

---

# Un outil de dépouillement de sources archivistiques basé sur des technologies XML

Emmanuel Château-Dutier\*<sup>1</sup> and Josselin Morvan\*<sup>2</sup>

<sup>1</sup>Université de Montréal (CRIHN) – Canada

<sup>2</sup>ANR Experts – Centre National de la Recherche Scientifique - CNRS – France

## Résumé

Depuis le Moyen Âge et probablement plus tôt, les autorités publiques confient à des personnes qu'elles estiment et qualifient compétentes l'action d'émettre un avis sur le savoir technique et scientifique, que ce soit dans le domaine gracieux comme contentieux. Cette recherche conduite dans le cadre d'un projet d'ANR vise à examiner, à partir d'un secteur économique majeur – celui du bâtiment à l'époque moderne –, le mécanisme de l'expertise : comment la langue technique régulatrice et maîtrisée des experts s'impose à la société, comment leur compétence technique se convertit en autorité, voire parfois en "abus d'autorité" ? L'existence d'un fonds d'archives exceptionnel (A.N. Z1J) qui conserve l'ensemble des procès-verbaux d'expertise du bâtiment parisien de 1643 à 1792 nous permet de lancer une enquête pluridisciplinaire (juridique, économique et architecturale) de grande envergure sur la question de l'expertise qui connaît, à partir de 1690, un tournant particulier. En effet, les experts se divisent alors en deux branches différentes exerçant deux activités concurrentes, parfois complémentaires : les architectes et les entrepreneurs. La base de notre travail consistera d'abord à établir parallèlement deux corpus : d'une part, l'établissement d'un dictionnaire prosopographique des 234 experts exerçant de 1690 à 1790 à partir de sources diverses (manuscrites et imprimées) ; d'autre part, l'inventaire et l'analyse des procès-verbaux d'expertise sur la même période.

Ce travail repose principalement sur la constitution de deux corpus : d'une part, le dépouillement systématique de dix années de procès-verbaux conservés dans la sous-série Z1J des Archives nationales, d'autre part la constitution d'une base de données prosopographique sur les experts parisiens pour toute la période étudiée. La base de données prend la forme d'une application web réalisée avec des technologies XML libres et ouvertes qui soutient l'ensemble du travail du projet. Plusieurs formulaires dynamiques ont été développés avec XForms afin de permettre la mise à jour des dépouillements, de pouvoir renseigner la prosopographie et pour disposer de formulaires spécifiques pour le traitement de certaines sources telles que les inventaires après décès ou des annuaires. Cette application permet à l'équipe de partager les données au fur et à mesure de la réalisation du travail et prochainement de publier les résultats avant d'avoir terminé les dépouillements. L'utilisation de la base de données permet également de faire des requêtes riches dans les données pour des usages statistiques, des analyses de réseau ou pour produire des visualisations dynamiques.

La constitution de cet outil réclamait une réflexion approfondie sur la modélisation afin de produire des données qui soient pérennes et interopérables. Dans le cadre d'un partenariat avec les Archives nationales de France, l'équipe souhaite alimenter des instruments de

---

\*Intervenant

recherche à partir des dépouillements réalisés pour lesquels il est nécessaire d'utiliser le standard ISAD-G et de communiquer les données au format XML-EAD du Conseil international des archives. Cependant, les questionnements spécifiques de notre recherche ne pouvaient pas être traités de manière satisfaisante avec ce format. Nous avons développé un schéma spécifique, compatible avec EAD, pour les dépouillements qui a fait l'objet de plusieurs ajustements à partir du travail sur les sources avant de pouvoir être implémenté informatiquement. Ce partenariat suggérait également d'emblée d'avoir recours au format XML-EAC-CPF issu du monde archivistique (norme ISAAR-CPF) pour la description des entités historiques en raison de sa cohérence avec EAD et des opportunités d'échange d'information possibles avec les Archives nationales qui utilisent actuellement ces standards pour ses référentiels d'acteurs.

