function or process is performed to achieve a particular goal, and *capability maturity models* are a specific type of maturity model.

Ideally, such a model can be used to systematically evaluate current levels of maturity, identify the strengths and weaknesses of particular capabilities, and indicate criteria for improvement through specific goals that are necessary to achieve a higher level of overall maturity. In this sense, maturity models can be designed to identify progressive paths leading to improved performance. Evaluating an organization within this framework allows for diagnostic assessment while supporting continuous improvement (Mettler & Rohner, 2009). To achieve this in practice, methods of applying a maturity model to conduct an organizational assessment are necessary to ensure valid and comparable results. The SEI has developed a comprehensive and rigorous framework for application of the CMM outlining very specific requirements for different methods of assessment with increasing degrees of rigor and reliability (SCAMPI Upgrade Team, 2011).

Maturity models themselves are thus just one central component of a comprehensive *assessment framework* consisting generally of three building blocks: a structured **model**, which is a “formal description of some aspects of the physical or social reality for the purpose of understanding and communicating” (Mylopoulos, quoted in Mettler & Rohner, 2009); a **method**, i.e. a systematic means for applying the model and achieving an objective with it; and a **tool**, a concrete or abstract thing that enables the completion of a particular task (OED, n.d.). The elements comprising each framework may vary, with the model alone serving as the minimum. Frameworks are also supported by documentation for one or all of these components.

Maturity models have been developed for domains ranging from business process management and enterprise content management to IT governance and information systems (IT Governance Institute, 2007; Object Management Group (OMG), 2008; Pelz-Sharpe, Durga, Smigiel, Hartman, & Byrne, 2009). Research interest in maturity models has increased in many other fields, provoking surveys of existing models (Jokela, Siponen, Hirasawa, & Earthy, 2006; Wendler, 2012). Since these models can be viewed in terms of their use across different organizational environments (rather than one single case), much of the knowledge base on maturity model development in Information Systems is heavily informed by Design Science Research. This form of research employs a conceptual framework of guidelines for the rigorous creation of IT artifacts as a way to solve problems at the confluence of people, technology, and organizations (Hevner, March, Park, & Ram, 2004). A typical research methodology adapted from this

framework, such as that described by Peffers et al. (2007), uses an iterative process to construct and evaluate practically relevant artifacts according to rigorous procedures. This methodology defines a series of phases for the design process, including the demonstration of the designed artifact in a real-world context and an evaluation of the artifact's effectiveness that informs the next cycle of design and development. Researchers developing and applying maturity models in other fields have adopted a Design Science perspective on the design and use of models. Taken together, this research that focuses on the designed model as an object of study forms a knowledge base for maturity models that can be applicable in many different domains. We draw on this research here, including initial work that identified the main phases of maturity model development and proposed design process models (De Bruin, Freeze, Kaulkarni, & Rosemann, 2005). This has been extended to the creation of a procedure model for systematically creating maturity models that outlines a rigorous approach to the development of models, specifying the key steps in the design process, as well as the documentation of key decisions at each step (J. Becker, Knackstedt, & Pöppelbuß, 2009). Further research has developed a set of general design principles for maturity models which can be broadly applied and used both to inform the development of maturity models, as well as to evaluate or compare models (Pöppelbuß & Röglinger, 2011).

Wendler (2012) provides an overview of these different kinds of research activities on maturity models across different domains through a systematic literature review, in order to determine “the maturity of maturity model research.” This work is significant in that it provides a synthesis of research on maturity models generally, and identifies common research topics, methods and trends in the literature over time. For our purposes, the findings on the “structure of research topics” provide a useful framework for analysis of research on assessment within DP. Research topics are categorized as either research *with* models, including the development, application, and validation of an individual model, or research *about* models, such as literature reviews, process models, and other theory about maturity models. These distinctions will be further described in the analysis of types of research below.

Development of Organizational Assessment in Digital Preservation

To address our first research question, how organizational assessment has developed in the field of DP, we present an overview of the existing models available, and the history of these developments. As a relatively young field, DP has not yet brought forth a solid body of research on organizational assessment. However, the relevance of the topic was recognized early on, and considerable attention has been paid to it in recent years.

The need for certification was recognized in the Garrett and Waters (1996) report on the activities of the Task Force on Archiving of Digital Information of the Research Libraries Group (RLG), and the subsequent RLG-OCLC report on Trusted Digital Repositories (TDR) Attributes and Responsibilities (Research Libraries Group, 2002). This focus on certification and compliance is also evident in the development of the *Open Archival Information System (OAIS) Reference Model* (ISO, 2012). Published initially in 2003 and later revised in 2012, the OAIS model does not address the concept of maturity, but has influenced many subsequent developments in assessment.

The period from 2006-2010 saw further collaborative work and developments, building on the foundations of the TDR Attributes and Responsibilities report and the OAIS model. Two influential documents were introduced: the NESTOR *Catalogue of Criteria for Trusted Digital Repositories* in 2006

in Germany and the *Ten Principles* for DP repositories, developed jointly in 2007 by the U.K.'s Digital Curation Centre (DCC), DigitalPreservationEurope (DPE), NESTOR and the Center for Research Libraries (CRL).

As there are now more options available for organizational assessment, we conducted a survey to help practitioners navigate the different choices, and determine if any of the options achieved a balance between simplicity in the assessment process and trustworthiness in the results (Maemura et al., 2015). We reported on fourteen identified assessment frameworks and compared features of each model, including the available tools or guidance for conducting an assessment, the intended audience for the results, and the mode of application (such as self-assessment or conducted by a third party or certified auditor). We noted an increase in the number of available assessment frameworks in recent years.

Building on this, Figure 1 presents a chronological view of these assessment frameworks, distinguishing between those focused on certification (bottom section of the figure) and those focused on planning and improvement (top section of the figure).

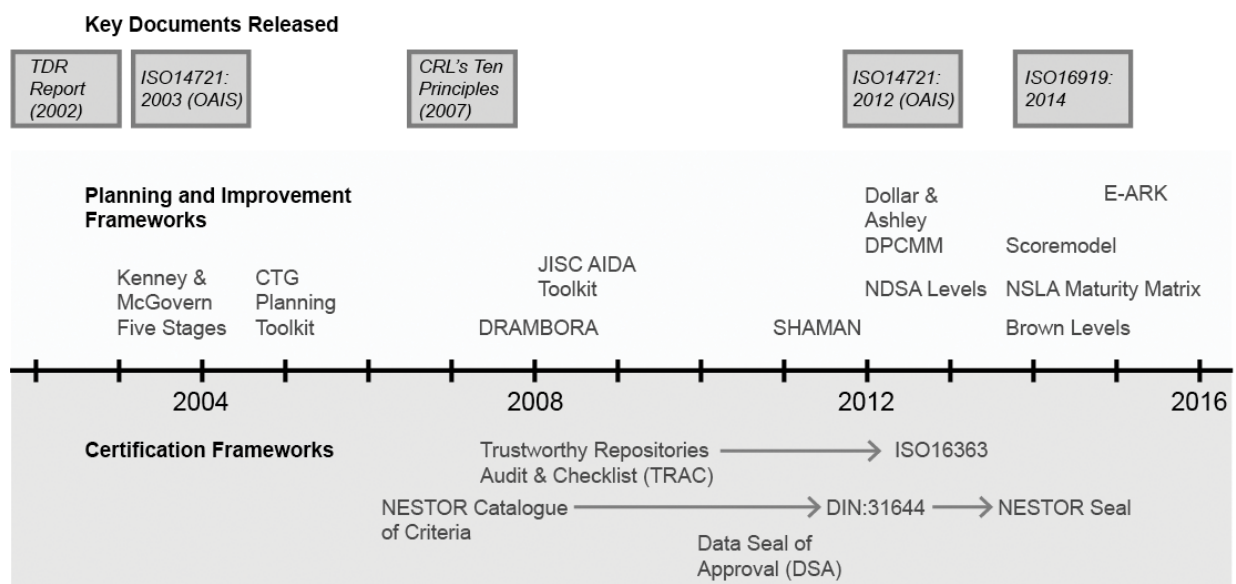


Fig. 1: Timeline of assessment frameworks

Early Work

Kenney and McGovern (2003) were the first to address organizational stages of maturity for DP in their essay the *Five Organizational Stages of Digital Preservation*. The *Stages* are presented in an easily accessible form, targeting self-assessment for audiences internal to an organization. As an early framework, the *Stages* remains high level, with little to address the details in carrying out an assessment, but influenced the development of frameworks such as the **Capability Assessment and Planning Toolkit** from the Center for Technology in Government (CTG) in 2005 (Pardo et al., 2005) and the **JISC Assessing Institutional Digital Assets (AIDA) Toolkit** from 2008 (University of London Computer Centre, 2009b).

The **Digital Repository Audit Method Based On Risk Assessment (DRAMBORA)** released in 2007 focuses on a situated analysis of risks and organizational readiness specific to an organization's assets and

activities, rather than based on a standardized model of these aspects (McHugh, Ross, Ruusalepp, & Hofman, 2007). This model was influenced by the *Ten Principles* and draws on the functional model of the OAIS, using the same categorization of activities.

Developments Toward Certification

The **Trustworthy Repository Audit and Certification (TRAC) Criteria and Checklist** was a significant step toward certification standards (OCLC, CRL, 2007). Development began in 2003, with pilot audits conducted during 2005-2006 (“TRAC Metrics | Center for Research Libraries (CRL),” n.d.). The criteria and checklist were first published in 2007. TRAC is notable as the first criteria set for audits of digital repositories, defining a series of items related to organizational infrastructure, digital object management, and technologies, security and technical infrastructure. Reports on successful audits have been published by CRL, and the TRAC criteria have informed many subsequent developments.

Following the standardization of the *OAIS Reference Model* in ISO14721:2012, the TRAC Criteria and Checklist was revised and formalized as **ISO16363** in 2013. ISO16363 draws strongly on the concepts outlined in ISO14721, and uses this as the basis for determining the trustworthiness of digital repositories. It takes into account different factors of the organization, as well as its infrastructure, security, and management. Certification is currently available through the Centre for Research Libraries (CRL) and the Primary TDR Authorization Body (PTAB). A related standard outlining requirements for certifying bodies based on ISO16363 was released in 2014 as ISO16919 (ISO, 2014). Since audit and certification are often long-term goals for repositories, there may be many steps along the path toward certification. In recognition of this, tools and methods have been developed to support self-assessment prior to a full audit; one example is MIT’s web-based self-audit tool (“Internal audit tool” n.d.).

The **Data Seal of Approval (DSA)**, introduced in 2010 by the Dutch-based DANS, is notable as an assessment framework that supports optional certification through an organized peer-review process (Harmsen, 2008). DSA takes a simpler approach, requiring organizations to comply with 16 guidelines.

A third option for certification was introduced in 2013 with the **NESTOR Seal** (NESTOR, 2013). This certification, based on the German DIN31644 standard, is envisioned as a middle-ground between the lightweight assessment of the Data Seal of Approval and the intensive auditing required by ISO16363 (Giaretta, Harmsen, & Keitel, 2010). The NESTOR Seal is awarded through a process of extended self-assessment then reviewed by NESTOR. Together, the three frameworks – ISO16363, NESTOR Seal, and DSA – comprise the European Framework for Audit and Certification.

Recent Developments - Alternatives Approaches to Certification

As the field matured, moving from initial proof-of-concept systems to full-fledged operations, research interest in assessment has increased. Additional assessment models were developed, accompanied by broader discussions on the challenges and effectiveness of assessment for certification or other purposes. It has been acknowledged that the full audit required by ISO16363 is time- and resource-intensive, demanding extensive preparation and documentation to execute the audit (Houghton, 2015; Rosenthal, 2014; Jackson, n.d.). Discussants at the Capability Assessment and Improvement Workshop (CAIW) at IPRES 2013 identified the need to find a balance between model flexibility and ease-of-use, and the trustworthy results of a rigorous audit (C. Becker & Cardoso, 2014).

