Developing tools and methods for Sharepoint to archive process

Osmo Palonen, Mikkeli University of Applied Sciences; Mikkeli, Finland; Juha Anttila, iITC Oy; Maaninka, Finland;

Abstract

Microsoft Sharepoint is getting more and more popular as a document management tool in private enterprises and it is often used in the public sector as well. Sharepoint can have different roles in the corporate IT roadmap, as its best supporting teamwork and intranet. However, Sharepoint has never included tools in a proper archiving. Mikkeli University of Applied Sciences (MUAS) with partners developed and tested tools and methods in Sharepoint to include necessary metadata needed in digital archiving and a process to ingest the documents in the archive. To utilize Sharepoint in proper document or records management, additional tools are needed. In this paper, a case study to add-on these functionalities is presented..

Archiving Services for Enterprises

In a R&D project Archiving Services for Enterprises, Mikkeli University of Applied Sciences (MUAS) and The Central Archives for Finnish Business Records (ELKA) studied the demand for digital archive service within the Finnish enterprises and organizations. In a questionnaire included this project the popularity of Sharepoint in document management was noteworthy. As archiving service providers for private sector Elka and MUAS have to consider how the contents created and held in this environment can be ingested and preserved in the long-term and permanent archive. IITC had a consultant role in this project.

Information Control System TOJ

The starting point was to find tools to enrich the metadata in Sharepoint environment by providing application service that extends the metadata options for the Sharepoint users. A Finnish company Innofactor PLC had developed an application TOJ (Information control system) to manage records and documents based on hierarchical classification and process phase description. TOJ also manages rules for user rights control, retention and preservation. The system is planned to complement both the EDRMS systems provided by Innofactor and other systems holding record information alike. When starting the project at the end of 2010 Innofactor had already made testing with Sharepoint, but not with an external independent repository.

TOJ has been developed in the first place for the Finnish public sector, where this kind of functionality is one of the qualifications for the digital long-term preservation of records. In the public sector TOJ is used to gather the required metadata about the organization, task and other context information automatically. Rules configured in the TOJ the access rights as well as the process of the record are defined.

Data model unification

Second phase in the project was to unify the data model in TOJ and digital archive to make successful transfers possible. Most of the metadata fields were just cross-connected, only some new fields were added in the standard interface used in the archival ingest.

The fact that the metadata model of TOJ is based on the national metadata standards SÄHKE and JHS (Public sector recommendation) and the metadata model used by ELKA and MUAS included most of those metadata fields for the archiving made the task easy to harmonize. Some differences were in the terminology also: those were solved adding eight fields using the flexibility of the digital archive software. In the TOJ version there were 27 metadata fields whereas archiving software had 140. Those eight fields which are mainly used in the records and case management in public sector were:

- PublicityClassChange
- InformationSystem (Registration)
- SecurityPeriodRule
- ProtectionLevel
- protectionClass
- RetentionPeriodRule
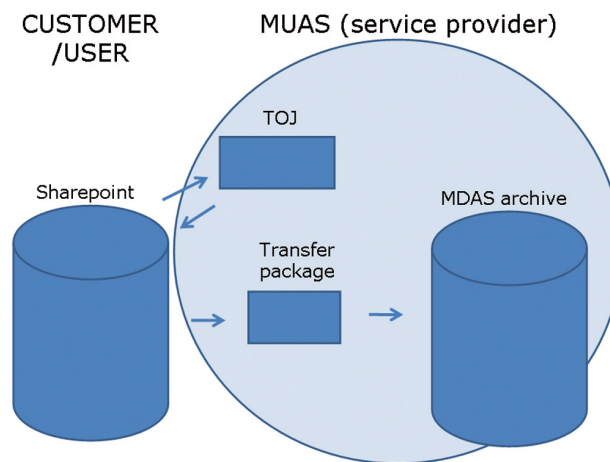- InstructionalData
- VersionRetentionRule

Figure 1: The communication principle.

