



HAL
open science

Quand l'informatique soulève des questions épistémologiques dans le domaine de la littérature de langue bretonne : l'exemple de la base de données PRELIB

Nelly Blanchard, Jean-Baptiste Pressac, Mannaïg Thomas

► **To cite this version:**

Nelly Blanchard, Jean-Baptiste Pressac, Mannaïg Thomas. Quand l'informatique soulève des questions épistémologiques dans le domaine de la littérature de langue bretonne : l'exemple de la base de données PRELIB. *La Bretagne Linguistique*, 2017, 21, pp.297-317. hal-01547673

HAL Id: hal-01547673

<https://hal.univ-brest.fr/hal-01547673v1>

Submitted on 1 Sep 2020

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.



Quand l'informatique soulève des questions épistémologiques dans le domaine de la littérature de langue bretonne : l'exemple de la base de données PRELIB

When computer science raises epistemological questions in the domain of Breton literature: the case of the PRELIB database

Nelly Blanchard, Jean-Baptiste Pressac et Mannaig Thomas



Édition électronique

URL : <http://journals.openedition.org/lbl/315>

Éditeur

Université de Bretagne Occidentale – UBO

Édition imprimée

Date de publication : 1 mai 2017

Pagination : 297-317

ISBN : 979-10-92331-31-8

ISSN : 1270-2412

Référence électronique

Nelly Blanchard, Jean-Baptiste Pressac et Mannaig Thomas, « Quand l'informatique soulève des questions épistémologiques dans le domaine de la littérature de langue bretonne : l'exemple de la base de données PRELIB », *La Bretagne Linguistique* [En ligne], 21 | 2017, mis en ligne le 01 mai 2020, consulté le 25 juin 2020. URL : <http://journals.openedition.org/lbl/315>



La Bretagne Linguistique est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International.

Nelly BLANCHARD*, Jean-Baptiste PRESSAC**,
Mannaig THOMAS***

Quand l'informatique soulève des questions épistémologiques dans le domaine de la littérature de langue bretonne : l'exemple de la base de données PRELIB

PRELIB¹ propose une nouvelle approche de la littérature de langue bretonne dans le cadre d'un projet de base de données en sciences humaines et sociales. Par une collaboration interdisciplinaire alliant littérature, sciences sociales et science informatique, il s'agit de mener l'étude des relations entre la littérature bretonne et d'autres littératures, notamment la littérature de langue française, et des relations internes au champ littéraire breton. Il s'agit notamment d'identifier les lieux de sociabilité et les réseaux dans lesquels s'inscrivent les acteurs et producteurs du champ – d'où la nécessité de la mise à disposition de données sur les trajectoires sociales et littéraires (institutionnelles ou non) de ces acteurs, ouvrant la voie à des analyses à la fois littéraires et socio-historiques.

* Maître de conférences de breton et celtique (HDR), UBO/CBRC (EA 4451)

** Ingénieur d'études CNRS, UBO/CBRC (UMS 3554)

*** Maître de conférences de breton et celtique, UBO/CBRC (EA 4451)

1. PRELIB : Projet de recherche en littérature de langue bretonne. La base de données fait partie d'un projet plus large, voir https://www.univ-brest.fr/crbc/menu/Projets_de_recherche/Le_point_de_vue_peripherique___Programme_de_recherche_en_litterature_de_langue_bretonne (site consulté le 22 septembre 2016).

Lors de l'élaboration d'un tel projet de compilation et d'analyse de données numériques, le travail ne coule pas toujours à débit régulier et les nombreux obstacles rencontrés donnent naissance à plusieurs types d'interrogations. La toute première que nous gardons toujours en tête, à partir de laquelle nous doutons et nous enthousiasmons, est celle de la juste place à trouver au curseur entre l'idéal d'une mise à plat sous forme de faits d'un ensemble exhaustif de données et la réalité de la disponibilité partielle des sources et de la difficulté, voire l'impertinence, de catégoriser les sources du domaine observé. Nous ne sommes ni aveugles ni découragés par le caractère infini du travail que cela représente, mais voulons tenter cette innovation pour ouvrir de nouvelles voies de compréhension du domaine que nous étudions. Cette question « existentielle » mise à part, d'autres types de questions s'enchevêtrent : elles sont d'ordre technique, d'ordre définitionnel superficiel et d'ordre épistémologique.

