The Network Electronic Archiving Concept for Organizing Digital Preservation for Small Archives

Kathrine Hougaard Edsen Johansen
Copenhagen City Archives/NEA
Københavns Rådhus,
1599 København V
Denmark
f27b@kff.kk.dk

Kim Schou Nørøxe
Ballerup Municipal Archives/NEA
Hold-An Vej 7
2750 Ballerup
Denmark
kin@balk.dk

ABSTRACT

This paper presents details about a network organization that enables small public archives in Denmark to do professional digital preservation that both lives up to established best practices and is economically feasible. The Network Electronic Archiving (NEA) is built around extensive collaboration and sharing with archival infrastructure and expert staff at its core.

To demonstrate how NEA works in practice, this paper includes the case of one of the small public archives that implement NEA support and services into their day-to-day work. From the point of view of a small archive this case illustrates how sharing and collaboration via NEA has allowed them to do full-scale digital preservation in accordance with OAIS.

KEYWORDS

Small archives, feasibility in digital preservation, community based approach to digital preservation, business models for digital preservation, economies of scale in digital preservation, collaboration and sharing, sustainable digital preservation

1 INTRODUCTION

1.1 'The Danish Model' - Extensive Archival Legislation

Denmark has extensive archival legislation. It specifies in detail what, how and when to archive and it gives far-reaching powers to the archives. This is particularly true for born-digital material. [1]

A national strategy for appraisal, transfer and preservation of data form the public sector sets the frame. It is based on early identification and migration with normalization before ingest. The strategy is set out by the National Archives and manifested in the laws and decrees that make up the legislative framework for archiving of born-digital materials.

The legislation for appraisal maps out the public sector and in detail specifies what data from the public administration must be transferred to archives for preservation. [2] Producers are obliged to submit this data to archives at their own expense by a deadline set by the archives. Public archives perform a thorough appraisal of all born-digital material created by the public sector's administrative bodies to identify what data must be archived. The national strategy is based on early identification. Appraisal is often done as early as when producers commission an IT system.

Further the law stipulates precisely how data should be submitted to archives. [3] This includes the specific data structure, data types, file formats and metadata requirements that submissions must comply with.

The legislation places the responsibility for and cost of submission on the producer. This is exploited to the benefit of archives by defining a SIP format that is practically identical to the AIP format. This effectively means that producers are responsible for the normalization before ingest s defined in the national strategy. The receiving archive validates submissions and does thorough quality assurance. It ensures compliance with the SIP format defined by law and checks the submission agreements for individual submissions. Any diversions must be corrected at the expense of the producers.

When an archive concludes that a submission lives up to requirements, the status of the Information Package effectively changes from SIP to AIP. No changes in structure or format are needed in the SIP to AIP transformation. However, minor adjustments due to packaging issues can be made.

1.2 Public Archives

All public archives are subject to the national laws and decrees for archiving of born-digital materials. This means that all public archives are mandated to enforce archival interests by requiring submissions and defining when and how data is submitted. This empowerment comes with obligations. Archives must do their part, which is no small task. It includes doing appraisal, making submission agreements, validating transferred Information Packages and ingesting them into archival storage, doing preservation planning, managing collections, providing access, etc. In short, they must run a full-scale digital archive in accordance with OAIS.

This is required from all public archives, which in a Danish context means every archive that operates under the Danish Archives Act [4] This includes The Danish National Archives, the largest public archive in Denmark, but the landscape of public archives also covers several small public archives at the municipal level. These are run independently by each municipality and often with a limited budget, and only a handful of employees.

2 NETWORK ELECTRONIC ARCHIVING

2.1 Meeting the Challenge Together

Running a digital archive is demanding for all archives and even more so for a small archive with a limited budget and size. This is a common challenge for small archives. NEA was created in 2007 to overcome this challenge. NEA, which is short for Network Electronic Archiving, is a not-for-profit network organization through which small archives collaborate closely and take advantage of economies of scale. [5]

In contrary to most other similar organizations like e.g. MetaArchive [6], NEA is more than a common storage facility and bit repository. The basic idea of NEA is that rather than having each archive employing specialized staff and building its own infrastructure for preservation, archives share them via the NEA. Each archive is run individually and is responsible for its own business. However, the day-to-day operations are supported by NEA to the extent needed by each archive. This makes it possible for small public archives to collect and preserve born-digital material, which ultimately means they can meet the demands placed on then by the national archival legislation.

