

Représenter et étudier les individus dans un corpus numérique : le cas de la correspondance d'Henri Poincaré

Representing and studying people in a digital corpus: the case of the Henri Poincaré correspondence

Nicolas Lasolle^{1,2,*}, Laurent Rollet¹

¹Université de Lorraine, CNRS, Université de Strasbourg, AHP-PreST, F-54000 Nancy, France

²Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

Abstract

Henri Poincaré (1854-1912) is a French scientist known for his major contributions to the fields of mathematics, physics and philosophy of science. His correspondence consists of more than 2000 letters, exchanges involving numerous correspondents from his private and professional sphere. Within the *Archives Henri-Poincaré* laboratory, several works are dedicated to the study and the valorization of this historical corpus; with a particular interest in digital publishing and in the proposal of advanced search tools relying on Semantic Web technologies. This article focuses on the question of individuals within a historical corpus, based on the case of this correspondence corpus. Methodological aspects relating to the numerical representation of actors and their relationships are presented. We then discuss methods and tools dedicated to the exploration of the corpus. In particular, recent work has led to the creation of a flexible querying system which can facilitate the discovery of connections between items in the corpus and which can support the identification of networks of people. By presenting historical researches focused on actors, we try to show the interest of digital approaches, in particular to support a study of the corpus from below, which brings to the fore actors and institutions usually considered as secondary.

Keywords

history, digital humanities, Semantic Web, knowledge base, flexible querying

1. Introduction

Henri Poincaré est né à Nancy en 1854, et est décédé à Paris en 1912. Tour à tour étudiant de l'École polytechnique (1873-1875) et de l'École des mines (1875-1878), il a obtenu un doctorat ès sciences mathématiques en 1879 à la faculté des sciences de Paris et s'est rapidement inscrit au sein d'un réseau scientifique et administratif important, comme en témoigne sa correspondance active et passive. Constitué d'environ 2200 lettres et allant de son entrée à l'École polytechnique en 1873 jusqu'à son décès en 1912, ce corpus témoigne de l'évolution professionnelle et personnelle de ce scientifique. L'étude et la valorisation de ce corpus est un projet de longue date qui

. *Workshop on Digital Humanities and Semantic Web*

*. Corresponding author.

. ✉ nicolas.lasolle@univ-lorraine.fr (N. Lasolle); laurent.rollet@univ-lorraine.fr (L. Rollet)

. 🌐 <https://github.com/nlasolle> (N. Lasolle);

<https://poincare.univ-lorraine.fr/fr/membre-titulaire/laurent-rollet> (L. Rollet)

. 🆔 0000-0002-1253-649X (N. Lasolle)

a conduit à la publication de quatre volumes thématiques. Le premier regroupe les échanges avec le mathématicien suédois Gösta Mittag-Leffler¹, échanges réguliers s'étalant durant la majeure partie de la carrière académique d'Henri Poincaré [1]. Le deuxième volume concerne les échanges avec des physiciens, des chimistes et des ingénieurs, [2]. Le troisième constitue la correspondance entre Henri Poincaré, des astronomes, et les géodésiens [3]. Le volume le plus récent publié à ce jour est dédié à la correspondance de jeunesse d'Henri Poincaré [4] lorsqu'il était en formation à l'École polytechnique puis à l'École des mines (1873-1878). Deux autres volumes sont en préparation au moment de la rédaction de ce document et devraient clôturer ce projet d'édition. L'un est dédié à la correspondance avec les mathématiciens et l'autre concerne la correspondance administrative et privée.

L'étude de ce corpus historique présente des enjeux scientifiques variés, de par les nombreux thèmes abordés au sein des échanges tenus par Poincaré [5]. Tout d'abord, l'étude d'une correspondance telle que celle d'Henri Poincaré permet d'apporter des éléments de contexte relatifs à l'élaboration de théories scientifiques. En effet, à la fin du XIX^e siècle, la lettre était un vecteur non négligeable de l'information scientifique [6, 7]. Au cours de sa carrière, Henri Poincaré a tenu une correspondance active avec nombre de ses confrères français et étrangers. Parmi ces correspondants scientifiques, nous pouvons citer le mathématicien suédois Gösta Mittag-Leffler, l'Italien Giovanni Battista Guccia ou encore le mathématicien allemand Felix Klein. L'étude de la correspondance d'Henri Poincaré peut également éclairer le fonctionnement de différentes institutions et sociétés savantes de son époque de par les échanges réguliers qu'il tenait avec plusieurs autres membres. Explorer cette correspondance permet également de mieux comprendre le contexte politique, social et culturel de la France de la fin du XIX^e siècle. Enfin, ce corpus présente Henri Poincaré sous un jour différent de ce qui est mis en avant par l'étude de ses travaux et permet ainsi d'introduire de nouveaux éléments biographiques. Ces échanges épistolaires peuvent révéler plusieurs points de la personnalité de Poincaré relatifs à sa méthode de recherche, ses relations, sa vie privée, etc.

Le corpus rassemble de nombreux acteurs, appartenant au cercle privé et professionnel de Poincaré. On dénote environ 500 correspondants et plus de 1000 personnes citées. Ces individus sont majoritairement français, mais de nombreux acteurs de premier plan sont des scientifiques ou académiciens étrangers tels que Gösta Mittag-Leffler, Felix Klein, ou Giovanni Battista Guccia. Dans le cadre de travaux historiques, de nombreuses recherches peuvent également amener les chercheurs à s'intéresser à des acteurs généralement considérés comme secondaires. Nous proposons de discuter des éléments relatifs au concept d'histoire par en bas avant de présenter le contexte numérique des travaux relatifs au corpus de la correspondance d'Henri Poincaré. En particulier, nous présenterons comment des outils numériques peuvent aider les chercheurs menant des recherches sur le corpus en s'appuyant sur les acteurs. Dans ce contexte, nous présenterons des travaux relatifs à la formalisation et à l'utilisation d'un système d'interrogation flexible pour le Web sémantique. Sans détailler les aspects techniques, nous proposons d'illustrer son fonctionnement et son intérêt au travers de plusieurs exemples issus du corpus de la correspondance d'Henri Poincaré. Ce système a notamment été intégré à un outil de navigation

1. Gösta Mittag-Leffler (1846-1927) est notamment reconnu pour avoir introduit un théorème (portant désormais son nom) relatif à la représentation des fonctions méromorphes par des séries de fractions rationnelles. En 1882, il crée la revue *Acta Mathematica* où publieront plusieurs mathématiciens de l'époque comme Georg Cantor, Sofia Kovalevskaja et Henri Poincaré.

permettant une exploration interactive du corpus.

2. Du mode d'existence des individus dans le corpus poincaréien

Le projet d'édition de la correspondance de Poincaré a été fondé dans les années 1990 en faisant le choix de publier l'ensemble des lettres, c'est-à-dire à la fois les lettres relevant de l'activité scientifique et intellectuelle du mathématicien et celles relevant de la sphère privée et administrative. Une telle décision n'allait pas forcément de soi à cette époque. Ce corpus de plus de 2000 lettres contenait en effet — outre des échanges épistolaires avec des grands noms de la science et de la scène intellectuelle française et étrangère — plusieurs centaines de lettres d'acteurs peu connus ou peu étudiés voire totalement inconnus : des membres de la famille élargie, des amis d'enfance, des amateurs, des journalistes, des inventeurs, des fous littéraires ou scientifiques, etc. Parallèlement à ce travail éditorial les chercheurs ont exploré systématiquement tous les services d'archives susceptibles de contenir des lettres envoyées par Poincaré afin de reconstituer au mieux des échanges épistolaires homogènes². Ces recherches ont ainsi permis de retrouver aux Archives nationales près d'une centaine de traces d'échanges entre Poincaré et les différents ministères dont il a dépendu durant sa carrière.

Que faire de ces documents parfois très riches en termes de contenus mais difficiles à exploiter en raison même des difficultés d'identification des acteurs et de leur mise en situation dans un contexte temporel ou thématique ? En effet de nombreuses lettres reçues par Poincaré n'étaient ni signées ni datées, certaines avaient même été retrouvées sous la forme de projets de brouillons dans les dossiers de carrière de Poincaré sans que l'on puisse disposer de la lettre définitive reçue par lui. De plus, lorsque Poincaré s'adressait, par exemple, au ministre des Travaux publics il était facile de déterminer l'identité du ministre destinataire mais, dans la mesure où la lettre de réponse était rédigée par un fonctionnaire de son cabinet et qu'on ne connaît pas le nom de son auteur, se posait alors la question de savoir s'il fallait attribuer cette lettre au ministre lui-même. La tentation était grande de ne faire porter les efforts de recherche et d'édition que sur les acteurs « intéressants » ayant déjà fait l'objet d'une forme de patrimonialisation historique (les savants de premier ordre, les ministres, les membres de l'Académie française). Le choix a finalement été tout autre : décision a été prise de prendre en compte toutes les lettres quels que soient leurs auteurs ou leurs contenus pour restituer l'image la plus complète possible des travaux, de la personnalité et du rayonnement social, public et mondain du mathématicien. Une telle démarche revenait non pas à considérer que tous les correspondants devaient être mis sur un même pied d'égalité mais qu'ils devaient au moins bénéficier d'une égale considération en termes d'analyse et de traitement éditorial. Plutôt que de se concentrer sur quelques acteurs majeurs pour rendre compte de la vie et de l'œuvre de Poincaré — Paul Appell, Pierre et Marie Curie, Lazarus Fuchs, Gösta Mittag-Leffler, Paul Painlevé pour n'en citer que quelques-uns — nous avons préféré une approche ouverte dans laquelle Poincaré était considéré comme le centre d'un réseau épistolaire dans lequel il pouvait compter d'une manière ou d'une autre pour ses correspondants et dans lequel ces acteurs pouvaient également avoir de l'importance pour

2. Le corpus contient en effet plus de 1000 lettres écrites de la main de Poincaré et près de 1000 qui lui été adressées.

lui. Cette démarche prenait appui sur un parti pris méthodologique en phase avec l'évolution récente de l'historiographie en histoire des sciences.

