



HAL
open science

DEF19: un défi technique?

Viera Rebolledo-Dhuin, Jean-Charles Geslot

► **To cite this version:**

Viera Rebolledo-Dhuin, Jean-Charles Geslot. DEF19: un défi technique?. 24ème conférence annuelle de SHARP (Society for the History of Authorship, Reading and Publishing): "Les langues du livre", Jul 2016, Paris, France. , 2016. halshs-01348146

HAL Id: halshs-01348146

<https://shs.hal.science/halshs-01348146v1>

Submitted on 29 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

DEF19 : Un défi technique ?

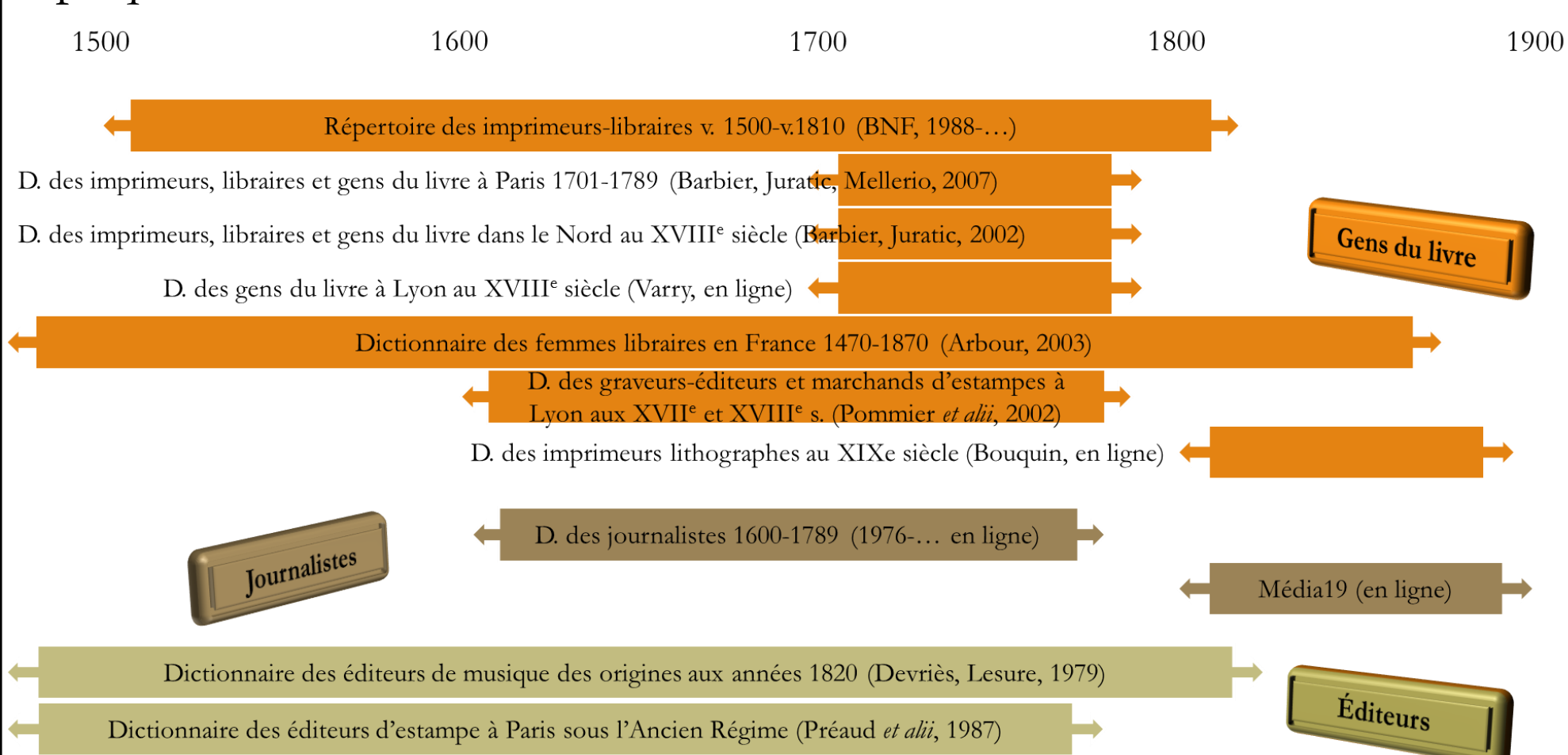
Un dictionnaire prosopographique des éditeurs français du XIX^e siècle à l'heure de la « 3^e révolution du livre »



Groupe DEF19 : Nathalie Aguirre (BnF), Marie-Claire Boseq (UVSQ), Nicole Brondel (AN), Céline Clouet (UVSQ), William Demet (UVSQ), Jean-Charles Geslot (UVSQ), Jean-Dominique Melot (BnF), Jean-Yves Mollier (UVSQ), Viera Rebolledo-Dhuin (UVSQ), Patricia Sorel (UVSQ), Aleth Tisseau Des Escottais (UVSQ), Hélène Veillard (UVSQ), Marion Veyssière (AN).

