PROJECT MUSE®
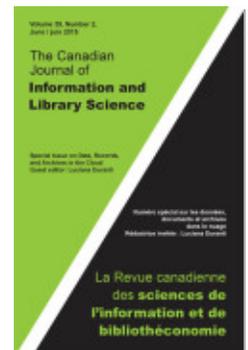
Public Cloud Archives: Dream or Reality? / Les archives publiques dans le nuage informatique: Rêve ou réalité ?

Anna Sobczak

➡ *For additional information about this article*
https://muse.jhu.edu/article/590942

Public Cloud Archives: Dream or Reality?

Les archives publiques dans le nuage informatique : Rêve ou réalité ?

Anna Sobczak
Szczecin University
amsobczak@gmail.com

Abstract: The German Landesarchiv Baden-Württemberg (State Archives of Baden-Württemberg) developed its own software called Digitales Magazin (Digital Storeroom) to appraise, acquire, manage, describe, and provide access and long-term preservation of different kinds of electronic records. In an effort to develop the system further and to offer it to other state archives as well as small public or private ones, several new cooperation possibilities were formulated, involving the use of the software among other in a software as a service or developer model.

Keywords: Germany, state archives, small archives, cloud computing, services, non-commercial software, SaaS

Résumé : Les archives d'État du Land de Bade-Wurtemberg en Allemagne (Landesarchiv Baden-Württemberg) ont développé leur propre logiciel appelé Digitales Magazin (Entrepôt numérique) pour évaluer, acquérir, gérer, décrire, fournir un accès et l'archivage à long terme de différents types de documents électroniques. Dans un effort pour poursuivre le développement du système et de l'offrir à d'autres archives d'État ainsi qu'à d'autres archives publiques ou privées de petite taille, plusieurs possibilités nouvelles de coopération ont été formulées, impliquant l'utilisation du logiciel dans un modèle SaaS (software as a service) ou dans un modèle développeur.

Mots-clés : Allemagne, archives d'État, petites archives, nuage informatique, services, logiciels non commerciaux, SaaS

The aim of this article is to present models of partnerships in digital preservation and building digital archives designed by the Landesarchiv Baden-Württemberg (State Archives of Baden-Württemberg). In the Federal Republic of Germany, there are more than 2,300 archives. They can be divided into the following categories: national, municipal, city, academic and research institutions, mass media, church, political organizations, parliament, business, private, and family/house. This article focuses mainly on state archives, which form part of the national archives group. Based on the administrative organization of Germany, there are sixteen such institutions, each located in one state (*land*).[1]

The Landesarchiv Baden-Württemberg is located in the southwest of Germany and is one of the most information technology (IT)-oriented archives in the country. It can be treated as one of the leaders in digital preservation, together with other archives, such as the Federal Archives (Bundesarchiv). It is responsible for records produced by state public agencies and the local community. The archives possess the oldest electronic holdings in the German Federal Republic (Naumann 2007, 53), which include data (statistics, pictures, and procedures) of the local census dated from 1961 to 1970.[2]

The problem of electronic records already began to emerge in the 1960s in both the Federal Republic of Germany and the German Democratic Republic. At that time, archivists in the German Democratic Republic were contemplating which storage medium would be best for data preservation,[3] while those in the Federal Republic of Germany started to consider how to preserve and store data created in a digital environment. Ideas ranged from printing everything out or reproducing it on microfilm,[4] but it was finally decided to preserve the material in the native form—that is, digitally. The Federal Archives put forward a proposal to acquire all digital records produced, but a strong federalist spirit made this idea impossible to realize. However, since the time that electronic records began to be transferred to archives, archivists started to develop electronic tools to manage them and protect them from obsolescence (Ullmann 1998, 599–600). Later, they also realized that archivists needed to be in contact with each other and with the administration when the records were being created and even during the design and implementation of systems to come to an agreement on technical issues, such as file format, depth of metadata description, and so on (Kluttig 2006, 56–58).[5]

Some of the first repositories for electronic records[6] were created in Germany by the Landesarchiv Baden-Württemberg (State Archives of Baden-Württemberg) and the Bundesarchiv at the beginning of the twenty-first century. The design was based on the Open Archival Information System (OAIS) model, and it covered the following areas: ingest, storage, management, administration, access, and preservation of trusted data (Keitel and Lang 2010, 55; Keitel 2013b, 148–49). However, the beginning of digital repositories, in this author's opinion, can be traced back to the first databases in which the descriptive metadata of records were collected; these later developed into much more advanced data systems.

Generally speaking, there are three concepts of digital preservation known to German archivists: by the creator of the records, by a third party (service provider), or by public archives (National Library of Australia 2003, 58–59). The first idea was rejected at the very start of discussions and the last one has gone on to become the most common in Germany.