Avant de se constituer en tant que discipline au début du xx^e siècle, l'histoire des sciences s'est d'abord construite sur le mode des éloges académiques et funéraires des grands savants. Et pendant longtemps elle a gardé de cette filiation une propension très forte à ancrer ses sujets d'études sur les acteurs de premier plan et à se focaliser essentiellement sur leurs activités relevant des sciences. L'ancrage social, voire politique des acteurs, leur inscription dans des réseaux scientifiques, intellectuels et mondains, la diversité des processus d'institutionnalisation des sciences ont souvent été écartés au profit d'approches *internalistes* basées, implicitement ou explicitement, sur le postulat que le travail scientifique peut être analysé indépendamment (ou presque) des contextes sociaux, économiques ou politiques qui le rendent possible. L'émergence des approches sociales et sociologiques des sciences et techniques a partiellement rebattu les cartes méthodologiques. Depuis quelques décennies, les historiens des sciences ont profondément revisité leurs méthodes et leurs sujets d'études : afin de mieux articuler approches internalistes et externalistes ils sont devenus très sensibles à la l'élaboration d'une *histoire par en bas* ou d'une *histoire populaire* des sciences qui donne la une place plus équilibrée aux acteurs secondaires : artisans, techniciens, ingénieurs, amateurs, acteurs professionnels des sciences, enseignants, militaires des armes savantes, enseignants, étudiants, etc.³

L'intérêt de cette approche est qu'elle invite à faire un pas de côté et à dépasser l'image parfois caricaturale de la *Science* au profit de la perspective plus large des *savoirs scientifiques en sociétés* ; elle invite également à décentrer le regard en considérant avec sérieux des acteurs qui bien qu'ils n'aient pas toujours bénéficié d'un fort capital académique ou scientifique ont pu jouer un rôle non négligeable dans le champ scientifique et intellectuel⁴.

La mise en œuvre de ce type d'approche dans le cadre de l'édition de la correspondance de Poincaré n'est pas sans conséquences sur le travail à réaliser. Tel qu'il a été conçu, que ce soit dans les ouvrages publiés ou en cours d'édition ou sur la plate-forme numérique dédiée, ce projet vise à « épuisier » les contextes de Poincaré et d'élaborer une encyclopédie totale sur cet acteur de premier plan en exploitant toutes les traces et sources disponibles⁵. Il est conçu dans une logique de valorisation et de vulgarisation scientifique avec pour perspective finale l'élaboration d'une biographie du mathématicien. Pour ce faire, un important travail doit être réalisé sur chacun des individus – correspondants, personnes citées dans les lettres et dans les apparats critiques – pour caractériser *a minima* leur identité sociale et professionnelle et déterminer les modalités de leurs relations avec Poincaré. La base de connaissances contient en effet des milliers de noms d'acteurs qui peuvent être des correspondants (pas toujours facilement identifiables), des amis, des « ennemis », des relations familiales, de vagues connaissances, des inconnus (non

3. Un bon exemple de cette approche est le livre de Clifford D. Conner *Histoire populaire des sciences* [8]. Celui-ci s'inspire d'ailleurs des approches développées en histoire générale, notamment par Howard Zinn dans son *Histoire populaire des Etats-Unis. De 1492 à nos jours* [9]. Pour une approche similaire en France, voir le livre de Gérard Noiriel *Histoire populaire de la France, de la Guerre de Cent Ans à nos jours* [10].

4. On ne donnera ici que quelques exemples représentatifs. La Société astronomique de France créée en 1882 par l'astronome et vulgarisateur Camille Flammarion a joué un rôle essentiel pour la structuration des sciences astronomiques mais elle comptait parmi ses membres, outre les grands noms de ces disciplines, des centaines d'amateurs ou de non professionnels, dont un nombre significatif de femmes.

5. Le site Web de la correspondance n'est d'ailleurs qu'une partie d'un écosystème numérique qui inclut d'autres plates-formes dédiées à ses publications, à sa biographie et aux documents iconographiques le concernant.

seulement des historiens mais de Poincaré lui-même), des collègues, des personnalités politiques, des philosophes, des acteurs de l'histoire politique ancienne ou récente, des auteurs d'œuvres littéraires⁶.

Pour ne donner qu'un exemple, la lettre relativement anodine de Poincaré ci-dessous, datant probablement de mars 1874, fait le récit d'une conversation lors d'un voyage en train de Nancy vers Paris avec des camarades de promotion à l'École polytechnique. Elle porte sur la beauté de jeunes femmes de la bonne société nancéienne puis sur les classements intermédiaires après les examens de février 1874. On y trouve près d'une vingtaine de références à des individus. Pour chacun d'entre eux, il a été nécessaire d'effectuer des recherches approfondies afin de les identifier et de déterminer la nature du lien relationnel qu'ils entretenaient avec Poincaré. On constate la présence d'au moins 5 strates de personnes qui renvoient à différents univers sociaux : des camarades polytechniciens ayant fait leurs études avec Poincaré au lycée de Nancy (les 2 premières lignes de la lettre), un groupe de 4 jeunes femmes nancéiennes, parfois aristocrates, qui renvoient à l'univers amical⁷, familial et mondain de la famille Poincaré à Nancy (lignes 2 et 3), une référence à un très proche ami d'enfance alors élève à l'École militaire de Saint-Cyr (Élie Rinck, lignes 3 et 4), un ensemble de camarades de promotion de l'École polytechnique (second paragraphe de la lettre) et une référence à une dame nancéienne, Mme Dreyfuss, qui conduisait le cotillon (le bal), lors d'une soirée mondaine. En outre, au-delà des noms évoqués par Poincaré, cette lettre fournit d'intéressants indices sur la personnalité et la psychologie de Poincaré dans sa jeunesse et met bien en évidence l'univers social et mondain dans lequel il évoluait alors.

La question qui se pose à l'historien éditeur de la lettre est alors de déterminer de quelle manière caractériser toutes ces personnes, pour autant qu'il soit possible de le faire. Cela revient à se poser des questions que seules une approche biographique et une connaissance approfondie des sources biographiques peuvent éclairer : ces personnes comptaient-elles pour Poincaré et à quel titre ? Quel était le degré de proximité de Poincaré avec elles ? Poincaré comptait-il réciproquement pour ces personnes ? S'agissait-il de personnes sans importance dans son vécu social ou jouaient-elles un rôle quelconque dans son parcours de vie ? Se pose également la question des sources mobilisables : il est possible de s'appuyer sur des données « objectives », à savoir d'autres lettres, des réseaux de citations, des photographies, des documents d'archives, des récits familiaux comme le journal rédigé par la sœur de Poincaré [11]. Il est également possible d'exploiter des données « indirectes », sujettes à des interprétations et à des hypothèses, telles que les connexions généalogiques identifiables, d'autres fonds de correspondance, des témoignages de seconde main, des bases de données en ligne telles que celle recensant tous les élèves passés par l'École polytechnique. Mais la grande difficulté est d'équilibrer l'investissement en termes de recherche d'informations en fonction de l'importance réelle ou supposée de ces personnes vis-à-vis de Poincaré. Dans le cadre de cette lettre, il apparaît qu'au regard de la biographie poincaréienne, les noms de Rinck, Bonnefoy, Petitdidier et Corps sortent nettement du lot car les sources épistolaires attestent l'existence d'une relation suivie entre Poincaré et ces

6. Ainsi durant ses années d'études à l'École polytechnique et à l'École des mines (1873-1878), Poincaré allait très souvent au théâtre et à l'opéra et il adorait les pièces d'opéra-bouffe ou les vaudevilles. Sa correspondance de jeunesse abonde de références aux auteurs de ces œuvres ainsi qu'aux comédiens et comédiennes qu'il voyait régulièrement sur scène [4]

7. Mademoiselle Jacquinet était ainsi la fille du recteur de l'Académie de Nancy et une amie proche de la sœur de Poincaré [11].



LA CORRESPONDANCE D'HENRI POINCARÉ

[ACCUEIL](#)[LETTRES](#)[RECHERCHE](#)[INDEX](#)

Lettre : Henri Poincaré à Eugénie Poincaré - 11 ou 12 mars 1874

[Modifier l'item](#)[Scan](#)[Transcription](#)[Métadonnées](#)[Citer ce document](#)

[11 ou 12 mars 1874]

Ma bonne maman,

J'ai fait un voyage très agréable dans un compartiment d'*x* ; nous en avons encore ramassé en route : Barré, Maniguet, Gilliot, Weiss¹, Bailliot², Ringenbach et moi. Du reste rien de particulier. Dreyfuss³ a complètement oublié M^{lle} de Pruneuf ; il maintient sa cote 15 pour M^{lle} Clarinval ; 20 pour M^{lle} Norberg, 18 pour M^{lle} Matuszinska⁴ et Jacquinet⁵. Rinck⁶ cote 20 M^{lle} Matuszinska ; 17 M^{lle} Norberg et 16 M^{lle} Jacquinet. Quant à moi je maintiens mes cotes. Élie trouve que M^{lle} Norberg a les jambes trop courtes.

