Preservica
Active digital preservation

"We had intensive contact with the designers of Preservica and the developers who implemented our solution, and we benefited a lot from that. Communication paths were short, and Preservica was also very proactive. We dealt with very competent people at every stage of the project."

Urs Meyer
IT Architect, SFA

Swiss Federal Archives

Recognising early the need for an all-encompassing end-to-end digital preservation system, the Swiss Federal Archives (SFA) started a collaborative process by engaging with Preservica as one of its expert partners. The aim was to establish an archive to secure, preserve, manage and access the SFA's digital records into the future. The solution that the SFA finally chose is based on the tried-and-tested Preservica Enterprise Edition system, originally developed in partnership with The National Archives of the UK. The system was established on time and within budget at the end of 2008.

Business Situation

The Swiss Federal Archives (SFA) are located in the national capital, Berne. The role of the institution is to advise the federal administration on records management and to appraise, to secure and to provide access to archive-worthy records of the Swiss Parliament, Government and Administration. It would need a bookshelf 60 kilometres long to hold the entire archives, which includes original documents such as the constitution and holds data in various formats such as, paper documents, photos, films, recordings and databases.

As well as storing millions of physical documents, the SFA contains a growing number of electronic records. In fact, by the end of 2008, the digital archive had grown to 12.5 terabytes (TB) in size – the paper equivalent of half-a-million trees or the capacity needed to store over 100,000 feature-length movies.

According to Preservica director Jon Tilbury, data storage and preservation has become a critical issue for many large organisations. "All organisations realise that they must organise their data – ingest it into an organised data structure. The biggest risk they face at the moment is losing it before it's even managed."

In order to increase efficiency and performance, the SFA's digital archive is based now on an application that manages the whole process from beginning to end. A key priority was preserving the integrity of the file structure, given that many of the digital files being sent for archiving had been stored in well organised record management systems.

"More and more data is now being kept in records management systems that are well structured. From 2012 on, the whole federal administration will have to change to a records management system that integrates already key prerequisites for digital

archiving. We need to preserve that structure too in order to guarantee the traceability and comprehensibility of the affairs," notes Urs Meyer, IT architect of the SFA.

The SFA further had to take into account that the digital archives will be expanding rapidly in future, as more and more government business is conducted electronically. To deal with this and other parameters, the SFA developed a digital management strategy and a policy that includes processes and guidelines.

In the beginning was the choice of the right partner. Preservica was well known within the European archiving community as an archiving specialist that had designed and built digital archives for a number of other national archives including the Dutch National Archives and The National Archives in the UK. What's more, Preservica and the SFA knew each other already from their collaboration in the 6th Framework-funded EU project called Planets, which was investigating basic techniques required for longterm information preservation.

After the SFA conducted a market study, Preservica stood out as the best candidate and was therefore mandated. Key was Preservica's experience with archives, because of their unique requirements among record-keeping institutions: not only do they need to keep records permanent and safe; each file they hold is unique – unlike books in a library, for example, which generally exist in multiple copies. Preservica has a good reputation as technology leader in digital preservation.

Project

As part of its brief to help build a Digital Information Repository (DIR) for the SFA, Preservica was contracted to undertake two key tasks:

- Write the software specification for the DIR based on the initial conceptual work that had been done by the SFA.
- Implement the archiving component of the DIR, based on its preservation technology.

Solution

The SFA's digital archive is powered by Preservica's Enterprise Edition, co-developed with The National Archives in the UK and used by The National Archives as the basis of its award-winning digital archive system. A key feature of the solution is Active Preservation technology. This means that the information contained within a record can be preserved even if the record is migrated to a new technology format.

"This is done by measuring the properties of those records before migration, migrating them, and then measuring their properties again when they've been imported into the new technology. We then compare the two sets of properties to ensure no information has been lost during the migration," explains Rob Sharpe, head of archiving solutions at Preservica.

Security of the archives is a key priority. The server(s) reside in a protected network zone. Access is possible only with ICA protocol (Citrix Terminal Server with encrypted communication). Data is stored three times: on a tape library system (two tape libraries from ADIC, now Quantum, at two separate locations) and a disk-based system (NAS Filer from Netapp).

Customisation

Preservica Enterprise Edition is an evolving technology platform that is capable of being tailored to the needs of each client. In the SFA's case, there was a requirement to modify the system in line with the existing security protocol. The software is designed such that all the files in an archival information package (AIP) – the name given to a group of related files submitted for storage – are held separately and it is the metadata describing those files that links them all together. However, for security reasons, the SFA wanted all of the files within the AIP to be held in one place, such that in the unlikely event of the software becoming corrupted, the SFA would know which files went together.

"For us, it was important that the unit of an AIP survives for safety reasons and Preservica did that modification for us," says Meyer.

The SFA also wanted to add new functionality to the software around identifying file formats during the ingestion process. The Droid software to identify file formats was already part of the system and this is one of the automated tests that is performed during the file ingest. But the SFA let refine and extend the number of tests, such as structural integrity tests that would, for example, verify whether metadata matched physical files and vice-versa.

Meyer judged the collaboration with Preservica as positive: "We had intensive contact with the designers and developers who implemented our solution, and we benefited a lot from that. Communication paths were short, and Preservica was also very proactive. We dealt with very competent people at every stage of the project."

Deployment

Preservica's involvement with the project formally began in August 2007. Both tasks – writing the software specification and implementing the archive – were completed on schedule and within budget by the end of 2008. The Preservica application was physically deployed in the first half of 2009.

Benefits

- Format management

The DIR provides a simple user interface that lets SFA personnel add or remove formats from the master list with ease. Furthermore, submitted file packages have a well defined structure. Once a packet is built, users know exactly what is in there and it is described properly in the metadata.

- Integration

The digital archive developed by Preservica integrates key applications already used by the SFA, such as its Archival Information System (AIS), based on scopeArchiv by Scope Solutions. Metadata that contains relevant data for the Scope system is extracted automatically during the ingestion process in the format required by AIS (XML). This eliminates the need for manual transfer of data which in turn bears the potential for error and imprecise recording.

- Archive design

The software specification written by Preservica has provided the blueprint for the efficient ingestion, storage, archiving and access-provision of archived material. The software specification maps the SFA's archival processes and supports them, as Preservica implemented an according workflow system."

- Security

The implementation of the archive is based on a robust security framework that includes defined roles. The SFA has defined roles for the main actions, that is, ingest, access, preservation, and administration. The system supports and enforces the concept of roles which is key to guarantee traceability and authenticity of data.

Summary

Preservica Enterprise Edition implemented for the Swiss Federal Archives, and the software specification that accompanied it, are vital elements of its digital archive and supports the SFA in the challenging task of preserving the valuable digital records today and for generations to come.