

Current and Anticipated Digitization Competencies

Sari Järn, South-Eastern Finland University of Applied Sciences, Mikkeli, Finland

Stina Westman, South-Eastern Finland University of Applied Sciences, Mikkeli, Finland

Abstract

Digitization is a key process in preserving, protecting and providing continued long-term access to archive materials. The competencies required in digitization pose challenges to individuals and organizations engaging in digitization activities. We take a competency mapping approach to support digitization skills anticipation. Current and anticipated digitization competencies in public and private organizations in Finland were surveyed using a digital preservation competency framework.

Results show that organizations see digitization requiring a wide set of competencies ranging from policies and legal issues to practical organization and technical delivery of digitization. There is anticipated need to optimize organizational and service capabilities in near future. The desired target state is quite advanced overall, regardless of current capabilities. The largest self-identified competency gaps exist in strategic and technical approaches to digital archiving and digital access to archives. Our results inform continued professional development and planning of future digitization efforts.

Motivation

The need to digitize existing analog materials is evident as materials need to be preserved and made available for citizens, research use, and to develop new business. The sheer volume of historical materials to be digitized demands competencies to plan and manage digitization projects as well as to conduct the practical work to digitize the materials and ensure access to them. So far there has been no systematic analysis on digitization competencies and competency gaps. Therefore, organizations may find it difficult to assess the competencies needed currently and in the future, and target their development activities and resources accordingly.

The developing digitization technologies (e.g. mass digitization solutions, imaging techniques, software and workflow solutions) pose novel needs for knowledge, skills and competencies in organizations. The development of digitization relies on skilled experts and developers and requires managing the existing and future competencies underlying digitization workflows. This is the case both in organizations engaging in digitization and organizations providing tools and support for it.

One practical obstacle to digitizing collections is the thus the shortage of knowledgeable staff and lack of continuous education as e.g. digitizing technology develops. Digitization competencies may become critical for utilizing novel tools, and in effect, benefitting

from the development of digitization. It should be a joint effort of research organizations, industry and educational institutions to support continuous learning for digitization. Currently, for example the Nordic countries do not have any public digitization education offering [1] so skills are acquired on the job.

Most often digitization skills are bundled within the skill set of digital preservation and no further details are given on what skills and knowledge do digitization competencies entail. Digitization has been seen as a key competency area for preservation professionals in generalist positions [2] and specific digitization competencies have been studied by some researchers [3]. Based on case studies digitization requires a multidisciplinary skill set and various kinds of competencies, including project planning, grant writing, project management, metadata, digital capture and digital asset management [3]. These competencies are key to if and how we are able to utilize and benefit from the advances made in digitization techniques, intelligent tools as well as workflows and modern access methods to digitized content.

Various efforts have been made to structure competencies and curricula for digital preservation, including the Matrix of Digital Curation Knowledge and Competencies [4] and the DigCurV Curriculum Framework [5], as well as a practice models for digital preservation, including the Digital Preservation Coalition's Rapid Assessment Model [6]. These provide a frame of reference on what types of knowledge, skills and competencies are needed for digitization practitioners, experts and managers.. The frameworks express the diversity of contexts in which digitization activities are planned, implemented and managed in organizations.

The most recent framework for digitization competencies is the Digital Preservation Competency Framework [7]. This model was chosen for the study because it covers potential competency areas widely, treating digitization as a specific form of digital preservation [2]. It also offers ready-to-use survey instruments. The framework's main structure includes 1) five high-level competency areas that offer an overview of and quick reference to the broad range of competencies required to undertake or support digital preservation work, 2) twenty-eight skill elements organized in groups under the competency areas, which break down the competencies and five skill levels against which an organization or individual might rate their competency with regards to a particular skill element. Alongside the framework, there is also a rapid assessment model, which is a digital preservation maturity modelling tool that has been designed to enable rapid benchmarking of an organization's digital preservation capability.