Nous n'évoquerons pas ici les questions superficielles que nous tâchons de résoudre au fur et à mesure des saisies et des petits obstacles rencontrés. Nous ne mentionnerons que très brièvement quelques réflexions techniques avant d'aborder ce qui est au cœur du présent article : les questions épistémologiques ou questions de fond concernant des définitions à la base même de la connaissance en littérature et sociologie de la littérature. Autrement dit, nous allons proposer de retracer quelques chemins tortueux que l'analyse de données à grande échelle, confrontées à la réalité des productions littéraires en langue bretonne, fait emprunter à des questions théoriques sur ce qu'est une œuvre et ce que sont un auteur et une relation entre individus.

Une base de données ouverte²

Dès le début du projet en 2012, la base de données PRELIB a été conçue dans la perspective d'une mise à disposition libre et gratuite des données depuis internet. Cette mise à disposition ne concerne pas uniquement la création de pages internet sur laquelle il est possible de consulter les données concernant les individus, les collectifs, les œuvres littéraires, etc. Elle concerne également la possibilité d'interroger la

2. La base de données est accessible à l'adresse suivante : <http://mshb.huma-num.fr/prelib/> (site consulté le 22 septembre 2016).

base afin de récupérer des jeux de données en vue d'une analyse à l'aide de logiciels de statistiques ou de visualisation de données.

Aussi, la base de données PRELIB repose sur MySQL³, un logiciel open-source de gestion de base de données relationnelle couramment utilisé pour la conception de sites internet. Le terme « relationnel » ne vient pas du fait que les données saisies concernent des relations entre des individus, des œuvres, des collectifs : les informations saisies dans les bases de données relationnelles s'apparentent à des tables où les enregistrements apparaissent en ligne et leurs attributs en colonne. Le terme relationnel vient du mot « relation » qui sert à désigner ces tables dans la théorie des bases de données relationnelles introduite par l'informaticien Edgar F. Codd en 1970⁴.

La base de données ainsi que le site internet permettant la saisie et la consultation des données sont hébergés depuis 2012 sur un serveur web mis à disposition par la TGIR Huma-Num⁵, une unité mixte de service, dépendant du CNRS, dont la mission est de faciliter le tournant numérique de la recherche en sciences humaines et sociales. La saisie des données peut donc être effectuée de manière centralisée avec un simple navigateur web. Le site de consultation est régulièrement mis à jour en fonction des évolutions de la structure de la base de données (fig. 1).

Les données « brutes » mises à disposition sont alignées avec d'autres référentiels de données, tel que IdRef de l'Abes⁶ ou le projet Wikidata⁷. L'alignement consiste, par exemple, à saisir dans la base de données que la personne « Pierre-Jakez Hélias » portant l'identifiant 757 dans PRELIB correspond à la notice d'autorité du Sudoc n° 026918064⁸ ou que la ville de Pouldreuzic dans laquelle est né Pierre-Jakez Hélias

3. Voir <https://fr.wikipedia.org/wiki/MySQL> (site consulté le 10 novembre 2016).

4. Voir https://fr.wikipedia.org/wiki/Modèle_relationnel (site consulté le 10 novembre 2016).

5. TGIR : Très grande infrastructure de recherche, voir <http://www.huma-num.fr> (consulté le 22 septembre 2016).

6. Voir <https://www.idref.fr/autorites/autorites.html> (consulté le 22 septembre 2016).

7. Voir https://www.wikidata.org/wiki/Wikidata:Main_Page (consulté le 22 septembre 2016).

8. Voir <http://www.idref.fr/026918064> (consulté le 22 septembre 2016).

correspond à l'élément Wikidata Q496294⁹. Cet alignement permet à la fois de lever des ambiguïtés et de récupérer des données non présentes dans PRELIB. C'est-à-dire d'attester que PRELIB désigne bien le même « Pierre-Jakez Hélias » que celui identifié par la notice d'autorité 026918064 du Sudoc ou que « Pouldreuzic » désigne bien la commune du Finistère dont le code commune INSEE est le 29225. Wikidata mettant à disposition les coordonnées spatiales des villes, il est alors possible d'afficher, par exemple, des données géolocalisées sur un fond de carte.

