The Nordic Model of Digital Archiving

Edited by Greg Bak and Marianne Rostgaard

First published 2024

ISBN: 978-1-032-35122-3 (hbk)
ISBN: 978-1-003-32540-6 (ebk)

Chapter 8

From national to international standards in Norway

From documents to data?

Martin Ellingsrud

DOI: 10.4324/9781003325406-10

The funder for this chapter is Aalborg University.

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Chapter 8

From national to international standards in Norway

From documents to data?

Martin Ellingsrud

Today, standardisation drives the archival world. Through the development and uptake of international, national, and community records management and archiving standards, archivists have moved towards global shared knowledge, and perhaps eventually will move from a document focus to a data focus. In Norway the power of standardisation was recognised by archivists and records managers earlier than elsewhere, driving Norway to develop national standards in advance of international standards, for example in the areas of records management and database archiving. In the absence of international standards, national solutions to global problems became prevalent.

This situation has changed, and current Norwegian practitioners are a more interlinked part of the international profession today than before. International thinking and standards have reached Norway, thus beginning a new phase in standards usage in Norway and engaging with the emerging discourse of "archives as data" (Mordell 2019). A data-driven age for archives promises much, including new insights for researchers using digital archives as data sets and new approaches to arrangement and description for archivists (Mordell 2019, 140).

Following a brief introductory discussion of standards and standardisation, this chapter has two main parts. Part One treats Noark (*Norsk arkivstandard*[1]), the Norwegian records management standard that predated ISO 15489 by almost two decades, starting with a broad historical overview and then considering to what degree Noark shifted Norwegian practice towards an "archives as data" paradigm. I do not suggest that the data view is somehow superior to the document view. The intention in this chapter is to show how certain standards fit an "archives as data" approach better than others, not to say that the document view is no longer necessary.

Part Two considers the trend towards international standards through the *digital archival package structure* (DIAS) project (2010–2012) and through simultaneous use in database archiving of the Norwegian standard ADDML and the internationally accepted Swiss Federal Archives standard *Software Independent Archiving of Relational Databases* (SIARD).

DOI: 10.4324/9781003325406-10

Standards can be said to be an agreed-upon common practice. When standards are widely used, they become "well known and associated with a specific, predictable behaviour that people come to rely on" (Hamill 2017, 40). Therefore, standards can be used to make communication, interaction, and exchange of information across domains easier. For example:

> Once a researcher familiarizes herself with a finding aid prepared using descriptive standards, she has the ability to predict the kind of information she will find in the next finding aid, where it is likely to be located, and how it might be organized (Hamill 2017, 40).

Standardisation has negative aspects as well. It sets rules and makes tidy boxes, allowing practitioners and researchers to overlook factors that do not neatly fit the model and perhaps oversimplifying others. To standardise is to generalise, and standards limit one's options and viewpoints. Standardisation can therefore potentially hinder new ways of thinking that the standard was not adopted to support. Standards can lead to orthodoxy and stale professional development if they are not continually reviewed. A standard has a normative point of view. As Duff and Harris remind us, standards are:

> a set of agreed-upon rules spanning more than one community of practice or site of activity and enduring over time; and the deployment of these rules to make things work together over distance and heterogeneous modes of measurement and description. The wider the span, the greater the distance, the more heterogeneous the modes, then the greater the violence done to the local, the individual, the eccentric, the small, the weak, the unusual, the other, the case which does not fit the conceptual boxes that are unavoidable in any form of standardization (Duff and Harris 2002, 281).

Standards, then, can stifle creativity and compel the profession to see the world through a specific viewpoint, which can hinder alternative views. Since it can be expensive to change work already done in accordance with a standard, standards create path dependency. Just as there are legacy systems, one can also have legacy standards. In some cases, even after a new standard has been put in place, the old standard may live on through work done in the past. Change in society, or specifically archives and records management in this instance, can sweep away the underlying assumptions that standards rely on (Bell 2011, 33–35).

Community of practice is a key concept here. Standards, whether from ISO, Library of Congress, or ICA, can be a common ground for a profession, a *lingua franca* if one will. By being a common ground, standards can also drive professionalisation, thus making (for example) archivists a separate and distinct group from other information professionals. While digital archival standards are not arcane, hidden lore being guarded by archivists, they are

part of professional discourse, a part of the profession's knowledge base, and they can become a barrier for those outside of the profession by building on specialised concepts and terms. Codifying practice, standards can form a foundation for a profession. As a nation's constitution can provide common understandings of law and society within the nation, standards offer professions, at least to some degree, an agreed-upon base of specialised knowledge for new recruits. Additionally, standards like METS and PREMIS are broadly used among digital preservation specialists, including archivists. Standards can thus provide a common ground across professions and act as bridges spanning professional borders even as, by introducing narrowly defined common language and concepts, they can alienate those not engaged in the field. They work towards unification and exclusion, simultaneously.

Part One: Noark – a "highly specific" Norwegian standard

Noark 1 to 4: Developing a national standard

Noark 1 was released in 1984 as a joint effort between the Norwegian Rationalization Directorate and the National Archives, to develop a requirement for electronic registry systems in the Norwegian national administration. Located under the finance department, and later the department for administrative affairs, the Rationalisation Directorate's main responsibility was to provide services that supported renewal in the state administration. Noark 1 was based on the National Archives' *Guidelines for Electronic Registries* (1984), which were themselves based on the directorate's study of the electronic registration of records in registry databases, a project that ran from 1979 to 1984 (Arkivverket 2018a; Valderhaug 2011, 110).