Recent developments have presented alternatives to intensive audits, including the *Levels of Preservation* from the National Digital Stewardship Alliance (NDSA), and the *Digital Preservation Environment Maturity Matrix* from the National & State Libraries of Australasia (NSLA), both introduced in 2013. These models present simple frameworks for assessments that can be completed in a short period of time. Both models were developed by consortia to help their members assess their activities and initiatives.

In contrast, two other recent models were the result of development by individuals. As part of their consulting services, Dollar and Ashley created the *Digital Preservation Capability Maturity Model (DPCMM)* assessment model. Introduced in 2012, their DPCMM is largely based on compliance with the ISO14721 standard, and references concepts from other CMMs (Dollar & Ashley, 2014c). The DPCMM can also be used for self-assessment with the DigitalOK online tool introduced in 2014 (Ashley, 2014; “DigitalOK,” 2014). Another maturity model, inspired by P2MM (Williams, 2013) from the field of project management, is presented by Brown in his book *Practical Digital Preservation* (2013). However, there is no detailed specification of the model, no description of the methodology, or examples to support its application. Scoremodel, another framework released in 2013, is an online tool for self-assessment (“Scoremodel,” 2013).

Recently, capability models have been developed as part of larger projects. The SHAMAN project developed a capability-based model of DP, based on Enterprise Architecture concepts, with the SHAMAN Reference Architecture in 2012 (Antunes et al., 2012). The European Archival Records and Knowledge Preservation (E-ARK) released a Maturity Model for Information Governance in February 2015 and an Archiving Maturity Model with pilot assessments and evaluations in October 2015 (Proença, Vieira, & Borbinha, 2015a, 2015b).

Summary of Model Developments and Approaches

The number of assessment frameworks in the DP community has increased significantly in recent years, with 9 of the 14 frameworks developed between 2012 and the present. However, the increase in options makes it difficult for practitioners to compare and select one for use.

Comparing the three primary scenarios for undertaking an assessment (initial planning, improvement, and certification), audit and certification have been seen as strong motivations for undertaking an assessment and 3 models support certification specifically: TRAC/ISO16363, NESTOR Seal and the Data Seal of Approval (DSA). Development of models for certification has centered on compliance with the OAIS Reference Model (ISO14721). Many models targeting improvement are also often aligned with the concepts and terminology presented in ISO16363 and OAIS.

While these development trends are significant, it is unclear how the models are supported by research on assessment. The following sections investigate the types of research on the development or use of the assessment frameworks presented above to develop a comprehensive view of how those models are applied, validated, or discussed through research and other literature.

Literature Search

Our second research question asks: what types of research exist on assessment frameworks in DP? To characterize the types of research on assessment models in DP, we conduct a systematic literature search

drawing on the process developed by Wendler (2012) for maturity models broadly across domains. We begin with an overview of the search process followed by a summary of the findings.

Search process

Reflecting on the search process, Wendler notes that while some academic databases were particularly fruitful in returning search results, there was no single source covering all publications (2012). It follows that a search process aimed at returning maturity models research must consider multiple venues and incorporate different channels as part of an overall search strategy. We therefore adopted this approach for our literature mapping, in order to determine if the kinds of research found in the specific domain of DP mirrors the findings for maturity model research more broadly, and to identify existing gaps in research in DP. Our search process was conducted across five different stages, each involving several steps. The application of each step varied accordingly, as described below. The initial search (Stage 1) was completed in June 2015, while subsequent searches (Stages 2-5) were completed December 2015 and February 2016.

Stage 1: Journal databases

The search began with a key term search of two academic journal databases: Scopus and JSTOR (**Step 1**). Since these databases are broad in scope and coverage, our search terms at this step included domain terms in order to focus on results using maturity models for DP. Our searches used permutations of two main components: *topic terms* (e.g. capability, maturity, assessment, improvement and audit), and *domain terms* (e.g. ‘digital preservation,’ ‘digital curation,’ ‘digital stewardship,’ ‘digital continuity,’ and ‘archives’). The term ‘digital archives’ was used in the JSTOR search to reduce the number of unrelated returns caused by the high volume of humanities literature in that database. The full list of search terms used for this stage is presented in Table 1. JSTOR allows for full text search while Scopus was restricted for the Article Title, Abstract and Keywords. These searches returned 456 results (249 from Scopus and 207 from JSTOR).

Table 1: key search terms

Topic terms	capability OR maturity
AND (activity)	assess* OR improve* OR audit
AND (domain)	Digital Preservation OR Digital Curation OR Digital Stewardship OR Digital continuity OR Digital Archives

After combining the results from JSTOR and Scopus, and removing duplicates, 133 articles were eliminated, reducing the total article count to 323 (**Step 2**). Reading the journal title, we applied inclusion criteria in order to limit the results to journals from the fields of information systems, archival science, library science, computer science, information management, and records management (**Step 3**). Some results, such as full conference proceedings and works not in English, were removed based on their format (**Step 4**). A final set of inclusion criteria was applied after assessing the article title and abstract (**Step 5**), and limiting our scope to articles:

- explicitly focused on the domain of DP (as defined by the concerns of OAIS and TDR);
- concerned with assessment at the organizational level (rather than digital objects, collections of digital objects, or specific functions); and,

- focused on capability, maturity, or a similar conception of the organization’s DP abilities as a whole.

Through these criteria, 80 items were removed based on domain, and 62 because they did not focus on organizational assessment concepts. This limited results to only 11 relevant papers for review from these two databases.

Stage 2: Discipline-Specific Venues

The second stage focused on two relevant sources in the domain of DP that are not indexed in large databases. A key term search (**Step 1**) was conducted in the International Journal of Digital Curation (IJDC) and in the proceedings of the International Conference on the Preservation of Digital Objects (iPRES). Presence in these publications was taken to indicate relevance to the domain of DP, so inclusion and exclusion Steps 3 and 4 were omitted, and initial key search terms were limited to the five topics shown in the first two rows of Table 1. The names of 14 models we previously surveyed (Maemura et al., 2015), and the acronyms by which they are known, were added as optional search terms. These are listed with their full name and source in Table 2 below.

Table 2: Model-based search terms for Stages 2 and 3

Model and Reference	Search Terms
Five Organizational Stages of Digital Preservation (Kenney & McGovern, 2003)	“five stages”, “five organizational stages”, “Cornell model”, “five organizational stages of digital preservation”*
Capability Assessment and Planning Toolkit (Pardo et al., 2005)	“Capability Assessment and Planning Toolkit”
Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) (McHugh, Ross, Ruusalepp, & Hofman, 2007)	DRAMBORA
JISC AIDA Toolkit (University of London Computer Centre, 2009b)	AIDA
Data Seal of Approval (Harmsen, 2008)	“Data Seal of Approval”, DSA*
ISO 16363 / TRAC (OCLC, CRL, 2007)	16363, ISO 16363, ISO16363, TRAC
Digital Preservation Capability Maturity Model (DPCMM) (Dollar & Ashley, 2014c)	DPCMM
SHAMAN Capability Model (Antunes et al., 2012)	SHAMAN
Maturity Levels (Adrian Brown) (2014)	“Adrian Brown”
NDSA Levels of Digital Preservation (Phillips, Bailey, Goethals, & Owens, 2013)	NDSA, “NDSA Levels”*
DIN31644 / NESTOR Seal (NESTOR, 2006, 2013)	NESTOR, “NESTOR Seal”*, “DIN31644”*
NSLA Maturity Matrix (Pearson & Coufal, 2013)	NSLA, “Maturity Matrix”*
Scoremodel (Gillesse & Vanstappen, 2014)	Scoremodel
e-ARK (Proença et al., 2015b)	e-ARK, EARK, E-ARK

*Search term added in Stage 3

These terms were used to search all indexed categories in IJDC, including titles, abstracts, indexed terms, full text, and supplementary files. Similar searches were conducted on the Phaidra repository where iPRES papers and proceedings are hosted (<https://phaidra.univie.ac.at/>). Here all available fields were searched, including owner, full-text, title, abstract, and identifier. The term “iPRES” was used in conjunction with the specific search term, in order to separate items from the conference from other items in the repository. This step of the search process resulted in an initial list of 280 items (131 from IJDC and 149 from iPRES).

Step 2 of this stage involved the removal of duplicate items retrieved within this set and compared with results of Stage 1. Three articles appeared in more than one location with the same title: two iPRES papers and one paper discovered in Stage 1 were later published in IJDC. Both versions of these were included. This resulted in 183 items. From here, the inclusion criteria from Stage 1 were applied (**Step 5**). These were modified slightly in that presence in either IJDC or iPRES was taken to indicate adequate relevance to the domain of DP. The remaining three inclusion criteria were applied as they had been in Stage 1 by reading the title and abstract for each article. This resulted in the removal of 122 items because they were not about organizational assessment, eight items that addressed assessment but not organizational DP capability as a whole, and six items from IJDC that focused on other areas besides DP, as a result of IJDC’s broader scope. This step further reduced the total item count to 47. At this point, the exclusion criterion from step 4 of Stage 1 was applied, along with an additional exclusion criterion in order to ensure internal consistency (**Step 6**):

- Content in the form of posters, presentation slides, or video lectures was excluded

10 items from the iPRES proceedings were removed by these criteria, resulting in 37 final items for this stage.

Stage 3: Non-Academic Search

The limited number of research articles from these academic sources suggests three explanations: that research on maturity models in DP exists but is not available from these sources; that preliminary work with maturity models in DP is not presented as peer-reviewed research; or that little to no work with maturity models exists in DP.

To address all three possibilities, we developed a supplemental search strategy to identify work beyond academic journals and conferences, including publications such as books, dissertations, reports or other grey literature by using highly focused Google search terms. The key terms used were drawn from the names of the models and their variations in Table 2 (**Step 1**). As the principle documentation of these models was already known, the terms “case” and “application” were added in order to capture any case studies or applications that may have been conducted. In some cases additional variations of the model names were added to return more targeted search results. The modified terms are indicated in Table 2 above. The terms “digital preservation” and “digital curation” were used in conjunction with the following model names and acronyms to increase the number of results relevant to the domain: AIDA, DSA, E-ARK, TRAC, and SHAMAN; while the term “maturity” was added to “Adrian Brown”. The searches were done outside of a Google account to minimize personalization of search results. The first three pages of the search results were considered. Inclusion criteria used in Step 5 of Stage 1 were again

used, and the full content had to be available in English through the links provided by Google (**Step 5**). The application of the exclusion criteria from Stage 2 (**Step 6**) resulted in 24 items added from this stage.

Stage 4: Personal resources

To ensure as comprehensive a review as possible, we added items not retrieved through stages 1-3 through four additional means:

Seventeen items were obtained directly from their authors, the framework creators, or project websites. This includes material obtained through correspondence with framework creators as part of our earlier survey (Maemura et al., 2015) and subsequent empirical studies. Authors’ knowledge resulted in five additional items. Consideration of previous work co-authored by one member of the research team led to two additional items. Finally, three items included in this literature review were discovered due to their publication in the same venues or journal issues that arose from our other search results.

A total of 27 items were added at this stage from these four different approaches.

Stage 5: Citations

The first four stages of searches resulted in 99 items for analysis. As a final stage, the citation list from each of these 99 items was scanned for additional relevant items. Only the “References” or similarly titled section at the end of each text was considered. Footnotes on individual pages and any references within the text itself were ignored. In order to be included, references had to be complete citations rather than URLs. Inclusion criteria used in **Step 5** were re-applied. This process resulted in 25 items, to which we applied the exclusion criteria from **Step 6**, with one additional criterion: Items without their full text available were excluded. Six items were removed at this point – a video lecture, a set of presentation slides, a set of workshop notes, and three informal documents not publicly available. The remaining 19 items increased the final list to a total of 118 items considered in this literature review.

The full item count for each step in each of the five stages is listed below in Table 3.