On a affiché le classement. Bonnefoy est second, Petitdidier troisième ; Debray 8, Belleville 6, Monestier⁷ 4 ; Maniguet 70 ; Herpin 12 ; Pinat 9, Mauger 85. Corps 44 ; Ruault 108 ; Le Mahieu⁸ 212 ; Barré 70 ; Gilliot 20 ; Millot 45.

C'est M^{me} Dreyfuss qui conduisait le cotillon⁹.

Figure 1 : Transcription d'une lettre d'Henri Poincaré à Eugénie Poincaré, telle que présentée sur le site <http://henripoincare.fr>.

personnes.

Face à ces milliers de personnes l'enjeu est donc de constituer une base de connaissances qui permette d'identifier ces individus et de les caractériser en termes sociaux sans tomber dans le piège de l'exhaustivité. Il ne s'agit donc pas d'élaborer des biographies individuelles précises — les biographies sur les grands correspondants de Poincaré ne manquent pas — mais de leur attribuer *a minima* une identité personnelle, sociale et professionnelle en n'oubliant pas que cette caractérisation a pour destination l'explicitation du corpus poincaréen. Cela revient ainsi à définir ces acteurs par rapport à une typologie de métiers (professeur, homme politique, actuaire, ingénieur, etc.), à des attaches institutionnelles (École polytechnique, Académie des sciences, Académie des sciences, Banque de France, Société mathématique de France, etc.) ou à des appartenances disciplinaires (mathématicien, physicien, philosophe, etc.). Une telle entreprise peut sembler relativement simple mais peut s'avérer complexe.

Comment ainsi caractériser l'épouse de Poincaré, Louise Poulain d'Andecy (1857-1834)? La facilité serait sans doute de la définir comme « femme au foyer et épouse de Poincaré » car elle n'eut manifestement aucune activité professionnelle [12]. Mais les sources nous indiquent qu'elle fut tout à la fois pour Poincaré sa secrétaire, sa « chargée de relations » et la

relectrice et correctrice de ses manuscrits (tout comme d'ailleurs ses enfants). Dans le même ordre d'idée, caractériser les attaches professionnelles de Poincaré, ou d'un très grand nombre de ses correspondants, s'avère complexe dans la mesure où beaucoup étaient membres de plusieurs dizaines d'académies et de sociétés savantes⁸. Par ailleurs, s'agissant des disciplines, un problème similaire se pose : faut-il caractériser Poincaré comme mathématicien, comme philosophe, comme physicien, comme astronome, comme administrateur de la recherche ? Tous ses qualificatifs peuvent s'appliquer à son parcours de vie et de carrière et ils posent la question redoutable de la profondeur temporelle que l'on veut attribuer à la base de connaissances des individus — Poincaré n'était pas mathématicien à 2 ans et il n'a publié des articles relevant de la philosophie qu'à partir du début des années 1890. Enfin, le statut des femmes est à considérer avec minutie : Eugénie Launois (1830-1897) était la mère de Poincaré et de sa sœur Aline ; elle était également l'épouse du père de Poincaré, Émile Léon Poincaré (1828-1890) et avait donc pour nom de mariage « Eugénie Poincaré ». Cette simple caractérisation ouvre une infinité de possibles : Eugénie Launois *alias* Eugénie Poincaré *alias* l'épouse du père de Poincaré *alias* la femme d'un professeur à la faculté de médecine de Nancy *alias* la mère de Poincaré *alias* la belle maman de Louise Poulain d'Andecy (l'épouse de Poincaré) *alias* la maman d'Aline Poincaré, sœur de Poincaré et future épouse du philosophe Émile Boutroux, une influence philosophique majeure pour Poincaré...

Pour finir, le traitement des individus dans un corpus de correspondance s'avère complexe lorsqu'il s'agit de les intégrer dans une base de connaissances. Il oblige à mener en amont une réflexion sur la notion d'identité sociale et sur ses transformations à travers le temps de vie d'un acteur. Il nécessite aussi d'analyser différents modes d'existence des acteurs : ce qu'ils sont par et pour eux-mêmes, ce qu'ils sont par et pour les autres, ce qu'ils déclarent / croient / aimeraient être, ce qu'ils sont au sein (ou en dehors) d'un groupe social, ce que les chercheurs et historiens déclarent ce qu'ils sont, ce qu'ils sont dans une représentation centrée sur le corpus poincaréen, ce qu'il sont lorsqu'on les caractérise à une époque spécifique en lien avec un échange épistolaire spécifique. Et ce ne sont que quelques possibles dont l'historien et l'informaticien doivent tenir compte. Dans la suite de cet article, nous proposons de détailler les approches numériques mises en place pour l'exploitation de ce corpus.

3. L'édition numérique

Depuis plusieurs années, les Archives Henri-Poincaré mènent différents travaux numériques autour du corpus de la correspondance d'Henri Poincaré. Le site Web utilisé pour éditer et publier le corpus de la correspondance d'Henri Poincaré est géré grâce au gestionnaire de contenus (CMS) Omeka S⁹ [13]. Ce système permet la mise en valeur numérique de collections liées au patrimoine culturel, qu'elles soient issues de musées, bibliothèques ou centres d'archives.

Le site Web lié au corpus est accessible à l'adresse <http://henripoincare.fr>, et se compose de 4 sous-sites : l'un dédié au corpus de la correspondance ; un deuxième qui s'articule autour de la bibliographie sur et par Poincaré ; un troisième qui présente des éléments biographiques relatant le parcours de Poincaré et un dernier proposant une iconographie regroupant des dizaines

8. Au cours de sa carrière Poincaré fut élu dans 45 académies françaises et étrangères.

9. <https://omeka.org/s/>.

LETTRE : Gösta Mittag-Leffler à Henri Poincaré - 22 mai 1881

Transcription Métadonnées Citer ce document

Helsingfors 22 mai 1881
Finlande

Monsieur,

Permettez-moi d'abord de vous remercier cordialement de votre lettre aimable¹ datée le 22/4 et du cadeau de votre thèse.² Je n'ai pas eu le temps encore d'étudier sérieusement celle-là mais je l'ai parcourue à la hâte ce qui m'a suffi pour voir combien des choses nouvelles vous y donnez et le premier moment que j'aurai libre, je veux employer à en approfondir l'étude.

Monsieur **Hermite** m'a envoyé votre travail : "*Sur les fonctions à espaces lacunaires*"³ et il m'a prié de le présenter dans son nom et le votre à notre société des sciences. La société a été très sensible de ce cadeau et m'a prié de vous présenter ses remerciements. Je vous envoie une épreuve en deux exemplaires en vous priant de vouloir bien me renvoyer l'une après y avoir fait les changements que vous trouverez convenables.

Et permettez-moi de vous dire franchement et loyalement que je trouve que vous devez faire ressortir les rapports que votre travail a avec les recherches de Monsieur **Weierstrass** publiées dans le "*Berliner Monatsbericht*" Août 1880 sous le titre "*Zur Functionenlehre*".⁴ Votre manière de définir une fonction — page 3 — est exactement la même que Monsieur **Weierstrass** emploie depuis 30 ans déjà, et vous trouvez les mêmes idées clairement développées dans le mémoire : "*Zur Functionenlehre*", page 12.⁵ C'est sur cette définition même que Monsieur **Weierstrass** a construit tout ce système sublime qu'il développe dans son cours à l'université de Berlin et qui embrasse la théorie générale des fonctions, la théorie des fonctions elliptiques, la théorie des fonctions Abéliennes et bien d'autres choses encore.⁶ Vous avez tort quand vous dites que Monsieur **Hermite** a mis le premier en lumière l'existence des fonctions à "espaces lacunaires".⁷

Vous ne pouvez pas savoir que Monsieur **Weierstrass** a parlé de telles fonctions depuis des années dans son cours mais dans le travail : "*Zur Functionenlehre*" il en donne l'exemple et met en lumière justement cette propriété. Les deux fonctions représentées par la série

$$\sum_{\nu=0}^{\infty} \frac{1}{x^{\nu} + x^{-\nu}}$$

— voir les pages 5, 13, 14 en "*Zur Functionenlehre*"⁸ — sont des telles fonctions à "espaces lacunaires"⁹ et la fonction remarquable

$$\sum_{\nu=0}^{\infty} b^{\nu} x^{a^{\nu}}$$

où b est un nombre positif plus petit que 1, a un nombre entier inégal¹⁰ et positif et

$$ab > 1 + \frac{3}{2}\pi$$

est aussi une telle fonction — voir les pages 26 et 27 en "*Zur Functionenlehre*"¹¹ —. Votre fonction

$$1 + \frac{1}{2}x^3 + \frac{1}{2^2}x^{3^2} + \dots + \frac{1}{2^n}x^{3^n} + \dots$$

Figure 2 : Extrait de la transcription d'une lettre telle que présentée sur le site henripoincare.fr.

d'images présentant Poincaré, sa famille, ses collègues, etc. Cet ensemble de sites Web est hébergé par Huma-Num [14] qui propose des services adaptés aux projets d'humanités numériques. La figure 2 présente une capture du site, correspondant à un extrait de la transcription d'une des premières lettres envoyées par le mathématicien suédois Gösta Mittag-Leffler à Henri Poincaré, le 22 mai 1881. Cette lettre traite notamment des travaux de thèse effectués par Poincaré et des liens soulignés par Mittag-Leffler avec les travaux du mathématicien allemand Karl Weierstrass, l'une des figures majeures de « l'École de Berlin », qui a apporté des contributions fondamentales dans plusieurs domaines tels que la théorie des fonctions analytiques et le calcul des variations.

