

Tags: Archiv Bewertung,

ArchiSelect, ou quand l'évaluation s'automatise

Si les fondements de l'évaluation archivistique ne changent pas avec l'avènement du numérique, les conditions matérielles de son déroulement demandent que l'archiviste puisse s'appuyer sur un accompagnement logiciel et technique *ad hoc*.

Le concept AENeas est la réponse neuchâteloise au défi de l'archivage numérique. Il propose notamment plusieurs outils (la «SuiteArchi») qui accompagnent le cycle de vie documentaire et le processus d'archivage, de la création d'un plan d'archivage à la mise à disposition des archives pour la consultation.

Cinq outils archivistiques ont été définis:

- ArchiClass, qui sert à l'élaboration de plans d'archivage (arborescence de classement des dossiers d'activité assorti de métadonnées).
- ArchiSelect, qui doit servir à accompagner l'archiviste dans l'évaluation des dossiers qui lui sont proposés pour versement et/ou élimination.
- ArchiVision, qui permettra de suivre la supervision des entités soumises à la loi sur l'archivage.
- ArchiPeren, qui sera le dépôt numérique et qui assumera également le conditionnement numérique des archives à verser.
- ArchiInfo, qui servira à la fois à élaborer et publier les inventaires, ainsi qu'à mettre les archives numériques à disposition du public.

Illustration 1: Schéma du processus d'archivage selon le concept AENeas.

Illustration 2: Outils définis dans la SuiteArchi.

Ces outils sont acquis sur le marché si des solutions satisfaisant aux besoins existent déjà; dans le cas contraire, ils sont développés spécifiquement. Le premier d'entre eux, ArchiClass, a été développé par la société Objectis à Yverdon-les-Bains, sur mandat des Archives de l'État de Neuchâtel. Depuis 2016, ce logiciel est utilisé dans les administrations cantonale et communales neuchâteloises, ainsi que par de nombreuses entités hors du canton.

ArchiSelect

ArchiSelect est actuellement au centre des efforts de développement de la part des Archives de l'État de Neuchâtel, avec le concours de la Haute école de gestion de Genève (par les professeurs Basma Makhoul Shabou et Arnaud Gaudinat et leurs assistants) et des sociétés Telono (Genève) et Contreforme (Neuchâtel). Ce logiciel doit permettre d'accompagner l'archiviste dans la macro- et la micro-évaluation de dossiers d'activité et de lots de données numériques, proposés par l'administration ou par des privés. Dans le monde papier, l'archiviste, grâce à son expérience et à ses connaissances, peut évaluer des dossiers plus ou moins bien constitués, voire du vrac. Il n'en est pas de même dans le monde numérique, où il est illusoire de vouloir évaluer «manuellement» d'énormes masses de répertoires, fichiers, courriels, etc.

Les administrations publiques, depuis deux à trois décennies, utilisent l'informatique au quotidien dans l'exécution de leurs tâches. Si, ces dernières années, une tendance à prévoir des systèmes documentaires réfléchis et suivant les recommandations archivistiques se fait jour, il n'en est pas de même pour la documentation plus ancienne, souvent déposée de manière informelle sur des serveurs ou d'autres supports. Ainsi, il importe de pouvoir traiter non seulement des dossiers d'activité constitués selon un plan d'archivage préétabli (en particulier dans ArchiClass) – la voie la plus simple – mais aussi du «vrac» – la voie la plus complexe.

Évaluation en masse de lots de données numériques

Dans ce dernier cas, la fouille de données («data mining», grâce à des algorithmes) devient un allié incontournable pour l'évaluation. L'archiviste doit pouvoir s'aider d'un système ad hoc afin d'analyser les documents, qui pourront être filtrés en fonction de toute une série de «métriques» classiques ou moins classiques, que ce soit le format (pdf, doc, xls, ...), l'auteur, le type de document (procès-verbal, pièce comptable, ...) ou encore la date. Les données devront pouvoir être regroupées en fonction des requêtes faites par l'archiviste (clustering) et représentées de différentes manières (graphiques, nuages de mots, ...). Le système pourra également indiquer à l'archiviste des degrés de similarité entre les documents (plusieurs états successifs d'un rapport, ou sa version finale dans deux formats différents par exemple).