Table 3: Literature search results

	Search Steps	Stage 1: Journal Databases	Stage 2: Discipline- Specific	Stage 3: Non- Academic	Stage 4: Personal Resources	Stage 5: Citations
1	Key term search	456	280	n/a	n/a	n/a
2	De-duplication of Search Results	323	183	n/a	n/a	n/a
3	Apply inclusion criteria based on journal title	159	n/a	n/a	n/a	n/a
4	Exclude non- English articles and full conference proceedings	153	n/a	n/a	n/a	n/a
5	Apply inclusion	11	47	24	27	25

	criteria based on article title and abstract					
6	Apply exclusion criteria based on format	n/a	37	24	n/a	19
7	Total Items from each Search:	11	37	24	27	19
	Merged Search Results:	11	48	72	99	118
	Final Item Count:					118
	Relevant after scanning whole content:					80

Final Exclusion Based on Content

The 118 items resulting from this 5-stage search process were filtered for exclusion based on a closer analysis of content. The initial inclusion criteria of explicit presence in the DP domain, concern with assessment at the organizational level, and the focus of assessment on capability, maturity or similar concept, were re-applied to the full contents of the articles. At this stage, articles from IJDC that were exclusively concerned with research data management and did not explicitly address DP concerns were excluded, as were articles that focused solely on the evaluation of software frameworks, or infrastructures, or that presented tools without discussing an underlying assessment model. At this point, 38 items were excluded, bringing the total number of items for categorization to 80.

Literature Search Findings

Initial analysis of the literature search reveals two findings. First, Figure 2 shows that the number of items published on assessment in DP has been increasing substantially over time. There was a significant increase in the number of papers published in 2010 and 2011, with a general trend of steady growth from 2010 to the present, culminating in a total of 13 items in 2014. (The search process returned 5 items published in 2015, but given the timing of the search, that number cannot be seen as representative of the current state of research and has been omitted from this chart.)

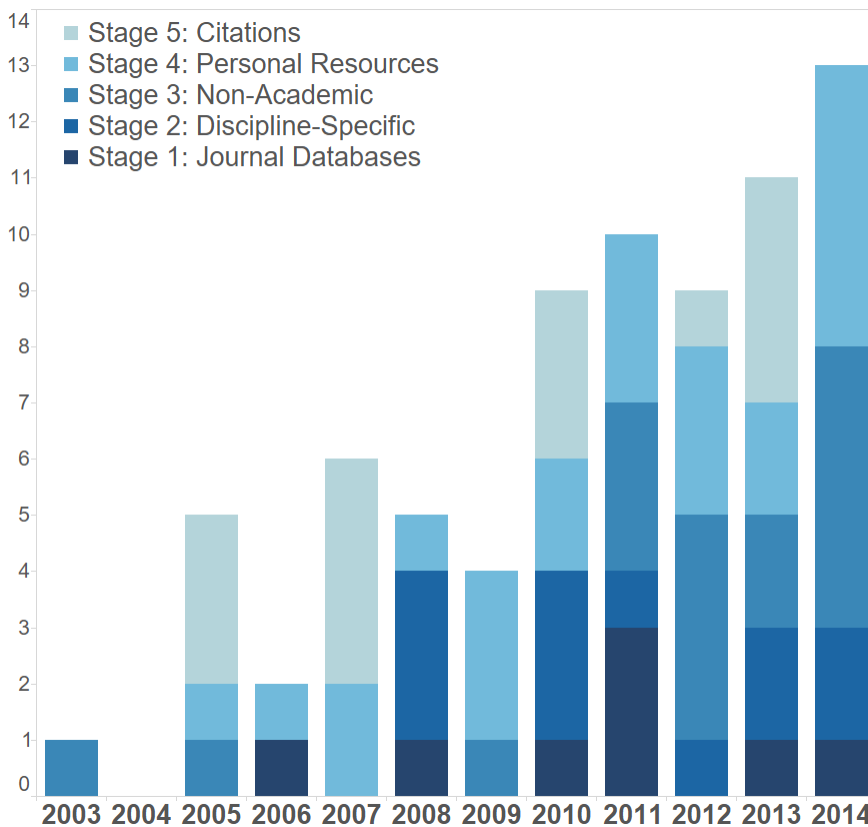


Fig. 2: Total publications per year by search stage

Little published research was discovered through formal publication venues, and the results were substantially increased by drawing on the network of authors and collaborators. Our initial search of databases provided only 9% of total literature discovered, and our search of domain specific peer-reviewed venues in Stage 2 provided less than a third. This indicates that the work on organizational assessment in DP is largely published for a non-academic audience. However, over 80% of the peer-reviewed items (as returned by the first two searches) were published during or after 2010, which indicates increasing interest in organizational assessment by scholars in the field.

Literature Mapping and Analysis

Mapping categories

We conducted a mapping of the relevant literature from this search, based on the classification scheme developed by Wendler (2012) for the different types of maturity model research. His analysis categorizes articles into four mutually exclusive aims, grouped into two distinct streams:

1. **Research *with* models** includes maturity model **development**, maturity model **application**, and maturity model **validation**. These types of research often take a single model as the object of study. These three steps - development, application, and validation – form a cycle of design, empirical testing and evaluation which then informs the next iteration of the design. This aligns with the design process described in Design Science research, whose principles and methodologies include iteration through this cycle to lead to more effective models (Hevner et al., 2004; Peffers et al., 2007).

2. In contrast, **research about models** does not take a single model as the object of study, but addresses broader questions of model creation and use. This category of ‘**Meta**’-articles includes conceptual work such as literature reviews, theoretical reflections, and process models to inform development or application.

Our analysis thus starts from these four categories: maturity model **Development**, **Application**, **Validation**, and **Meta**-Articles that reflect on maturity models.

To characterize the diversity of results found in our search process, we also categorized the venue of publication as either journal, conference proceedings, book, or report. Figure 3 maps the four types of research into publication venues. It reveals that a great deal of work on model development and application is published in reports and conference proceedings, with varying degrees of peer review.

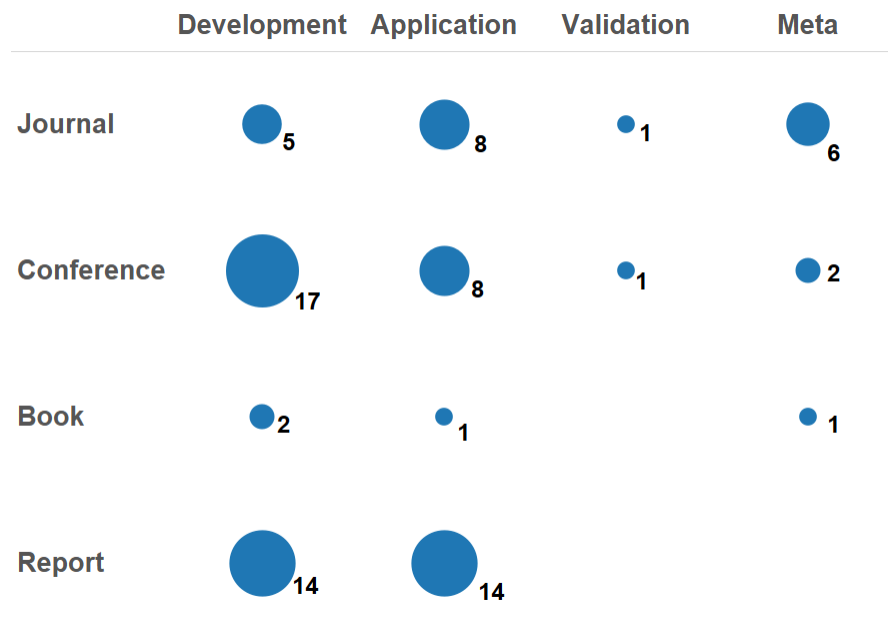


Fig. 3: Research type by publication venue

Development

We divided the category of **Development** into three distinct sub-categories:

- **Model Development** includes work that introduces an assessment framework or provides full documentation of a model. This sub-category also includes work that substantially builds upon an existing model, such as the introduction of a new version with significant changes.
- **Proto-Development** includes work that lays the conceptual groundwork or vision for a model that is not yet developed.
- **Model Presentation** includes work that presents a previously developed model, introducing it to a new community or in a new venue.

The majority of publications in the **Development** category fall under the main sub-category of **model development**. Some of these present a model in full, and include explanatory notes on each element of

the framework such as TRAC (OCLC, CRL, 2007), the NESTOR Seal (NESTOR, 2013), DPCMM (Dollar & Ashley, 2014c), and AIDA (University of London Computing Centre, 2009b). This also includes publications that present additional descriptions of the development process by model creators, such as the CTG report on their approach (Cresswell, Pardo, & Canestraro, 2006). Works categorized as Development often includes some aspect of application of the model to a real organization, through pilot studies like those used with E-ARK (Proença et al., 2015b). Papers like these were determined to be model development since the model was applied by the creators during its production. Five papers fell in the subcategory of **proto-development** and eight in **model presentation**. The full mapping of these categories and references to specific papers can be found in Appendix 1, Table 8.

Application

The majority of articles that fall under the **Application** category reported on results of an assessment within a specific organization. These include formal reports and documentation developed through certification processes, such as the six reports published by the CRL on successful audits using TRAC (Center for Research Libraries (CRL), 2010, 2011, 2012, 2013, 2014, 2015). Also included here are reports from consultants such as the report on the DPCMM applied with Council of State Archivists (Council of State Archivists, 2012), as well as a number of articles describing individual self-assessments (Downs & Chen, 2010; Elstrøm & Junge, 2014; Quisbert, 2008). One case study notes the experiences of application towards certification, including details of the time and resources required (Mitcam & Hardman, 2011). Four of the papers in this category reported on the application of more than one model to an organization (Elstrøm & Junge, 2014; Farrington, Fischer, Hanus, & Wopschall, 2013; Giarretta et al., 2012; Massol, Rouchon, & Bechard, 2013). These were not considered model validation, since the comparison was not conducted with the intention to evaluate the models themselves, but instead to provide different perspectives on evaluating the organization.

Our interpretation of the Application category differs somewhat from Wendler’s initial categorization since our search includes broader sources beyond academic journals, while his excludes “non-academic, journalistically written articles, opinion and experience papers, as well as simple success stories” (2012, p. 1322). With this in mind, we sought to explore how models have been adopted by others beyond their creators, and if applications of models were simply experience papers and success stories or if there was critical evaluation of the models through these applications. Table 4 shows all models for which third-party applications were found, listing each paper that reports on this model and summarizing any insights on the assessment framework noted by the authors. No third party application was found for any of the other fourteen models. Generally, certification frameworks have been more widely adopted and applied than improvement frameworks. Few of these third-party applications went very far in terms of critical discussions or evaluation of the model itself, and instead focus on presenting the details and context of a particular organization.