Problem

This study is an empirical implementation of the Digital Preservation Competency Framework [7] to digitization competencies. We apply the competency framework to organizations engaging and supporting digitization efforts, enabling them to identify and describe the skills, knowledge, and competencies required for successful digitization. We report on findings from both public and private organizations aiming to gain insight for developing further education and training on digitization.

Our research questions are as follows:

- What competencies and capabilities are recognized as key to undertaking and supporting digitization activities?
- Where do current competency gaps exist in organizations?
- How are competency needs expected to change in the near future?

Approach

The study was conducted as a self-assessment survey utilizing both a questionnaire and semi-structured interviews. Quantitative and qualitative data were collected according to the framework [7] from various types of organizations (N = 12), e.g. national, local and private GLAM organizations (libraries, archives, museums), local public authorities and private companies. Organizations were free to define what scope of their operations and digital assets they wanted to assess in the survey. The organizations were recruited via the Finnish Memory Campus ecosystem which brings together memory organizations with the aim of strengthening cooperation and commercialization of new innovations. Depending on the organization, one or more persons participated in the survey. The respondents represented roles such as: Digitization specialist, Researcher, Curator, Planning officer, Records and archives manager, Information management and development manager, and Archive director.

The Competency Audit Toolkit (DPC CAT) from the Digital Preservation Coalition (DPC) was used in the survey. It includes tools for surveying both the competencies and capabilities related to digital preservation, wherein digitization falls under. Current digitization competency needs were identified and described along the distinct 28 competency areas in the Competency Framework [7]. A competency relates to a combination of skills, knowledge, and behaviors that, when combined, allow an individual to perform the duties of their role. A skill refers to a specific ability that can be applied to complete a particular task or to reach a certain outcome. Competency in a particular area is therefore achieved through gaining and being able to apply related skills.

Additionally, the Rapid Assessment Model (DPC RAM) maturity modelling tool was also utilized as part of the toolkit for modelling the organizations' capabilities in digitization. These capabilities refer to organizational capacity to deploy resources, such as 'bundles' of skill, know-how, attitudes, behavior, structures and processes [8]. The maturity modelling was done on the RAM scale (0 = minimal awareness, 1 = awareness, 2 = basic, 3 = managed and 4 = optimized) for both current and target states.

The surveys were conducted online via MS Teams in March and April 2023. Each organization was reserved 90 minutes for the survey and they were able to continue later if the survey was not completed within the allotted time. The survey was conducted in Finnish utilizing the original English framework materials, terms and descriptions. All but one respondent gave their consent to have the interview recorded and transcribed. Both researcher notes and transcription data were used in the qualitative data analysis.

In the meeting, the researcher interviewed the respondents, starting with the 28 skill elements in the CAT tool (Figure 1). First, the respondents chose elements were relevant for their digitization work. Only the relevant elements were discussed further within their survey. Respondents were asked in a thematic interview, for each relevant skill element, "What kind of competencies does your organization currently have?" and "What kinds of competencies does your organization need in the future (three to five years)?" The goal was to elicit descriptive information on the current and anticipated competencies. Finally, the RAM scales (Figure 2) were filled out reflecting current and target capability levels at large. The capability maturity levels (minimal awareness, awareness, basic, managed, optimized) were surveyed according to each organization's self-reported status at the time of data collection and as their anticipated needs, i.e. target level in 3-5 years. The researcher facilitated the self-assessment by connecting the maturity levels to examples given by respondents in the earlier interview.

Competency Area	Skill Element No.	Skill Element
Governance, Resourcing, and Management	1	Policy Development
	2	Risk Management
	3	Resource Management
	4	Staff Management
	5	Strategy and Planning
	6	Analysis and Decision-Making
Communications and Advocacy	7	Effective Communication
	8	Collaboration and Teamwork
	9	Stakeholder Analysis and Engagement
	10	User Analysis and Engagement
	11	Advocacy
	12	Training
	13	Producing Documentation
Information Technology	14	General IT Literacy
	15	Computer Programming
	16	System Procurement
	17	Storage Infrastructures
	18	Information Security
	19	Workflow Development and Implementation
Legal and Social Responsibilities	20	Legal and Regulatory Compliance
	21	Environmental Impact
	22	Inclusion and Diversity
	23	Ethics
Digital Preservation Domain Specific	24	Metadata Standards and Implementation
	25	Information Management Principles
	26	Approaches to Preservation
	27	DP Standards and Models
	28	Managing Access