Omeka S permet de représenter les données sous la forme de triplets dans la logique du

```
<lettreA expéditeur henriPoincaré>
<lettreA destinataire marieSklodowskaCurie>
<lettreA rédigéÀ Paris>
<lettreA thème "Prix nobel">
<lettreA dateDeRédaction 1911-12-11>
```

Figure 3 : Un exemple de triplets décrivant partiellement une lettre.

modèle RDF¹⁰ [15], modèle principal pour représenter les données du Web sémantique. Un ensemble de triplets RDF forme un graphe composé de trois types de nœuds : des *ressources nommées*, des *ressources anonymes* et des *littéraux*. Une *ressource nommée* est identifiée par un IRI et permet de décrire une classe (p. ex. *Personne*, *Mathématicien*, *Lettre*, etc.), une propriété (p. ex. *expéditeur*, *destinataire*, etc.) ou une instance (p. ex. *henriPoincaré*, *lettre11*, etc.)¹¹. Une *ressource anonyme* représente une ressource qui n'est pas explicitement identifiée (*nœud vide*). Un littéral correspond à une valeur constante d'un type donné (entier, chaîne de caractères, date, etc.), c'est un objet atomique et primitif.

Il est possible de définir des relations entre les nœuds du graphe par l'utilisation de propriétés décrivant les ressources les composant. Ces relations sont caractérisées par des triplets de la forme *< sujet prédicat objet >*. Le sujet représente la ressource (nommée ou anonyme) à décrire. Le prédicat est une propriété qui décrit cette ressource. L'objet est la valeur associée à la propriété et peut être une ressource nommée, une ressource anonyme ou un littéral. La figure 3 donne un exemple de description partielle d'une lettre de la correspondance d'Henri Poincaré.

Bien qu'Omeka soit un outil s'inscrivant dans le mouvement du Web sémantique¹² il présente plusieurs limites [16]. En effet, il n'est pas possible d'intégrer plusieurs éléments introduits par le langage RDFS tels que des hiérarchies entre des classes (par ex. une lettre autographe est une forme de lettre qui est une forme de document) et entre des propriétés (un destinataire est un correspondant). Ce type de relations est le support de mécanismes d'inférences permettant de tirer parti des connaissances associées à un domaine. De plus, dans le contexte du Web sémantique, le langage SPARQL permet de formuler des requêtes expressives afin d'exploiter les liens entre les ressources [17]. Les données Omeka S sont stockées à l'aide d'une base de données MySQL dédiée qui ne peut pas être directement interrogée par des requêtes SPARQL.

Ces différents éléments ont conduit à la création d'une base au format RDF, alimentée quotidiennement par un export et une conversion des données du site Omeka S¹³. Les données de cette base sont exposées au travers d'un point d'accès SPARQL, qui permet l'interrogation du graphe RDF. Par exemple, la requête Q (représentée de façon informelle avec Q) peut être formulée pour interroger cette base.

10. *Resource Description Framework*.

11. Dans un souci de lisibilité, les ressources nommées ne sont pas représentées en utilisant des IRI complets dans cet article.

12. Omeka S est une version « sémantique » du logiciel Omeka, qui propose l'intégration d'ontologies et la possibilité de lier des ressources, comme le fait le modèle RDF.

13. Ce script, développé avec le langage Python, est générique et peut être réutilisé pour toute base Omeka S. Le code source et des détails sur son utilisation sont disponibles sur un dépôt GitHub public (<https://github.com/nlasolle/omekas2rdf>).

$$Q = \left\{ \begin{array}{l} \text{SELECT ?\ell} \\ \text{WHERE } \{ \\ \quad \text{?\ell a Letter .} \\ \quad \text{?\ell sentBy henriPoincaré .} \\ \quad \text{?\ell cite ?x .} \\ \quad \text{?x a Mathematician .} \\ \quad \text{?\ell subject "geometry"} \\ \} \end{array} \right.$$

$$Q = \left\{ \begin{array}{l} \text{Donner les lettres envoyées par Henri Poincaré} \\ \text{qui citent un mathématicien et traitant de géométrie.} \end{array} \right.$$

4. Représentation des correspondants et des personnes citées

Lors de l'édition numérique du corpus, les acteurs ont tout d'abord été décrits assez brièvement par l'édition du nom, des prénoms, des dates de naissance et de décès, ainsi que par une courte description lorsque suffisamment d'information était disponible. Récemment, la volonté d'ajouter des détails pour les acteurs du corpus est apparue, notamment afin de mettre en avant leur place dans la chronologie poincaréenne, et dans l'optique de mettre en avant des réseaux de personnes, qu'ils soient déjà connus ou non. Ces réseaux peuvent correspondre à des réseaux privés, par le biais de relations familiales ou amicales ; à des réseaux scientifiques, par le biais de travaux communs, d'échanges scientifiques ; à des réseaux académiques, au travers des liens étroits qu'entretenait Poincaré avec certains enseignants et membres d'établissements de formation. Il est un enjeu majeur que de rendre compte des rôles tenus par les correspondants au sein de ce corpus.

Pour cela, les acteurs sont décrits par un ensemble de métadonnées précisant des informations à leur propos. Cela comprend tout d'abord un lieu de naissance indiquant la ville et le pays de naissance¹⁴. La ou les nationalités attachées aux personnes est indiquée en complément de ces informations.

Afin de rendre compte des activités des acteurs, un ensemble de classes correspondants à des situations sociales et professionnelles ont été créées. Ce travail est délicat, car il nécessite d'effectuer des simplifications de la réalité, qui peuvent parfois nous sembler inadaptées. Il peut néanmoins encourager la prise de recul et force parfois à se replonger dans leur histoire et à lever certaines ambiguïtés lorsque c'est possible. Plusieurs groupes principaux ont émergé, tels que l'ensemble des universitaires, des scientifiques, des ingénieurs, des académiciens, des

14. Se pose la problématique des évolutions administratives relatives aux pays. Pour certains acteurs, le pays de naissance n'existe plus en tant que tel. C'est notamment le cas pour Marie Skłodowska-Curie, née en 1869 à Varsovie dans le royaume du Congrès, qui était alors une province de l'empire Russe, et située géographiquement dans l'actuelle Pologne. Dans ces situations, le choix a été effectué de renseigner le pays de naissance tel qu'il existe actuellement. Cela est nécessaire afin de créer des relations entre le graphe du corpus de la correspondance et le graphe public GeoNames, où les pays sont listés selon les entités administratives associées telles qu'elles existent de nos jours. À chacune de ces entités sont associées des coordonnées géographiques.

éditeurs scientifiques, etc. Certains acteurs sont également associés à un rôle déterminé par leur profession principale, telle que la profession d'écrivain, de journaliste, d'avocat, de traducteur, etc. Il est important de souligner que la plupart des acteurs ne sont pas limités à une seule description, et que nombre d'entre eux appartiennent à plusieurs groupes.

De nombreux acteurs sont également indissociables de leurs rôles au sein d'administrations et de sociétés savantes. Il a été choisi de n'indiquer que les liens les plus significatifs, pouvant apporter un éclairage important aux échanges tenus avec Poincaré. Parmi les institutions les plus souvent citées, nous pouvons évoquer l'Académie des sciences, la *Royal Society*, le Cercle mathématique de Palerme, ou encore le Bureau des longitudes.

Pour certains acteurs, les échanges et les relations avec Poincaré sont dissociés de leurs fonctions ou des rôles qui leur sont généralement attribués en dehors du corpus. C'est pourquoi il est apparu nécessaire de distinguer des informations générales d'informations spécifiques au corpus. Dans ce contexte, un ensemble de propriétés ont été définies pour décrire la façon dont Poincaré connaît et interagit avec ces acteurs. Cinq types de réseaux ont été identifiés pour ce corpus : le réseau familial, le réseau amical, le réseau scientifique, le réseau mondain et le réseau polytechnicien. Encore une fois, de nombreux acteurs sont liés à plusieurs de ces réseaux. Par exemple, le mathématicien Paul Appell, camarade de Poincaré au lycée impérial de Nancy, collègue avec qui il a scientifiquement échangé tout au long de sa carrière qui était également un ami proche¹⁵.

