

On the Breakdown of the Controlled Environment Paradigm in Norwegian Archival Repositories

Herbjørn Andresen¹[0000-0003-4137-4263]

¹ OsloMet – Oslo Metropolitan University, Oslo, Norway
Herbjorn.andresen@oslomet.no

Abstract. The rules and arrangements that govern transfer of digital records to the archive repositories in Norway, rely on a controlled environment paradigm. This paradigm is the basis for assuming the authenticity and evidential values of the archives. The concept of a paradigm is borrowed from the theory of science, but it can also be relevant to fields of practice. In the theory of science, it denotes distinct concepts, methods and thought patterns that guide what contributions are perceived as valid within a field. If or when a paradigm ceases to provide adequate guidance in a field, it may break down, possibly leaving the field in need of a new paradigm. The discussion in this paper apply a theory of different responses to paradigm breakdowns in order to explore stakes and opportunities at a point of crisis for the current paradigm on transfer of born-digital records to archive repositories.

Keywords: Digital Repositories, Archives Life Cycles, Theory of Science.

1 Introduction

The incitement for this short paper is a recent Norwegian proposition for new legislation on records and archives [1]. The proposition report contains bold changes in many areas, most of them left out from this paper. One of the areas that is discussed, albeit very briefly, is a possible need to change the ways transfer of born-digital records to the repositories is arranged and controlled. The current rules require the semantics of an electronic archive to be interpreted, tested and determined at transfer time. Later use of the archives must, and do, rely on the transformations that have been carried out earlier. The discussion, contained in a brief paragraph on the concept of data lakes, suggests transformations, and with them interpretations and verifications, might better be postponed to the time of use. This would have an impact on the entire basis for assuming authenticity and evidential value of records in the archive repositories. That is probably a good thing.

2 Theoretical Perspective, Responses to Paradigm Breakdowns

The rudimentary discussion in this paper labels the existing rules and assumptions on transfer to repositories in Norway a “controlled environment paradigm”. An account of

its main ingredients is given below in section 3. The discussion itself draws on the concept of paradigms, known from the theory of science. In Kuhn's famous book on the structure of scientific revolutions, paradigms are defined as "universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners" [2]. Following such an achievement, further contributions are made as "puzzle-solving" within the confines of the paradigm that was established. A working paradigm is a mechanism for building cumulative knowledge effectively. The paradigm provides concepts, methods and thought patterns that guide what is seen as legitimate contributions to the knowledge base, and what is not. However, the paradigms do not last forever. The clause element "for a time" is a crucial part of Kuhn's definition. When new results or knowledge do not fit well into the paradigm, they are perceived as anomalies, often not valued on their own merits, but instead judged as incompatible with the paradigm. If anomalies are persistent, and throw important assumptions of a paradigm into doubt, the paradigm may break down. Crises are a necessary precondition for the emergence of novel theories, or paradigm shifts, in Kuhn's terminology.

Kuhn includes in his writing some reflections on responses to crises, intertwined in historical examples. For the purpose of this paper, responses to paradigm breakdowns are instead taken from another paper, by E.R. Alexander, that develops a more stringent typology of different responses to paradigm breakdowns [3].

Alexander labels the first type of response 'ritual response'. This is a pattern of responding in denial of the breakdown. Even though the evidence of an untenable paradigm emerges, the ritualist will ignore the crisis and cling to the existing paradigm.

A second type is the 'avoidance response'. Avoiding is not denial. The evidence of a crisis in the current paradigm is acknowledged, by accepting that it is happening. An avoidance response will often even emphasize the anomalies, but keep perceiving them as anomalies viewed from within the paradigm. Though the breakdown as such is recognized, the practical implications are ignored.

A third type is the 'abandonment response'. The breakdown is recognized, and current models are abandoned as they appear unnecessary and dysfunctional. An abandonment response could for example be framed as a determination to "adopt the intuitive wisdom of the practitioner" and stay away from grand theories. By rejecting the need for model problems and solutions, the mechanism for judging new contributions will also be lacking.

The fourth type of response is the 'search response'. The starting point for this response is to recognize the breakdown for what it is, and at the same time acknowledge the need for new model problems and solutions. Even though, as according to Kuhn's definition, one is aware that new paradigms are also only for a time.

3 The Controlled Environment Paradigm in Norway

Rules on how government agencies should create and organize their records have been around in Norway since early 18th century. At the heart of these rules was the *registraturprinzip*, a system of creating current registries organized by cases, and archiving

according to the original order. This principle has remained prevalent for most of the public records practices.