NEA was created when 7 small public archives joined forces. The network has since grown and today NEA counts 28 public archives. The network is hosted by Copenhagen City Archives, which is the largest member of NEA.

2.2 Common Archival Infrastructure

NEA provides a shared archival infrastructure, which means that archives can avoid creating their own set-up. Ingest, Archival Storage, Preservation Planning and Management are handled centrally in NEA, but access is clearly separated. Each archive only has access to its own data. The handling of data is contractually agreed on between NEA and individual member archives in separate data processing agreements.

The common facilities are evaluated internally to ensure they live up to national and international best practices and an information security certification process with the ISO 27000 series of standards was started recently. There are no immediate plans for other certifications.

The common infrastructure comprises of

- **Ingest Platform:** Set-up for validation and quality assurance of SIPs and subsequent ingest to the archival storage
- **Bit Repository:** A common bit repository for archival storage
- **Access Platform:** Archives access their own data through this platform. The platform can also be used to retrieve data for dissemination to end users
- **Data Management Platform:** This is the central system that supports the archives in everything from appraisal to administration. This platform is also used by NEA staff for administration and preservation purposes.

NEA does not offer a finding aid. Instead NEA collaborates with a separate community called STARBAS, which offers a finding aid service to small archives. This community is also hosted by Copenhagen City Archives and many NEA archives are also part of the STARBAS community. [7]

The infrastructure is mainly proprietary and developed in house. The is because NEA was created in 2007, when the marked for digital preservation tools was still smaller. Furthermore, the founding archives had limited knowledge of existing tools at the time.

NEA has recently stated a shift towards reusing existing and preferably open source tools rather than developing tools in house. The first tool has been replaced by an open-source alternative originating from the E-ARK project [8] and the replacement of other tools is being planned.

2.3 Shared Expert Staff

At the core of NEA is a shared group of experts that have the skills, expertise and experience needed to do OAIS-compliant archiving. Members draw on the expertise and knowledge of this shared staff as needed on a consultancy basis. NEA staff support member archives in all aspects of their day-to-day operations, including appraisals, negotiations with producers, validation & quality assurance, ingest and access.

The shared expert group currently counts 10 people with different areas of expertise. It includes IT-professionals, archivists, preservation professionals and developers. Together the staff cover the skills, expertise and experience needed to support member archives.

2.4 Community

The archives in NEA have a strong member community. Within the network, archives collaborate, share experiences and lessons learned, and support each other. Because all member archives act under the same legislation and use the same archival infrastructure, outcomes and lessons learnt by one archive are often directly applicable at others. The community does not only offer colleagues and peers to discuss challenges and concerns with. It gives coherence and a sense of not being alone.

2.5 NEA Services

Each archive in NEA is run independently. Each archive is responsible for its own business and is accountable for living up to the legislative demands placed on archival archives. The role of NEA is solely to support member archives in running their individual businesses. The responsibility placed on archives is law-based and cannot be outsourced to NEA.

To meet the individual needs of member archives, the services and support provided by NEA is flexible and adaptable. NEA offers different 'service packages' with varying levels of support, which allows member archives to tailor services to their individual situations depending on their ambitions and their own level of expertise.
The services offered are updated regularly as the needs of member archives evolve. NEA currently offers six service packages:

- **NEA Member:** Access to the common archival infrastructure, online tools, training, NEA R&D and administration.
- **NEA Support:** Individual support on all matters related to digital archiving on a consultancy basis. Member archives acquire support from the shared staff as needed.
- **NEA Storage:** Secure archival storage in NEAs repository including ongoing Preservation Planning and Data Management
- **NEA Data:** Export of data on demand in DIP format
- **NEA Access:** Online access to collections via the access platform
- **NEA KMD:** Joint yearly submissions of data from intermunicipal IT-systems from the supplier KMD. Data is separated upon ingest.