Digitales Magazin (DIMAG) was started in 2006 as a project called Concept of a Digital State Archives (Konzeption für ein digitales Landesarchiv),[7] shortly after the first electronic records were transferred to the archives, and some reflections on digital archiving were made in 2002 by Christian Keitel.[8] At that point, there were no preservation systems for archives, and archivists only heard of electronic repositories in libraries. Thus, these were used as a basis

for further discussion on the matter, despite differences existing on the preservation of digital material between the two types of institutions. In the following years, archivists were trying to define how to preserve a new kind of material, while maintaining its connection to the analogue-based primary holdings, and they considered which metadata standards should be used (for example, OAIS, METS, Nestor manuals, ISAD (G), EAD, PREMIS, or the National Library of New Zealand Metadata Model). The initial step was research, which led to the identification of all kinds of records created by the local public administration. After this very important stage, the production of digital repositories went on track. At the beginning, DIMAG was equipped with two modules: ingest and registration of records. The access point module was added later, and further file formats were recognized and supported. The final version featured the following sections: appraisal, description, access, storage, and preservation. The hierarchical structure of the storeroom allowed searches based on its tectonics: fonds, series, subseries, units, objects, and their properties defined in technical metadata.

This software has been in use since 2006 and presently can operate with materials of different file formats (limited) and functionalities such as records, intranet sites, websites, audio-visual, and data extracted from databases (for example, land register files).[9] Materials stored in the repository can be digitally born, therefore created natively in an electronic environment or digitized from analogue material. Apart from establishing the repository, several important documents were prepared, such as *Records Digital Preservation Metadata* (*Metadaten für die Archivierung digitaler Unterlagen*)[10] and other tools for automated batch identification of file formats and properties, called IngestList.[11] As a result of cooperation with other archives (Hessische Staatsarchive, die Staatliche Archive in Bayern), new modules, such as the automatic Submission Information Package for ingest were planned (Keitel 2013a, 54; Keitel 2013b, 147–51; Kemper and Naumann 2013, 86).

DIMAG was meant to be easily accessible via a standard Internet browser and configurable for different needs and users. It meets the purpose of acquiring archival material from several archives in one storage space, makes unauthorized access impossible, and can be offered as a final product for self-installation and use or can also be used as a service to other archival institutions, in particular smaller ones (Keitel 2013a, 56–60).[12]

Both concepts of sharing software or providing it as a service are not new but are still not very common. These concepts were developed in Germany in relation to the requirements of archival science, its methodology and the national law, which determines how the software needs to function. There are some examples, even worldwide, of archivists sharing software, such as open-source International Council on Archives (ICA) AtoM[13] web-based description software developed by Artifactual Systems in collaboration with the ICA and other international parties. The success of sharing this software is based on the application of international descriptive standards and interface multilingualism.

The idea of offering software as a service (SaaS) is an opportunity for institutions that are required to preserve digital records but have small budgets for

IT solutions and have no or poor know-how.[14] This service is based on principles of cloud computing.[15] It generally means that clients receive their virtual space that can be accessed remotely. This solution is of course connected with some regular costs, but it is cheaper than creating one's own infrastructure and providing for its own long-term maintenance (Keitel n.d., 1–2).

The many inquiries about digital preservation received by the Landesarchiv Baden-Württemberg demonstrate that not every archive can afford, or would want to create, its own digital repository in the future, mainly due to financial reasons. They also proved that German archivists need to agree on approaches on collaborative preservation and license the software widely in the German archival world. Not only would this give an opportunity to test DIMAG, but it would also help to find co-developers. First attempts to promote the software and find potential collaborators began in 2009. One year later, the software was used by the Hessisches Hauptstaatsarchiv Wiesbaden (Hessian State Archives Wiesbaden) and, in 2012, by the Staatliche Archive Bayern (Bavaria State Archives), including five more state archives from Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, and Schleswig-Holstein in 2014.[16] At present, there are four kinds of partnerships: repository (*Magazinpartnerschaft*), development (*Entwicklungspartnerschaft*), support (*Supportpartnerschaft*), and provider (*Dienstleistungspartnerschaft*) partnership. The first one is dedicated to archives that have no resources to maintain the software or IT infrastructure on their own but are able to pay a fee for space in the digital storerooms of other archives. The second one is for those who are ready to co-develop the software, for example, by programming new modules. The third kind of partnership offers software for installation on an archives' own IT infrastructure with a limited period of technical support (including installation and updates). The last one suits those archives that want to use digital preservation as a service delivered by external data processing centres. It is very similar to the first option.

The partnerships are planned as business models. Software is never given for free and always involves legal issues. Participating archives have to contribute to the costs of building DIMAG and at the very minimum accept the main technical solutions. They are also not allowed to decompile the source code, license it, pass it on to third parties, or do anything that could be perceived as competitive behaviour (Keitel 2013a, 55; Keitel 2013b, 152–55). As developers, the Hessian and Bavarian state archives create tools for ingest and access.