In Norway, the duty to maintain a registry of communications and other records has long existed in the public sector, mandated by royal decree in 1740 for letters sent to agencies (Vistdal 2011, 4).[2] This practice links Norway to a larger Germanic or northern European tradition within the archival world, where Prussia already had a registry in the seventeenth century (*Geschäftstagebuch*) (Miller 2003, 50). The registry itself is a list or a table where key metadata about each record, such as title and date, is registered in chronological order. The registry makes it possible to preserve basic contextual links in the archive, to preserve the archival bond. Today, search capabilities in electronic registries can allow users to disregard the chronological order in which metadata is registered.[3] Registries are an internal records management tool for the agencies, but redacted registries are also made public. Thus, citizens can use registries to discover records of interest and then send freedom of information requests to public agencies asking for the records.

Noark 1 was a requirement specification for electronic registry systems, which most often were simple database systems with tables containing metadata about

each record and the case file it belonged to. For the Rationalizing Directorate, the purpose for Noark 1 was two-fold: Noark 1 was to be a standard for its own registry system, but also a guide for other information technology developers and government agencies, thus spreading the seeds of standardisation. A standardised electronic registry would in the end simplify case control/overview and searching (Valderhaug 2011, 111). For the National Archives, this standardisation would simplify transfer and reuse of records and data from the agencies, since the stated principle was that registry databases were to preserved for disposition to the National Archives in an electronic format, even though the guidelines themselves never specified how such a transfer would occur, and stipulated that for the time being registries were to be printed and submitted on paper (Valderhaug 2011, 110–111).

Noark 1 was part of a regulation phase in the 1980s, as new rules and regulations came to control records management, including the already mentioned guidelines for registries, appraisal rules for state government agencies (1989), and rules for the usage of microfilm (1988). These initiatives were a direct consequence of a review that had documented several challenges, such as appraisal, when it came to records management in state agencies. The *Freedom of Information Act* (1971) was a key driver, since public information that is well-structured is more discoverable. Stated briefly, the authorities wanted a good regime for records in place. As with much of this work, Noark 1 was a home-grown Norwegian product. The relevant authorities did not look outside to find inspiration, as there were few, if any, foreign models. Electronic records management was still in its formative period. Noark 1 drew inspiration from the longstanding Norwegian registry tradition instead.

Noark 1 quickly became a *de facto* standard or requirement specification for the Norwegian public sector. Even though it was supposed to be used by the national administration only, it became a guiding standard for system development also in the municipal sector. Several new versions were developed: Noark 2 (1987), Noark 3 (1994), Koark (specifically for the municipal sector; 1995), and Noark 4 (1999). In 1990 the National Archives assumed full responsibility for the standard, as it best fit that agency's responsibilities.

Noark 3 made possible an integration between the registry and an electronic archive or *fonds*, creating a *de facto* electronic case handling system and records management system, in Norwegian known as the *sak/arkiv*-system, *sak* roughly meaning case. Cases, case files, and case handling have a long history in Norway, starting with paper-based registry systems that linked related incoming documents. Through Noark 3 a tight integration between the registry and the administration of cases and case files in the creating agencies was made possible. The system was not only a registry, but also a tool for access, workflow administration, and the production of records and cases or case files. Noark 3 finally also specified a format for delivery of electronic registry databases to the National Archives. Where Noark 1 had simply stated that transfers of registries should happen, Noark 3 specified

how an electronic transfer of a digital registry was to happen. Nonetheless, for security and legal reasons, records themselves were to continue to be preserved in paper format for the long term (Valderhaug 2011, 103, 112–113; Vistdal 2011, 4). So, in the agencies there were two *fonds* mirroring each other, one digital and one paper. Transfers of digital records from an agency to a repository came first with Noark 4.

In 1995, Koark, an adaptation of Noark for the municipal sector, was released (Norwegian: *Kommunal standard for EDB baserte sak-/arkivsystemer*). The Central Union of Municipalities (Norwegian: *Kommunenes sentralforbund*), the interest and employer organisation of municipalities, developed Koark as Noark did not cover case handling for elected municipal public bodies. It also attempted to integrate a Noark-based case handling system with non-Noark-based business systems, such as those used within certain functions of municipalities, such as health and social services. Of course, business systems also exist in the state sector. However, the *de facto* situation was that such integrations did not really happen, whether under Noark or Koark (Valderhaug 2011, 113).

Many Noark systems are general case file systems; it was only with Noark 5 this changed to a reasonable degree (see below). Since the cases the systems handle vary by agency, it is not possible to make a few fixed work and documentation flows to cover all processes and case matters. Business systems, on the other hand, are developed exactly in those instances where there is a more specialised, predictable, and standardised process that an agency needs a system to handle, mainly because of the sheer volume of cases. In many instances, these systems were developed without due attention to archival demands, leading to a precarious situation when it came to their preservation (Fonnes 2009, 184–185). Noark and broader archival principles were simply not considered in the development of business systems.