Table 4 : Third Party Applications and Lessons Learned

Framework	Third-Party Applications	Lessons Learned
Primary Purpose: Initial Planning		
NDSA	(Goethals, 2013)	Assessment process took approximately two hours with a knowledgeable assessor (Goethals).
Primary Purpose: Improvement		
AIDA	(Miller, Blake, & Sorsby, 2012)	Too much room for interpretation, elements in

		different legs can be too similar, and there is too much text on each page. Flexibility was AIDA's strongest feature.
DRAMBORA	(Elstrøm & Junge, 2014) (Farrington et al., 2013) (Karlach, 2010) (Massol et al., 2013) (National Technical Library, 2010) (Vaska & Pejsova, 2012) (Quisbert, 2008)	The DRAMBORA online tool did not support the Czech language, creating barriers to use (Karlach, NTL, Vaska). Guidelines for DRAMBORA need to be adapted for specific cases or circumstances (Quisbert).
Primary Purpose: Certification		
DSA	(Giaretta et al., 2012) (Massol et al., 2013) (Mitcam & Hardman, 2011) (Schumann, 2012) (Schumann & Mauer, 2013) (Vardigan & Lyle, 2014)	Assessment with DSA is lightweight and straightforward (Vardigan & Lyle) DSA self-assessment forces institutions to reflect on their own operations (Schumann & Mauer) The time frame for the assessment depends on two factors: how far the repository is from meeting the criteria and the extent of documentation it has (Mitcam & Hardman)
ISO 16363 / TRAC	(Bote, Termens, & Gelabert, 2011) (Downs & Chen, 2010) (Elstrøm & Junge, 2014) (Farrington et al., 2013) (Ferreira, Faria, Hahn, & Duretec, 2014) (Giaretta et al., 2012) (Houghton, 2015) (Kirchhoff, Fenton, Orphan, & Morrissey, 2010) (Rodrigues et al., 2014) (Schultz & Gore, 2010) (Termens & Locher, 2012) (Quisbert, 2008)	Self-assessment with ISO16363 demands a lot of documentation. Being only available in English, there can be language barriers for non-native speakers. It was hard to define the level of detail for each metric (Elstrøm & Junge). ISO16363 has ambiguous terminology and reflects the theoretical (abstract) nature of OAIS. Assessment with this model is time-consuming and resource-heavy (Houghton). Focus on centralized repositories and infrastructure does not adapt well to assessing distributed archives. Inaccurate assumptions about ingest, ignores the possibility of dark archiving as legitimate activity (Schultz and Gore). TRAC needs to be adapted for specific cases / circumstances. No method exists for assessment (Quisbert).
NESTOR / DIN 31644	(Giaretta et al., 2012) (Quisbert, 2008)	The NESTOR criteria needs to be adapted for specific cases / circumstances. No method exists for assessment (Quisbert).

Validation and Meta

As highlighted by Figure 4, very little research exists in the categories of validation and research *about* assessment frameworks. Only one paper, published in two versions, could be seen as addressing model **validation** to some degree. It focuses on DRAMBORA and discusses the outcomes of empirical testing through pilot case studies. However the findings from these studies primarily informed the subsequent design iteration and development of the DRAMBORA interactive tool.

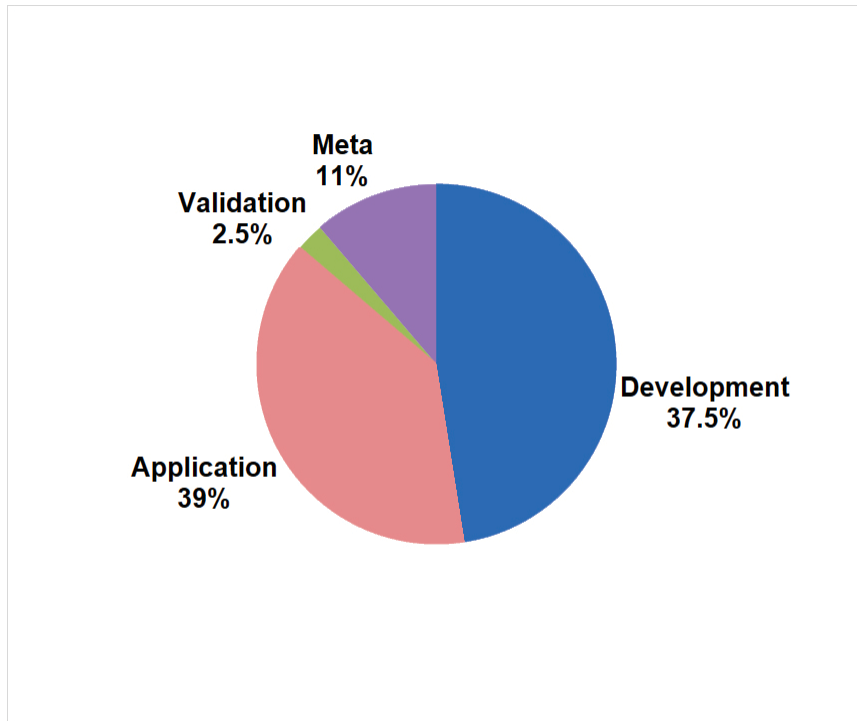


Fig. 4: Publications by research type

Table 5 summarizes for each of the frameworks the level of evidence that is publicly available from applications. It reveals that in particular for improvement-oriented models, there is very limited empirical evidence from assessment cases available for validating the models themselves.

Table 5: Evidence available for each Framework

Framework	Empirical Evidence
Primary Purpose: Initial Planning	
Five Organizational Stages	Examples from Cornell, no further case studies.
NDSA	Content-based case studies
Primary Purpose: Improvement	
AIDA	Multiple applications, but little documentation or evidence available
Brown	None
CTG	Extent of use is unclear, limited evidence.
DPCMM	Model has been applied, but no documentation or evidence available

DRAMBORA	Extensive and well documented pilot assessments and database of completed assessments.
e-ARK	Results of the pilot assessments are available.
NSLA	Only the initial study for which the tool was created.
Scoremodel	Some previous assessments can be seen.
SHAMAN	None
Primary Purpose: Certification	
DSA	Many applications, publicly available documentation, some published case studies
ISO16363	Many applications and case studies
NESTOR	Multiple applications, but limited evidence

Finally, the **Meta** category includes a handful of papers that address broad discussions of the purpose of assessment. This includes the report on the Capability Assessment and Improvement Workshop at iPRES 2013 (C. Becker & Cardoso, 2014), as well as earlier discussions of certification (Dale, 2005; Dobratz & Schoger, 2005).

Figure 5 maps frameworks, sorted by purpose of Certification or Improvement, to the types of research discussed above. It is clear that a large focus in the literature is placed on certification with ISO16363/TRAC, with many references specifically describing its application. Since both TRAC and DRAMBORA were released in 2007, their relative age may account for their wider dissemination, and greater number of references. Besides DRAMBORA, very little work is seen in applications for the other improvement-oriented framework. More strikingly even, not one publication on certification frameworks was considered to conduct a validation of a framework.

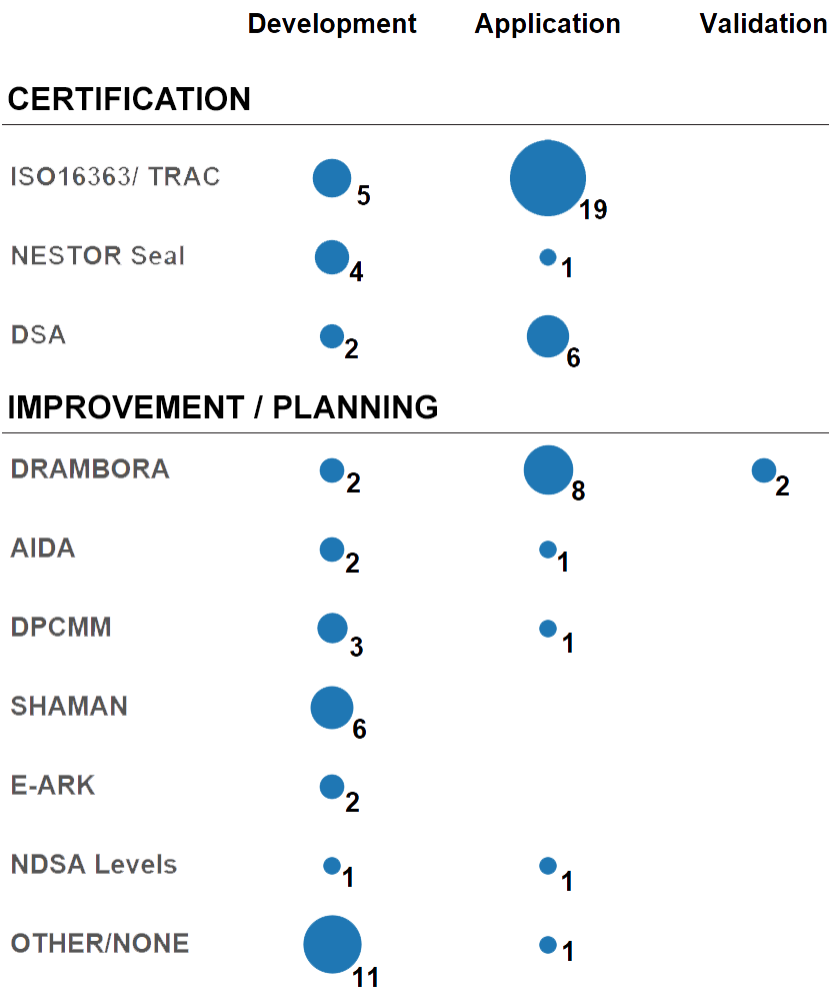


Fig. 5: Model references by type of research

Mapping Summary and Trends

The analysis of the model development and the literature mapping in DP presented above reveals several findings. Assessment frameworks in DP have developed into two broad groupings based on purpose: (1) certification, and (2) improvement or planning. A small set of models is now available for certification, with most work focused around the ISO16363/TRAC standard. A growing number of models have been developed recently to support improvement. Most of the recent work is focused on model development, while research on application of models is growing. There remains comparatively little work on validation of the models, methods of assessment, or reflection on the design and purposes of assessment frameworks in meta-theoretical articles.

The general trends of publications in each of these categories are seen in Figure 6. Work in Development has been increasing steadily, and is followed by increases in publications on Application. In contrast, very little work exists addressing validation or meta-theoretical reflections on the concepts and foundations underlying assessment; these categories have not seen a similar increase over time.

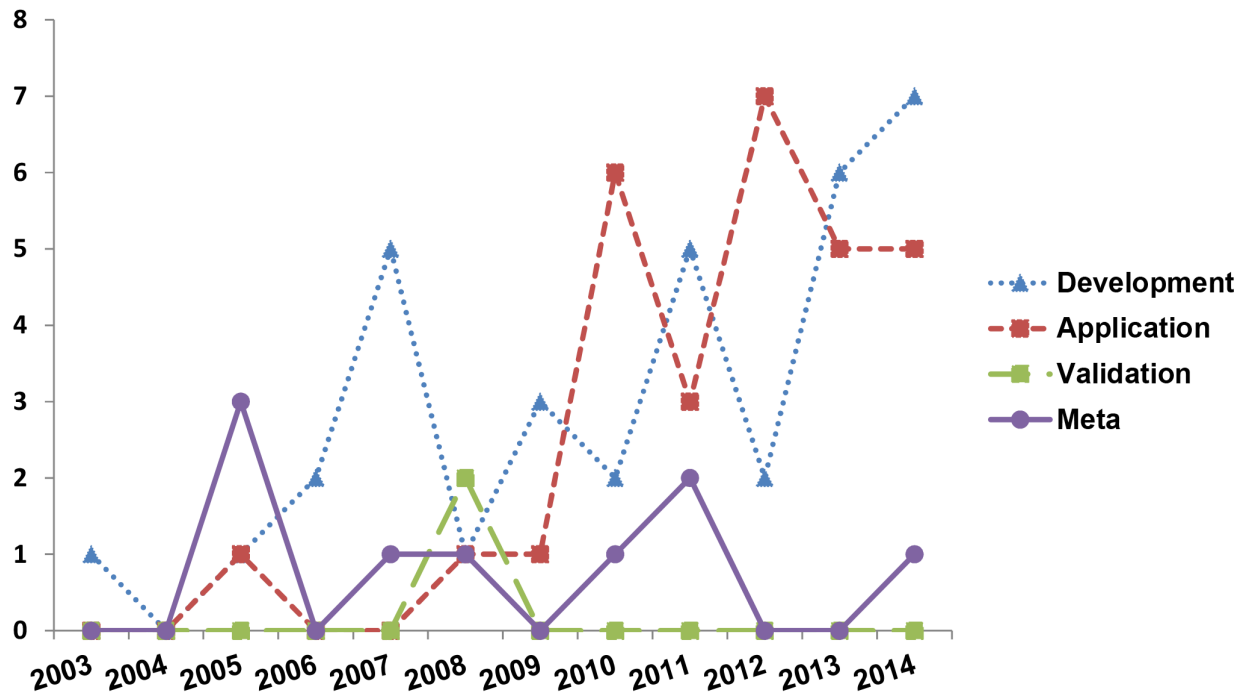


Fig. 6: Research type over time

These findings may also be compared to maturity model research more broadly. Wendler (2012) finds that across domains, the majority of research is focused on development, followed by application, with significantly lower numbers of articles that focus on validation of models or meta-theoretical discussion. However, in our mapping of research in DP, we find that there is a significant gap in validation compared with research in maturity models generally, where 14% of the literature reviewed was found to address validation, compared to the 2.5% (2 of 79) articles in our set (see Table 6).