To put the idea of DIMAG into action as a service, cooperation with central and local data centres belonging to the Data Processing Network Baden-Württemberg (Datenverarbeitungsverbund Baden-Württemberg) was initiated, and it brought about some promising results. Data centres provided their IT infrastructure and technical support, and the State Archives of Baden-Württemberg delivered specialized support on archival issues related to the software. This idea is still in progress and so far has been presented, among others, to municipal archives at various conferences. Archivists expressed their interest, and the archives intended to finalize the cooperation with three local data centres. The first test project was started with the Municipal Data Processing Region

Stuttgart (Kommunale Datenverarbeitung Region Stuttgart) and university archives (Keitel 2013a, 56–57).[17]

To summarize, the archivists at the Landesarchiv Baden-Württemberg are of the opinion that all kinds of cooperation result in success, but, in practice, however, it can be difficult to find someone to work with, and to convince, a partner to cooperate (Keitel 2013a, 54). These archivists were, and still are, very lucky to have such partners on board as the Data Processing Network Baden-Württemberg and especially the Municipal Data Processing Region Stuttgart or the Karlsruhe Institute for Technology (Karlsruher Institut für Technologie) and seven out of sixteen state archives that want to work together on developing the software. This has developed into a phenomenon on a global scale. One might reflect on the fact that outsourcing solutions for archives had been considered as early as the 1990s, but they had not been implemented because of legal constraints related to the transfer, security, and access of data by third parties (Ullmann 1998, 601). Furthermore, as can be seen from the German example, public archives in the cloud are close to becoming a reality in many states, cities, and municipalities, with universities and churches also being interested in using DIMAG.

Notes

1. Federal Republic of Germany, as a federal parliamentary republic, consists of sixteen regions with limited autonomy. They are called *land*—that is, state.
2. More about these data (Keitel 2004).
3. This kind of the approach shows that the German Democratic Republic archivists treated archival material in a traditional way. More attention was paid to the data carrier than to information itself—the opposite of what it is like in modern digital preservation attempts.
4. However, this idea of preserving digital archival material on microfilm was successfully developed by, among others, the State Archives of Baden-Württemberg. More about this project can be found in ''ARCHE: Projekt mit Förderung durch—'Förderung von innovativen Netzwerken' des Bundesministeriums für Wirtschaft und Arbeit,'' Landesarchiv Baden-Württemberg, http://www.landesarchiv-bw.de/web/46253. ''Projekt Ausbelichtung von Farbdigitalisaten mit dem ARCHE-Laserbelichter. Erprobung des Echtbetriebs,'' Landesarchiv Baden-Württemberg, http://www.landesarchiv-bw.de/web/49137.
5. In Baden-Württemberg, for example, local authorities' involvement in implementing archiving systems is guaranteed by the local archiving law.
6. ''Electronic records'' is a term used to refer to all records in digital form, transferred to, and accessioned in, the archives. It does not matter if they are born digital or digitized (Keitel 2010).
7. ''Baden-Württemberg, Hessen und Bayern kooperieren bei der Archivierung digitaler Unterlagen. Bundesweit einmalige Drei-Länder-Kooperation zur Software-Fortentwicklung,'' http://www.landesarchiv-bw.de/highlight_hp3.php?hl_link=http://www.landesarchiv-bw.de/web/bundesweit_einmalige_drei-laender-kooperation_zur_software-fortentwicklung/53471&q=kooperation.
8. More about the Projekt 'Konzeption für ein digitales Landesarchiv'. Projekt mit Förderung durch das Land Baden-Württemberg,'' see Landesarchiv Baden-Württemberg. http://www.landesarchiv-bw.de/web/44346.

9. It is an archives for land register records, so there were some changes in the primary DIMAG done. For more about this issue, see Lang (2013).
10. Germany, Landesarchiv Baden-Württemberg (Keitel, Naumann, and Lang 2008).
11. See IngestList Beta, http://sourceforge.net/projects/ingestlist/.
12. More about the technical site of the project can be found in Keitel and Lang (2010).
13. See ICA AtoM, https://www.ica-atom.org/.
14. Definition of software as a service (SaaS) by the National Institute of Standards and Technology (NIST 2011, 2).
15. Definition of cloud computing by the National Institute of Standards and Technology (NIST 2011, 2–3).
16. Elektronische Archivierung im Digitalen Archiv Nord (DAN), Landesamt für Kultur und Denkmalpflege Mecklenburg-Vorpommern, http://www.kulturwerte-mv.de/cms2/LAKD1_prod/LAKD1/de/Landesarchiv/Elektronisches_Landesarchiv/Elektronische_Archivierung_im_Digitalen_Archiv_Nord/index.jsp.
17. Digitale Langzeitarchivierung, Kommunale Informationsverarbeitung Reutlingen-Ulm, http://www.rz-kiru.de/,Lde/Startseite/Service/DIMAG.html.