In 1999, Noark 4 replaced both Noark 3 and Koark. By this time electronic records could be transferred in electronic format to either municipal archives or the National Archives (Valderhaug 2011, 113–115), ending, in theory, the need for paper records. ISO formats were to be the preferred archival (ingest) formats, so here one sees a slight move towards more international thinking. ISO formats were likely seen as regulated and maintained by an international standardisation agency, thus making them properly documented and robust. Noark 4 coincided with Norwegian archival legislation coming into effect in the same year, providing Norway with a legal basis for records management and archives (Vistdal 2011, 9), and making it legally required for all government agencies, including municipalities, to use Noark-based registries.

Noark 4 had its own weaknesses, combining requirements specifically centred around records management with those around case handling. The line between the two became blurred. Noark 4 became a requirement specification for a complete case handling and records management system, thus

increasing systems complexity, and becoming too costly to integrate with specialised business systems (Valderhaug 2011, 115, 127).

Noark 5: influence from abroad

Around year 2000 Norwegian archival discourse was not much influenced by international thinking. Norwegian professional discourse did not deal with the quality of archives as transactional evidence. The word transaction does not appear in Noark 4. Standardisation work in Norwegian records management was driven by a pragmatic goal: easier preservation (Valderhaug 2011, 124).

Meanwhile, much of the rest of the world was beginning to coalesce around a few important standards. The American records management standard DoD (Department of Defence (US)) 5015.2-STD (1997), influenced by the UBC Project (forerunner to the deeply collaborative and international series of InterPARES projects), was developed by the U.S. Department of Defence, and in 1998 NARA authorised it for use in the U.S. federal sector. Although it was launched before Noark 4 was finished, the American standard did not exert any influence on the Norwegian standard. ISO 15489 (2001) was too late for consideration during Noark 4 development (Sirevåg 2014, 23). Standardisation work in Norway, up until about 2000, was by and large an insular process. In May 2002 the first version of MoReq was introduced in Barcelona and was compared to DoD 5015.2-STD and Noark 4; in this comparison it was noted that Noark 4 was "highly specific to Norway"[4] (Sirevåg 2014, 24). Development of the Norwegian standard was driven by pragmatic experimentation, building on already existing traditions in the public sector, like registries and case handling. The context was national and local.

This was soon to change. International standards played a larger role in Noark 5, released in 2008, whose stated goal was to better integrate so-called specialised business systems and general case handling systems.[5] This new version strictly separated records management requirements from other requirements by placing all records management relevant demands in a section dealing with the *inner core,* which can roughly be viewed as a *fond* structure with registry, appraisal, and extraction functionality. This last requirement was vital for creating *Open Archival Information System* (OAIS)-style Submission Information Packages (SIPs). The *outer core* were *inner core* requirements for optional systems/modules that were to interact with the *inner core*. *Complete Noark 5,* corresponding to a general case handling and a record management system, specified requirements and recommendations for all modules and systems (Riksarkivet 2016, 23–24; Valderhaug 2011, 127–128). Noark 5 did to a certain extent address the old business system problem. Some business systems eventually became Noark 5 compliant thanks to integration with a core. Although many such systems continued without one, at last the possibility of an integration was established in the standard.

Noark 5 evidently drew inspiration from relevant international standards (Arkivverket 2018b, 9; Valderhaug 2011, 127). Direct influence from MoReq can be seen in certain areas. For example, the ingest format is based on objects, an aggregation of fields from different tables, not whole database tables (Haugen 2014, 10). Perhaps surprisingly, the influence went both ways. MoReq 2010, the latest edition of that standard, was influenced by Noark 5. MoReq 2010 has, like Noark 5, an inner and outer core. The reasoning here is once again to separate records management requirements from other requirements, and to integrate RM functionality with other functions or systems (Sirevåg 2014, 24). Influence from ISO 15489 and ISO 30300 can also be seen. For example, Noark 5 directly references ISO 30300 (and thus indirectly ISO 15489, since ISO 30300 uses it as a source) when it comes to the properties a record must have, namely, reliability, authenticity, integrity, and usability (Arkivverket 2018b, 10–11).

So, why was Noark 5 more influenced by international trends than earlier versions of the standard? International standards finally existed. ISO 15489 came out in 2001, two years after Noark 4. By 2008, standards such as MoReq, DoD 5015.2, and ISO 15489 were available, better known, and seen to be useful. Additionally, international standards, and in particular ISO standards, have a certain weight of authority. Such standards are made and reviewed by experts through extensive processes; they are understood to be state-of-the art products backed by expertise from around the globe. National standards based on international standards can at least partially be said to draw upon this international authority and expertise.

Professionalisation is also relevant to this discussion. By drawing on international standards, Norwegian archivists were part of a broader international community of practice. Norway's archivists could be part of broader professional trends, perhaps leading to greater cooperation outside the national borders. International standardisation can be seen as legitimising, enshrining professional values, and providing archivists and records managers with authoritative documentation that lends power to their institutional roles (Bell 2011, 38).