Figure 1 Competency Audit Toolkit (CAT) applied in the study [7]

Skills gaps were analyzed from the data based on differences of respondents' target state and current state. This was done both at the organizational level and across all respondents, reflecting the overall status of digitization competencies available in the ecosystem. Quantitative and qualitative data were triangulated across the instruments used in the survey.

Organizational capabilities		
A	Organizational viability	Governance, organizational structure, staffing and resourcing of digital preservation activities.
B	Policy and strategy	Policies, strategies, and procedures which govern the operation and management of the digital archive.
C	Legal basis	Management of legal rights and responsibilities, compliance with relevant regulation and adherence to ethical codes related to acquiring, preserving and providing access to digital content.
D	IT capability	Information Technology capabilities for supporting digital preservation activities.
E	Continuous improvement	Processes for the assessment of current digital preservation capabilities, the definition of goals and the monitoring of progress.
F	Community	Engagement with and contribution to the wider digital preservation community.
Service capabilities		
G	Acquisition, transfer and ingest	Processes to acquire or transfer content and ingest it into a digital archive.
H	Bitstream preservation	Processes to ensure the storage and integrity of digital content to be preserved.
I	Content preservation	Processes to preserve the meaning or functionality of the digital content and ensure its continued accessibility and usability over time.
J	Metadata management	Processes to create and maintain sufficient metadata to support preservation, discovery and use of preserved digital content.
K	Discovery and access	Processes to enable discovery of digital content and provide access for users.

Figure 2 Rapid Assessment Model (RAM) applied in the study [6]

Results and discussion

Four out of the five main competency areas outlined by the framework (CAT tool) were found highly relevant for digitization. Figure 3 displays the share of skills elements within the area selected as relevant across all respondents.

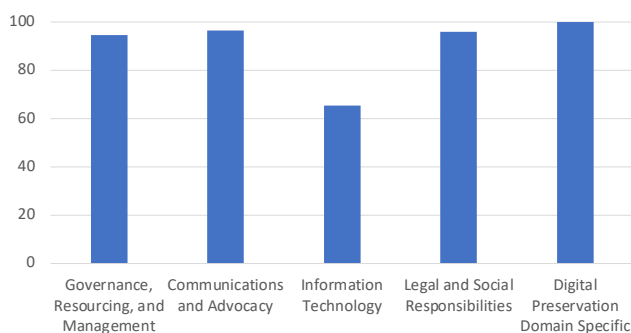


Figure 3 Percentage of skills elements chosen as relevant across all organizations

The fifth competency area *Information Technology* divided the respondents, with most choosing related skills elements as relevant for them. Two organizations deemed the whole competency area irrelevant for their digitization operations and further two did not select any skills elements relevant beyond General IT literacy. This

finding indicates that the competency areas outlined by the framework are overall relevant for digitization, but digitization competency profiles differ.

Evaluations on the organizational capability were visualized to show key capabilities and major gaps. The results show that across all organizations surveyed, there were digitization capability needs across all areas of the framework (RAM tool). Figure 4 shows the average target and current capability levels with the associated standard deviation.

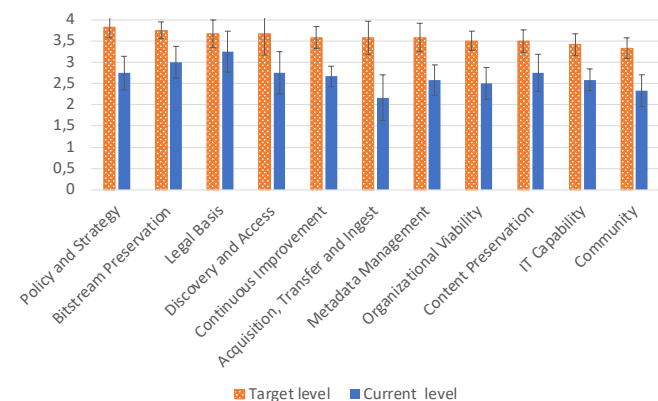


Figure 4 Average current and target capability levels of across all organizations, with one standard deviation