Mener cette édition numérique implique de devoir effectuer de nombreux choix pour représenter les données du corpus. Bien que ces choix puissent parfois être contraignants, car ils correspondent à une simplification de la réalité et ne traduisent que partiellement certains faits, cet exercice peut cependant encourager la prise de recul sur les objets étudiés et peut ainsi contribuer à renouveler les recherches. Dans ce contexte, il est important de rendre compte de ces choix pour éviter les ambiguïtés, pour renforcer la cohérence des travaux, et pour présenter la méthodologie de recherche utilisée qui pourrait inspirer d'autres chercheurs. À titre d'exemple, le tableau 1 regroupe des statistiques relatives aux cinq correspondants avec lesquels Poincaré a le plus échangé.

Dans la suite de cet article, nous présentons la façon dont les données du corpus peuvent être exploitées au travers d'outils numériques.

5. Un outil d'interrogation flexible pour explorer le corpus

5.1. L'idée de transformation de requêtes

Au sein du site dédié à Henri Poincaré, une interface utilisant l'outil Solr [19] permet d'effectuer des recherches plein texte pour les lettres ayant une transcription¹⁶. Par exemple, une recherche après la saisie du terme « géométrie » retourne un ensemble de 26 lettres rédigées ou reçues par Henri Poincaré. Bien qu'il soit utile dans de nombreux cas, cet outil ne profite pas des technologies du Web sémantique. Le langage SPARQL est plus expressif et permet de formuler des requêtes complexes. Cependant, l'utiliser nécessite de maîtriser une syntaxe particulière

15. Paul Appell a d'ailleurs publié en 1925 un ouvrage biographique afin de rendre hommage à son ami Henri Poincaré [18].

16. Environ 60% des lettres sont associées à une transcription rédigée en \LaTeX puis convertie en HTML.

Table 1

Statistiques sur les cinq correspondants avec lesquels Poincaré a le plus échangé.

Personne	Lettres échangées	Citations (lettre)	Citations (apparat)	Description
Mittag-Leffler, Gösta (1846-1927)	262	54	179	Mathématicien suédois, créateur d' <i>Acta Mathematica</i>
Guccia, Giovanni Battista (1855-1914)	72	7	34	Mathématicien italien, éditeur des <i>Rendiconti del Circolo Matematico di Palermo</i>
Darwin, George Howard (1845-1912)	50	2	5	Mathématicien et astronome britannique
Klein, Felix (1849-1925)	41	37	48	Mathématicien allemand
Saint-Arroman, Raoul de (1849-1915)	39	2	2	Écrivain, journaliste et haut fonctionnaire français

qui n'est pas facile d'utilisation pour tous les historiens et visiteurs du site. Dans ce contexte, des collaborations entre historiens et informaticiens s'intéressent à l'élaboration et aux usages d'outils de recherche interactifs pour explorer le corpus de la correspondance d'Henri Poincaré. En particulier, des travaux visent à proposer des outils basés sur un mécanisme d'interrogation flexible, s'intégrant dans le cadre du Web sémantique [20].

Imaginons un historien des sciences à la recherche d'informations concernant la *mécanique rationnelle* dans la correspondance d'Henri Poincaré. Un point de départ possible serait de s'intéresser aux échanges avec Paul Appell qui a rédigé plusieurs traités de mécanique rationnelle, notamment un premier volume paru en 1893 [21]. Soit \mathcal{Q} une requête informelle¹⁷ formulée par l'historien :

$$\mathcal{Q} = \left\{ \begin{array}{l} \text{Donner les lettres envoyées entre 1890 et} \\ \text{1895 par Paul Appell à Henri Poincaré et} \\ \text{qui mentionnent des travaux en mécanique.} \end{array} \right.$$

L'exécution de cette requête sur la base RDF du corpus de la correspondance pourrait retourner des résultats qui ne sont pas satisfaisants pour l'historien. Les raisons suivantes de cette insatisfaction peuvent être considérées :

- l'ensemble des résultats est trop grand : l'historien voudrait spécialiser sa requête ;
- l'ensemble des résultats est vide ou trop petit : il souhaiterait généraliser sa requête ;
- les résultats obtenus n'apportent pas de réponse à la problématique initiale : il devrait envisager de nouvelles pistes de recherche.

Une approche pour remédier à ce problème consiste en la définition et l'application de règles de transformation de requêtes. Ces règles peuvent être générales ou dépendantes d'un domaine. Par exemple, voici quatre nouvelles requêtes qui pourraient résulter de la transformation de \mathcal{Q} :

$$\mathcal{Q}_1 = \left\{ \begin{array}{l} \text{Donner les lettres envoyées entre 1890 et} \\ \text{1895 par Henri Poincaré à Paul Appell et} \\ \text{qui mentionnent des travaux en mécanique.} \end{array} \right.$$

17. Les requêtes utilisées dans ce document sont présentées de façon informelle dans un souci de lisibilité mais elles correspondent toutes à une requête SPARQL.

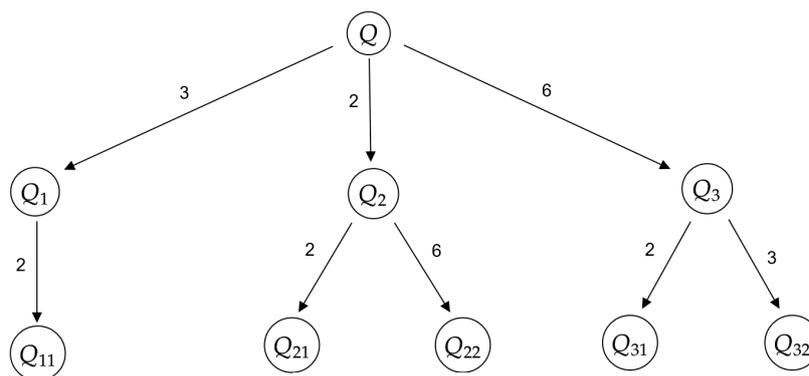


Figure 4 : Exemple d'un arbre de recherche tronqué à la profondeur 2.

$$\mathcal{Q}_2 = \left\{ \begin{array}{l} \text{Donner les lettres envoyées entre 1890 et} \\ \text{1895 par un mathématicien à Henri Poincaré et} \\ \text{qui mentionnent des travaux en mécanique.} \end{array} \right.$$

$$\mathcal{Q}_3 = \left\{ \begin{array}{l} \text{Donner les lettres envoyées entre 1890 et} \\ \text{1895 par } \underline{\text{Émile Picard}} \text{ à Henri Poincaré et} \\ \text{qui mentionnent des travaux en mécanique.} \end{array} \right.$$

$$\mathcal{Q}_4 = \left\{ \begin{array}{l} \text{Donner les lettres envoyées } \underline{\text{après 1895}} \\ \text{par Paul Appell à Henri Poincaré et} \\ \text{qui mentionnent des travaux en mécanique.} \end{array} \right.$$

\mathcal{Q}_1 est générée en appliquant une règle d'échange de l'expéditeur et du destinataire de la lettre. \mathcal{Q}_2 est générée en appliquant une règle visant à remplacer une instance de classe par n'importe quel membre de cette classe. Dans notre cas, la ressource décrivant Paul Appell fait partie de la classe des mathématiciens. \mathcal{Q}_3 correspond à l'application d'une règle visant à remplacer l'un des correspondants par une personne avec laquelle il a collaboré. \mathcal{Q}_4 correspond à l'application d'une règle visant à modifier les bornes temporelles liées à la date d'écriture de la lettre. D'autres générations de requêtes peuvent être imaginées et sont dépendantes des règles existantes et des données de la base RDF. Il est possible de combiner plusieurs transformations de requêtes tant qu'un coût maximal n'a pas été atteint. En effet, dans le fonctionnement actuel de l'outil, un coût (défini comme un nombre positif) est associé à chaque règle de transformation de requêtes. Ainsi, l'application successive de règles de transformation correspond à l'exploration d'un arbre de recherche comme l'illustre la figure 4.

L'objectif de ce mécanisme de transformation est double. Tout d'abord, il permet de présenter des résultats sémantiquement proches de ceux correspondant aux critères de recherche et qui peuvent apporter une réponse au problème de recherche pour lequel la requête initiale avait été formulée. Mais ce mécanisme peut aussi faire émerger de nouvelles connaissances pour le domaine. En effet, dans le contexte de la correspondance d'Henri Poincaré, il peut permettre de

dégager de nouveaux liens ou d'affiner des liens existants entre des personnes, des institutions, etc.

Ce système d'interrogation flexible a fait l'objet de plusieurs publications, qui précisent notamment sa formalisation informatique en expliquant la syntaxe et l'algorithme permettant l'application des règles de transformation [20, 22, 23]. Nous proposons de nous intéresser à plusieurs cas d'utilisation, en présentant des exemples issus du corpus de la correspondance d'Henri Poincaré.