Perhaps no standard represents the international community of digital preservation specialists as much as OAIS, whose influence is particularly obvious in Noark 5, which relies on OAIS terminology like SIP, when it deals with extractions from Noark 5 systems. Nevertheless, OAIS and Norwegian traditions are not a perfect fit. This can be seen in the OAIS conceptualisation of Preservation Description Information (PDI), "[t]he information which is necessary for adequate preservation of the Content Information" (CCSDS 2012, 1–14). In OAIS, PDI narrowly focuses on preservation history, not records management metadata, but provides no way of capturing records management metadata other than as PDI. This is contradictory to Norwegian thinking, where the registry is a document in its own right, and merely a source of metadata for other documents. The point here is that the OAIS model did not give enough leeway to view registry data, multifaceted as

it is, both as a discrete set of information content *and* records metadata. In the end, the choice was made that registry reports (both the full and redacted versions) were viewed as content information (Riksarkivet 2010, 36). This is a small discrepancy, but it is worth highlighting as an example of how international standards and local context can clash. This will be discussed further below with reference to METS and PREMIS. The Norwegian registry tradition "did not fit the conceptual boxes" (Duff and Harris 2002, 281) of OAIS, resulting in practical choices that emphasised the national tradition rather than the international standard. This is, however, a rare example of clashing viewpoints in the Noark standard. It seems that Norwegian archivists saw OAIS as a standard that largely worked when it came to key archival concepts, and to a certain extent it does, seeing how it has conquered the archival world around the globe.

Noark 5 was intended to simplify the transfer of digital records, including registries, to the National Archives, a process described largely with reference to OAIS; in Noark 5 OAIS terms such as content information, representation information, and Preservation Description Information (PDI) with its subcategories appear regularly (Arkivverket 2018b, 94–95). The standard maps OAIS SIPs to Noark 5 extractions by, for example, identifying logs and tracking metadata in Noark 5 as PDI (Arkivverket 2018b, 95). In Norwegian archival discourse PDI is often called "technical metadata" in official reports and standards, referring to metadata that is necessary to interpret, understand, and use electronic content (Arkivverket 2018b, 95). This use of the term technical metadata seems to favour one form of representation information over another: structural metadata over semantic. This is further emphasised in two reports published by the National Archives in which technical metadata is explicitly defined as information necessary to render the files (format, structure of the file) (Riksarkivaren 2012, 13; Riksarkivet 2010, 13). One explicit technical metadata element in Noark 5 is the accompanying XML schema files that define the structure of XML files inside the submission package (Arkivverket 2018b, 95). Once again there is a focus in the Norwegian standard on the structural element. This does not mean that Norwegian archivists are not aware of semantic information and its importance; they do preserve such information (Riksarkivarens forskrift 2017, chapter 5). However, it does suggest an emphasis on the more structural elements of digital data. In a way, this conforms to a clichéd view of the archivist's focus on what "composes archives, on internal relations among descriptive items, on the structural datum" (Feliciati and Alfier 2013, 85).

The focus on structural representation information, however, can also be seen as a matter of priority. For a long time, the focus has been simply to extract content from systems, especially systems that were not standardised, such as specialised business systems. As discussed above, these systems were not created with standardisation and preservation demands in mind (Fonnes 2009, 185). So, it seems likely that the driving factor became a sort of rescue

operation in which the focus is on simple preservation. Therefore, the priority was to extract sufficient technical metadata so the records at least could be rendered.

This discussion of technical metadata, including elements defining digital structures, is representative of the kind of massaging that must happen when international standards like OAIS are drawn upon while creating national standards like Noark 5. It is included to illustrate the kind of work that must happen when international standards are built upon in local contexts.

Towards archives as data?

Data, we are told, are the new oil, an idea introduced by mathematician Clive Humby in 2006 (Viernes 2021), promising new insights through data analytics. To many Norwegians, the comparison can be bittersweet. While oil income has provided for much of the modern welfare state through an increase of Norway's wealth, the age of oil seems to be ending as the realities of climate change becomes clearer. As oil seems to enter its twilight years, a new age of (big) data seems poised to emerge, even as all nations around the world struggle to reconcile energy consumption and digital infrastructure. For Norway, as for all nations, the abundance of data can raise concerns about how to preserve data in a manner adapted to a world in need of green solutions (Pendergrass et al. 2019).

However, this chapter seeks to determine if standards are adapted to an "archives as data" viewpoint. What about Noark? Noark is based on a traditional archival hierarchy, and to all intent and purposes, it is document-centric. In this chapter, data are considered to be the discrete pieces of information that together make up a document. A document, or record, is understood to be a piece of evidence that retells or explains something, such as a minute from a meeting. If a completed form represents a record or document, the individual pieces of information submitted to fill out the form represent data.

To put it bluntly, given that Noark is document-centric, it does not lend itself easily to an archives-as-data paradigm. The Noark hierarchy stops at the document level. There is no "splitting of the atom" to access the world one step down, the data level. The registry, Noark 5's genesis, handled (paper) documents. As a result, Noark 5 is focused on documents and not data. The content in Noark 5 extractions is mostly PDF/A files, converted from Word documents or e-mails. The fact that e-mail formats, so rich in (meta)data, are "flattened" into PDF/A speaks volumes about the prevalence of the document view and the challenges of shifting to a data-centric approach in light of legacy standards like the Noark series.

One past attempt to make Noark 5 more data friendly was to experiment with a Noark 5 ontology in Resource Description Framework (RDF). This was part of a larger project that investigated new ways of records delivery to

the National Archives. The project was a joint effort by the National Archives, the Norwegian Directorate of Cultural Heritage (who provided the pilot *fonds*), and external consultants from Bouvet (Isaksen, Urtegaard and Frodesen 2015).

