“We were looking for an IT specialist to provide us with more than just a machine with software. A flexible attitude and thinking along with us from our situation were the main conditions for cooperation. Hitachi Data Systems appeared to be the only supplier that really understood our problems.”

Bert Nijhoff
ICT Team Leader
Stadsarchief Amsterdam

Stadsarchief Amsterdam (Amsterdam City Archive)

INDUSTRY Government: Public Services

SOLUTIONS File and Content Services, Replication, Green Solutions, Storage Management
Hardware — Hitachi Content Archive Platform
Services — Provided by Hitachi TrueNorth Channel Partner Data Matters
Stadsarchief Amsterdam Secures Historic Inheritance of Amsterdam with the Hitachi Content Archive Platform

With 22 miles of archives — a historic-topographic collection with millions of maps, drawings and pictures, a library, extensive sound, movie and photo archives, and more than eight million digital archive documents — Stadsarchief Amsterdam is the biggest city archive in the world. Besides official and public documents, it also stores private collections and archive documents from companies. To prevent these documents from getting lost, Stadsarchief Amsterdam turned to Hitachi Data Systems for a solution to digitize its archives, and ensure longevity and scalability.

At this moment, Stadsarchief Amsterdam manages over eight million scans. This mainly concerns scans made from paper archive documents. These scans are also called “digital duplicates” and increase with a weekly average of 10,000 scans. Next to these analog documents, Stadsarchief Amsterdam also deals with digitally created documents. The number of these digitally created documents is increasing rapidly.

For all these digital archive documents or objects, it is very important to determine how the authenticity and integrity can be guaranteed. In other words, one hundred years from now, a document has to come out of the archive in the same condition as when it was stored. To achieve this, Stadsarchief Amsterdam looked for a digital archiving system that would fulfill these requirements. This so-called Digital Storage Depot (DSD) would also be required to possess an adequate capacity to optimally prepare Stadsarchief Amsterdam for the future.

Requirements Program

According to the Dutch Archiving Legislation, archives have to comply with strict rules. In 2008, Bert Nijhoff, ICT Team leader of Stadsarchief Amsterdam, in cooperation with the government’s Expertise Center, created a “Requirement Program” for replacing the old storage depot.

Nijhoff explains, “Most archives start such a project at the beginning of the archiving process. Thus, at the intake side: how do archive documents have to be delivered; how can we check their completeness? Because we wanted to ensure the durability of the documents, we started in the middle, meaning at the storage. This was a new approach and, as far as we know, not done before. This is why we started to look for an IT specialist to provide us with more than just a machine with some software. A flexible attitude and thinking along with us from our situation were important conditions for cooperation with the suppliers. In spite of all the understanding displayed by the other four parties, only Data Matters and Hitachi Data Systems appeared to really understand our problems.”

Automatic Verification

Data Matters and Hitachi Data Systems implemented the Hitachi Content Archive Platform as a Digital Storage Depot for

"With Hitachi Content Archive Platform, we have an extremely scalable digital archiving system. Based on the current disk standards we can extend Content Archive Platform up to 247PB. Even if we scan very intensively, I will not live to see this limit reached.”

Bert Nijhoff
ICT Team Leader
Stadsarchief Amsterdam
digitized as well as digitally created objects. A control mechanism in the Content Archive Platform automatically and continuously checks the digital archive documents for authenticity. Each object in the DSD has a code, which is also called a HASH value. As soon as this value no longer corresponds with the original HASH value (for example, due to failing electronics) object is restored based on the always-present backup copy. These repairs are recorded in the Content Archive Platform log. The digital archiving system also gives notice when the storage period for an object has expired, so it might be destroyed, if desired.

An important aspect of the DSD design is the link with ScopeArchiv. This is a metadata management system that stores information about archive documents, such as owner, archive location and intake date. To create a smooth exchange of information between the DSD and ScopeArchiv, Data Matters developed a Web interface based on open standards. Nijhoff clarifies, “We do not know what programs for managing archives’ metadata will be used one hundred years from now. Therefore, it is important not to completely integrate a metadata management system like ScopeArchiv, but instead to exchange information through an open interface. The technology for archive storage systems will also greatly evolve in the coming year. It is therefore important that checking in and out of archive objects through open standards is possible, to allow migration to another system.”

Scalable Archive Platform

With Hitachi Content Archive Platform, Stadsarchief Amsterdam has an extremely scalable digital archiving system. “Today, the digital archive documents take up 18TB. Based on the current disk standards, we can extend Hitachi Content Archive Platform up to 247PB. Even if we scan very intensively, I will not live to see this limit reached,” Nijhoff says.

At the end of 2008, the parties involved started with the implementation of the Content Archive Platform digital archive system. Due to the innovative nature of the project and the progressing insights during the project, new problems arose in the course of the implementation process. These problems, mostly related to the cooperation between Content Archive Platform and ScopeArchiv, have been solved by using an additional tool set through a Web interface.

To optimally protect the valuable digital archive documents, Data Matters and Hitachi Data Systems installed a second digital archive system, at a second location, which is an exact duplicate of the DSD. Thanks to permanent synchronization between both locations, the documents in the archive are safe. This way, Stadsarchief Amsterdam hopes to be finally freed from making conventional tape backups. “It is a fact that a tape backup cannot guarantee the authenticity,” Nijhoff explains.

Result

At the end of May 2009, the platform became operational. After a few weeks of testing, the digital archive system now runs at full capacity and everybody is enthusiastic about it. Says Nijhoff, “Everything runs according to expectation. Now we have to design surrounding processes and get it out of the project atmosphere. Of course, this also means a different way of working for our employees…. No longer do they need to know what kind of paper must be stored in an acid-free box, but instead they need to know in what format the documents are best stored in: PDF, ODF, TIF, JPEG and what the compression factor has to be. To prepare them for this, we are organizing training courses.”

According to Nijhoff, the success of this project mainly stems from the informal way of working. “Our organization is characterized by innovative and creative dealing with the limited means we have to safeguard our historic inheritance for the future,” he explains. “For this, short communication lines are needed. Therefore, we consciously chose a small team, consisting of representatives of Stadsarchief Amsterdam, one project manager from Data Matters and one from Hitachi Data Systems. Thanks to the flexibility of both suppliers, it really became a shared project with room for creative solutions. It was not at all a routine job, but we can already say that it has been successful,” he concludes.

About Data Matters

For more than eight years, Data Matters has been an acknowledged specialist in data storage and data management for the government as well as the business world. Data Matters advises, designs, implements and maintains entire infrastructure solutions for data storage and management. Based on specialist knowledge and years of expertise, combined with the application of high quality, certified products from leading manufacturers, Data Matters realizes reliable, scalable and cost-efficient storage solutions. Data Matters provides futureproof solutions, optimally attuned to the needs and desires of the customer and completely integrated with existing ICT and business environments.

For more information, visit www.datamatters.nl.