À l'origine du projet DEF19 : une lacune, deux thèses

Initialement DEF19 vise à combler une lacune, celle de l'absence de dictionnaire ou répertoire systématique des éditeurs à l'époque contemporaine, contrairement à ce qui est à la disposition des chercheurs pour l'époque moderne.



Et s'appuie sur deux thèses, dirigées par Jean-Yves Mollier (UVSQ-CHCSC), qui portent le projet :

- celle de Viera Rebolledo-Dhuin, « La librairie et le crédit. Réseaux et métiers du livre à Paris (1830-1870) », 2011.
- celle de Marie-Claire Boseq, « La librairie parisienne sous surveillance (1814-1848). Imprimeurs en lettres et libraires sous les monarchies constitutionnelles », 2013.

qui ont permis l'édification de deux bases de données pionnières sur les éditeurs parisiens du XIX^e siècle à partir du fonds des faillites d'une part, et des dossiers de brevet d'autre part.

Sources : histoire du livre et prosopographie

Trois corpus de sources principaux constituent l'armature de la base de données DEF19 :

- Les notices d'autorité de la BnF qui comportent des informations sur l'état civil, les adresses fréquentées et des notes sur les activités et/ou titulaires des personnes physiques et/ou des entités morales (1 300 acteurs actuellement).



- Les dossiers de brevet conservés au sein de la série F18 aux AN.

De 1810 à 1881, tout libraire et imprimeur doit détenir un brevet ou une autorisation pour exercer son activité. Les dossiers de brevet, qui contiennent des pièces administratives (acte de naissance, certificats de moralité et de capacités) et des enquêtes policières inscrivant les acteurs dans leur environnement social, ont fait l'objet de deux recensements archivistiques non publiés à ce jour :

Libredit pour les gens du livre parisiens (plus de 5 000 brevets) et **Gutenberg** pour ceux de la France métropolitaine (environ 13 000 brevets), qui seront intégrés dans la base DEF19.

- Les dossiers de faillites, conservés dans les centres des Archives départementales, doivent faire l'objet d'un recensement systématique au sein des départements sinon au sein des grandes villes françaises du livre. Le fonds des faillites est, comme l'ont montré les historiens économistes, riche d'informations quantitatives (taille et durée de vie des entreprises) et qualitatives (histoire familiale et entrepreneuriale, réseaux de crédit). Il reste pourtant sous-exploité en histoire méso-culturelle. C'est ce à quoi entend remédier la base de données DEF19. Par ailleurs, le fonds de faillite s'avère être une entrée pour des recherches prosopographiques s'ouvrant à d'autres corpus de sources.

DEF19 : un double objectif qui s'inscrit dans la révolution des Humanités numériques et de l'Open data

DEF19 vise à créer une base de données relationnelles prosopographiques des éditeurs français du XIX^e siècle, publiable sous deux formes :

- Un dictionnaire prosopographique « papier » sur les éditeurs de 1810 à 1881, utile aux chercheurs en histoire du livre.
- Un « dictionnaire en ligne », portant sur une période plus large (1800-1914) et impliquant un plus grand nombre d'entités (parents non éditeurs, bailleurs de fonds, etc.) qui puisse constituer une véritable base de données interactive et ouverte permettant des analyses qualitatives et quantitatives variées pour situer les éditeurs dans leur environnement (entrepreneurial, familial, social et géographique).

Cela suppose la mise en place d'un outil facilitant la compilation et la saisie de diverses sources, ainsi que la structuration, la consultation et l'analyse des données.

Le programme DEF19 mobilise différents acteurs de l'histoire du livre et utilise les outils actuels de la recherche en Sciences humaines et Sociales (SHS) en s'appuyant sur les infrastructures proposées par les Humanités numériques (HN).

Le choix d'Omeka : valoriser l'ouverture, l'interopérabilité et l'exploitation des données

Après une phase de test de différentes solutions payantes ou gratuites, hébergées en local ou en intranet (MySQL, QWAM, ATOM), nous avons choisi de développer notre base de données sous Omeka S, hébergée par Huma-Num, avec le soutien de Limonade & Co.



