

Modelling Time, Places, Agents for the HyperOtlet research project

Henri Sergent ¹

¹HyperOtlet Project, Centre Maurice Halbwachs (ENS)

ABSTRACT

HyperOtlet is a transdisciplinary project about a major work in the history of information sciences: the *Traité de documentation*, published in 1934 by Paul Otlet (1868–1944). A passionate advocate of a new field, documentation, he created original tools and practices and thus, a new knowledge tradition. In its own structure, the 431 pages *Traité* is built on an original information architecture designed by Paul Otlet and meant to facilitate information and knowledge retrieval. Otlet wrote the different entries of the *Traité* on index cards then bundled them in a book which form is very close to an encyclopedia (Van Acker, 2018). He also envisioned a utopia of remote universal access to structured knowledge, books, press, images and audiovisual content.

Our project is combining different profiles (historians, information science researchers, librarians) from various teams and institutions:

- [Centre Maurice Halbwachs](#)
- [MICA](#)
- [ENSSIB](#)
- [MSH Paris Nord](#)
- [Mundaneum](#)

The ANR-funded HyperOtlet project will articulate a set of historical and documented research with a collaborative digital platform and elaborate an augmented corpus around 4 axes:

- new contextualisation of digital culture “documentary regime” (Müller, 2011);
- discussions about current issues of documentation in digital humanities (Le Deuff, 2018), of “documentarity” (Pédauque, 2007) and “documentality” (Ferraris, 2012);
- promote the creation of an epistemic community able to represent a European documentary tradition;
- development of new ways to consult, read and navigate within written works.

In our project we will go beyond this landmark monument (monolithic and finite book) of information science to build and open it to linked information in the form of an hyperdocument (Rayward, 1994). Doing so we, as Paul Otlet foretold, will “abandon resolutely the traditional form of the bound volume, isolated and complete” (Otlet, 1911). In *Monde : Essai D’Universalisme (1935)* Otlet envisioned a knowledge organisation system based on *Temps*, (Time) *Lieux* (Places) *Organisations*, *Personnes*. We aim to model a coherent and reusable data framework based on the possibilities of Omeka S and the particular unique nature of this analogue organisation system from the 1930s.

To do so, we used the recently crowdsourced Wikisource digital version of the *Traité* and split it using a regular expression pattern matching the naming system designed by Otlet himself. Deconstructing the book into these “cards” will ease future text mining and natural language processing operations such as Named-Entity Recognition, and collaborative annotation process by our scientific team. This level of granularity will allow us to tag parts of the *Traité* and extract knowledge out of the dense textual content in a “redocumentarisation”

(Salain, 2007) process. We want to extract most of the named entities (all the persons names, places and organisations, and historical events) and build a directed graph using the same state-of-the-art information architecture but with many more possibilities in terms of design, navigation, information retrieval through a faceted search engine, collaborative trans-media annotation.

This data structure we modelled is very close to the legacy knowledge structure used in the *Traité and other* Otlet's Works itself and will allow easier integration of attributes such as semantic metadata (using controlled vocabularies Dublin Core FOAF, SKOS, BIO, BIBO and existing databases: Geonames). A search engine interface will be part of the application to allow queries and view *Traité de documentation* entries from different points of view (faceted search), thus enabling future data visualisation use cases. We plan to present this ongoing project and its progression during the *Modelling Time, Places, Agents* and we will welcome any insights on our process and be glad to hear from other research approaches as well.

ACKNOWLEDGMENTS

This project is linked to an [Agence Nationale de Recherche](#) funding.

REFERENCES

- Epron, Benoît, and Muller, Catherine. May 2018. “De l’Abécédaire à HyperOtlet – L’Abécédaire des mondes lettrés, un outil d’écriture collaborative savante” *ÉCRIDIL 2018– Écrire, éditer, lire à l’ère numérique –Université de Montréal*. Montréal; Canada. <https://dliis.hypotheses.org/files/2018/10/EcrdilPdf-Muller-Epron.pdf>
- Ferraris, Maurizio. 2012. “Documentality: Why It Is Necessary to Leave Traces”. New York: Fordham University Press.
- Le Deuff, Olivier. May 2018. “HyperOtlet: entre héritage et design prospectif pour les humanités digitales”. May 13–15, 2015, *5th Hyperheritage International Symposium Heritage and experience design in the digital age*. University of Constantine 3 | Constantine, Algeria. <https://f-origin.hypotheses.org/wp-content/blogs.dir/4632/files/2018/11/Ledeuffhyperotlet.pdf>
- Müller, Bertrand. 2011. De l’archive au document. Remarques sur l’évolution des régimes documentaires entre le XIXe et le XXIe siècle. *Territoires contemporains*, Centre Georges Chevrier (CGC), Université de Bourgogne, <https://halshs.archives-ouvertes.fr/halshs-01395409>
- Otlet, Paul. 1911. “L’avenir du livre et de la Bibliographie”. Bruxelles; Paris; Zurich: Institut International de Bibliographie.
- Otlet, Paul. 1934. “Traité de documentation: le livre sur le livre : théorie et pratique”. Bruxelles: Editions Mundaneum-Palais Mondial. https://fr.wikisource.org/wiki/Trait%C3%A9_de_documentation
- Pédauque, R. T. 2007. “La redocumentarisation du monde”. Toulouse: Cépuadès-Éditions.
- Raymond, Christian. Fayolle, Julien. 2010 “Reconnaissance robuste d’entités nommées sur de la parole transcrite automatiquement”, In *Proceedings of "17e conférence sur le Traitement Automatique des Langues Naturelles" (TALN' 10)*, July 2010, Montréal, Québec, Canada. <https://hal.inria.fr/inria-00561732>
- Rayward, W. Boyd. 1994. “Some Schemes for Restructuring and Mobilising Information in Documents: A Historical Perspective.” *Information Processing & Management* 30: pages 163-75. [https://doi.org/10.1016/0306-4573\(94\)90062-0](https://doi.org/10.1016/0306-4573(94)90062-0)
- Rayward, W. Boyd. 1994. “Visions of Xanadu: Paul Otlet (1968-1944) and Hypertext.” *Journal of the American Society for Information Science* 45: pages 235-50. [https://doi.org/10.1002/\(SICI\)1097-4571\(199405\)45:4%3C235::AID-ASI2%3E3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199405)45:4%3C235::AID-ASI2%3E3.0.CO;2-Y)
- Salaün, Jean.-Michel. 2007. “La redocumentarisation, un défi pour les sciences de l’information”. *Etude de communication*, (30) <https://journals.openedition.org/edc/428>
- Van Acker, Wouter. 2018. “Rethinking the Architecture of the Book: Unbinding the Spine of Paul Otlet’s Positivist Encyclopaedism.” *Knowledge Organization* 45(4): 281–291. <https://doi.org/10.5771/0943-7444-2018-4-281>
- Van Den Heuvel, Charles, Rayward, W. Boyd. 2011. “Facing Interfaces: Paul Otlet’s Visualizations of Data Integration”. *Journal of the American Society for Information Science and Technology archive*. Volume 62 Issue 12, Pages 2313-2326 . <https://doi.org/10.1002/asi.21607>
- Vitali-Rosati, Marcello, 2014. “Qu’est-ce que l’éditorialisation ? Suite”. *Culture numérique*. <http://blog.sens-public.org/marcellovitalirosati/quest-ce-que-leditorialisation-suite/>