Table 6: Distribution of Type of Research Compared to Related Fields

	Our mapping (in DP)	Maturity Models generally (Wendler, 2012, Figure 5, p. 1324)
Development	47.5% (38 papers)	46%
Application	39% (31 papers)	35%
Validation	2.5% (2 papers)	14%
Meta	11% (9 papers)	6%

Research gaps and opportunities

Building on this literature mapping, we assess gaps in the current landscape of research on assessment in DP and identify needs and opportunities for further research.

Adoption of foundations. Many publications we reviewed refer to frameworks such as the de-facto standard SEI CMM and suggest they are building on these (e.g. Dobratz, Rödiger, Borghoff, Rätzke, & Schoger, 2010; Dollar & Ashley, 2014c). However, key concepts and terminology from these frameworks seem to be adopted uncritically and sometimes incorrectly. For example, the DPCMM claims to follow the CMM model and support ‘benchmarking’ of organizations. However, its alignment with the structure

of the CMM framework does not extend beyond the high-level terminology and decomposition of a capability into “key process areas”. It does not address the assessment method nor attempt to demonstrate how such an assessment method, and the evidence gathered throughout its application, would address the concerns and criteria outlined in the Appraisal Requirements for the CMM as articulated by the SEI (SCAMPI Upgrade Team, 2011).

Models, tools and methods. Table 7 provides an overview of frameworks grouped according to their primary focus (initial planning, improvement, or certification). For each framework, it indicates whether it contains a model, a method, and a tool; and the level of guidance available for applying the framework.

Table 7: Overview of models and their primary purpose

Model	Method	Tool	Guidance available
Primary Purpose: Initial Planning			
Five Organizational Stages	No	n/a	Limited. Key indicators note high-level processes.
NDSA Levels of Digital Preservation	No	n/a	Set of practical recommendations for use exists
Primary Purpose: Improvement			
CTG	No	Paper Template	Significant guidance for application through workshops.
AIDA	No	Electronic Forms / Online Tools	Limited to instructions for tool; recommendations and feedback provided by project team
DRAMBORA	Yes	Electronic Forms / Online Tools	Guidance documents are available, very detailed results
Brown	No	n/a	Limited. High-level processes identified
SHAMAN	Yes	n/a	
DPCMM	No	Online Tools	Limited to the description of the model
NSLA	No	Paper Template	Some guidance is provided for use of the tool. Results are limited to identifying areas of weakness.
Scoremodel	No	Online Tools	Limited recommendations both for use and in results.
e-ARK	No	n/a	Limited to the description of the model and the pilot assessments.
Primary Purpose: Certification			
DSA	Yes	Online Tools	Guidance documents are available for applying for seal
ISO16363	No	Online Tools	Guidance documents are available for conducting audit
NESTOR	Yes	Electronic Forms	Little guidance beyond addressing documentation to provide for seal

Adapted from (Maemura et al., 2015).

Many current frameworks focus only on the model itself, described in a report targeted at decision makers, while some provide a tool, web- or document-based, to support the capture of information according to the set of criteria. However, few frameworks were found that: specify a concrete method to follow for conducting the assessment; articulate clearly who should conduct this assessment and how; or provide concrete guidance on how to interpret the results, and how these results may be used to develop concrete steps toward improving evaluated organization's abilities. As the literature on process assessment in other fields demonstrates, these are critical components for the successful application of such models in organizational environments: Without systematic and repeatable methods of carrying out an assessment, results are unreliable and cannot be compared across time or between different organizations.

Benchmarking and comparability. There is an understandable desire to 'benchmark' organizations to obtain comparable results across the discipline, and 27 of the 80 publications mention the term *benchmark*. However, attention has focused more on standardizing models to do so, and less on rigorous methods to obtain reliable evidence and evaluate each individual case. While many refer to the influential role of the CMM in providing such an instrument, few point out that the CMM framework comes with a set of very specific requirements for the assessment method applied, and concrete restrictions on what constitutes reliable assessment. The Appraisal Requirements (SCAMPI Upgrade Team, 2011) specify three levels of assessment methods:

1. The lowest level, C, refers to rather light-weight self-assessment methods for internal diagnosis. Methods at this level must not be used to calculate a summative score.
2. The mid-level, B, comes with more specific requirements. For example, it requires the participation of a certified assessor with minimum experience requirements and collection of supporting evidence and documentation. Similarly to level C, organizations that apply methods complying with the requirements at this level are prohibited from reporting a summative score.
3. The usage of any summative evaluation result, including the judgment of maturity level, is restricted to assessment methods complying with the highest level, A. It comes with stringent criteria for the qualifications of the assessors, specific criteria for the evidence captured as part of the assessment process, minimum requirements for the expertise and experience of the assessment team, minimum criteria for the evidence examined during the assessment, and correspondingly high costs (SCAMPI Upgrade Team, 2011). The restrictions have been introduced based on extensive experience of the inherent challenges of reliability in assessment and an appreciation of the dangerous tendency to rely on summary evaluation scores for comparison across organizations - even in cases where scoring is unreliable or potentially biased.

Some of these restrictions can be seen as strategic barriers at least partially designed to open up opportunities for consulting services. Nevertheless, it seems prudent to remain cautious about the feasibility and benefits of comparisons across organizations. For many scenarios in which assessment is sought, such comparison also seems much less relevant than the situated awareness and identification of specific improvements toward organizational goals that can be the focus of improvement-oriented assessment. However, this also means that certification-focused assessment models, such as ISO 16363 and assessment frameworks heavily aligned with it, should not be assumed to be equally effective for an organization's improvement, for the following reasons:

1. They emphasize very particular ways of working in compliance with the OAIS model, which has been criticized as unhelpful (Kulovits, 2013; Rosenthal, 2014; Schultz & Gore, 2010). Even in capability assessment initiatives focused on the CMM, naïve reliance on prescriptive frameworks can be actively harmful (Fayad & Laitnen, 1997).
2. They focus solely on comparability across organizations. This is not necessarily important or beneficial for individual organizations, especially considering the lack of appreciation of organizational culture in the OAIS model (Foscarini & Oliver, 2012).
3. Finally, this focus on certification also brings a corresponding overhead of effort (Elstrøm & Junge, 2014; Houghton, 2015; Rosenthal, 2014).

A critical perspective on design, evaluation and validation: Similar to a phenomenon noted by Wendler, application reports often take the form of “success stories” described by the authors of models. As noted above, the level of scrutiny that assessment frameworks have undergone, and the level of evidence that is available for others to evaluate and scrutinize, is surprisingly low. Similarly, notions of “best practice” are adopted uncritically across the board; while the term features 129 times across 50 of the publications surveyed, it is never itself defined, discussed or reviewed. The amount of evidence available to demonstrate that the practices espoused as “best practice” are consistently showing results superior to other practices is unclear. The question is, then, how valuable an assessment can be for an organization if the foundation of the model used, the clarity of the concepts employed, the suitability of the conceptual structure, and the effectiveness of the assessment method are unclear.

Each of these aspects (design, evaluation and validation) can benefit from the rigorous approach of design science research. Applying design principles (Pöppelbuß & Röglinger, 2011), and guidelines (J. Becker et al., 2009) can ensure that the purpose, audience, terminology, concepts and foundations of a model are clearly defined and documented. Structured design methods (Hevner et al., 2004; Peffers et al., 2007) stipulate rigorous application and validation as a part of the design process. Further, as validation is identified as a clear gap in the mapping presented here, researchers could benefit from the tools and approaches to validation that these frameworks provide. Presently it is unclear to what degree existing frameworks have been designed with these principles and requirements in mind.

Conclusions

Assessment is a crucial concern for managing digital repositories and archives. It is often carried out as input for strategic decisions and resource investment. Certification may be a key motivation for assessment, as an external warrant of trustworthiness. However, assessment for improvement is just as central to the continued improvement of a repository’s ability to preserve. The difference in focus for improvement-oriented assessment means a shift in perspective and raises rather different requirements for the assessment framework.

This review of assessment frameworks in Digital Preservation has cast a wide net to incorporate research from a range of research activities and provide a systematic overview of the landscape of prior research in this area of growing interest to the information professions. We showed that while academic research interest and publication output has been increasing steadily over the past decade, much additional work has been published - primarily for non-academic audiences. The stream of activity has thereby followed a similar pattern as in other fields studied by Wendler, but shows even larger gaps in particular aspects.

Whether assessment is done with the purpose of certification or improvement, organizations applying a framework need to be able to trust the diagnosis structured by the assessment model and delivered through application of the assessment method. In such situations, misleading assurance about the reliability of the diagnosis is as harmful as lack of specific guidance for improvement and/or lack of well-defined assessment methods. The review highlighted gaps in the existing research landscape and argued that the development of assessment models needs strengthening with rigorous, well-founded research that provides robust, well-designed methods and models for assessment and scrutinizes these approaches critically for the benefit of the professional community.

Ongoing work will address the key gaps identified above by applying the abovementioned principles, frameworks and methods to conduct a critical evaluation of selected assessment frameworks in theory and practice. The aim is to provide a perspective that unifies insights into the design process (J. Becker et al., 2009; De Bruin et al., 2005; Hevner et al., 2004; Peffers et al., 2007) and the resulting assessment framework (Pöppelbuß & Röglinger, 2011) with rigorous evaluation of the quality of the product in use through case study research (Yin, 2009). The objective is to provide a replicable design method for assessment frameworks in digital preservation and neighboring domains.

Acknowledgements

Part of this work was supported by the Vienna Science and Technology Fund (WWTF) through the project BenchmarkDP (ICT12-046). The authors thank Jess Whyte for technical contributions to the coding and analysis of mapping results.

Appendix 1: Mapping Results

Table 8: Mapping Results

Development papers
Subcategory: Model Development
Ambacher, B. I. (2007). Government Archives and the Digital Repository Audit Checklist. <i>Journal of Digital Information</i> , 8(2). Retrieved from https://journals.tdl.org/jodi/index.php/jodi/article/view/190
Antunes, G., Barateiro, J., Becker, C., Borbinha, J., Proença, D., & Vieira, R. (2012). <i>SHAMAN Reference Architecture v3.0</i> . SHAMAN (Sustaining Heritage Access through Multivalent Archiving).
Antunes, G., Proença, D., Barateiro, J., Vieira, R., Borbinha, J., & Becker, C. (2012). Assessing Digital Preservation Capabilities Using a Checklist Assessment Method. In <i>Proceedings of the 9th International Conference on Preservation of Digital Objects</i> . Toronto, ON, Canada. Retrieved from http://www.isis.tuwien.ac.at/node/16203
Becker, C., Antunes, G., Barateiro, J., & Vieira, R. (2011). A capability model for digital preservation. In <i>Proceedings of the 8th International Conference on Preservation of Digital Objects</i> (pp. 01–04). Singapore.
Becker, C., Antunes, G., Barateiro, J., Vieira, R., & Borbinha, J. (2011a). Control objectives for DP: Digital preservation as an integrated part of it governance. <i>Proceedings of the American Society for Information Science and Technology</i> , 48(1), 1–10.
Brown, A. (2013). <i>Practical digital preservation: A how-to guide for organizations of any size</i> . London: Facet Publishing.
Colet, L., & Renault, A. (2013). Process Assessment for Digital Records Preservation and Information Security. Presented at the <i>Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)</i> , Lisbon, Portugal. Retrieved from http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013colet.pdf
Consultative Committee for Space Data Systems (CCSDS). (2011). <i>Audit and Certification of Trustworthy Digital Repositories, Magenta Book</i> .
Coufal, L., & Pearson, D. (2013). <i>Digital Preservation Environment Matrix</i> . National & State Libraries Australasia. Retrieved from http://www.nsla.org.au/publication/digital-preservation-environment-maturity-matrix
Cresswell, A. M., Pardo, T. A., & Canestraro, D. S. (2006). Digital capability assessment for government: a multi-dimensional approach. In <i>Electronic Government</i> (pp. 293–304). Springer.
Dollar, C., & Ashley, L. (2014c). <i>Digital Preservation Capability Maturity Model (DPCMM) Background and Performance Metrics Version 2.6</i> .
Gillesse, R., & Vanstappen, H. (2014). The Case For Yet Another Digital Preservation Evaluation Tool. In <i>Archiving 2014 Final Program and Proceedings</i> (Vol. 2014, pp. 25–27). Berlin, Germany: Society for Imaging Science and Technology.
Kenney, A. R., & McGovern, N. Y. (2003). The five organizational stages of digital preservation. <i>Digital Libraries: A Vision for the 21st Century. A Festschrift for Wendy Pratt Lougee on the Occasion of Her Departure from the University of Michigan</i> .
McHugh, A., Ross, S., Ruusalepp, R., & Hofman, H. (2007). <i>The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)</i> . Digital Curation Centre. Retrieved from http://eprints.gla.ac.uk/33617/

NESTOR Working Group Trusted Repositories - Certification. (2006). *NESTOR Catalogue of Criteria for Trusted Digital Repositories*.