Omeka est un logiciel créé par le Centre Roy Rosenzweig pour l'histoire et les nouveaux médias, également fondateur du logiciel bibliographique Zotero. C'est un logiciel créé pour les chercheurs en SHS afin de faciliter entre autres la mise en place d'expositions virtuelles. Il est aujourd'hui utilisé par de nombreuses institutions françaises et donne lieu à une communauté d'utilisateurs en expansion.

La version sémantique d'Omeka 2.4, nommée Omeka S, se fonde sur le langage RDF (Resource Description Framework) qui optimise l'interopérabilité et les ontologies sémantiques des données. Par ailleurs, il propose un ensemble de modules facilitant l'intégration de référentiels externes (de géolocalisation ou bibliographiques par exemple).

Structuration des données de la base DEF19

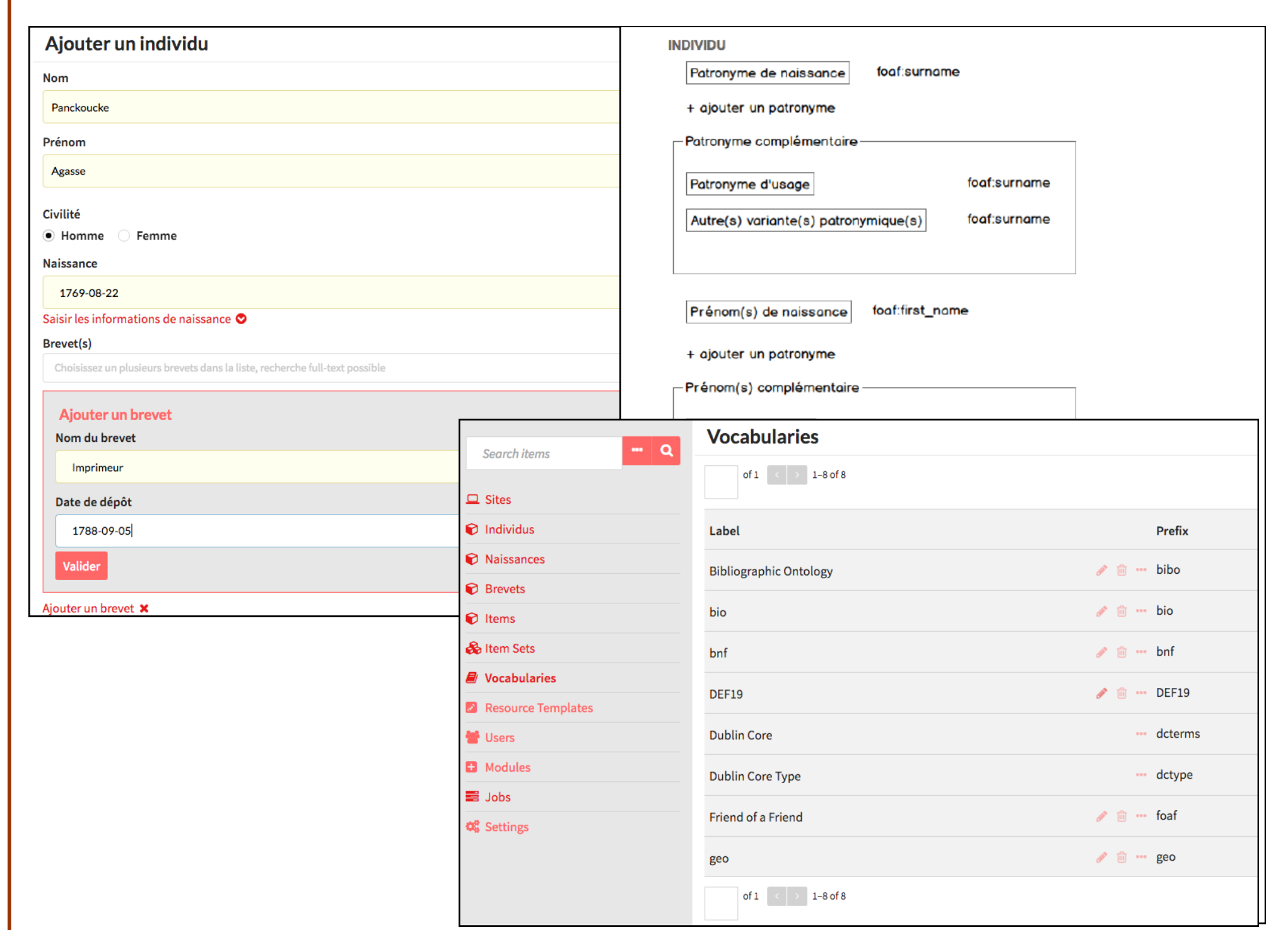
La structuration des données d'un tel programme nécessite de prendre en compte à la fois les sources, la nature des résultats des analyses envisagés et les contraintes induites par le choix du logiciel. Il s'agit de faciliter :