5.2. Quelles règles de transformation pour quels usages ?

Certaines règles génériques, qui exploiteraient les relations entre les éléments de l'ontologie, peuvent présenter un intérêt pour proposer des alternatives lors des recherches, notamment par la généralisation de la requête initialement formulée. Le système SQTRL permet également de définir des règles dépendantes du domaine d'application, et peut ainsi encourager l'exploitation de connaissances spécifiques. Bien que les règles de transformation n'aient pas à être rédigées par les experts du domaine, il est important qu'ils participent activement à la réflexion sur leur formalisation et leurs possibles applications pour la recherche, ici, dans un cadre historique¹⁸. Afin d'exploiter les différentes données relatives aux acteurs, plusieurs règles de transformation ont été proposées. L'exploitation de ces informations permet d'offrir de nombreuses directions de recherche à un utilisateur explorant le corpus. Il est intéressant de noter qu'au-delà de l'étude des travaux de Poincaré, le corpus peut éclairer les travaux scientifiques de certains acteurs ayant interagi avec Poincaré. Dans certains cas, les correspondances avec des acteurs sont conséquentes et fournissent une base de travail intéressante. Dans d'autres cas, cet ensemble de lettres peut n'avoir été que partiellement constitué, ou être inexistant, ou bien ne pas contenir les informations pouvant apporter des éléments de réponse aux problématiques de recherche. Cependant, il peut exister un nombre important de lettres mentionnant ce correspondant, et au sein desquelles des éléments historiques et scientifiques pertinents peuvent éclairer le questionnement initial. Dans ce contexte, une règle est proposée pour remplacer un correspondant en tant que personne citée dans le corps de la lettre dans une requête SPARQL. Il est également intéressant de proposer la fonctionnalité inverse qui remplace une personne citée pour qu'elle soit un correspondant des lettres recherchées. Il existe également des versions alternatives de ces règles qui s'appuient sur la propriété utilisée pour indiquer les personnes citées dans l'apparat critique.

Imaginons un chercheur s'intéressant aux travaux du mathématicien allemand Karl Weierstrass (1815-1897). Ce dernier, lauréat de la médaille *Copley* de la *Royal Society* de Londres en 1897, a proposé des travaux de premier plan en analyse qu'il a appliqués pour des contributions significatives dans le champ du calcul des variations [24]. Au cours de sa carrière, il a eu pour disciples de nombreux mathématiciens avec lesquels Henri Poincaré a échangé au cours de sa vie parmi lesquels nous pouvons citer Georg Cantor, Lazarus Fuchs et Sofia Kovalevskaja. Au sein du corpus, il existe uniquement 4 lettres avec Karl Weierstrass défini en tant que correspondant. Cependant, nous remarquons qu'il existe des lettres où ce mathématicien est

18. En effet, l'idée du système SQTRL est notamment d'encourager la collaboration entre des chercheurs spécialistes d'un domaine, et des informaticiens qui seraient à l'aise avec la syntaxe des règles et le cadre technologique du Web sémantique.

explicitement cité dans le corps (49 lettres) ou dans l'apparat critique (64 lettres), principalement dans des échanges avec Gösta Mittag-Leffler. Dans ce genre de situations, une application de règle pourrait transformer la requête informelle Q en Q' .

$$Q = \left| \begin{array}{l} \text{Donner les lettres échangés entre} \\ \text{Henri Poincaré et Karl Weierstrass} \\ \text{qui traitent d'analyse.} \end{array} \right. \overset{\text{subsCorr}}{\rightsquigarrow} Q' = \left| \begin{array}{l} \text{Donner les lettres dont Henri} \\ \text{Poincaré est correspondant citant} \\ \text{Karl Weierstrass et qui traitent d'analyse.} \end{array} \right.$$

Dans une autre démarche, plusieurs règles ont été définies, suivant le même motif, afin d'exploiter les métadonnées décrivant les acteurs. Si un acteur est explicitement mentionné dans le corps d'une requête SPARQL, il est possible de généraliser cette requête en s'appuyant sur une de ses caractéristiques. Par exemple, la requête suivante recherche les lettres échangées avec Giovanni Battista Guccia :

$$Q = \left| \text{Donner les lettres échangées avec Giovanni Battista Guccia.} \right.$$

À partir des informations le concernant, le système peut proposer plusieurs transformations dont celles générant les quatre requêtes suivantes :

$$Q_1 = \left| \begin{array}{l} \text{Donner les lettres échangées avec} \\ \text{un mathématicien.} \end{array} \right. \quad Q_2 = \left| \begin{array}{l} \text{Donner les lettres échangées avec} \\ \text{un italien.} \end{array} \right.$$

$$Q_3 = \left| \begin{array}{l} \text{Donner les lettres échangées avec un} \\ \text{membre du cercle mathématique de Palerme.} \end{array} \right. \quad Q_4 = \left| \begin{array}{l} \text{Donner les lettres échangées avec un} \\ \text{éditeur scientifique.} \end{array} \right.$$

Des motifs de co-occurrences pourraient appuyer des transformations de requêtes lors de recherches sur le corpus. Par exemple, une possibilité serait de s'intéresser aux liens entre les correspondants et les thèmes associés aux lettres. Lors d'une recherche impliquant un correspondant, il peut être pertinent de proposer de rechercher des lettres avec un autre correspondant, qui a échangé avec Poincaré sur un même thème.

Une autre forme de transformation pourrait s'appuyer sur des liens entre une période temporelle et un thème. Partant d'un thème indiqué comme critère d'une recherche sur la correspondance, il pourrait être intéressant de proposer une nouvelle requête avec des critères de bornes temporelles construits à partir de la période temporelle où apparaît ce thème. Par exemple, la requête suivante s'intéresse au thème « fonction fuchsienne ».

$$Q = \left| \text{Donner les lettres traitant de fonctions fuchiennes.} \right.$$

Au sein de la correspondance, 12 lettres, rédigées entre avril 1881 et décembre 1882, traitent des fonctions fuchiennes. En appliquant la transformation évoquée plus haut, la requête deviendrait :

$$Q' = \left| \text{Donner les lettres rédigées entre le 11 avril 1881 et le 05 décembre 1882.} \right.$$

6. Des systèmes pour l'exploration du corps

Dans la première partie de ce chapitre, nous avons évoqué plusieurs règles de transformations et leurs possibles applications pour aider à mener des recherches historiques sur le corpus de la

correspondance d'Henri Poincaré. Nous proposons de décrire deux interfaces de recherches qui s'appuient sur l'utilisation de ce mécanisme. La première est un système de recherche s'appuyant sur des formulaires. Le deuxième est un outil de navigation qui exploite les similarités entre des ressources d'un graphe RDF. D'autres outils, qui ne s'appuient pas sur le mécanisme d'interrogation flexible seront également brièvement décrits.

6.1. Interface fondée sur des formulaires

Une première application du mécanisme d'interrogation flexible se retrouve dans un système de recherche simple, qui permet de générer des requêtes SPARQL via une interface fondée sur des formulaires. Le formulaire affiché s'adapte au type de ressources recherchées (articles, documents, lettres, personnes), plusieurs champs sont proposés à la saisie. Pour plusieurs d'entre eux, un mécanisme d'autocomplétion est utilisé, qui s'appuie sur l'utilisation des étiquettes (*label*) associés aux identifiants de la base. Par exemple, dans la base du corpus, Henri Poincaré est associé à l'identifiant `henripoincare.fr/api/items/843` mais l'interface affiche la chaîne "Poincaré, Henri (1854-1912)". À tout moment, l'utilisateur peut visualiser dans un champ dédié la requête SPARQL qui est générée. Une fois une requête exécutée, les résultats correspondants sont affichés dans un tableau et peuvent être exportés au format CSV. Le système propose également des requêtes alternatives s'appuyant sur l'application de règles de transformation. Sélectionner une transformation va permettre de générer une nouvelle. À tout moment, l'utilisateur peut visualiser la requête initialement formulée ainsi que la requête générée après application d'une règle de transformation. Les différents résultats sont agrégés dans le même tableau, en distinguant les résultats additionnels des résultats initiaux par l'utilisation de couleurs. Cet outil permet à un utilisateur de formuler des requêtes SPARQL sans aucune connaissance technique tout en permettant l'utilisation du système d'interrogation flexible.

6.2. Outil de navigation

Une autre forme d'exploration du corpus est relative à la notion de navigation qui part du constat qu'il est fréquent de débiter par l'étude d'une ressource du corpus avant de s'intéresser à d'autres ressources liées ou similaires. Dans ce contexte, un outil de navigation a été développé pour ce corpus. Le principe est de partir d'une lettre du corpus et de retrouver des ressources présentant des similarités. Pour cela, après la sélection de la ressource initiale, le système propose automatiquement plusieurs conditions de recherche que l'utilisateur peut sélectionner comme critère de recherche (positivement ou négativement). Pour l'exemple présenté en figure 5, cela conduit à la génération de sept conditions relatives à l'expéditeur, au destinataire, aux thèmes et aux personnes citées. À côté de chacune d'entre elles, il peut également retrouver un entier indiquant le nombre de ressources correspondant à ce critère dans le graphe. Après sélection des conditions qui l'intéressent, il a la possibilité de générer et d'exécuter une requête SPARQL. Les résultats apparaissent alors dans la partie droite de l'interface, avec des détails pour chaque lettre retrouvée. Il est possible de filtrer les résultats selon la date de rédaction des lettres, de manière à s'intéresser à une période temporelle spécifique. Des fonctionnalités annexes comme l'affichage d'un graphique correspondant à la distribution temporelle des résultats et l'export des données au format CSV sont également proposés. À partir des résultats, l'utilisateur peut