References

Keitel, Christian. n.d. *Zugänglichkeit contra Sicherheit? Digitale Archivalien zwischen Offline-Speicherung und Online Benutzung.* http://www.staatsarchiv.sg.ch/home/auds/06/_jcr_content/Par/downloadlist_3/DownloadListPar/download_2.ocFile/Text%20Keitel.pdf.

——. 2004. "Erste Erfahrungen mit der Langzeitarchivierung von Datenbanken. Ein Werkstattbericht." In *Digitales Verwalten: Digitales Archivieren, Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt,* Band 19, ed. Rainer Hering and Udo Schäfer, 71–81. Hamburg: Hamburg University Press. http://www.staatsarchiv.sg.ch/home/auds/08/_jcr_content/Par/downloadlist_2/DownloadListPar/download_9.ocFile/Text%20Keitel.pdf.

——. 2010. "Digitale Archivierung beim Landesarchiv Baden-Württemberg." *Archivar* 1: 23.

——. 2013a. "Dienstleisterpartnerschaften mit DIMAG." In *Das neue Handwerk. Digitales Arbeiten in kleinen und mittleren Archiven.* Vorträge des 72. Südwestdeutschen Archivtags am 22. und 23. Juni 2012 in Bad Bergzabern, ed. P. Müller and K. Naumann, 54–57. Stuttgart, Germany: W. Kohlhammer.

——. 2013b. "DIMAG-Kooperationen." In *Digitale Archivierung in der Praxis.* 16. Tagung des Arbeitskreises "Archivierung von Unterlagen aus digitalen Systemen" und nestor-Workshop "Koordinierungsstellen," Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg Serie A, vol. 24. ed. Ch. Keitel and K. Naumann, 147–55. Stuttgart, Germany: W. Kohlhammer.

Keitel, Christian, and Rolf Lang. 2010. "DIMAG und IngestList. Übernahme, Archivierung und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg." In *Archivische Informationssysteme in der digitalen Welt, Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg Serie A*, vol. 23. ed. G. Maier and T. Fritz, 55–60. Stuttgart, Germany: W. Kohlhammer.

Keitel, Christian, Kai Naumann, and Rolf Lang, eds. 2008. *Metadaten für die Archivierung digitaler Unterlagen.* http://www.landesarchiv-bw.de/sixcms/media.php/120/48392/konzeption_metadaten10.28354.pdf.

Kemper, Joachim, and Kai Naumann. 2013. "Selbermachen! Praktische Tipps zur Archivierung digitaler Unterlagen, Digitalisierung und Öffentlichkeitsarbeit im Netz." In *Das neue Handwerk. Digitales Arbeiten in kleinen und mittleren Archiven,* ed. P. Müller and K. Naumann. Vorträge des 72. Südwestdeutschen Archivtags am 22. und 23. Juni 2012, Bad Bergzabern, 86. Stuttgart, Germany: W. Kohlhammer.

Kluttig, Thekla. 2006. ''Strategies of German State Archives for the Preservation of Electronic Records.'' In Archives in the New Age: The Strategic Problems of the Automatization of Archives Information. *Papers of the International Conference Warsaw*, September 28–29, 2001, Colloquia Jerzy Skowronek dedicata, ed. A. Biernat and W. Stępniak, 56–59. Warsaw: Naczelna Dyrekcja Archiwów Państwowych – Wydział Wydawnictw.

Lang, Rolf. 2013. ''Die elektronische Grundakte in G-DIMAG.'' In Digitale Archivierung in der Praxis. 16. Tagung des Arbeitskreises ,,Archivierung von Unterlagen aus digitalen Systemen'' und nestor-Workshop ,,Koordinierungsstellen, Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg Serie A Heft 24, ed. Christian Keitel and Kai Naumann, 129–41. Stuttgart, Germany: W. Kohlhammer.

National Institute of Standards and Technology. 2011. *The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology.* Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

National Library of Australia. 2003. *Ochrona dziedzictwa cyfrowego. Zalecenia.* Warsaw: Naczelna Dyrekcja Archiwów Państwowych.

Naumann, Kai. 2007. ''Älteste digitale Archivquelle der Bundesrepublik gesichert: Daten der Volkszählung von 1961 für das Land Baden-Württemberg übernommen und aufbereitet.'' *Der Archivar* 60 (1): 53.

Ullmann, Angela. 1998. ''Der EDV-Ausschuß der Archivreferentenkonferenz des Bundes und der Länder 1972–1994.'' *Der Archivar* 51 (4): 587–607.