- la saisie et la consultation des données par les chercheurs tout en liant les données entre elles.
- l'import et l'export des données en vue de l'intégration des données existantes et des analyses propres aux axes des chercheurs grâce à d'autres logiciels spécialisés.

Ainsi, quatre thématiques structurent la base de données DEF19 :

- Les informations biographiques (état civil, mariage, etc.),
- Les informations professionnelles et éditoriales (brevet, adresses, spécialités éditoriales, etc.),
- Celles relatives à l'inscription sociale (sociabilités politiques, économiques, de fortune; contentieux, etc.) des éditeurs,
- et aux références archivistiques et/ou bibliographiques.

Limonade & Co met actuellement en place une interface de visualisation et de gestion facilitant la saisie à la volée de nouveaux éditeurs tout en structurant les ontologies sémantiques qui seront aussi utiles à l'analyse des données.



Un partenariat entre quatre institutions du livre, soutenu par l'ANR (Agence nationale de la Recherche)

Le projet DEF19 se fonde sur le travail commun de différents acteurs (chercheurs, archivistes, bibliothécaires, informaticiens) appartenant à des institutions spécialistes de l'histoire du livre :

- Le Centre d'histoire culturelle des sociétés contemporaines de l'Université de Versailles-Saint-Quentin-en-Yvelines (UVSQ-CHCSC),
- Le pôle Métiers du livre de l'Université de Nanterre Paris Ouest La Défense, situé à Saint-Cloud (UPOND-IUT Ville d'Avray-Saint-Cloud-Nanterre),
- Les Archives nationales (AN) qui conservent notamment la série F18, relative aux brevets des gens du livre (libraires, imprimeurs, taille-doucier)
- La Bibliothèque nationale de France (BnF) dont le service de l'inventaire rétrospectif est chargé de la publication des notices d'autorité.

Soutenu financièrement par l'Agence nationale de la Recherche (ANR) depuis juillet 2014 pour une durée de quatre ans, DEF19 est un programme d'ampleur nationale appelant à d'autres collaborations d'envergure internationale.