Même en l'absence d'un plan d'archivage préétabli, les données peuvent être structurées, au moins partiellement, par dossiers. L'analyse de la structure de classement, ainsi que des éventuels identifiants et titres des documents, peut permettre d'identifier des dossiers d'activité ou des séries organiques, préservant ainsi le lien archivistique existant.

Cette analyse multiforme et multicritère permettra à l'archiviste de déterminer la valeur archivistique des documents en s'appuyant notamment sur les dimensions d'évaluation apparaissant dans la norme ISO 15489: valeur probante (fiabilité, authenticité, preuve historique), exploitabilité (accessibilité technique, cognitive et juridique) et représentativité (contexte organisationnel et socio-culturel).

Évaluation de dossiers d'activité d'après un plan d'archivage préexistant

Lorsque les dossiers d'activité proposés sont classés selon un plan d'archivage prédéfini, l'archiviste devra essentiellement confirmer les sorts finaux pressentis définis dans ArchiClass. Au besoin, il devra pouvoir consulter simplement les dossiers d'activité, voire procéder à une analyse plus fine du même type que celle nécessaire pour les lots de données. Dans le cas de dossiers en série dont seul un échantillon sera versé aux Archives, l'archiviste pourra s'assurer de la bonne représentativité en fonction des dates, par exemple.

À l'issue du processus d'évaluation, ArchiSelect produira des bordereaux de versement et d'élimination. Le traitement effectif de l'élimination et/ou du versement sera cependant effectué hors du logiciel, par le producteur ou le service informatique, chargés de s'assurer de l'effacement irréversible des données, ou par le service d'archives, responsable du conditionnement numérique des documents.

Grâce à ArchiSelect, l'archiviste disposera d'un outil efficace d'aide à la décision dans le traitement des dossiers d'activité et lots de données numériques qui lui sont proposés; il restera maître de l'évaluation, qui sera semi-automatisée et ne saurait se passer du regard humain.

Ces outils sont acquis sur le marché si des solutions satisfaisant aux besoins existent déjà; dans le cas contraire, ils sont développés spécifiquement. Le premier d'entre eux, ArchiClass, a été développé par la société Objectis à Yverdon-les-Bains, sur mandat des Archives de l'État de Neuchâtel. Depuis 2016, ce logiciel est utilisé dans les administrations cantonale et communales neuchâteloises, ainsi que par de nombreuses entités hors du canton.

ArchiSelect

ArchiSelect est actuellement au centre des efforts de développement de la part des Archives de l'État de Neuchâtel, avec le concours de la Haute école de gestion de Genève (par les professeurs Basma Makhoul Shabou et Arnaud Gaudinat et leurs assistants) et des sociétés Telono (Genève) et Contreforme (Neuchâtel). Ce logiciel doit permettre d'accompagner l'archiviste dans la macro- et la micro-évaluation de dossiers d'activité et de lots de données numériques, proposés par l'administration ou par des privés. Dans le monde papier, l'archiviste, grâce à son expérience et à ses connaissances, peut évaluer des dossiers plus ou moins bien constitués, voire du vrac. Il n'en est pas de même dans le monde numérique, où il est illusoire de vouloir évaluer «manuellement» d'énormes masses de répertoires, fichiers, courriels, etc.

Les administrations publiques, depuis deux à trois décennies, utilisent l'informatique au quotidien dans l'exécution de leurs tâches. Si, ces dernières années, une tendance à prévoir des systèmes documentaires réfléchis et suivant les recommandations archivistiques se fait jour, il n'en est pas de même pour la documentation plus ancienne, souvent déposée de manière informelle sur des serveurs ou d'autres supports. Ainsi, il importe de pouvoir traiter non seulement des dossiers d'activité constitués selon un plan d'archivage préétabli (en particulier dans ArchiClass) – la voie la plus simple – mais aussi du «vrac» – la voie la plus complexe.