The target level for capabilities was high, averaging over 3.5 (managed to optimized) in most areas. The high target level was found to be true regardless of current level of capabilities. A two-tailed Student's t-test showed no difference in target levels ($t(10) = 1.963$, $p > 0.078$) when comparing those respondents whose current level was below average ($M = 3.40$, $sd = 0.45$) to those whose current level was above average ($M = 3.84$, $sd = 0.23$). The largest capability gaps existed in the areas of *Acquisition, Transfer and Ingest* (average gap 1.4 competency levels) as well as *Policy and Strategy* (average gap 1.1 competency levels). The target level was closest in *Legal Basis* (average gap 0.4 competency levels), where current capabilities were evaluated highest.

Based on the qualitative data, the anticipated needs related to *Acquisition, Transfer and Ingest* capabilities deal with skills related to e.g. moving to digital archiving, procuring and using new equipment, information and cybersecurity knowledge, quality management and compliance issues as well as technology neutral workflows. Based on the qualitative data, the anticipated needs related to *Policy and strategy capabilities* deal with e.g. costing and funding of digitization services, knowledge sharing within the archival community, evaluating maturity of innovations and technology as well as evolving value appraisal and societal impact of digitization. The anticipated competency needs for *Legal basis* mainly dealt with continued monitoring of the changes in legislation and related good practice in the digital preservation community.

The responding organizations' capability profiles varied. The results indicate that organizations were targeting either a "managed" or

“optimized” level in organizational capabilities. The differences were due to for example some organizations not seeking a prominent role in the digitization community nor in-house IT capabilities. Due to the small sample size, no statistical analysis could be conducted but a variety of profiles is shortly described here.

All surveyed public organizations expressed high target capability levels but identified varying current levels, as you can see by comparing Figures 5 and 6, showing the RAM profiles of a national cultural heritage institution and a local public authority.

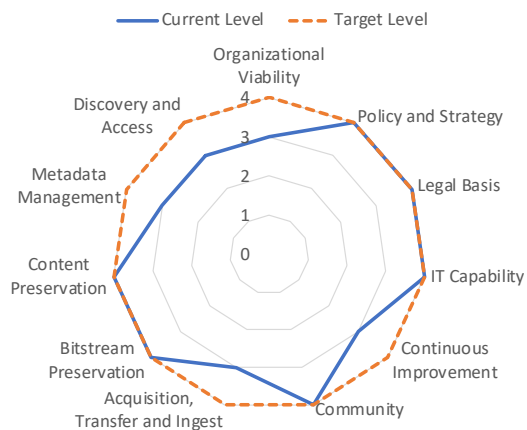


Figure 5 Capability profile of a national cultural heritage institution

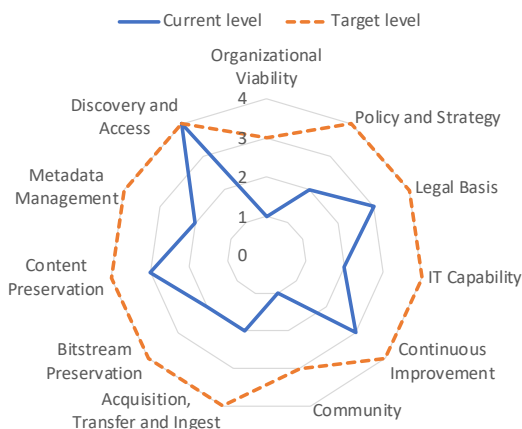


Figure 6 Capability profile of a local public authority

The national cultural heritage institution reflected on their capabilities by benchmarking to others in their operational environment, describing their role as one with leadership. The public authority’s target levels were high and known capability gaps were ascribed mostly to a lack of resources and a recent reorganization where digitization operations were revamped.