From an early start in the mid-20th century, the National Archives has also, increasingly, given guidance and exercised powers of control and auditing over the records practices in government agencies. Therefore, a regime where the different government agencies adhered to rules and conventions imposed by the National Archives had already been established when the first version of a national standard for electronic records management systems was issued in 1984 [4]. This implied an expectation that electronic archival materials ending up in their repositories would be based on structures and functionalities that safeguarded reliable content, of known provenance and internal structure, and with a high perceived evidential value. 1984 appears to be really early for a standard on electronic records systems, but the first versions only covered registry information. The actual documents were still transferred on paper, until the 1999 version of the standard allowed for transferring electronic documents, either as scanned picture files or as output files from word processing systems or similar.

Also, in 1999, to maintain the control over records creation in electronic systems, the National Archives introduced a scheme for approving electronic records management systems to be used by government agencies. Each agency can buy a system from some vendor in the market, or develop the records systems themselves, but they will have to keep paper files unless the system they use have been approved.

Even before the OAI reference model [5], the National Archives had decided the transfer of electronic records to the repository should only be a transfer of the data content and context information, and not a transfer of systems hardware or executable program needed to display or render the records. The government agencies, when they are prompted to transfer records, produce an extract from the system conforming to extraction requirements in the national standard [4]. In the vocabulary of OAI [5], they produce a SIP, a submission information package.

After the SIP has been transferred, the National Archives performs a testing procedure, to verify if the submitted package conforms to the prescribed structure and metadata requirements. After passing these test procedures, the National Archives approves the transferred archival unit, and accept a conferred responsibility for the further maintenance and adherence to access rights and limitations and whatever else an archival institution assumes responsibility for.

The controlled environment paradigm builds in part on the Jenkinsonian notion of an unbroken chain of custody as a warrant for the evidential value of the archives [6]. The requirements pertaining to born-digital materials transferred to the National Archives repository are modelled on a similar assumed need for an end-to-end control throughout a linear life cycle of the records. Jenkinson did not, however, support the view that archival institutions should intervene in the records management activities before transfer. Other theorists have spoken in favour of such practices, for instance Terry Cook, "...no reliable record will even survive to be available to the archivist to preserve in the traditional way unless the archivist intervenes in the active life of the record, sometimes before it is even created" [7].

4 Anomalies in The Controlled Environment Paradigm

The controlled environment paradigm, as ordained in Norway, puts a huge demand for quality and compliance in a long series of complex processes. The first severe signs of anomalies showed up in the middle of the chain, as errors in the run of the test procedures to verify the submitted transfers. There were more often than not a high number of errors, such as missing metadata, inconsistent data, structural flaws and so on. Quite often the errors were not imposed by wrong implementations of the records management systems or extraction procedures, but in the recordkeeping practices.

One example of perceived errors in the records, that were often discovered at the time of testing transfers, was case files “left open” in the records management system. A demand imposed by the ancient *registraturprinzip* is that case files must be closed. This is a step that users of the system might not even be aware of. The immediate interpretation of such errors was that the systems, even though they adhered to the national standard, were used in a wrong way. When it turned out such errors were frequent and pervasive, it was no longer tenable to view them as mere user errors. In some cases, it might even be counter-intuitive to close case files. The interpretation of such errors encountered in the testing of transfers shifted gradually, from perceiving them as non-conformance to perceiving them as possible anomalies to the controlled environment assumption.

The disheartening test results caused time-consuming, and therefore expensive, iterations of sending new corrected versions of the extracted transfer files back and forth. A very strict adherence to the commitment to an unimpeded controlled environment made it virtually impossible to get archives through to a conferred status. A possible side effect of the slow transfer and acceptance process may have been to slow down the deployment of more innovative digital preservation technologies, therefore what might be viewed as an avoidance response may in fact be influenced by the available tools. A part of the National Archives’ solution to the acknowledged anomalies was to relax the testing procedures, in two aspects. First, the requirements for consistency and accuracy in the records metadata could be a bit more flexible than the requirements laid down in the standard [4]. Second, and slightly more heretical to the controlled environment paradigm, minor and obvious errors in the extractions could easily be corrected by the National Archives testing unit themselves, instead of requiring the record creating agency to submit a revised version of the extract. The limit to what corrections may be made by the testers without impeding the warrant of authenticity is hard to define in a principled way. Instead, there has been some practice-based rules of thumb to draw this line.