NESTOR Working Group Trusted Repositories - Certification. (2013). *Explanatory notes on the NESTOR Seal for Trustworthy Digital Archives*.

OCLC Online Computer Library Center, The Center for Research Libraries (CRL). (2007). *Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist*. OCLC Online Computer Library Center, The Center for Research Libraries (CRL). Retrieved from http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

Pardo, T., Cresswell, A. M., Dawes, S. S., Burke, B., Dadayan, L., Embar, S., & Kwon, H. (2005). *Building State Government Digital Preservation Partnerships: A Capability Assessment and Planning Toolkit, Version 1.0*. Center for Technology in Government, University at Albany, SUNY.

Phillips, M., Bailey, J., Goethals, A., & Owens, T. (2013). The NDSA Levels of Digital Preservation: An Explanation and Uses. In *IS&T Archiving 2013 Final Program and Proceedings*, Washington, USA.

Proença, D., Vieira, R., & Borbinha, J. (2015a). *D7.1 E-ARK - A Maturity Model for Information Governance - Initial Version* (Project Deliverable No. D7.1). Retrieved from <http://www.eark-project.com/resources/project-deliverables/19-d71-e-ark-a-maturity-model-for-information-governance-initial-version>

Proença, D., Vieira, R., & Borbinha, J. (2015b). *D7.2 Archiving Maturity Model - Initial Assessment and Evaluation* (Project Deliverable No. D7.2). Retrieved from <http://www.eark-project.com/resources/project-deliverables/46-d72initassess>

Slade, S., Pearson, D., & Coufal, L. (2014). A Digital Preservation Environment Maturity Matrix for NSLA Libraries. In *Proceedings of the 11th International Conference on Preservation of Digital Objects*. Melbourne, Australia. Retrieved from https://www.nla.gov.au/sites/default/files/pages_from_ipres2014-proceedings-slade_pearson_coufal.pdf

University of London Computer Centre. (2009a). *Assessing Institutional Digital Assets (AIDA): A JISC Final Report*.

University of London Computer Centre. (2009b). *The AIDA self-assessment toolkit Mark II*. Retrieved from <http://aida.jiscinvolve.org/wp/toolkit/>

Vieira, R., Cardoso, E., & Becker, C. (2014). A traceable maturity assessment method based on Enterprise Architecture modelling. In *Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International* (pp. 245–253). IEEE.

Subcategory: Proto-Development

Antunes, G., Barateiro, J., Becker, C., Borbinha, J., & Vieira, R. (2011). Modeling Contextual Concerns in Enterprise Architecture. In *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2011 15th IEEE International* (pp. 3–10). <http://doi.org/10.1109/EDOCW.2011.9>

Becker, C., Antunes, G., Barateiro, J., Vieira, R., & Borbinha, J. (2011b). Modeling Digital Preservation Capabilities in Enterprise Architecture. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 84–93). New York, NY, USA: ACM. <http://doi.org/10.1145/2037556.2037570>

Dobratz, S., Rödig, P., Borghoff, U. M., Rätzke, B., & Schoger, A. (2010). The use of quality management standards in trustworthy digital archives. *International Journal of Digital Curation*,

5(1), 46–63.

Kulovits, H. (2013). A perspective of the Austrian State Archives on the systematic assessment and improvement of Digital Preservation activities, processes and systems. Presented at the *Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*, Lisbon, Portugal. Retrieved from <http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013kulovits.pdf>

Proença, D. (2013). A systematic approach for the assessment of digital preservation activities. Presented at the *Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*, Lisbon, Portugal. Retrieved from <http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013proenca.pdf>

Subcategory: Model Presentation

Dillo, I., & de Leeuw, L. (2015). Ten Years Back, Five Years Forward: The Data Seal of Approval. *International Journal of Digital Curation*, 10(1), 230–239.

Dobratz, S., & Schoger, A. (2007). Trustworthy Digital Long-Term Repositories: The NESTOR Approach in the Context of International Developments. In L. Kovács, N. Fuhr, & C. Meghini (Eds.), *Research and Advanced Technology for Digital Libraries* (pp. 210–222). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-74851-9_18

Dollar, C., & Ashley, L. (2014a). *Assessing Digital Preservation Capability Using a Maturity Model Process Improvement Approach*. Retrieved from http://www.nycarchivists.org/Resources/Documents/Dollar-Ashley_2013_DPCMM%20White%20Paper_NAGARA%20Digital%20Judicial%20Records_8Feb2013-1.pdf

Dollar, C., & Ashley, L. (2014b). Building the Business Case Using the Digital Preservation Capability Maturity Model. Presented at the *International Council on Archives Conference*, Girona, Spain. Retrieved from <http://www.securelyrooted.com/dpcmm/>

Giaretta, D. (2010). How does one know which repository is worth its salt? *ERCIM News*, 80. Retrieved from <http://ercim-news.ercim.eu/en80/special/which-repositories-are-worth-their-salt>

Giaretta, D., & Lambert, S. (2009). International Audit and Certification of Digital Repositories. Presented at the *PV 2009 Conference*, Madrid, Spain: European Space Astronomy Centre (ESAC). Retrieved from http://www.sciops.esa.int/SYS/CONFERENCE/include/pv2009/papers/35_Giaretta_InternationalAuditAndCertificationOfDigitalRepositories.pdf

Harmsen, H. (2008). Data Seal of Approval, assessment and reviewing of research data repositories. In *Proceedings of the Fifth International Conference on Preservation of Digital Objects*. London, UK: British Library.

McHugh, A., Innocenti, P., Ross, S., & Ruusalepp, R. (2007). Risk Management foundations for digital libraries: DRAMBORA (Digital Repository Audit Method Based on Risk Assessment). In *DELOS 2nd Workshop on Foundations of Digital Libraries*. Budapest, Hungary.

Application Papers

Beedham, H., Missen, J., Palmer, M., & Ruusalepp, R. (2005). *Assessment of the UKDA and TNA compliance with OAIS and NETS standards*. UK: UK Data Archive, The National Archives.

Bote, J., Termens, M., & Gelabert, G. (2011). Evaluation of Healthcare Institutions for Long-Term Preservation of Electronic Health Records. In M. M. Cruz-Cunha, J. Varajão, P. Powell, & R. Martinho (Eds.), *ENTERprise Information Systems* (pp. 136–145). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-24352-3_15

- Centre for Research Libraries. (2010). *Report on Portico Audit*. Retrieved from <http://www.crl.edu/sites/default/files/reports/CRL%20Report%20on%20Portico%20Audit%202010.pdf>
- Centre for Research Libraries. (2011). *Report on Hathitrust Audit*. Retrieved from <http://www.crl.edu/sites/default/files/reports/CRL%20HathiTrust%202011.pdf>
- Centre for Research Libraries. (2012). *Report on Chronopolis Audit*. Retrieved from http://www.crl.edu/sites/default/files/reports/Chron_Report_2012_final_0.pdf
- Centre for Research Libraries. (2013). *Report on Scholars Portal Audit*. Retrieved from <http://www.crl.edu/reports/scholars-portal-audit-report-2013>
- Centre for Research Libraries. (2014). *Report on CLOCKSS Audit*. Retrieved from http://www.crl.edu/sites/default/files/reports/CLOCKSS_Report_2014.pdf
- Centre for Research Libraries. (2015). *Report on the Canadiana.org Digital Repository*. Retrieved from <http://www.crl.edu/reports/certification-report-canadianaorg-digital-repository>
- Council of State Archivists. (2012). *SERI Phase One Report: Report from the Council of State Archivists' State Electronic Records Initiative (SERI) Committee*. Council of State Archivists. Retrieved from http://www.statearchivists.org/seri/phase_one_report.htm
- Downs, R. R., & Chen, R. S. (2010). Self-Assessment of a Long-Term Archive for Interdisciplinary Scientific Data as a Trustworthy Digital Repository. *JoDI: Journal of Digital Information*, 11(1), 4.
- Elstrøm, G. V., & Junge, J. G. (2014). Self-assessment of the Digital Repository at the State and University Library, Denmark-a Case Study. In *Proceedings of the 11th International Conference on Digital Preservation*. Melbourne, Australia.
- Farrington, K., Fischer, M., Hanus, R., & Wopschall, M. (2013). *Assessing the iSchool Digital Repository and Digital Archaeology Lab*. Retrieved from <https://ford.ischool.utexas.edu/xmlui/handle/2081/30582?show=full>
- Ferreira, M., Faria, L., Hahn, M., & Duretec, K. (2014). *SCAPE: report on compliance validation*. SCAPE Project. Retrieved from <http://repositorium.sdum.uminho.pt/handle/1822/30689?locale=en>
- Giaretta, D., Schrimpf, S., L'Hours, H., Dillo, I., Doorn, P., Rouchon, O., ... Salza, S. (2012). *Report on Peer Review of Digital Repositories (Part B of D33.1)* (No. D33.1 (Part B)). Alliance for Permanent Access to the Records of Science Network. Retrieved from http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D33_1B-01-1_0.pdf
- Goethals, A. (2013). An Example Self-Assessment Using the NDSA Levels of Digital Preservation. Presented at the *Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*, Lisbon, Portugal. Retrieved from <http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013goethals.pdf>
- Houghton, B. (2015). Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. *D-Lib Magazine*, 21(3), 5.
- Innocenti, P., & Vullo, G. (2009). Assessing the preservation of institutional repositories with DRAMBORA: case studies from the University of Glasgow. *Bollettino AIB*, 49(2), 139–156.
- Karlach, P. (2010). An Audit of the National Repository of Grey Literature using the DRAMBORA tool. In *Grey literature repositories* (p. 126). Prague, Czech Republic: National Technical Library,

VeRBuM publishing house.