Our sample contained multiple types of private organizations. Figure 7 shows the RAM profile of a company engaging in digitization business, describing their own capabilities as supportive

to those of their customers. Figure 8 is the profile of a private archive that serves dedicated stakeholders, digitizing domain content by demand alongside other customer service tasks.

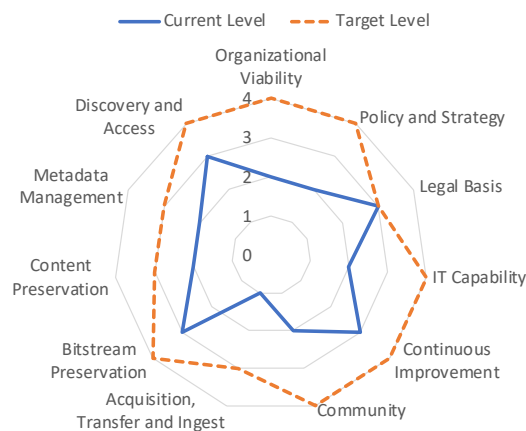


Figure 7 Capability profile of a private company

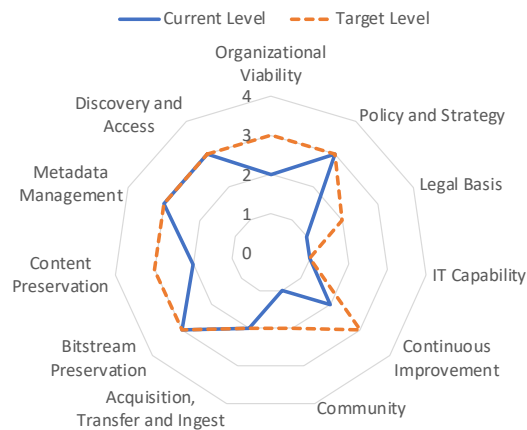


Figure 8 Capability profile of a private archive

A further exploration of the qualitative data along the skill areas (CAT tool) sheds light on the anticipated future (3-5 years) competencies. Tables 1 through 5 display examples of the competencies referred to by respondents in the distinct competency areas. Note that the sample size varies, reflecting the number of organizations who chose the skills area relevant to them.

When discussing the strategies for digitization, various organizations reflected to the purposes of digitization. Evaluating the wider societal impact of digitization efforts was noted as a novel type of competency need, though related to value appraisal. The policies and management of digitization were discussed in light of respondents engaging in varying models of digitization (on demand services, mass digitization, project-based digitization, voluntary efforts etc.). The key staff competencies discussed included a combination of business and user understanding, technical skills and domain knowledge related to the materials to be digitized.

Table 1 Governance, Resourcing, and Management examples

Skill	Anticipated needs (N = 12)
Policy Development	- Digitization policy updates - Balancing systematic and agile modes of digitization work
Risk Management	- Shared management principles - Managing legacy systems
Resource Management	- Business models for digitization - Ensuring continuity of operations
Staff Management	- Working with lack of resources - Evolving work profiles - Succession planning
Strategy and Planning	- Impact of digitization - Capturing tacit knowledge
Analysis and Decision-Making	- Value of use and value to users - Benchmarking and scenarios

Within their stakeholder domain, various organizations stressed the archives' role as research infrastructures and required competencies to serve the research community. They also described benefits from the latest knowledge and testing of new solutions. Ensuring continuity of digitization under project-based funding was seen as a common goal for developing advocacy. For some organizations, competence development was an ecosystemic effort. However, not all were active participants of the digital preservation community.