La volonté de produire des données pérennes et interopérables et cette nécessité de fournir des notices au format XML nous ont naturellement guidés vers ces technologies pour le développement de notre plateforme de travail. L'application a donc été développée avec XQuery, un langage normalisé par le W3C, et la base de données XML native, libre et open source, BaseX. L'utilisation de XQuery et de son extension RESTXQ nous a permis de créer une application web dynamique, sans faire appel à aucun autre langage de programmation. C'est un langage qui peut facilement être maîtrisé par les chercheurs pour exprimer leurs requêtes à la base de données. Pour aller au bout de notre démarche tournée vers les technologies XML, nous avons fait le choix d'utiliser XForms pour l'intégration de tous nos formulaires (avec XSLTForms).

En l'absence de modèles adaptés, nous avons développé notre propre modèle pour le dépouillement des procès-verbaux d'expertise et des inventaires après-décès. Après avoir évalué les différentes solutions existantes pour le volet prosopographique, et en particulier l'extension de CIDOC-CRM, BioCRM, développée dans le cadre de Reassembling the republic of letters, nous avons finalement choisi d'utiliser Ric-CM pour la structuration de nos données, un modèle conceptuel développé par l'ICA (International Council on Archives). Malgré l'attrait de sa simplicité, l'ontologie BioCRM imposait de définir un vocabulaire pour la description des relations sociales (le travail autour de CRMsoc débutant seulement). Ric-CM présentait l'intérêt de s'inscrire dans la continuité des quatre standards archivistiques internationaux existants (ISAD-G, ISAAR-CPF, ISDF, ISDIAH) et d'être très très proche du format XML-EAC-CPF destiné à la description des acteurs historiques (personnes, institutions et groupes) et de leurs relations, tout en étant plus complet. Moyennant quelques aménagements, le format EAC-CPF nous permet de produire des données sémantiques en utilisant les XSLT développées par les Archives nationales de France.

Les trois formulaires pour le dépouillement des expertises, la saisie des données prosopographiques et l'analyse de inventaire après-décès sont maintenant pleinement opérationnels. Plusieurs milliers de procès-verbaux ont été dépouillés et sont reliés à des données prosopographiques sur plus d'une centaine d'experts. Le fait d'employer une pile de technologies unitaire a beaucoup facilité leurs développements. Surtout, nous disposons d'une entière maîtrise sur le produit, ce qui facilite beaucoup son adaptation aux besoins de la recherche. XQuery est un langage de choix pour préparer facilement toutes sortes de jeux de données en vue de l'analyse dans différents formats. Nous l'utilisons pour produire l'interface de consultation mais aussi des ressources au format GEXF ou JSON qui permettent de construire dynamiquement des analyses et des visualisations de réseau à partir de l'application avec des bibliothèques JavaScript comme D3.js ou bien Sigma.js. L'exploitation des données nous permet ainsi de préciser les logiques de répartition des affaires entre experts mais surtout de mener une étude diachronique des deux groupes professionnels, du réseau des acteurs et de sa constitution en rapport aux affaires.

## Bibliographie

- Barbot, Michela, et Robert Carvais. 2020. "Des archives pour analyser la ville et pour dessiner ses territoires". *Histoire urbaine* 59 (3) : 63-84. <https://www.cairn.info/revue-histoire-urbaine-2020-3-page-63.htm>.