By buying into service packages, member archives can get the expertise and facilities they need to create and complete their own OAIS. The network does not monitor or evaluate member archives' compliance with OAIS. Each archive is responsible for its own business and decides for if and how OAIS compliance is evaluated.

3 CASE: BALLERUP MINICIPAL ARCHIVE

3.1 A Small Public Archive

The municipality of Ballerup is a suburban area outside Copenhagen, covering 34 km^2 and just over 48.000 citizens. [9] The municipality as an administrative unit was created in 1953 but the history of the area dates back to the 13^{th} century.
Ballerup Municipal Archives, founded in 1979, are the official archives of Ballerup. [10] The collection covers analogue and digital materials. The analogue collection includes 4 kilometers of paper, an audiovisual collection and a collection of photos. The digital collection includes digitized versions of selected materials from the analogue collection and 2.6 TB of born-digital material. The oldest material in the collection is from 1801 and the newest is from 2017. The staff counts one archivist, one archives assistant, a project employee, two regular volunteers working at the archives and several loosely connected volunteers.
The archive serves a dual purpose. It is partly the archive for the local history of the Ballerup area and partly the public archive for the municipality of Ballerup. In its function as archive for local history, various materials documenting the local history are collected and preserved. Submissions happen when individuals, local organizations and companies submit their personal archives on a voluntary basis. In this function the archive is not subject to the Archives Act. As a public archive, documentation from the public administration of the municipality that has historical value or serves as evidence for matters of significant administrative or legal significance for citizens and authorities is collected and preserved. In this capacity the archive acts under the Archives Act and has the authority to enforce archival interests along with the obligation to run a full-scale digital archive in accordance with the national strategy and legislative framework.
Ballerup Municipal Archives uses support from NEA in its function as a public archive and consequently the following only concerns this function.

3.2 Identification and Appraisal

As a public archive, Ballerup Municipal Archives collects and preserves data from the public administration of the municipality. The law clearly states that the overall responsibility for submitting data to the archive lies with the administration. It also specifies what data from the public administration must be archived. As a public archive, the Ballerup Municipal Archives assist the administration in carrying this responsibility by identifying data that must be preserved and setting deadlines for submissions.
To ensure that all data is transferred to the archive as prescribed by law, the archive monitors data creation in the administration and closely follows the lifecycle of IT systems in use. This requires collaboration with various functions in the administration, including the records managers, IT services and contract managers. When a new IT system is commissioned the archive is notified allowing for early appraisal based on the legislation for appraisal and the archive's detailed knowledge of the data creation in the administration. When appraisals are complex, Ballerup Municipal Archives draw on expertise from NEA. To ensure timely submission of data the archive creates a submission plan deciding when each dataset shout be submitted. Normally data is submitted every 5 years or at the time a system is shut down. If a system is shut down earlier than expected, the archive is notified allowing the submission plan to be revised accordingly.

3.3 Submission and Preservation

The archive notifies the administration when it is time for submission. Shortly thereafter, the archive, assisted by NEA, negotiates with the administration and the supplier. NEA brings expert knowledge of the submission format and legislation to the negotiations. Negotiations result in a contract with the supplier and a submission agreement specifying how and when data is to be submitted.
As producers are obliged by law to submit data to archives at their own expense, cost is a focal point for the administration. The archive has this in mind when balancing costs, quality and archival interests in the negotiations, while striving to get favorable conditions in the contracts. NEA experts support the negotiations and decision making, the community also represents an asset. Particularly when it comes to costs and contracts, the community is useful as a forum for comparing and discussing offers and contracts with other archives.