Table 2 Communications and advocacy examples

Skill	Anticipated needs (N = 12)
Effective Communication	- Solutions for peer learning - Marketing and negotiation skills
Collaboration and Teamwork	- Specialization within digitization - Mentoring
Stakeholder Analysis and Engagement	- Building partnerships, consortia - Understanding research use and building research collaboration
User Analysis and Engagement	- Data analytics competencies - Understanding uses and affordances of digitized materials
Advocacy	- Ensuring funding - Highlighting reuse
Training	- Sharing practices and experiences
Producing Documentation	- Documenting for development, outsourcing, continuity

All organizations did not deem IT skills as relevant. Their reasoning was that as they had outsourced the digital capture, they did not seek to build in-house competencies in this area. However, they discussed needing competencies to plan and procure technical services as well as manage quality. Others noted explicitly that even if services were obtained via partnerships, the organization should have enough IT capability to successfully plan system projects, run procurements, steer development, capitalize on the services, etc. Those building their own IT capabilities reflected on developing and testing services alongside process development. Some IT needs emerged due to the current geopolitical situation with the respondents reflecting upon the evolving information and cybersecurity requirements and related competencies.

Table 3 Information technology examples

Skill	Anticipated needs (N = 10)
General IT Literacy	- Novel equipment - Requirements specification
Computer Programming	- Role of AI in digitization and overall in preservation
System Procurement	- Steering procurements and guiding IT partners
Storage Infrastructures	- Growing data volumes - Move to cloud storage
Information Security	- Holistic approaches to security - Effects of AI to security
Workflow Development and Implementation	- Optimizing production - Digital-only workflows - International standardization

The respondents described the environment of digitization as ever-changing, and expressed needs to both impact better legislation, and transform ensuing compliance requirements into joint operational guidelines. The sustainability of digitization was reflected from various viewpoints as present in the framework.

Table 4 Legal and Social Responsibilities examples

Skill	Anticipated needs (N = 12)
Legal and Regularity Compliance	- Impacting and adapting to changing legislation - Joint compliance definitions
Environmental Impact	- Practical calculation of carbon footprint in digitization
Inclusion and Diversity	- Recognizing, communicating and balancing distortions in archives
Ethics	- Open data and sensitive data

Assessing skills for digital preservation led to discussions of practical implementations of principles and standards. Overall, both holistic understanding of the principles and methods as well as detailed practical skills were sought after. Artificial intelligence (AI) was mentioned both a valuable tool for preservation and a threat to the reliability of digital information.

Table 5 Digital Preservation Domain Specific examples

Skill	Anticipated needs (N = 12)
Metadata Standards and Implementation	- Evolving standards - Changing platforms - Contributing to development
Information Management Principles	- Interoperability in practice - Management for research use - Reliability and authenticity of information in the age of AI
Approaches to Preservation	- Digital preservation solutions for different types of content
DP Standards and Models	- Detailed technical specifications - Practical implementation
Managing Access	- Dealing with technical debt - Ensuring usability

Conclusions

Our results shed light on the digitization competencies recognized and anticipated by organizations engaging in digitization activities. A wide array of competencies were recognized as key to undertaking and supporting digitization activities. Understanding the users, current uses and potential uses of digitized materials was a key competency. Substance knowledge and digital competencies were seen as interlinked. This was complemented by needs for business understanding. All participating organizations were seeking to optimize their digitization service capabilities in the near future. The largest competency gaps reflect needs for new abilities in strategic and technical approaches to digital archiving and digital access to archives. These needs ranged from practical issues to existential matters, i.e. the changing role and relevancy of archives.

Our respondents could define competency needs for the near future. At the same time, often the description of the need was defined as “keeping up with progress”, reflecting a model of continuous learning. Respondents described planning and conducting competence development via multiple methods: formal education, learning on the job, mentoring, peer learning and benchmarking were all mentioned frequently. In many organizations, lack of resources was mentioned as a blocker to acquiring necessary skills or utilizing the existing capabilities. In terms of competency needs, this translated to negotiation and lobbying skills as well as transforming digitization business models and workflows.

Different types of organizations referred to different types of competencies, skills and knowledge in their assessments. Small public archives and businesses described their competency needs based on current operation models where digitization was mostly done on-demand, based on customer requests. The staff engaging in digitization often had other duties as well, and digitization could be a seasonal effort. These organizations reflected on holistic and processual competency needs in order to develop their own working models around digitization and gain knowledge from the digital preservation community. Large, public actors could rely on their strategy and set goals for digitization when describing both current and future needs. Their future needs dealt with both optimizing internal competencies and gaining further benefit from goal-oriented collaboration and research on digitization.