RDF stands for Resource Description Framework and makes use of a simple structure to interconnect data: subject, predicate, and object. A simple example can be Harald (S) *is king of* (P) Norway (O). This forms a triplet, since it consists of three parts. RDF forms the backbone of the semantic web and linked data. Thus, if implemented in RDF, Noark 5 might move away from hierarchical XML structures to a graph where each (meta)data element is a separate element interconnected with other elements through links and relations, as in ICA's *Records in Contexts* (RiC) (EGAD 2021). The Noark 5 standard would serve as the basis for an ontology in which entities and relations could be modelled in RDF.

As exciting as this is, Noark 5 RDF did not take off. As with any new technology, timing is everything. Perhaps this new approach to Noark simply was premature. Moreover, the regulations for submitting records to the National Archives were not changed to allow Noark 5 RDF SIPs; hierarchical XML files remain the explicit requirement in the standard. Although RDF could be a useful nudge towards a more data-oriented way of archiving, Noark 5 would require new regulations to fundamentally shift Norwegian practice.

In a fresh development, in spring 2023 the National Archives announced its intention to use RiC in its national portal, Digitalarkivet. This would appear to herald a shift to RDF not for transfer of records to the National Archives, as in the earlier Noark 5 RDF experiment, but rather for user access. There is unfortunately not more to add now, given how recent this still is. However, RDF seems poised to make a powerful "comeback" in Norwegian archives.

Less Noark, more experimentation?

In 2021 the National Archives decided that Noark would not receive any further development, just minor error corrections (in IT jargon: bug fixes). A more experimental phase was ushered in with *StandardLab*, where a range of stakeholders explore what kind of standardization work is needed (Arkivverket 2021). Noark 5 remains a valid standard and part of public sector archival regulation. Noark 5 systems are still in use, at least for the time being. However, from the National Archives' point of view, there will be no more development work.

Instead, agencies will have more flexibility to achieve the new concept of built-in or automatic record capture (archives by design), in which systems do not require human input to ensure that records are captured. This will require new systems and according to the National Archives, new methods that do not have to include Noark 5. There is also a push towards sharing of data "once only." Here the idea is that citizens should only give the same data one time to one agency, after which it can be shared by different agencies on an

as-needed basis. This vision requires new methods and, likely, new standards for seamless data sharing. Such reuse is permissible under the General Data Protection Regulation (GDPR), the European Union legislation regarding personal data. Although the general rule is that data are to be used for one purpose only, personal data can be reused if it is "based on the data subject's consent or on a Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard the objectives referred to in article 23(1)" (GDPR 2018, article 6), for instance protection of citizen's rights and interests (GDPR 2018, article 23). Such exceptions are not meant to be used gratuitously; juridical considerations must be followed to assure legal (re)use.

Part Two: The DIAS project and database archiving

The DIAS project

The DIAS (*digital arkivpakkestruktur*) project brought Norwegian digital archiving into line with the *Reference Model for an Open Archival Information System* (OAIS), the most important international standard for digital archiving, and one behind which the international digital archives community of practice has coalesced. The DIAS project's purpose was to define clearly information packages, a central concept in OAIS-compliant archiving, and to say which additional international standards, like METS and PREMIS, were to be used and how. It is interesting to note that Dissemination Information Packages (DIPs) are barely treated in DIAS, which is focused on a preservation perspective and not on dissemination of digital archives to researchers (Riksarkivaren 2012, 39–40). This focus on the mechanics of digital preservation over access is, perhaps at least in a historical sense, typical of Norwegian, and Nordic, digital archiving more broadly. The focus was to produce and test the extractions or SIPs; access was to be addressed later.

The DIAS project, which started in 2010 and issued its final report in 2012, had a huge impact on the Norwegian archival sector, cementing the use of OAIS and integrating other international standards including METS and PREMIS. DIAS had two precursors, Elmag 1 and Elmag 2 (Riksarkivet 2010, 4–5). While the Elmag projects introduced international standards including OAIS, METS, PREMIS, and TRAC (a precursor to ISO 16363, which concerns trustworthy digital repositories), it was DIAS that clearly defined how they were to be implemented. The Norwegian National Archives and its partners in the municipal sector did not have to look far for inspiration, drawing on work done by the Swedish National Archives (Riksarkivaren 2012, 10; Riksarkivet 2010, 37).

DIAS offers specific PREMIS and METS implementations, known as DIAS-PREMIS and DIAS-METS, largely based on Swedish practice, in another adaptation of international standards to local needs. METS, and by

extension DIAS-METS, promotes document-based thinking, through the metadata-encapsulation of traditional archival documents, and their arrangement into hierarchical structures.[6] PREMIS, on the other hand, lends itself to thinking based on data. It has its own Web Ontology Language (OWL) ontology, allowing all components to be linked to each other (for instance linking agents to preservation actions) and offering data interlinked by RDF triplets. This is more in line with the thinking behind ICA's RiC than Noark 5.

DIAS represents a Nordic and specifically Norwegian implementation of international standards. Drawing on OAIS, DIAS specifies how SIPs and Archival Information Packages (AIPs) are to be structured and introduces the concept of interlinked generations of AIPs, as content is continually migrated forward to keep up with evolving digital technologies. Each of these generations is linked as Archival Information Collections (AICs), with a number of XML schemas incorporated into METS files (Riksarkivaren 2012, 18–19). SIPs are preserved "as is," as AIP-0. DIAS further introduces a pragmatic variation on the OAIS concept of an Archival Information Unit (AIU). While an AIU is described in OAIS as containing "exactly one Content Information object (which may consist of multiple files) and exactly one set of PDI" (CCSDS 2012, 4–42), in DIAS the AIU is treated as metadata that is an addition to AIP-0. An AIU is to be based on DIAS-METS and DIAS-PREMIS, and other standards if relevant, and contain documentation relevant to ingest, such as repository testing (Riksarkivaren 2012, 17–19).