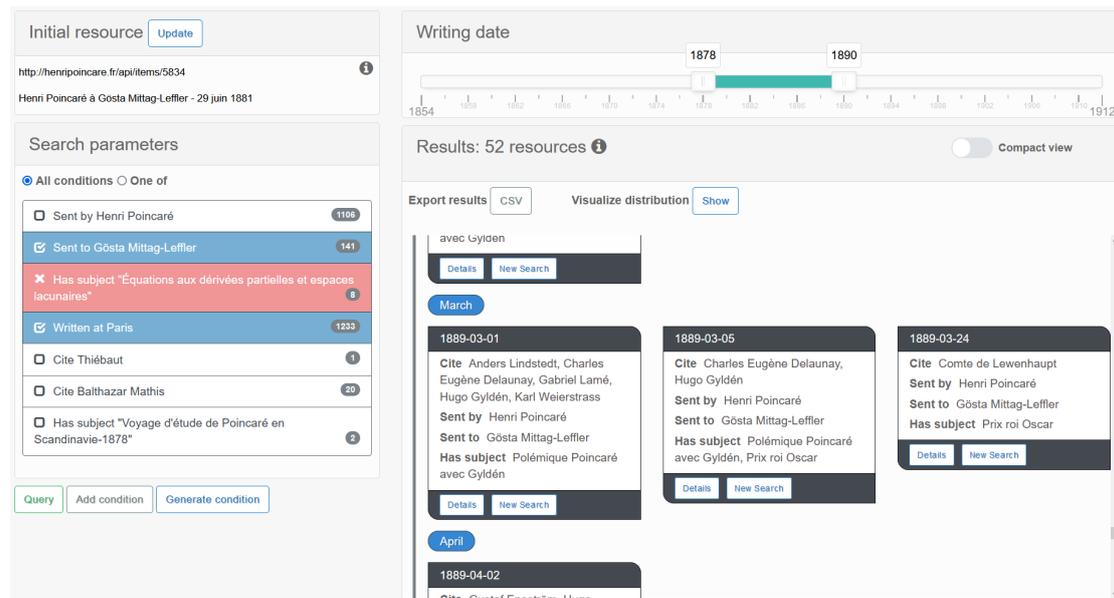


Figure 5 : L’outil de navigation en utilisation pour explorer le corpus de la correspondance d’Henri Poincaré.

sélectionner une nouvelle ressource qui sera le point de départ d’une nouvelle recherche. Les conditions de recherche seront alors mises à jour. Il est ainsi possible de naviguer de ressource en ressource.

Pour aller plus loin, le système propose la génération automatique de conditions additionnelles en appliquant des règles de transformation de requêtes. En effet, dans certaines situations, les conditions générées peuvent ne pas être suffisantes pour fournir des résultats intéressants ou surprenants.

Ce système peut être intéressant pour explorer les données d’une façon originale. Dans certaines situations, la plupart des ressources liées pourraient être déjà connues des historiens, mais dans d’autres cas, l’outil pourrait mettre en évidence des liens inattendus entre les ressources et être le point de départ de nouvelles considérations historiques. À titre d’exemple, ces nouveaux liens peuvent révéler certains motifs entre des thèmes scientifiques et des individus (en tant que correspondants ou personnes citées), en particulier pour les acteurs considérés comme mineurs et pour lesquels aucun travail de recherche n’a été initié.

Une vidéo de démonstration de cet outil est disponible en ligne¹⁹ et le code source est disponible sur un dépôt GitHub public²⁰. Cet outil a été pensé dès le départ, comme un système réutilisable et est donc accompagné d’un fichier de configuration qui permet de se connecter à d’autres points d’accès SPARQL. Les différents éléments présentés dans l’interface sont tous paramétrables. Une documentation relative à son installation et sa configuration, et la présentation d’essais avec des données de DBpedia sont notamment fournies sur le dépôt GitHub.

19. <https://videos.ahp-numerique.fr/videos/watch/f90ff003-39db-4b4c-ade6-fb18b86d9244>.

20. https://github.com/nlasolle/rdf_navigation_tool

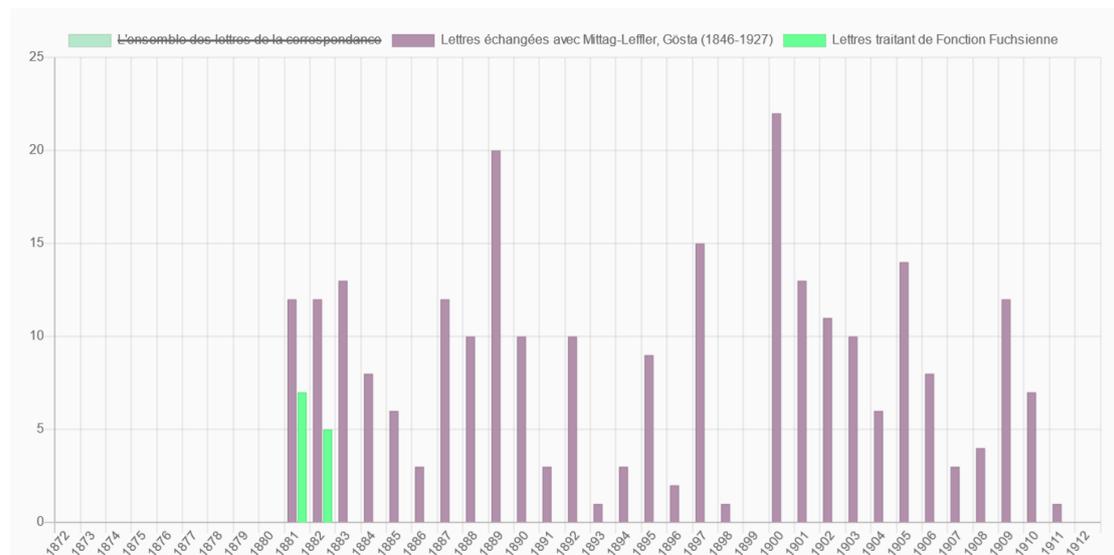


Figure 6 : Exemple d’affichages de distributions relatives à la rédaction de lettres respectant certains critères.

6.3. Des outils s’appuyant sur des statistiques et des données spatio-temporelles

D’autres outils de recherche ont été mis en place et sont à la fois à destination du grand public souhaitant découvrir le corpus et à destination des historiens souhaitant mener des études quantitatives. Le code source de ces différents outils est accessible sur un dépôt GitHub public²¹ et une vidéo de démonstration est disponible en ligne²². Contrairement à l’outil de navigation, il n’y a pas de système de configuration qui permet de se connecter à un autre corpus mais il est cependant possible de s’inspirer des outils proposés en réutilisant certaines parties de l’application.

Le premier outil permet à un utilisateur de sélectionner un correspondant ou un thème, et de visualiser la distribution des lettres associées. Pour cela, l’outil s’appuie sur l’année de rédaction des lettres. Il permet notamment de superposer différentes distributions, ce qui peut être utile pour rapidement mettre en évidence des motifs ou pour valider certaines hypothèses. Le graphique peut être exporté en tant qu’image et il est également possible d’exporter les données correspondantes au format CSV. Par exemple, la figure 6 superpose la distribution des lettres échangées entre Henri Poincaré et Thomas Craig et la distribution des lettres ayant pour thème la géométrie.

Le deuxième outil correspond à un simple tableau dynamique, qui regroupe des statistiques relatives aux personnes du corpus. Pour chacune d’entre elles, le nombre de lettres échangées avec Poincaré, le nombre de lettres la citant, et le nombre de lettres dont l’apparat la cite est indiqué. En cliquant sur une ligne du tableau, il est possible d’accéder à la page du site qui donne

21. <https://github.com/nlasolle/ahpo-data-exploration-tools>

22. <https://videos.ahp-numerique.fr/w/gjj2DJ9mZmVnKehwuDgWFk>

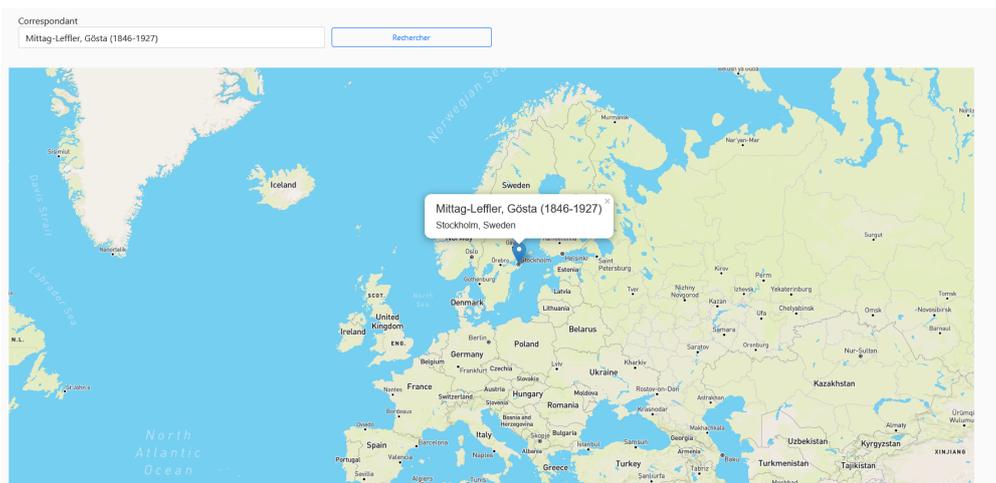


Figure 7 : Outil de visualisation des lieux de naissance associés aux personnes du corpus.

des détails sur la personne²³. Ce tableau s'appuie sur l'interrogation du point d'accès SPARQL qui permet une mise à jour continue des données présentées.