The problems of achieving an acceptable test result revealed further problems to the asserted controlled environment from records creation to continuous custody. When extracts were sent back and forth to make it pass the tests, the necessary corrections would have to be made either to the actual records, or to the records management system, or to the extraction procedure generating the extracted files to be transferred. Making changes to the actual records, for instance closing case files that the records manager had left open, would solve the problem of consistency, but at the cost of no longer reflecting the actual case-handling practices in an accurate way. If the problems were

either in the records management system or in the extraction program, it would be the same vendor who interpreted the error reports from the repository, and who tried to come up with solutions. The actions of the programmer are not necessarily transparent, even if the records management system as such has been approved at an earlier stage. It might be hard to tell whether any tweaking of the extraction procedure in order to make the transfer pass the test blurs the strict requirements on how the extract to be transferred should correspond to the actual records.

The anomalies as a fundamental problem proliferated mainly from the middle of the chain, verification of transfers, to the antecedents in the records management and extraction processes. However, there are also some concerns on the other side of the transfer process, to the usability of the transferred archives. Even though a number of extracted, born-digital archives by now have been received and verified, they appear to be hard to put into use beyond retrieving specific records residing in a known archival information package. The semantics of the records management systems have been evolving, even with a common standard for records management systems, different archives are not easily linked or compared. Semantic structures and metadata schemes that are interpreted, fixated and closed at transfer time do not provide much leeway for exploring possibly interesting combinations of data from different parts of the archives. The limits on usability have not been explored to the same extent as the anomalies of records management, extraction and transfer.

5 Perceiving and Responding to the Anomalies

The first perceptions of problems with the value chain for electronic archives were in the low efficiency of testing and verifying transfers. At first, this was handled within the paradigm, as a slow process, taking as long as it had to, in order to keep in line with the controlled environment paradigm. It was thought to be a learning process, soon ready for speeding up. The anomalies were perceived as such when decisions were made to relax the tests in order to speed up the transfer process. The profoundness of the problems was discovered and understood gradually. The anomalies were persistent and amounted to a crisis for the paradigm.

Responses to the crisis have been varying and have included all the four types listed in section 2 above. The ritualist response was seen when the problems were mainly dealt with at a department level, as a need for improvement, and taking the necessity of the controlled environment paradigm for granted. A choice between the three other different responses require an acknowledgement of the crisis, severe difficulties in maintaining the controlled environment from records creation to continuous custody for born-digital materials as a warrant for the evidential value. At the organization level, a project was established to present and compare different model concepts for the transfer and maintenance of archives [8]. The project looked into different ways to improve the transfer process, including relaxing requirements and accepting more diversified submissions. Still, none of the proposed model concepts were clear and outspoken departures from the notion of a controlled environment, where the evidential value depends on adherence to the tenets of this paradigm.

The model concept that was eventually chosen, was named the “zero plus”-alternative. Essentially, this alternative implies keeping the same processes going, the “plus”-part signifying a need for systematic work on improvement. The outcome of this evaluation process, viewed as a response to a paradigm breakdown, is the avoidance option.

6 Conclusion

The project for modernizing receipt and ingestion of born-digital materials was never in denial of the anomalies vexing the current paradigm. It was, in a way, a search response, ending in avoidance. This is understandable, and not really deserving of too much criticism. The archival community would need a very strong stomach to abandon a controlled environment paradigm, as it would put the justification of the whole endeavor in jeopardy. The only viable alternative to avoidance therefore seems to be the search response.

The proposed new legislation on archives in Norway, that was briefly mentioned in the introduction, is promising. Not because it necessarily provides clear solutions to how the breakdown of the controlled environment paradigm should be handled, but because it indicates that the search for a new and viable paradigm can still be an option.

References

1. Official Norwegian Report. NOU 2019:9 Fra kalveskinn til datasjø. Ny lov om samfunnsdokumentasjon og arkiver. [“From Vellum to Data Lakes”. Proposal for enactment of new legislation on societal documentation and archives]. 2019.
2. Kuhn TS. The structure of scientific revolutions. 2nd ed. Chicago: University of Chicago Press; 1970.
3. Alexander ER. After Rationality, What? A Review of Responses to Paradigm Breakdown. *Journal of the American Planning Association*. 1984;50(1):62-9.
4. National Archives Norway. NOARK Norsk arkivstandard [Norwegian national standard for public sector electronic records and archives]. (First edition 1984, 5th generation from 2008, current version is version 5.5).
5. ISO 14721:2012 Space data and information transfer systems – Open archival information system (OAIS) – Reference model. Geneva: International Organization for Standardization.
6. Jenkinson H. A manual of archive administration including the problems of war archives and archive making: Clarendon Press; 1922.
7. Cook T. What is past is prologue : A history of archival ideas since 1898, and the future paradigm shift. *Archivaria*. 1997(43):17-63.
8. National Archives Norway. Konseptutredning om modernisering av arkivvedlikehold og overføring til arkivdepot (MAVOD) [Report on evaluation of concepts for modernizing archives maintenance and transfer to archival repositories]. 2018.