In short, through DIAS Norwegian archivists were able to build upon international standards, and especially OAIS, METS, and PREMIS, without sacrificing local needs. DIAS offers implementations of these standards, building upon past work in Norway and Sweden, and serving as a bridge between Nordic and international traditions.

Database archiving: ADDML and SIARD

Archival Data Description Markup Language (ADDML) is a National Archives of Norway XML-based standard for technical description of database extractions or datasets. Its main purpose is to describe so-called "flat files," in which data are organised either by delimiter separation or fixed positioning. ADDML describes the structure of a dataset and is also a tool to test datasets against what they are supposed to contain (Arkivverket 2020).

ADDML was developed because, as with the records management standard Noark, the National Archives felt a need and perceived a gap in the available standards. Database extractions were steadily arriving at the archives, so something had to be done. ADDML had a long path from development to use, with all versions up to 7.2 designated as development versions that saw no actual usage. The standard was developed alongside Arkadukt, the software used in creating and editing technical metadata descriptions in ADDML. It was not until the year 2000 that the standard was implemented in practice. The

Swedish and Finnish National Archives subsequently implemented ADDML and joined in the development of the standard (Arkivverket 2019, 5).[7]

ADDML was for a time the only available standard for its purpose. However, an alternative came with SIARD. Published by the Swiss Federal Archives in English in 2008 (SFA 2008), it quickly become an international standard used in several countries. ADDML was not actively promoted by the National Archives internationally and was not picked up outside Norway except in Sweden and Finland. SIARD had English documentation early on, while documentation for ADDML was in Norwegian. Moreover, adopting SIARD was made easy thanks to the SIARD suite of software, provided freely by the Swiss Federal Archives. A SIARD dataset includes the content from all tables (one XML file for each table) as well as technical metadata about the tables (Library of Congress n.d.). With the SIARD suite, users can extract database content and add additional metadata for long-term description. SIARD has also been implemented in the software tools Database Preservation Toolkit and Spectral Core, further extending its support. While tools that implement ADDML were developed, like Arkade 5, a test programme for datasets (recently expanded to SIARD datasets), and the aforementioned Arkadukt, both these tools were developed to fit Norwegian needs.

Arne-Kristian Groven at the Norwegian National Archives played a major role in bringing SIARD to Norway through a project in 2012–2013. Municipal archives quickly began to look at SIARD with interest. Not long after the National Archives approved SIARD as a legal extraction format (Riksarkivarens forskrift 2017, §5–12 f). Here one can see push and pull factors towards SIARD and away from ADDML. The SIARD suite took a "straight from the box" approach that made the Swiss standard simpler for non-technical archivists to use than ADDML. While ADDML is very flexible, offering a range of implementation options, making use of these options requires deeper knowledge of digital archival practices and technologies than equivalent implementations of SIARD. What SIARD lacks in flexibility it makes up for in a very standardised approach, which anyone with some basic knowledge about databases can understand. Additionally, as an international standard SIARD opens new avenues for cooperation and exchange of information. ADDML continues to be supported for the time being, and its documentation has to a certain degree improved (Arkivverket 2019). ADDML is still an official DIAS standard, and it still has a role in managing Noark 5 datasets. Nonetheless, municipal archivists still prefer SIARD. To a certain extent, comparing ADDML and SIARD is like comparing apples and oranges. ADDML was constructed before there were any other alternatives, and has, thanks to its flexibility, a different and broader field of use than SIARD.[8]

Whatever their differences in usage and implementation, SIARD and ADDML share one important similarity: they are both data-oriented. Metadata about content data elements is found in many specialised business systems and databases and can be documented in ADDML or SIARD extractions. While

each system will not have the same metadata, many agencies or municipalities bought the same system from the same vendor. This led to the creation of Documaster Decom, a programme that creates standardised documentation of each system and its information elements in either SIARD or ADDML (Documaster 2019). Development of Documaster Decom was driven by municipal archives and the vendor Documaster. Whether this approach is something that can be exported to other countries remains to be seen. In Norway there are not many vendors that provide specialised business systems, or for that matter, Noark systems. Relatively limited competition made the creation of templates a manageable task.

In the end, it remains to be seen whether SIARD will completely take over from ADDML in Norway. Uptake of SIARD could become a self-strengthening circle: SIARD is being used, thus SIARD must be used more. However, SIARD is adapted to relational databases. The next generation may be graph databases for all we know. Perhaps ADDML, with a few adjustments, can again help preserve such database content? Or will new standards arise in Norway or elsewhere? The future is uncertain.

Conclusion

In the last decades of the twentieth century, Norway sought to realise the benefits of standardised records management and archiving in the public sector. Since there were not international standards available at that time, Norway created its own standards, including Noark and ADDML. Since then, robust international standards such as ISO 15489, SIARD, and the suite of standards associated with OAIS have come into common use among records managers and archivists. Norway is now faced with a choice: whether to maintain its national standards or to embrace international standards. Ultimately, the same motivations that drove the creation of national standards – including ease of information exchange among peer institutions and ease of access by members of the public – are now driving the adoption of international standards.