- Kirchhoff, A., Fenton, E., Orphan, S., & Morrissey, S. (2010). Becoming a certified trustworthy digital repository: The Portico experience. In *Proceedings of the 7th International Conference on Preservation of Digital Objects*. Vienna, Austria.
- Massol, M., Rouchon, O., & Bechard, L. (2013). Certification and Quality: A French Experience. In *Proceedings of the Tenth International Conference on Preservation of Digital Objects (iPres), 2013, Lisbon, Portugal*. Lisbon, Portugal. Retrieved from https://www.openaire.eu/search/publication?articleId=od_____1726::98dfb62f9e9932c5b7cc86cb91a31676
- Miller, L., Blake, M., & Sorsby, M. (2012). Evaluation of the Assessing Institutional Digital Assets (AIDA) Toolkit. *Science & Technology Libraries*, 31(1), 92–99.
- Mitcham, J., & Hardman, C. (2011). *ADS and the data seal of approval-a case study for the DCC*. DCC. Retrieved from <http://www.dcc.ac.uk/resources/case-studies/ads-dsa>
- National Technical Library. (2010). *Audit of the National Repository of Grey Literature (NRGL) in the NTK using the DRAMBORA tool, Second Audit*. Czech Republic. Retrieved from http://nrgl.techlib.cz/images/DRAMBORA_2010_EN.pdf
- Quisbert, H. (2008). Evaluation of a Digital Repository. In *IS&T Archiving 2008 Final Program and Proceedings* (Vol. 2008, pp. 120–124). Bern, Switzerland: Society for Imaging Science and Technology.
- Rodrigues, E., Ferreira, M., Carvalho, J., Príncipe, P., Faria, L., Silva, H., & Moreira, J. M. (2014). Large scale repository auditing to iso 16363. *Open Repositories 2014*, 1–4.
- Schultz, M., & Gore, E. B. (2010). The Importance of Trust in Distributed Digital Preservation: A Case Study from the Metaarchive Cooperative. In *Proceedings of the 7th International Conference on Preservation of Digital Objects*. Vienna, Austria.
- Schumann, N. (2012). Tried and Trusted. *IASSIST Quarterly*, 36.
- Schumann, N., & Mauer, R. (2013). The GESIS data archive for the social sciences: A widely recognised data archive on its way. *International Journal of Digital Curation*, 8(2), 215–222.
- Termens, M., & Locher, A. E. (2012). Digital preservation audit on spatial data: a practical experience. Presented at the *International Symposium on Service-Oriented Mapping (SoMaP 2012)*, Vienna, Austria.
- Vardigan, M., & Lyle, J. (2014). The Inter-university Consortium for Political and Social Research and the Data Seal of Approval: Accreditation Experiences, Challenges, and Opportunities. *Data Science Journal*, 13(0), PDA83–PDA87.
- Vaska, M., & Pejsova, P. (2012). Audit DRAMBORA for Trustworthy Repositories: A Study Dealing with the Digital Repository of Grey Literature. *The Grey Literature Journal*, 8(2), 96–105.

Validation Papers

- McHugh, A., Ross, S., Innocenti, P., Ruusalepp, R., & Hofman, H. (2008a). Bringing self assessment home: Repository profiling and key lines of enquiry within DRAMBORA. In *Archiving 2008 Final Program and Proceedings* (Vol. 2008, pp. 13–19). Bern, Switzerland: Society for Imaging Science and Technology.
- McHugh, A., Ross, S., Innocenti, P., Ruusalepp, R., & Hofman, H. (2008b). Bringing Self-assessment Home: Repository Profiling and Key Lines of Enquiry within DRAMBORA. *International Journal of Digital Curation*, 3(2), 130–142. <http://doi.org/10.2218/ijdc.v3i2.64>

Meta Papers

- Becker, C., & Cardoso, E. (2014). Report on the Capability Assessment and Improvement Workshop (CAIW) at iPres 2013. *D-Lib Magazine*, 20(3/4). <http://doi.org/doi:10.1045/march2014-becker>
- Dale, R., & Gore, E. (2010). Process Models and the Development of Trustworthy Digital Repositories. *Information Standards Quarterly*, 22(2), 14. <http://doi.org/10.3789/isqv22n2.2010.02>
- Dale, R. L. (2005). Making Certification Real: Developing Methodology for Evaluating Repository Trustworthiness. *RLG DigiNews*, 9(5). Retrieved from <http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file3047.html>
- Dobratz, S., & Schoger, A. (2005). Digital Repository Certification: A Report from Germany. *RLG DigiNews*, 9(5). Retrieved from <http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file3047.html>
- Dobratz, S., Schoger, A., Rätzke, B., Borghoff, U. M., & Rödiger, P. (2008). The Use of Quality Management Standards in Trustworthy Digital Archives. In *Proceedings of the Fifth International Conference on Preservation of Digital Objects*. London, UK.
- Giaretta, D. (2011). Who Is Doing a Good Job? Audit and Certification. In *Advanced Digital Preservation* (pp. 461–480). Springer.
- Hank, C., Tibbo, H. R., & Barnes, H. (2007). Building from Trust: Using the RLG/NARA Audit Checklist for Institutional Repository Planning and Deployment. In *Archiving 2007 Final Program and Proceedings* (Vol. 2007, pp. 62–66). San Antonio, Texas, USA.
- Klump, J. (2011). Criteria for the trustworthiness of data centres. *D-Lib Magazine*, 17(1), 6.
- Ross, S., & McHugh, A. (2005). Audit and Certification of Digital Repositories: Creating a Mandate for the Digital Curation Centre (DCC). *RLG DigiNews*, 9(5). Retrieved from <http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file3047.html>

References

- Antunes, G., Barateiro, J., Becker, C., Borbinha, J., Proença, D., & Vieira, R. (2012). *SHAMAN Reference Architecture v3.0*. SHAMAN (Sustaining Heritage Access through Multivalent Archiving).
- Antunes, G., Barateiro, J., Becker, C., Borbinha, J., & Vieira, R. (2011). Modeling Contextual Concerns in Enterprise Architecture. In *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2011 15th IEEE International* (pp. 3–10). <http://doi.org/10.1109/EDOCW.2011.9>
- Ashley, L. (2014, August 19). *Free Public DPC Self-Assessment Goes Live*. Retrieved from <http://www.securelyrooted.com/blog/2014/8/19/dpc-self-assessment-goes-live>
- Atkins, W., Goethals, A., Kussmann, C., Phillips, M., & Vardigan, M. (2013). *Staffing for Effective Digital Preservation: An NDSA Report*. National Digital Stewardship Alliance. Retrieved from <http://hdl.loc.gov/loc.gdc/lcpub.2013655113.1>
- Becker, C., & Cardoso, E. (2014). Report on the Capability Assessment and Improvement Workshop (CAIW) at iPres 2013. *D-Lib Magazine*, 20(3/4). <http://doi.org/doi:10.1045/march2014-becker>
- Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management - A Procedure Model and its Application. *Business & Information Systems Engineering*, 1(3), 213–222.
- Bote, J., Termens, M., & Gelabert, G. (2011). Evaluation of Healthcare Institutions for Long-Term Preservation of Electronic Health Records. In M. M. Cruz-Cunha, J. Varajão, P. Powell, & R. Martinho (Eds.), *ENTERprise Information Systems* (pp. 136–145). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-24352-3_15
- Brown, A. (2013). *Practical digital preservation: A how-to guide for organizations of any size*. London: Facet Publishing.
- Centre for Research Libraries. (2010). *Report on Portico Audit*. Retrieved from <http://www.crl.edu/sites/default/files/reports/CRL%20Report%20on%20Portico%20Audit%202010.pdf>
- Centre for Research Libraries. (2011). *Report on Hathitrust Audit*. Retrieved from

- <http://www.crl.edu/sites/default/files/reports/CRL%20HathiTrust%202011.pdf>
- Centre for Research Libraries. (2012). *Report on Chronopolis Audit*. Retrieved from http://www.crl.edu/sites/default/files/reports/Chron_Report_2012_final_0.pdf
- Centre for Research Libraries. (2013). *Report on Scholars Portal Audit*. Retrieved from <http://www.crl.edu/reports/scholars-portal-audit-report-2013>
- Centre for Research Libraries. (2014). *Report on CLOCKSS Audit*. Retrieved from http://www.crl.edu/sites/default/files/reports/CLOCKSS_Report_2014.pdf
- Centre for Research Libraries. (2015). *Report on the Canadiana.org Digital Repository*. Retrieved from <http://www.crl.edu/reports/certification-report-canadianaorg-digital-repository>
- Council of State Archivists. (2012). *SERI Phase One Report: Report from the Council of State Archivists' State Electronic Records Initiative (SERI) Committee*. Council of State Archivists. Retrieved from http://www.statearchivists.org/seri/phase_one_report.htm
- Cresswell, A. M., Pardo, T. A., & Canestraro, D. S. (2006). Digital capability assessment for e-government: a multi-dimensional approach. In *Electronic Government* (pp. 293–304). Springer.
- Crosby, P. B. (1979). *Quality is free : the art of making quality certain*. New York: McGraw-Hill.
- Curley, M., Kenneally, J., & Carcary, M. (2015). *IT Capability Maturity Framework (IT-CMF): The Body of Knowledge Guide*. Van Haren Publishing.
- Dale, R. L. (2005). Making Certification Real: Developing Methodology for Evaluating Repository Trustworthiness. *RLG DigiNews*, 9(5). Retrieved from <http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file3047.html>
- De Bruin, T., Freeze, R., Kaulkarni, U., & Rosemann, M. (2005). Understanding the Main Phases of Developing a Maturity Assessment Model. In B. Campbell, J. Underwood, & D. Bunker (Eds.), *Australasian Conference on Information Systems (ACIS)*. Sydney, Australia. Retrieved from <http://eprints.qut.edu.au/25152/>
- DigitalOK. (2014). Retrieved March 7, 2016, from <http://digitalok.org/Login.aspx>

- Dobratz, S., Rödiger, P., Borghoff, U. M., Rätzke, B., & Schoger, A. (2010). The use of quality management standards in trustworthy digital archives. *International Journal of Digital Curation*, 5(1), 46–63.
- Dobratz, S., & Schoger, A. (2005). Digital Repository Certification: A Report from Germany. *RLG DigiNews*, 9(5). Retrieved from <http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file3047.html>
- Dollar, C., & Ashley, L. (2014c). *Digital Preservation Capability Maturity Model (DPCMM) Background and Performance Metrics Version 2.6*.
- Downs, R. R., & Chen, R. S. (2010). Self-Assessment of a Long-Term Archive for Interdisciplinary Scientific Data as a Trustworthy Digital Repository. *JoDI: Journal of Digital Information*, 11(1), 4.
- Elström, G. V., & Junge, J. G. (2014). Self-assessment of the Digital Repository at the State and University Library, Denmark—a Case Study. In *Proceedings of the 11th International Conference on Digital Preservation*. Melbourne, Australia.
- Farrington, K., Fischer, M., Hanus, R., & Wopschall, M. (2013). *Assessing the iSchool Digital Repository and Digital Archaeology Lab*. Retrieved from <https://ford.ischool.utexas.edu/xmlui/handle/2081/30582?show=full>
- Fayad, M. E., & Laitinen, M. (1997). Process Assessment Considered Wasteful. *Communications of the ACM*, 40(11), 125–128. <http://doi.org/10.1145/265684.265701>
- Ferreira, M., Faria, L., Hahn, M., & Duretec, K. (2014). *SCAPE: report on compliance validation*. SCAPE Project. Retrieved from <http://repositorium.sdum.uminho.pt/handle/1822/30689?locale=en>
- Foscarini, F., & Oliver, G. (2012). The Information Culture Challenge: Moving Beyond OAIS. Presented at the *Open Research Challenges (ORC) Workshop at International Conference on Preservation of Digital Objects (iPRES 2012)*, Toronto, ON, Canada. Retrieved from <https://digitalpreservationchallenges.files.wordpress.com/2012/09/foscarini.pdf>

- Garrett, J., & Waters, D. (1996). *Preserving Digital Information. Report of the Task Force on Archiving of Digital Information (No. Pub63)*. Washington, DC: Research Libraries Group. Retrieved from <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>
- Giaretta, D., Harmsen, H., & Keitel, C. (2010, July 8). *Memorandum of Understanding to create a European Framework for Audit and Certification of Digital Repositories*. Data Seal of Approval, CCSDS Repository Audit and Certification Working Group, DIN Working Group “Trustworthy Archives – Certification.” Retrieved from <http://www.trusteddigitalrepository.eu/Memorandum%20of%20Understanding.html>
- Giaretta, D., Schrimpf, S., L’Hours, H., Dillo, I., Doorn, P., Rouchon, O., ... Salza, S. (2012). *Report on Peer Review of Digital Repositories (Part B of D33.1)* (No. D33.1 (Part B)). Alliance for Permanent Access to the Records of Science Network (APARSEN). Retrieved from http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D33_1B-01-1_0.pdf
- Gillesse, R., & Vanstappen, H. (2014). The Case For Yet Another Digital Preservation Evaluation Tool. In *Archiving 2014 Final Program and Proceedings* (Vol. 2014, pp. 25–27). Berlin, Germany: Society for Imaging Science and Technology.
- Goethals, A. (2013). An Example Self-Assessment Using the NDSA Levels of Digital Preservation. Presented at the *Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*, Lisbon, Portugal. Retrieved from <http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013goethals.pdf>
- Harmsen, H. (2008). Data Seal of Approval, assessment and reviewing of research data repositories. In *Proceedings of the Fifth International Conference on Preservation of Digital Objects*. London, UK: British Library.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105.