- Barbot, Michela, et Robert Carvais. 2019. " Les livres sur le toisé et l'estimation en France et en Italie (XVIe-XIXe siècle) : circulations, continuités, ruptures ". Dans *Le livre technique avant le XXe siècle : À l'échelle du monde*. Sous la direction de Liliane Hilaire-Pérez, Valérie Nègre, Delphine Spicq, et Koen Vermeir, 243-60. Histoire. Paris : CNRS Éditions. <http://books.openedition.org/editionscnrs/27736>.
- Carvais, Robert, et Valérie Nègre. 2015. " Parisian Surveyors (1690-1792): Founding An Expert Corps ". Dans *5ICCH Proceedings*. 383-94. Chicago. <https://structurae.net/en/literature/conference-paper/parisian-surveyors-1690-1792-founding-an-expert-corps>.
- Carvais, Robert. 2015. " Mesurer le bâti parisien à l'époque moderne. Les enjeux juridiques et surtout économiques du toisé ". *Histoire urbaine*, no 43 (novembre) : 31-53. <https://doi.org/10.3917/rhu.043.0031>.
- Carvais, Robert, et Valérie Nègre. 2018. " Experts and building assessments. An international comparison (thirteenth-twentieth century) ". *Verschiedene Materialien*. Dans *Building knowledge, constructing histories*. Sous la direction de Ine Wouters, Stephanie Van de Voorde, Inge Bertels, Bernard Espion, Krista de Jonge, et Denis Zastavni, *Proceedings of the 6th International Congress on Construction History (6ICCH 2018)*, July 9-13, 2018, Brussels, Belgium. Boca Raton London New York Leiden : CRC Press, Taylor & Francis Group.
- Carvais, Robert, Michela Barbot, Emmanuel Chateau-Dutier, et Valérie Nègre. 2018. " Pratiques des savoirs, entre jugement et innovation. Experts et expertises du bâtiment, Paris, 1690-1790. Premiers jalons pour une recherche ". *Projet spontané Rapport final*. Paris : Mission de recherche Droit & Justice. <http://www.gip-recherche-justice.fr/publication/pratiques-des-savoirs-entre-jugement-et-innovation-experts-expertises-du-batiment-paris-1690-1790-premiers-jalons-pour-une-recherche/>.
- Delpu, Pierre-Marie. 2015. " La prosopographie, une ressource pour l'histoire sociale ". *Hypothèses* 18 (1) : 263-74. <https://www.cairn.info/revue-hypotheses-2015-1-page-263.htm>.
- Dozo, Björn-Olav. 2008. " Données biographiques et données relationnelles. Notes théoriques pour une utilisation complémentaire des outils quantitatifs ". *COntEXTES*. *Revue de sociologie de la littérature*, no 3 (juin). <https://doi.org/10.4000/contextes.1933>.
- Hotson, Howard, et Thomas Wallnig, dir. 2019. *Reassembling the Republic of Letters in the Digital Age: Standards, Systems, Scholarship*. Göttingen : Göttingen University Press. <https://doi.org/10.17875/gup2019-1146>.
- Lemercier, Claire, et Emmanuelle Picard. 2012. " Quelle approche prosopographique ? " Dans *Les uns et les autres. Biographies et prosopographies en histoire des sciences*. Sous la direction de Laurent Rollet et Philippe Nabonnaud, 605-30. Presses Universitaires de Nancy ; Éditions Universitaires de Lorraine. <https://halshs.archives-ouvertes.fr/halshs-00521512/document>.
- Nègre, Valérie. 2015. " The Toisés and the Emergence of New Technicians in Eighteenth-Century France ". Dans *Der Lehrbuchdiskurs über das Bauen*. Sous la direction de Uta Hassler, 112-21. Zürich : ETH.
- Genet, Jean-Philippe. 1988. " L'informatique au service de la prosopographie : PROSOP ". *Mélanges de l'École française de Rome* 100 (1) : 247-63. <https://doi.org/10.3406/mefr.1988.2972>.
- Tuominen, Jouni. 2016. " EMLO Prosopographical Data Model: Towards a Biographical Conceptual Reference Model ". *STSM report, Cost Action IS1310*. *Reassembling the Republic of Letters*. <http://www.republicofletters.net/wp-content/uploads/2017/03/cost-stsm-report-tuominen-jouni-2016-04-28.pdf>.

## Références techniques

CIDOC-CRM <http://www.cidoc-crm.org>

BioCRM <http://seco.cs.aalto.fi/projects/biographies/>

CRMSoc <http://www.cidoc-crm.org/crmsoc/>

D3.js <https://d3js.org>

EAC-CPF <https://eac.staatsbibliothek-berlin.de>

EAD <https://www.loc.gov/ead/EAD3taglib/>

GEXF <https://gephi.org/gexf/format>

RESTXQ <http://exquery.github.io/exquery/exquery-restxq-specification/restxq-1.0-specification.html>

Ric-CM <https://www.ica.org/fr/records-in-contexts-modele-conceptuel>

Sigma.js <http://sigmajs.org>

XQuery 3.1 <https://www.w3.org/TR/xquery-31/>.

XForms <https://www.w3.org/TR/2009/REC-xforms-20091020/>.

XSLTForms <http://www.agencexml.com/xsltforms>