There is no single, clear-cut answer to the question of whether to adopt international standards or maintain national standards. As seen in Noark and the DIAS project, international standards like OAIS, METS, and PREMIS can be adapted to local needs. The result can be a kind of national "flavour" of the international standard. Moreover, continuing use of ADDML alongside SIARD suggests that international standardisation is not inevitable. Just as the DIAS project demonstrated the need for local implementations of international standards, continuing use of ADDML demonstrates a role for national standards.

Similarly, there is no single, clear-cut conclusion about whether Norway will embrace Mordell's archives-as-data paradigm. Noark and DIAS offer conflicting evidence of a sector-wide shift towards archives-as-data, as both in different ways continue document-centric traditions of archiving. On the other hand, ADDML and SIARD support such a development. Since they

deal with data from databases this is not surprising, and perhaps points towards the conclusion that Mordell's archives-as-data paradigm functions best when dealing with data that was created, managed, and archived as data, rather than data mined out of other kinds of records. There is little need to discard everything old. Change requires careful thought and experimentation to know what viewpoints and methods are still useful in our continually evolving digital world. At the same time, the National Archives seems recently to have edged closer to an archives-as-data, with its choice to build on ICA's RiC standard in its national portal, Digitalarkivet.

It is here that the National Archives own experiment comes in. Saying that Noark will no longer be updated ushers in a new era of experimentation. Public agencies, vendors, and the National Archives will together explore new methods of records management, not least in record capture and disposition. Where Noark was once required for everyone, now everyone can do as they please – within reasonable limits and including to continue with Noark. Perhaps this new approach will result in more sector or agency-specific solutions, but that remains to be seen. Standards will continue to play a role, not least because of the need to share data, but perhaps more standards, and different standards, will come into play. The public sector is freer than ever before to explore new ways, including data-driven ways, to create, capture, and use data and records. How it will use that freedom, and if that leads to more, or less, standardisation and a strengthened movement towards archives-as-data, remains to be seen.

Notes

1 It should be noted that in Norway the word *arkiv* is used both for archives and records management. There is no strict separation of the two as one might find in an Anglo-American context. Recently though, the term *dokumentasjonsforvaltning,* which can be translated to records management, has been introduced into Norwegian discourse.
2 This requirement extends to the private sector in limited instances – for example, when there is sufficient public ownership of a company, or when stipulated in contracts between government agencies and the private corporations that, increasingly, are involved with providing public services.
3 In digital registry systems used today, case handling and the records themselves are more intertwined and directly interlinked than in the paper world. For the general user, everything is in the same "place" and accessed through the same user interface.
4 Two specifically Norwegian functions in the standard were *avskrivning* and *periodisering*. *Avskrivning* is information that says how an incoming letter was handled, like for instance answered by an outgoing letter. *Periodisering* is dividing up the *fonds* in distinct time periods.
5 With Noark 5, a small change was made. Noark was hereby to be called *Norsk arkivstandard* (Norwegian archive standard), not *Norsk arkivsystem*, the name used in previous Noark versions, highlighting its change from a requirement specification to a standard.
6 This point is based on e-mail correspondence October 2021 with Herbjørn Andresen, professor of archival sciences, OsloMet.

7 I am uncertain about the extent of use ADDML has in these two countries today.
8 This point is based on e-mail correspondence from May 2021 with Thomas Sødring, associate professor of archival sciences at OsloMet.

References

Arkivverket. 2018a. *Tidligere NOARK-versjoner*. Last edited December 6, 2018. https://www.arkivverket.no/forvaltning-og-utvikling/noark-standarden/tidligere-noark-versjoner#!#block-body-1.

Arkivverket. 2018b. *Noark 5, versjon 5*. https://www.arkivverket.no/forvaltning-og-utvikling/noark-standarden/noark-5/noark5-standarden/_/attachment/download/f2f1eab5–55c4–475a-8f06-f7f645c4dd7a:0b3cb47721fd3e416629b3b21b913621c4028d2f/Noark%205%20v%205.pdf.

Arkivverket. 2019. *ADDML 8.3 Archival Data Description Markup Language, versjon 8.3. Generell del*. https://www.arkivverket.no/forvaltning-og-utvikling/regelverk-og-standarder/andre-arkivstandarder/addml-archival-data-description-markup-language/_/attachment/download/e3e195d0-aa21-4df3–99a8–31096f3e4525:c0a09d695060c1cae7e5b820feb3bca97c50d049/ADDML%208-3%20Generelt.pdf.

Arkivverket. 2020. *ADDML (Archival Data Description Markup Language)*. Last edited October 27, 2020. https://www.arkivverket.no/forvaltning-og-utvikling/regelverk-og-standarder/andre-arkivstandarder/addml-archival-data-description-markup-language.

Arkivverket. 2021. *Arkivverket avslutter utviklingen av Noark-standarden*. Published June 10, 2021. https://www.arkivverket.no/nyheter/arkivverket-avslutter-utviklingen-av-noark-standarden-copy.

Bell, Alan. 2011. "Standards and standards culture: Understanding the nature and criticisms of standardisation." *Comma* 2011(2), 25–39. 10.3828/comma.2011.2.02.

Consultative Committee for Space Data Systems [CCSDS]. 2012. *Reference Model for an Open Archival Information System (OAIS)*. Second Edition. https://public.ccsds.org/pubs/650x0m2.pdf.