- Houghton, B. (2015). Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. *D-Lib Magazine*, 21(3), 5.
- Internal audit tool - Archivemata. (n.d.). Retrieved February 21, 2016, from https://wiki.archivemata.org/Internal_audit_tool
- International Organization for Standardization (ISO). (2012). *ISO 14721:2012 Space data and information transfer systems - Open archival information system (OAIS) - Reference model*. International Organization for Standardization. Retrieved from http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284
- International Organization for Standardization (ISO). (2014). *ISO 16919:2014 Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories*. International Organization for Standardization. Retrieved from http://www.iso.org/iso/catalogue_detail.htm?csnumber=57950
- IT Governance Institute. (2007). *Control Objectives for Information and Related Technology (CobIT)*, 4.1th Edition. IT Governance Institute. Retrieved from <http://www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/COBIT-4-1.aspx>
- Jackson, A. (n.d.). *What are the benefits of certifying a repository as a "Trusted Digital Repository" (ISO 16363)*. Retrieved March 1, 2016, from <http://anjackson.net/zombse/032013%20Digital%20Preservation/static/questions/171.html>
- Jokela, T., Siponen, M., Hirasawa, N., & Earthy, J. (2006). A survey of usability capability maturity models: implications for practice and research. *Behaviour & Information Technology*, 25(3), 263–282. <http://doi.org/10.1080/01449290500168079>
- Karlach, P. (2010). An Audit of the National Repository of Grey Literature using the DRAMBORA tool. In *Grey literature repositories* (p. 126). Prague, Czech Republic: National Technical Library, VeRBuM publishing house.
- Kenney, A. R., & McGovern, N. Y. (2003). The five organizational stages of digital preservation. *Digital Libraries: A Vision for the 21st Century. A Festschrift for Wendy Pratt Lougee on the Occasion*

of Her Departure from the University of Michigan.

Kirchhoff, A., Fenton, E., Orphan, S., & Morrissey, S. (2010). Becoming a certified trustworthy digital repository: The Portico experience. In *Proceedings of the 7th International Conference on Preservation of Digital Objects*. Vienna, Austria.

Kulovits, H. (2013). A perspective of the Austrian State Archives on the systematic assessment and improvement of Digital Preservation activities, processes and systems. Presented at the *Capability Assessment and Improvement Workshop (CAIW) at the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*, Lisbon, Portugal. Retrieved from <http://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013kulovits.pdf>

Maemura, E., Moles, N., & Becker, C. (2015). A Survey of Organizational Assessment Frameworks in Digital Preservation. Presented at the *12th International Conference on Digital Preservation (iPRES)*, Chapel Hill, NC, USA.

Maier, A. M., Moultrie, J., & Clarkson, P. J. (2012). Assessing Organizational Capabilities: Reviewing and Guiding the Development of Maturity Grids. *IEEE Transactions on Engineering Management*, 59(1), 138–159.

Massol, M., Rouchon, O., & Bechard, L. (2013). Certification and Quality: A French Experience. In *Proceedings of the Tenth International Conference on Preservation of Digital Objects (iPres)*, 2013, Lisbon, Portugal. Lisbon, Portugal. Retrieved from https://www.openaire.eu/search/publication?articleId=od_____1726::98dfb62f9e9932c5b7cc86cb91a31676

McHugh, A., Ross, S., Ruusalepp, R., & Hofman, H. (2007). *The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)*. Digital Curation Centre. Retrieved from <http://eprints.gla.ac.uk/33617/>

McKinney, P., Benson, L., & Knight, S. (2012). From Hobbyist to Industrialist. Challenging the DP Community. Presented at the *Open Research Challenges (ORC) Workshop at International Conference on Preservation of Digital Objects (iPRES 2012)*, Toronto, ON, Canada. Retrieved

from <https://digitalpreservationchallenges.wordpress.com/2012/09/20/draft-workshop-agenda/>

Mettler, T., & Rohner, P. (2009). Situational Maturity Models as Instrumental Artifacts for Organizational Design. In *DESRIST '09*. Malvern, PA, USA.

Miller, L., Blake, M., & Sorsby, M. (2012). Evaluation of the Assessing Institutional Digital Assets (AIDA) Toolkit. *Science & Technology Libraries*, 31(1), 92–99.

Mitcham, J., & Hardman, C. (2011). *ADS and the data seal of approval-a case study for the DCC*. DCC. Retrieved from <http://www.dcc.ac.uk/resources/case-studies/ads-dsa>

National Technical Library. (2010). *Audit of the National Repository of Grey Literature (NRGL) in the NTK using the DRAMBORA tool, Second Audit*. Czech Republic. Retrieved from http://nrgl.techlib.cz/images/DRAMBORA_2010_EN.pdf

NESTOR Working Group Trusted Repositories - Certification. (2006). *NESTOR Catalogue of Criteria for Trusted Digital Repositories*.

NESTOR Working Group Trusted Repositories - Certification. (2013). *Explanatory notes on the NESTOR Seal for Trustworthy Digital Archives*.

Object Management Group (OMG). (2008, June). *Business Process Maturity Model (BPMM), Version 1.0*. Object Management Group (OMG). Retrieved from <http://www.omg.org/spec/BPMM/1.0/PDF>

Online Computer Library Center (OCLC), Center for Research Libraries (CRL). (2007). *Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist*. Retrieved from http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

Oxford English Dictionary. (n.d.). Retrieved April 20, 2015, from <http://www.oed.com/>

Pardo, T., Cresswell, A. M., Dawes, S. S., Burke, B., Dadayan, L., Embar, S., & Kwon, H. (2005). *Building State Government Digital Preservation Partnerships: A Capability Assessment and Planning Toolkit, Version 1.0*. Center for Technology in Government, University at Albany, SUNY.

- Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). *Capability Maturity Model for Software, Version 1.1* (Technical Report No. CMU/SEI-93-TR-024). Software Engineering Institute, Carnegie Mellon University.
- Pearson, D., & Coufal, L. (2013). *Digital Preservation Environment Matrix. National & State Libraries Australasia*. Retrieved from <http://www.nsla.org.au/publication/digital-preservation-environment-maturity-matrix>
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <http://doi.org/10.2753/MIS0742-1222240302>
- Pelz-Sharpe, A., Durga, A., Smigiel, D., Hartman, E., & Byrne, T. (2009, February 19). *ECM3: ECM Maturity Model, Version 1.0*. Wipro Technologies, CMS Watch, Smigiel Consulting Group, Hartman Communicatie. Retrieved from https://ecmmaturity.files.wordpress.com/2009/02/ec3m-v01_0.pdf
- Phillips, M., Bailey, J., Goethals, A., & Owens, T. (2013). The NDSA Levels of Digital Preservation: An Explanation and Uses. In *IS&T Archiving 2013 Final Program and Proceedings*, Washington, USA: Society for Imaging Science and Technology.
- Pöppelbuß, J., & Röglinger, M. (2011). What makes a useful maturity model? a framework of general design principles for maturity models and its demonstration in business process management. In *ECIS 2011 Proceedings*.
- Proença, D., Vieira, R., & Borbinha, J. (2015a). *D7.1 E-ARK - A Maturity Model for Information Governance - Initial Version* (Project Deliverable No. D7.1). E-ARK (European Archival Records and Knowledge Preservation). Retrieved from <http://www.eark-project.com/resources/project-deliverables/19-d71-e-ark-a-maturity-model-for-information-governance-initial-version>
- Proença, D., Vieira, R., & Borbinha, J. (2015b). *D7.2 Archiving Maturity Model - Initial Assessment and Evaluation* (Project Deliverable No. D7.2). E-ARK (European Archival Records and Knowledge

- Preservation). Retrieved from <http://www.eark-project.com/resources/project-deliverables/46-d72initassess>
- Quisbert, H. (2008). Evaluation of a Digital Repository. In *IS&T Archiving 2008 Final Program and Proceedings* (Vol. 2008, pp. 120–124). Bern, Switzerland: Society for Imaging Science and Technology.
- Research Libraries Group. (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. Mountain View, California, USA.
- Rodrigues, E., Ferreira, M., Carvalho, J., Príncipe, P., Faria, L., Silva, H., & Moreira, J. M. (2014). Large scale repository auditing to iso 16363. *Open Repositories 2014*, 1–4.
- Rosemann, M., De Bruin, T., & Power, B. (2006). A model to measure business process management maturity and improve performance. In J. Jeston & J. Nelis (Eds.), *Business Process Management*. Butterworth-Heinemann.
- Rosenthal, D. S. H. (2014, August 12). *TRAC Audit: Lessons* [Personal Blog]. Retrieved from <http://blog.dshr.org/2014/08/trac-audit-lessons.html>
- SCAMPI Upgrade Team. (2011). *Appraisal Requirements for CMMI Version 1.3 (ARC, V1.3)* (Technical Report No. CMU/SEI-2011-TR-006). Software Engineering Institute, Carnegie Mellon University. Retrieved from <http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=9959>
- Schultz, M., & Gore, E. B. (2010). The Importance of Trust in Distributed Digital Preservation: A Case Study from the Metaarchive Cooperative. In *Proceedings of the 7th International Conference on Preservation of Digital Objects*. Vienna, Austria.
- Schumann, N. (2012). Tried and Trusted. *IASSIST Quarterly*, 36, 23-27.
- Schumann, N., & Mauer, R. (2013). The GESIS data archive for the social sciences: A widely recognised data archive on its way. *International Journal of Digital Curation*, 8(2), 215–222.
- Scoremodel. (2013). Retrieved March 7, 2016, from <http://scoremodel.org/en/home>
- Sinclair, P., Billenness, C., Duckworth, J., Farquhar, A., Humphreys, J., Jardine, L., ... Sharpe, R. (2011). Are you ready? Assessing whether organisations are prepared for digital preservation.

- International Journal of Digital Curation*, 6(1). Retrieved from <http://www.ijdc.net/index.php/ijdc/article/view/178>
- Termens, M., & Locher, A. E. (2012). Digital preservation audit on spatial data: a practical experience. Presented at the *International Symposium on Service-Oriented Mapping (SoMaP 2012)*, Vienna, Austria.
- TRAC Metrics | Center for Research Libraries (CRL). (n.d.). Retrieved February 21, 2016, from <https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac>
- University of London Computing Centre. (2009b). *The AIDA self-assessment toolkit Mark II*. Retrieved from <http://aida.jiscinvolve.org/wp/toolkit/>
- Vardigan, M., & Lyle, J. (2014). The Inter-university Consortium for Political and Social Research and the Data Seal of Approval: Accreditation Experiences, Challenges, and Opportunities. *Data Science Journal*, 13(0), PDA83–PDA87.
- Vaska, M., & Pejsova, P. (2012). Audit DRAMBORA for Trustworthy Repositories: A Study Dealing with the Digital Repository of Grey Literature. *The Grey Literature Journal*, 8(2), 96–105.
- Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and Software Technology*, 54(12), 1317–1339.
- Williams, G. (2013). *PRINCE2 Maturity Model (P2MM), Version 2.1*. AXELOS Limited. Retrieved from https://www.axelos.com/Corporate/media/Files/P3M3Model/PRINCE2_Maturity_Model_P2MM.pdf
- Yin, R. K. (2009). *Case study research : design and methods* (4th ed.). Los Angeles, Calif.: Sage Publications.