Comme cette base de données porte sur la littérature, les tables principales sont celles qui concernent les œuvres et les auteurs comme le montre le schéma suivant :

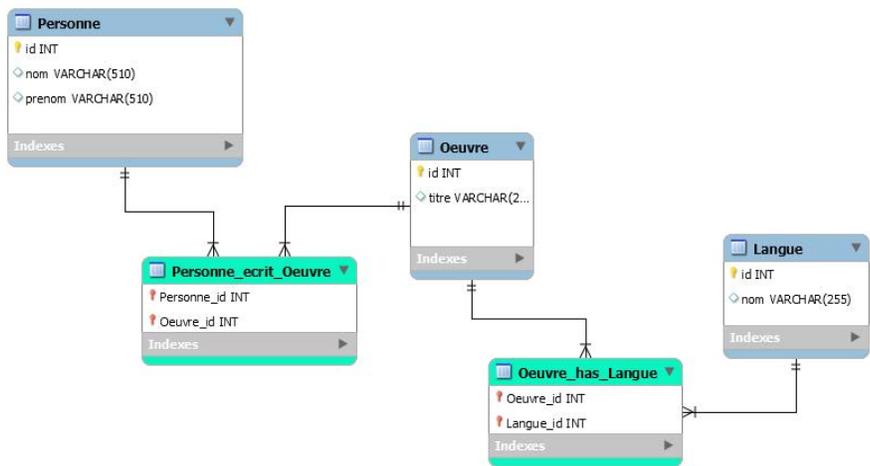


Fig. 2 : Schéma simplifié des tables « Personne » et « Œuvre » de la base de données PRELIB.

Qu'est-ce qu'une œuvre ?

Le premier montage de la base

Lorsque nous avons conçu la base PRELIB, les œuvres en langue bretonne faisaient bien évidemment partie des éléments fondamentaux, pertinents et indispensables à saisir : il s'agit, premièrement, de saisir le

9. Voir <https://www.wikidata.org/wiki/Q496294> (consulté le 22 septembre 2016).

titre des œuvres écrites en breton, les auteurs et éditeurs concernés, le lieu et la date d'édition ; et deuxièmement, d'ajouter des éléments qui nous ont paru également intéressants à renseigner :

- le type d'œuvres : nous sommes ici confrontés à l'élaboration d'une liste fixe qui serait ni trop vaste car non-informative, ni trop précise car elle ne permettrait pas l'analyse quantitative des données. Par exemple, faut-il une catégorie « Récit » ou différencier « Roman », « Nouvelle », « Conte »... ? Faut-il une catégorie « Chanson » ou faut-il différencier « Cantique », « *Gwerz* », « Cantate »... ? ;
- la langue d'écriture, puisque certaines œuvres peuvent être bilingues. Nous pensons, par exemple, à *Bepred Breizad*¹⁰ de François-Marie Luzel ou à l'anthologie de poésie bilingue *Défense de cracher par terre et de parler breton* de Yann-Ber Piriou¹¹ ;
- l'ajout potentiel d'un ou plusieurs contributeurs à l'œuvre, puisque certaines œuvres sont cosignées. Sur ce point, certains cas sont particuliers et demandent un temps de réflexion, par exemple, *E Bro Yvonne*¹², bio/autobiographie rédigée par Jean Roparz à partir des témoignages oraux d'Yvonne Riou : y a-t-il deux auteurs ou les deux personnes jouent-elles des rôles différents et comment les qualifier ?

Puis, comme c'est la notion de relation entre les personnes qui anime notre projet, d'autres éléments nous sont apparus rapidement avoir de l'importance et nous avons complété la table « Œuvre » par ces données. Il s'agit de signaler, s'ils existent :

- le lien d'inspiration avec d'autres œuvres, soit lorsque l'œuvre saisie est inspirée d'une autre œuvre, soit lorsqu'elle a inspiré une autre œuvre. À l'intérieur de cette question, se pose de manière plus précise ce qu'« inspiration » signifie : s'agit-il d'une traduction, d'une adaptation, d'une inspiration plus éloignée dont témoignerait ou non un exergue ou une dédicace, etc. ? La ques-

10. François-Marie LUZEL, *Bepred Breizad*, Morlaix, Haslé, 1965.

11. Yann-Ber PIRIOU, *Défense de cracher par terre et de parler breton. Poèmes de combat 1950-1970*, Paris-Honfleur, P.-J. Oswald, 1971.