Évaluation en masse de lots de données numériques

Dans ce dernier cas, la fouille de données («*data mining*», grâce à des algorithmes) devient un allié incontournable pour l'évaluation. L'archiviste doit pouvoir s'aider d'un système *ad hoc* afin d'analyser les documents, qui pourront être filtrés en fonction de toute une série de «métriques» classiques ou moins classiques, que ce soit le format (pdf, doc, xls, ...), l'auteur, le type de document (procès-verbal, pièce comptable, ...) ou encore la date. Les données devront pouvoir être regroupées en fonction des requêtes faites par l'archiviste (*clustering*) et représentées de différentes manières (graphiques, nuages de mots, ...). Le système pourra également indiquer à l'archiviste des degrés de similarité entre les documents (plusieurs états successifs d'un rapport, ou sa version finale dans deux formats différents par exemple).

Même en l'absence d'un plan d'archivage préétabli, les données peuvent être structurées, au moins partiellement, par dossiers. L'analyse de la structure de classement, ainsi que des éventuels identifiants et titres des documents, peut permettre d'identifier des dossiers d'activité ou des séries organiques, préservant ainsi le lien archivistique existant.

Cette analyse multiforme et multicritère permettra à l'archiviste de déterminer la valeur archivistique des documents en s'appuyant notamment sur les dimensions d'évaluation apparaissant dans la norme ISO 15489: valeur probante (fiabilité, authenticité, preuve historique), exploitabilité (accessibilité technique, cognitive et juridique) et représentativité (contexte organisationnel et socio-culturel).

Évaluation de dossiers d'activité d'après un plan d'archivage préexistant

Lorsque les dossiers d'activité proposés sont classés selon un plan d'archivage prédéfini, l'archiviste devra essentiellement confirmer les sorts finaux pressentis définis dans ArchiClass. Au besoin, il devra pouvoir consulter simplement les dossiers d'activité, voire procéder à une analyse plus fine du même type que celle nécessaire pour les lots de données. Dans le cas de dossiers en série dont seul un échantillon sera versé aux Archives, l'archiviste pourra s'assurer de la bonne représentativité en fonction des dates, par exemple.

À l'issue du processus d'évaluation, ArchiSelect produira des bordereaux de versement et d'élimination. Le traitement effectif de l'élimination et/ou du versement sera cependant effectué hors du logiciel, par le producteur ou le service informatique, chargés de s'assurer de l'effacement irréversible des données, ou par le service d'archives, responsable du conditionnement numérique des documents.

Grâce à ArchiSelect, l'archiviste disposera d'un outil efficace d'aide à la décision dans le traitement des dossiers d'activité et lots de données numériques qui lui sont proposés; il restera maître de l'évaluation, qui sera semi-automatisée et ne saurait se passer du regard humain.



Grégoire Oguey

Études aux universités de Neuchâtel et de Paris-IV-Sorbonne. Licencié ès lettres et sciences humaines de l'Université de Neuchâtel (2009). Membre de l'Institut suisse de Rome (2014-2015). Archiviste diplômé de la *Scuola vaticana di paleografia, diplomatica ed archivistica* (2015). Chercheur à l'Institut d'histoire de l'Université de Neuchâtel.



Pascal Schneiter

Études à la Haute Ecole d'Ingénierie et de Gestion du Canton de Vaud et de Paris-ESME-Sudria. Ingénieur en informatique logiciel (2004). Titulaire d'un postgrade EMBA (2009). Formation continue en archivistique (2016).

Abstract

Français

Les Archives de l'État de Neuchâtel sont en train de développer, avec le concours de la Haute école de gestion de Genève, un logiciel d'aide à la décision pour l'évaluation: ArchiSelect. Grâce à la fouille de données, ArchiSelect offrira différents filtres et regroupements de données à l'archiviste, afin de l'aider à déterminer les valeurs archivistiques, puis de produire les bordereaux de versement et d'élimination.

Deutsch

Das Staatsarchiv Neuchâtel entwickelt zur Zeit in Zusammenarbeit mit der HEG - Haute école de gestion de Genève eine Software zur Entscheidungshilfe bei der Evaluation: ArchiSelect.

Mit Hilfe von Datenmining bietet ArchiSelect dem Archivar verschiedene Filter und Daten an, um die Archivwürdigkeit der Akten zu prüfen und die Ablieferungs- und Aussonderungslisten zu erstellen.