The pilot approach

A pilot was realized in co-operation with an internationally operating Finnish industrial enterprise. First task was to define what kind of material can be used in the testing to make the pilot significant for all partners. Based on the negotiations with the industrial partner it was agreed that a production development project where building and engineering work has been documented can be a good example. First reason was: the material of a project
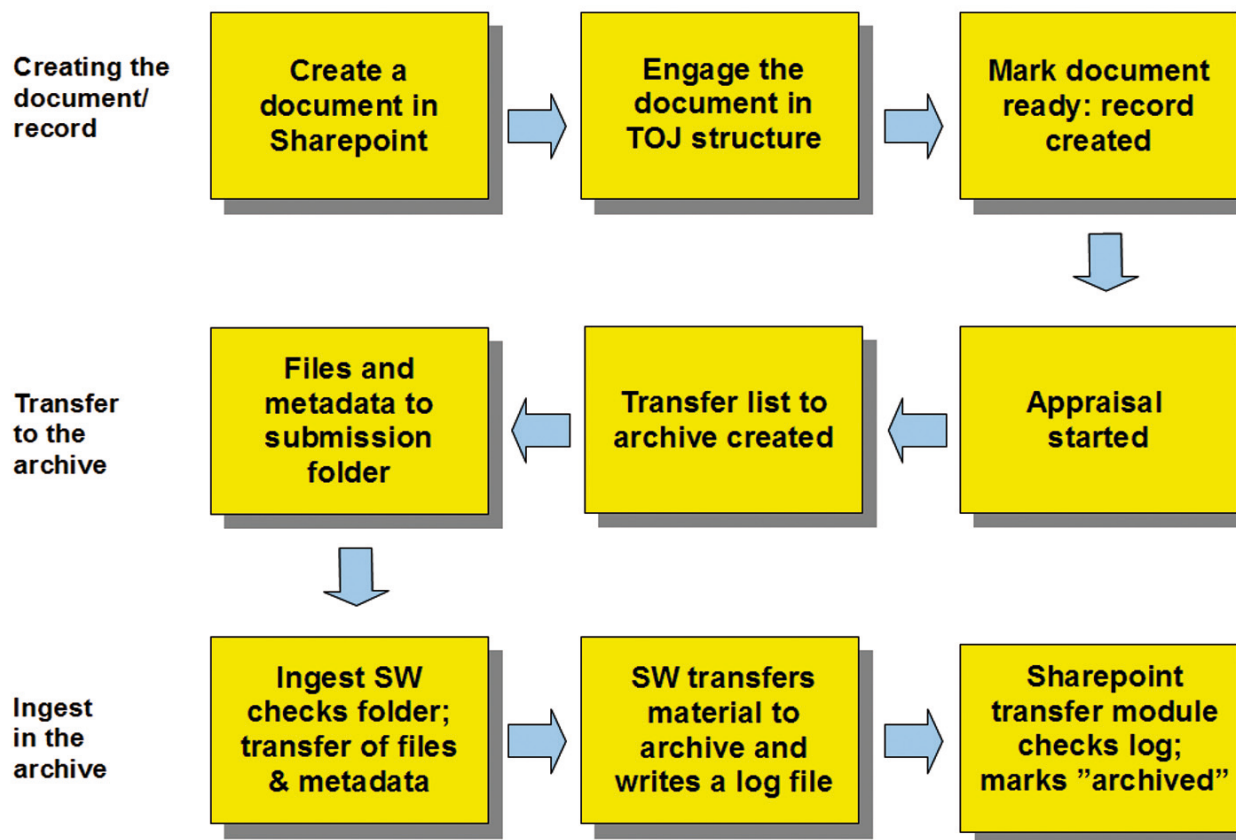
Figure 2: Process phases in the archiving process.

like that need to be kept safe and available for decades. That is a good reason for a proper digital archiving. The other was the complexity of the material: building and engineering drawings, financial material, memorandums, documents for the authorities and public as well as from contractors and third parties can be included, for example. Another benefit, when running a pilot lasting only a few months, a completed project was a better choice than continuing activity.

When using information that was already created, the personnel of the industrial partner was not very keen to use a wide range of additional tasks for archiving. There were only limited amount of new information that should be given manually. In the pilot the configuration of TOJ system was the key task, if that was done properly, the rest of the job was more or less to select classification to enrich the Sharepoint material to fit the archival requirements. TOJ as well as the archiving applications were provided for the pilot by MUAS IT-services via VPN network connection.