Documaster. 2019. *What is Documaster Decom?* Last edited August 26, 2019. https://decomsupport.documaster.no/nb-NO/support/solutions/articles/43000494288-what-is-documaster-decom-.

Duff, Wendy, and Verne Harris. 2002. "Stories and names: Archival description as narrating records and constructing meanings." *Archival Science 2*(3), 263–285. 10.1007/BF02435625.

Expert Group on Archival Description [EGAD]. 2021. *Records in Contexts—Conceptual Model, consultant draft version 0.2*. https://www.ica.org/sites/default/files/ric-cm-02_july2021_0.pdf.

Fonnes, Ivar. 2009. *Arkivhåndboken for offentlig forvaltning*. Second edition. Oslo: Kommuneforlaget.

Feliciati, Pierluigi, and Alessandro Alfier. 2013. "Archives on the web and users expectations: Towards a convergence with digital libraries." *Rev. Nat. Cent. Digitization 22*, 81–92. https://www.researchgate.net/publication/285906869_Archives_on_the_web_and_users_expectations_Towards_a_convergence_with_digital_libraries.

Forskrift om utfyllende tekniske og arkivfaglige bestemmelser om behandling av offentlige arkiver [Riksarkivarens forskrift]. 2017. https://lovdata.no/dokument/SF/forskrift/2017-12-19–2286.

Hamill, Lois. 2017. *Archival Arrangement and Description: Analog to Digital*. Lanham: Rowman & Littlefield.

General Data Protection Regulation [GDPR]. 2018. https://gdpr.eu/tag/gdpr/.

Haugen, John Atle. 2014. "En sammenligning mellom Noark 5 og MoReq2010—Og mellom norsk og internasjonal arkivtradisjon." In *Dokumentasjonsforvaltning og arkiv i det 21. Århundre*, edited by André Neergaard, 9–27. Bergen: Fagbokforlaget.

Isaksen, Øystein, Gunnar Urtegaard, and Kari Frodesen. 2015. *Rapport Samdok 2014 –Synkron avlevering (STAT) Riksantikvaren (RDF)*. https://docplayer.me/36825592-Samdok-samla-samfunnsdokumentasjon.html.

Library of Congress. No date. SIARD (Software Independent Archiving of Relational Databases) Version 1.0. *Sustainability of Digital Formats: Planning for Library of Congress Collections*. https://www.loc.gov/preservation/digital/formats/fdd/fdd000426.shtml.

Mordell, Devon. 2019. "Critical questions for archives as (big) data." *Archivaria 87*, 140–161. https://archivaria.ca/index.php/archivaria/article/view/13673.

Miller, Thea. 2003. "The German registry: The evolution of a recordkeeping model." *Archival Science 3*(1): 43–63. 10.1023/A:1024992405980.

Pendergrass, Keith, Walker Sampson, Tim Walsh, and Laura Alagna. 2019. "Toward environmentally sustainable digital preservation." *The American Archivist 82*(1), 165–206. 10.17723/0360-9081-82.1.165

Riksarkivaren. 2012. *Vel bevart! Rapport fra samarbeidsprosjektet DIAS – Digital arkivpakkestruktur*. https://www.arkivverket.no/forvaltning-og-utvikling/regelverk-og-standarder/dias-prosjektet-digital-arkivpakkestruktur/_/attachment/download/c1ce3bc3–91f7–434a-83f9–6053b78ff93e:8ce35b2ec0fdb612b7ce318d977eac9ba2c57fbd/DIAS_Sluttrapport-Del1.pdf.

Riksarkivet. 2010. *Digitalt og autentisk—Planlegging av ny depotløsning for Arkivverkets digitalt skapte arkivmateriale*. https://www.arkivverket.no/for-arkivsektoren/langtidslagring/_/attachment/download/46508a47–640d-4f22–839d-40de52941315:22bc01443f9eaa09d320a9d03c48941d1ba78e61/Elmagrapport%202010%20(fullversjon).pdf.

Riksarkivet. 2016. *Noark 5, versjon 4.0*. https://www.arkivverket.no/forvaltning-og-utvikling/noark-standarden/noark-5/noark5-standarden/_/attachment/download/4f704db6-fe38–40cc-ad61-8c2a3fe45e64:9ca43b8ae474015270f3ec76ece9dc455d4e76d3/Noark5v4%20Standard%20for%20elektronisk%20arkiv%20(01.12.2016).pdf.

Swiss Federal Archives [SFA]. 2008. *SIARD Format Description*. https://github.com/DILCISBoard/SIARD/blob/master/SIARD%201.0/format/2008-09-30/SIARD%2BFormat_en.pdf.

Valderhaug, Gunnar. 2011. *Fotnote eller tekst? Arkiv og arkivarar i det 21. hundreåret*. Oslo: ABM-Media.

Viernes, Francis Adrian. 2021. "Stop Saying 'Data is the New Oil'." *Geek Culture*. Published September 15, 2021. https://medium.com/geekculture/stop-saying-data-is-the-new-oil-a2422727218c.

Vistdal, Oskar. 2011. "Forvaltnings- og arkivrutiner i historisk perspektiv—Brudd eller sammenheng?." *Tidsskriftet Arkiv 2*, 1–11. 10.7577/ta.911.

Sirevåg, Trond. 2014. Utviklingen av Noark-standarden 1984 – 2008. *Tidsskriftet Arkiv 5*, 1–26. 10.7577/ta.1149.