Collaborations techniques du programme DEF19

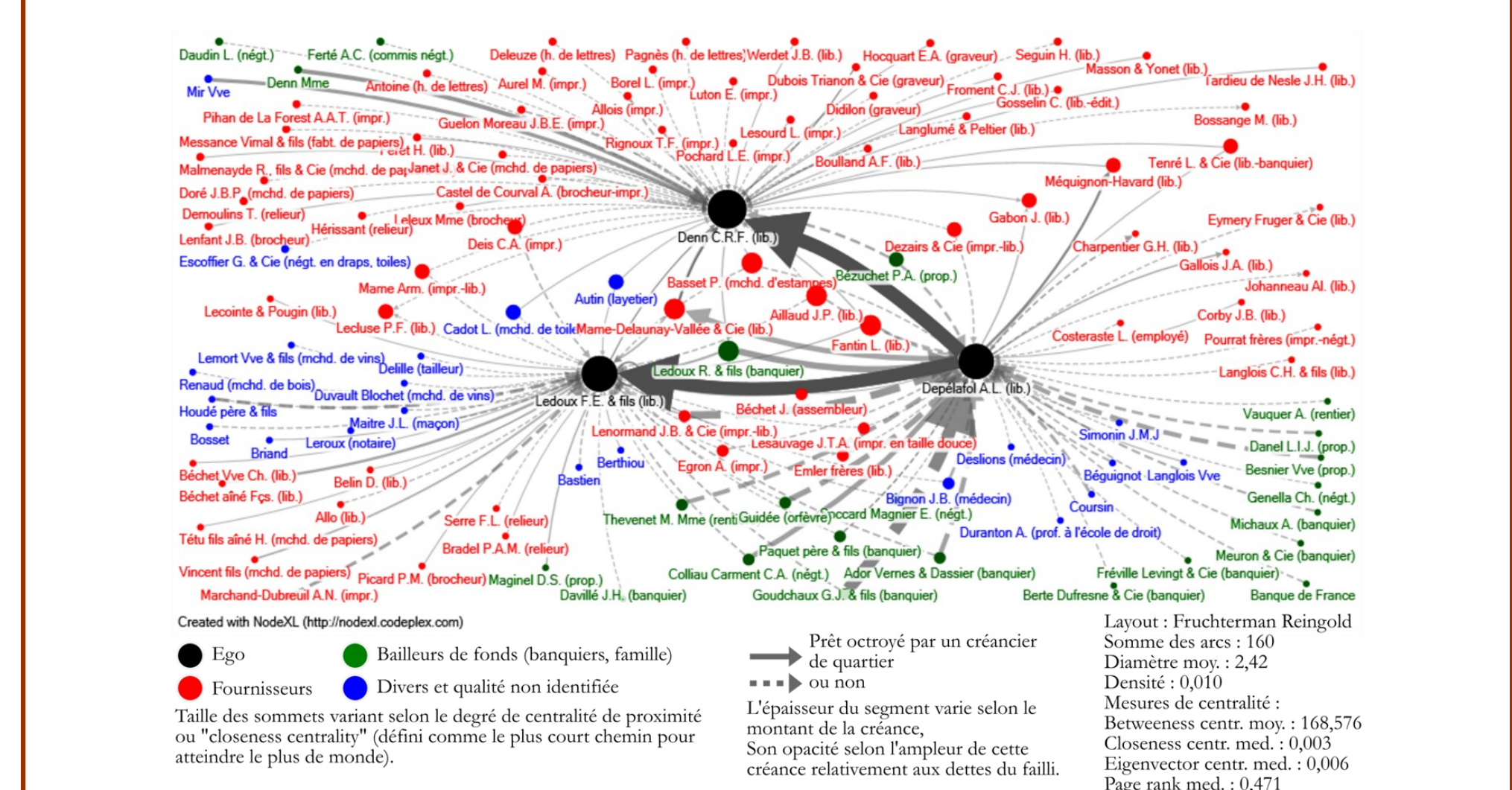
Le développement informatique de la base de données DEF19 implique par ailleurs directement :

- le pôle des services à la recherche de la Direction des Bibliothèques et de l'Information Scientifique et Technique (DBIST) de l'UVSQ (<http://www.bib.uvsq.fr/services-aux-chercheurs>) qui constitue une interface primordiale entre les différents acteurs impliqués directement ou indirectement dans le programme DEF19.
- la très grande infrastructure de recherche (TGIR) Huma-Num (<http://www.huma-num.fr/>), chargée de mettre en œuvre un dispositif humain et technologique à l'échelle nationale et européenne soutenant les projets de HN, qui hébergera la base de données DEF19.
- l'entreprise Limonade & Co (<http://www.limonadeandco.fr/>) qui travaille pour la valorisation des archives et du patrimoine avec de nombreuses institutions publiques et assure la mise en place de la base de données DEF19 sous Omeka.

Les fonctionnalités de DEF19 : ouverture vers d'autres collaborations scientifiques

La base de données DEF19 présentera plusieurs fonctionnalités :

- L'export des données en format texte ouvert sous forme tabulaire (CSV) afin de mener des analyses statistiques, des analyses de réseaux sous le logiciel libre de traitement des données et d'analyses statistiques R (© The R Foundation - <https://www.r-project.org/about.html>) ou des analyses des données généalogiques sous le logiciel de traitement des données individuelles et généalogiques Puck (© Groupe TIIP/CNRS - www.kintip.net et <https://www.kinsources.net/>) par exemple.



Prêt à recevoir par un créancier
de quartier
de zone
l'épaisseur du segment varie selon le montant de la créance.
Son opacité selon l'ampleur de cette créance relativement aux dettes du failli.

- Un plugin de géolocalisation historique qui s'appuiera sur les travaux de GeoHistoricalData (<https://geohistoricaldata.org/>), groupe d'une vingtaine de chercheurs attachés à une vectorisation collaborative de plans anciens de la France et de Paris à différentes échelles du XVIII^e au XX^e siècle.
- L'utilisation du RDF favorisera le dialogue avec l'Atelier Data.bnf.fr chargé de rendre les données de la BnF plus utiles sur le Web.

Pour suivre l'avancement du projet DEF19 et la communauté de chercheurs qui s'y rattache, consultez notre carnet de recherche : <http://def19.hypotheses.org/>



Poster réalisé par Viera Rebolledo-Dhuin et présenté à la 24^e conférence annuelle de SHARP (Society for the History of Authorship, Reading and Publishing) - 20 juillet 2016 - BnF/BULAC, Paris (France).