We also wanted to evaluate the suitability of the chosen framework for mapping digitization competencies. The framework covered the field of digitization well, as all organizations were able to conduct a facilitated self-assessment with it. The scope of the tool was thought to be wide, as expected. One respondent compared the scope to a quality standard they work with. Another noted that it might have even been too detailed for a small archive. Based on the interviews, the competency mapping would have benefitted from fine-grained examples of some technical skills related to digitization as part of the domain-specific skills. As it was, some organizations reflected only on the generic IT competencies rather than the technical facet of the digitization process itself. Another viewpoint to digital preservation that was not evident from the tool but quite prominently discussed by our respondents, was the research use of digitized materials and related workflows. One organization wished the tool included examples related to archive materials as research materials.

Based on the quantitative and qualitative data, organizations' profiles and goals for digitization competencies differ. In our small sample, the gap between large, well-established digitization organizations and small actors in the field was not very wide. It is possible that organizations had different interpretations of the capability levels, even after these were defined via examples. This should be considered when utilizing these results. It should also be noted that our sample only covered a small fraction of private sector companies and local authorities in Finland. The formulation of digitization capability profiles would benefit from replication with a broader set of organizations.

Our results will be used in developing digital preservation and digitization curricula on European and national levels. The results may also be used to facilitate e.g. planning of digitization projects, service development, structuring professional development and recruitment and writing grant proposals. Conducting a competency mapping can benefit institutions planning to embark upon a digitization project, upskilling their workforce or updating their digitization workflows and tools. The framework enables a structured way to set goals for and monitor competence development based on the recognized skill gaps.

References

- [1] H. Ristolainen, K. Saltiola, S. Järn, “Need for digitization professionals: creating new trainings in Finland”, *Xamk Beyond* 23 (2022), pg. 37-57. Available at: <https://urn.fi/URN:ISBN:978-952-344-458-4>
- [2] M.M. Miller, M. Horan, “Evolving roles of preservation professionals: Trends in position announcements from 2004 to 2015”, *Library Resources & Technical Services*, 61(4) (2017), pg. 183. Available at: <https://doi.org/10.5860/lrts.61n4.183>
- [3] G. O’Hara, E. Lapworth, & C. Lampert, C., “Cultivating Digitization Competencies”, *Information Technology and Libraries*, 39(4) (2020). Available at: <https://doi.org/10.6017/ital.v39i4.11859>
- [4] Lee, Christopher A., and Helen Tibbo, “Where’s the Archivist in Digital Curation? Exploring the Possibilities through a Matrix of Knowledge and Skills”, *Archivaria* 72, pg. 123-168 (2011). Available at: <https://archivaria.ca/index.php/archivaria/article/view/13362>
- [5] DigCurV Project, “Curriculum Framework for Digital Curation”, (2013). Available at: <https://digcurv.gla.ac.uk>
- [6] Digital Preservation Coalition, “Rapid Assessment Model”, Version 2 (2021). Available at: <http://doi.org/10.7207/dpcram21-02>
- [7] Digital Preservation Coalition, “DPC Competency Framework”, 1st Edition (2022). Available at: <https://doi.org/10.7207/dpccf22-01>
- [8] Wang, C.L., and Ahmed, P.K., “Dynamic Capabilities: A review and research agenda”, *International Journal of Management Reviews*, 9(1) 31-51 (2007).

Author Biography

Sari Järn (Cultural Manager, Bachelor's Degree) has worked for 20 years in digital archiving projects and trained more than 80 digitization experts. Currently she works as RDI Specialist at Digitalia, an RDI unit at South-Eastern Finland University of Applied Sciences.

Stina Westman (D.Sc. (Tech.)) has an extensive research and development background in user-centered research and service development in the areas of media, information systems, data management, open science as well as education. Currently she works as director of Digitalia, an RDI unit at South-Eastern Finland University of Applied Sciences.