Enfin, le dernier outil exploite les données géographiques relatives aux personnes du corpus. Le premier correspond à une interface où un utilisateur peut visualiser des lieux de naissance, sous la forme de marqueurs par rapport à une personne choisie (voir figure 7, p. 19). Par exemple, en sélectionnant Felix Klein, la carte met en évidence son lieu de naissance (Düsseldorf, Allemagne) et les villes d'exercice connues (plusieurs villes allemandes). Une autre carte propose d'afficher les lieux de naissance connus de l'ensemble des correspondants avec lesquels Poincaré a échangé.

7. Conclusion

Le corpus de la correspondance d'Henri Poincaré est un corpus riche, contenant de nombreuses informations pour des travaux relevant de l'histoire des sciences et de l'histoire sociale. L'usage d'outils numériques peut appuyer des recherches historiques s'y intéressant et encourager la découverte de liens. Un constat porté par l'équipe est que l'édition numérique du corpus est un travail conséquent qui nécessite une phase de réflexion préliminaire visant à définir les données d'intérêt, la structure de la base, les choix ontologiques. Il est nécessaire de déterminer les informations qui devront apparaître dans la base, quelle sémantique associer à chaque métadonnée, comment structurer la base, quelles ontologies réutiliser et sous quelles formes. Il y a ensuite le travail d'alimentation de la base qui consiste à éditer les faits sous la forme de triplets. Cela nécessite de s'accorder quant aux sources utilisées pour décrire les éléments du corpus et de déterminer jusqu'où aller dans les descriptions. Il peut être tentant d'être le plus exhaustif possible, mais cela peut parfois conduire à des approximations, ou à intégrer des données qui

23. Par exemple, la page <http://henripoincare.fr/s/correspondance/item/346> offre une synthèse des informations relatives au mathématicien allemand Felix Klein.

ne présentent qu'un intérêt mineur pour la recherche historique tout en nécessitant un travail conséquent pour retrouver les informations. Les contraintes des projets de recherche font que des tâches relatives à l'édition sont parfois limitées dans le temps, et il faut donc effectuer des choix en conséquence pour garantir un corpus numérique cohérent bien qu'incomplet.

Le mécanisme d'interrogation flexible peut présenter plusieurs intérêts dans le cadre de l'étude. Il peut tout d'abord permettre d'accéder plus rapidement à des informations utiles. Pour les personnes ne connaissant pas le corpus, cela peut être l'occasion d'être guidé dans leurs recherches pour comprendre son organisation. Ce système peut également s'inscrire dans une méthode heuristique en permettant de mettre en avant des liens qui, bien souvent pourraient être déjà connus mais dans certains cas être inattendus. Dans le cadre de l'étude d'un corpus historique tel que celui de la correspondance d'Henri Poincaré, il peut notamment mettre en avant des liens avec des acteurs parfois qualifiés de secondaires mais qui peuvent présenter un intérêt lors de l'étude de la vie et de l'œuvre de Poincaré, ou lors de l'étude d'institutions savantes de la fin du XIX^e siècle et du début du XX^e siècle.

Remerciements

Ce travail a bénéficié d'une aide de l'État, gérée par l'Agence Nationale de la Recherche, au titre du projet Investissements d'Avenir Lorraine Université d'Excellence, portant la référence ANR-15-IDEX-04-LUE.

Un grand merci à toutes les personnes ayant participé, de près ou de loin, à ce projet d'édition et d'exploitation numérique de la correspondance d'Henri Poincaré.

Références

- [1] P. Nabonnand (Ed.), *La correspondance entre Henri Poincaré et Gösta Mittag-Leffler*, Birkhäuser, 1998.
- [2] S. Walter, É. Bolmont, A. Coret (Eds.), *La correspondance entre Henri Poincaré et les physiciens, chimistes et ingénieurs*, Birkhäuser, 2007. doi :10.1007/978-3-7643-8303-9.
- [3] S. Walter, P. Nabonnand, R. Krömer, M. Schiavon (Eds.), *La correspondance entre Henri Poincaré, les astronomes, et les géodésiens*, Birkhäuser, 2016. doi :10.1007/978-3-7643-8293-3.
- [4] L. Rollet (Ed.), *La correspondance de jeunesse d'Henri Poincaré : les années de formation. De l'École polytechnique à l'École des Mines (1873-1878)*, Publications of the Henri Poincaré Archives, Birkhäuser, 2017. doi :10.1007/978-3-319-55959-9.
- [5] L. Rollet, P. Nabonnand, *Éditer la correspondance d'Henri Poincaré*, in : F. Henryot (Ed.), *L'historien face au manuscrit. Du parchemin à la bibliothèque numérique*, UCL-Presses Universitaires de Louvain, 2012, pp. 285–304. URL : <http://books.openedition.org/pucl/1282>.
- [6] J. Peiffer, *Faire des mathématiques par lettres*, *Revue d'histoire des mathématiques* 4 (1998) 143–157.
- [7] A. Gilroy, W. M. Verhoeven, *Epistolary histories : letters, fiction, culture*, University of Virginia Press, 2000.
- [8] C. D. Conner, *Histoire populaire des sciences*, Editions l'Echappée, Paris, 2011.

- [9] H. Zinn, *Une histoire populaire des États-Unis. De 1492 à nos jours*, Agone, Paris, 2002.
- [10] G. Noiriel, *Une histoire populaire de la France : De la guerre de Cent Ans à nos jours*, Agone, Paris, 2018.
- [11] A. Boutroux, *Vingt ans de ma vie, simple vérité : la jeunesse d'Henri Poincaré racontée par sa sœur (1854-1878)*, Histoire des sciences, Hermann, 2012.
- [12] L. Rollet, *Jeanne Louise Poulain d'Andecy, épouse Poincaré (1857-1934)*, Bulletin de la Société des amis de la Bibliothèque et de l'Histoire de l'École polytechnique (2012) 18–27.
- [13] C. Boulaire, R. Carabelli, *Du digital naïve au bricoleur numérique : les images et le logiciel Omeka*, in : E. Cavalié, F. Clavert, O. Legendre, D. Martin (Eds.), *Expérimenter les humanités numériques. Des outils individuels aux projets collectifs*, Les Presses de l'Université de Montréal, Montréal, Québec, 2017, pp. 81–103. URL : <http://www.parcoursnumeriques-pum.ca/du-digital-naive-au-bricoleur-numerique>.
- [14] N. Larrousse, J. Marchand, *A Techno-Human Mesh for Humanities in France : Dealing with preservation complexity*, in : Pierazzo, Elena and Ciotti, Fabio (Ed.), *DH 2019*, Utrecht, Netherlands, 2019. URL : <https://hal.archives-ouvertes.fr/hal-02153016>.
- [15] F. Manola, E. Miller, B. McBride, et al., *RDF 1.1 Primer*, 2014. URL : <https://www.w3.org/TR/rdf-primer>, dernière consultation : juillet 2021.
- [16] N. Lasolle, P. Willaime, *Sémantiser Omeka S : pourquoi et comment?*, in : *Omeka - Projets scientifiques, culturels et/ou documentaires*, Nancy, France, 2020. URL : <https://hal.univ-lorraine.fr/hal-02973316>.
- [17] S. Harris, A. Seaborne, E. Prud'hommeaux, *SPARQL 1.1 Query Language, recommandation du W3C*, 2013. URL : <https://www.w3.org/TR/sparql11-query/>, dernière consultation : juillet 2021.
- [18] P. Appell, *Henri Poincaré, Nobles vies - Grandes œuvres*, Plon, 1925.
- [19] T. Grainger, T. Potter, *Solr in action*, Manning Publications Co., 2014.
- [20] O. Bruneau, E. Gaillard, N. Lasolle, J. Lieber, E. Nauer, J. Reynaud, *A SPARQL Query Transformation Rule Language – Application to Retrieval and Adaptation in Case-Based Reasoning*, in : D. Aha, J. Lieber (Eds.), *Case-Based Reasoning Research and Development. ICCBR 2017, Lecture Notes in Computer Science*, Springer, Cham, 2017, pp. 76–91.
- [21] P. Appell, *Traité de mécanique rationnelle, volume 1*, Gauthier-Villars, 1893.
- [22] O. Bruneau, N. Lasolle, J. Lieber, E. Nauer, S. Pavlova, L. Rollet, *Applying and Developing Semantic Web Technologies for Exploiting a Corpus in History of Science : the Case Study of the Henri Poincaré Correspondence*, *Semantic Web – Interoperability, Usability, Applicability (2021)* 359–378. doi :10.3233/SW-200400.
- [23] N. Lasolle, O. Bruneau, J. Lieber, E. Nauer, S. Pavlova, *Assisting the RDF Annotation of a Digital Humanities Corpus Using Case-Based Reasoning*, in : J. Z. Pan, V. Tamma, C. d'Amato, K. Janowicz, B. Fu, A. Polleres, O. Seneviratne, L. Kagal (Eds.), *The Semantic Web - ISWC 2020*, Springer, Cham, 2020, pp. 617–633.
- [24] P. Dugac, *Eléments d'analyse de Karl Weierstrass*, *Archive for History of Exact Sciences* 10 (1973) 41–176. URL : <http://www.jstor.org/stable/41133363>.