In the actual submission and preservation process, Ballerup Municipal Archives relies heavily on support from NEA. Staff from NEA validates and quality assures submitted SIPs and ingests data to the repository as AIPs. NEA manages the archival storage in the common repository and ensures that proper actions are taken to ensure long-term preservation.
At the time of ingest, NEA staff creates a DIP that consists of all data in the AIP. The DIP is made available to the Ballerup Municipal Archives in the access platform.

3.4 Access

Holdings are displayed to end users through the STARBAS finding aid, which allows end users to search the collections and request access. Ballerup Municipal Archives access their own data via the access platform, through which data can be retrieved and disseminated to end users. If data is needed in another format than what is possible through the access platform, a special DIP can be created by NEA staff upon request.

3.5 Benefits of Joining NEA

To the Ballerup Municipal Archives, the primary benefit of NEA is the possibility of doing professional digital archiving at a low cost. Without NEA, the archives would employ staff to undertake all functions in house and create their own infrastructure, the price of which would by far exceed the cost of buying services from NEA. Some 15% of the budget is currently spent on services from NEA. Another major advantage of joining NEA is the community. It offers a formalized forum for exchanging ideas, collaborating and discussing issues. Something that would be time consuming and difficult for archivists at Ballerup Municipal Archives to find elsewhere. Lastly, NEA provides robustness to the archives. With only 3 employees, Ballerup Municipal Archives is vulnerable to changes in staff. By joining a network like NEA, the dependency on single persons is reduced because knowledge is spread out and shared within the network All in all this helps to ensure business continuity.

4 ACHIEVEMENTS

Since its creation in 2007, NEA has supported small public archives in doing professional digital archiving. More than 55 TB of born-digital materials, covering over a thousand Information Packages with data dating back to 1979, have been transferred to member archives for long-term preservation. As the case with Ballerup Municipal Archives demonstrates, this is possible because of the network organization and the extensive collaboration and sharing.
The approach has proved so beneficial that similar network organizations for small public archives have been formed in Denmark. The KOMDA network has 11 member-archives [11] and the EBA network has 5 member-archives [12]. This means NEA no longer is the only network of its kind in the country but it is the first and it is by far the largest. The three network organizations have a good, non-competing relationship and collaborate when relevant.
NEA, KOMDA and EBA all have slightly different business models, but they all build on the same basic concept. The reuse of this collaborative approach proves that the basic concept of sharing and taking advantage of economies of scale is attractive and viable.

5 CONCLUSION

The approach presented in this paper does not propose a universal solution. It is one way of lifting the difficult task small archives face; being obliged to preserve digital material in a way that lives up to established best practices while lacking the size, resources, expertise and skill to do it. This paper demonstrates that it is possible to do full-scale digital preservation at small institutions in a way that is manageable and economic feasible.
The case with the Ballerup Municipal Archives illustrates how NEA works in practice and demonstrates the difference NEA makes to a small archive. By buying into services and support form NEA, Ballerup Municipal Archives can do professional digital archiving at a low cost. Running a preservation practice of equal quality, professionalism and robustness at the same low cost would be impossible without this network organization.

REFERENCES

[1] www.sa.dk/en/appraisal-transfer/legislation
[2] Executive order 183 of 26. January 2018 regulates appraisal in municipalities www.retsinformation.dk/Forms/R0710.aspx?id=198687 Other executive orders exist for other parts of the public sector
[3] Executive order 1007 of 20. August 2010 regulates submission of data from public authorities www.retsinformation.dk/Forms/R0710.aspx?id=132898
[4] The Danish Archives Act www.retsinformation.dk/Forms/R0710.aspx?id=183862
[5] Johansen and Neuhard (2017): How collaboration and sharing has made OAIS compliant archiving possible at (very) small archives, poster, IPRES 2017, Kyoto, Japan
[6] www.metaarchive.org/
[7] www.starbas.net
[8] www.eark-project.com/resources/eark-tools
[9] www.ballerup.dk
[10] www.ballerupstadsarkiv.dk
[11] www.aalborgstadsarkiv.dk
[12] www.eba.esbjergkommune.dk