Extended metadata in Sharepoint

By using the TOJ, a special hierarchical classification and rules for the pilot case were created and the application was opened as an extension to Sharepoint. The users could then get the extended metadata via pull-down menus just by clicking. Based on

the selection TOJ gave a retention period and other essential information to the document and it was ready to transfer into the archive. The status information was marked "archived" and the document was frozen. After transfer to archive no changes in the document were possible.

The ERDM professionals within the industrial enterprise processed the archival material of the project selected for the pilot. Based on the hierarchical classification selected in TOJ, the metadata including the retention period given in the Information Control Plan (the plan behind TOJ configuration) was automatically inherited. At the end of this process, the documents were marked complete in Sharepoint and no changes were allowed. In the pilot, the documents marked complete, including the metadata, were transferred to the archive manually. In the standard production, the transfer should be done periodically at defined intervals by an automatic process.

Documents managed in Sharepoint and transferred in the MUAS digital archive were then available via customized archival web browser interface based on the standard in MUAS. The record managers in the enterprise had their special interface to manage the digital archive. In the pilot project there were no tools inside Sharepoint to access the contents in the archive, only those frozen documents were available. The project group was unanimous that if the digital archiving transfer is in standard production,

extensions to archive access direct from Sharepoint have to be developed. Technically, it was seen a possible solution, however in the pilot like this there were no resources available. Even when the pilot was not very extensive, it proved that it is possible to enrich metadata in the Sharepoint environment to include metadata required in the proper archiving processes. The technical solution is to build Sharepoint compatible extensions which will provide

Figure 3: The MUAS archive interface for archivist and records manager includes following main functions: Main - Search - Browsing - My settings - Add content - Organize archive - User management. There are nine tab sheets where the metadata is stored for the professional user.

the user tools to easily add the missing information. The project also showed the weakness of the solution of this kind: the corporate IT is not always willing to give permission to install these modules to implement a pilot like this in the internationally operating production environment of the enterprise. In this case, the Sharepoint installation was build in the cloud service of the project partners. Still, the project indicated that there are additional tools available for Sharepoint to preserve material in long-term and by using those contents managed by Sharepoint can be included in the trusted repository. It also showed that all the elements needed in the business archive can be implemented by Software as a service (SAAS) model.

The major problem with Sharepoint from the archiving and records management point of view has been the limited capabilities to manage information based on the ERDM and archiving standards. With tools like TOJ, this problem can be solved. The input from the ordinary users can be very limited when these tools are flexible and properly configured to meet the needs of different operations. The records managers and archivists can then be the only professionals who need to use the professional tools. The users within the enterprise may not even know that they are using document or record management software that is compliant with archiving. In the best case they just see that Sharepoint is running faster when the system is not loaded with old documents.

Partners in the project

This pilot project was a co-operative effort, the partners of MUAS were Innofactor PLC (www.innofactor.com) and Profium Oy (www.profium.com). EDRMS consultant IITC was appointed as specialist and project coordinator. Sharepoint pilot was a part of the project Archives for Enterprises organised by the MUAS. The project was co-funded by Research to business programme Tuli founded by Tekes, the R&D funding unit of the Finnish government.

Mikkeli University of Applied Sciences is a leading organization in education, research and development as well as service provider in electronic content and records management and digital archiving in Finland. The university is a limited company: owners are the City of Mikkeli as well as Towns of Pieksämäki and Savonlinna.

IITC (Industrial ITC Oy) is a provider of information technology services to fulfil the needs of Finnish companies and organizations. The main services are consulting, training and publishing in Information Management, Enterprise Content Management (ECM), Document Management (DM) and Product Data Management (PDM/PLM).

Author Biography

Osmo Palonen has MA degree in history and IT in University of Tampere. He has FKT diploma including graphic arts, IT and management in Helsinki Institute of Marketing. He worked 16 years as journalist, including project management for the editorial IT-systems. Before joining the Mikkeli UAS in 2003 he worked 14 years for Honeywell Industrial Automation. In Department of Electrical and Information Technology Palonen has been in charge of the project and ASP-service development and has now senior specialist role. He is 2012 the chairman of the board in the Association of Business Archives and been the Editor of the quarterly publication Faili since 2007.

Juha Anttila is consultant and CEO at IITC. He has 15 year experience on information management. He has written a book about document management (Dokumenttien hallinta, ITPress) and some other publications. He is also known speaker at information management events and a member of the board in the Association of Business Archives.