12. Yvonne RIOU et Jean ROPARS, *E bro Yvonne : enñorioù ur beisantez a Vro Leon*, Brest, Emgleo Breiz, 1995.

tion de la langue d'écriture de l'œuvre a alors pris de l'importance car la base ne contient pas que des œuvres en langue bretonne : les *Géorgiques* de Virgile, *Den lille Havfrue* (*The little mermaid*, *La petite sirène*) d'Andersen ou *Asterix et la rentrée gauloise* de Goscinny en font par exemple partie comme source d'inspiration ou objet de traduction ;

- la présence d'une dédicace imprimée : une dédicace témoigne en effet d'une marque de respect et de reconnaissance (donc d'un certain type de relation entre deux individus) sans que cela implique qu'une œuvre ait été inspirée par une autre œuvre. Les dédicataires ne sont d'ailleurs pas toujours des auteurs : l'abbé Le Clerc dédicace l'une de ses œuvres entre autres à ses frères et sœurs et à ses parents ; François Jaffrennou à un camarade de régiment d'Infanterie ; le colonel Troude à Lefournier, son éditeur qui vient de mourir avant l'impression de l'ouvrage ; Erwan Berthou à Jean Le Fustec, grand druide de la Gorsedd ; Yves Le Moal à François Vallée ; Paul Keineg, Erwan Evenou et Yann-Ber Piriou à leurs « camarades de l'UDB¹³ », etc. Il y a donc plusieurs types de relations qui se cachent derrière ces expressions de reconnaissance.
- le lien d'« hébergement » d'une œuvre dans une revue (ou sa présence dans un manuscrit) et les liens directs ou indirects que cela peut supposer avec les dirigeants de la revue. Par exemple, le conte *Toull al lakez*¹⁴ de Lan Inisan a d'abord été publié dans *Feiz-ha-Breiz* avant d'être publié sous forme d'ouvrage, cela témoignant de liens entre Lan Inisan et la revue *Feiz-ha-Breiz*, avec les dirigeants de la revue à son époque, etc.

La partie et le tout

Ce dernier point sur l'hébergement d'une œuvre dans un support hébergeur nous a emmenés à soulever la question de l'unité minimale en matière de création littéraire et donc à cerner plus précisément ce

13. Paol KEINEG, « Le poème du pays qui a faim », *Hommes liges des talus en transe*, Paris-Honfleur, P.-J. Oswald, 1969 ; Yann-Ber PIRIOU, *Ar mallozhioù ruz*, Paris-Honfleur, P.-J. Oswald, 1974 ; Erwan EVENOU, *Benn goulou deiz*, s. l., s. n., 1972.

14. Lan INISAN, *Toull al lakez*, Brest, Lefournier, 1878.

qu'est une œuvre : la partie ou le tout ? Si *Toull al Lakez* est une œuvre publiée dans une revue, pourquoi un poème publié dans une revue ne serait-il pas également une œuvre, comme un poème de Jakez-Erwan Mouton dans *Al Liamm* ou de Yann-Erwan Plourin dans *Planedenn* ? Dans ce cas, qu'en est-il d'un poème publié dans un recueil ? Ou d'une nouvelle, d'un conte, d'une pièce de théâtre publiés dans un recueil ? Chaque poème (nouvelle, conte, pièce...) est-il une œuvre et le recueil ne relève-t-il que de l'édition ? Dans ce cas, il faut saisir le recueil en temps qu'édition et tous les poèmes qui le constituent en tant qu'œuvres. À cette question épistémologique sont liées des questions techniques : outre la nécessité de retrouver le recueil pour dresser la liste complète des poèmes (ces informations n'étant bien entendu pas toujours disponibles dans nos sources ni dans les catalogues des bibliothèques), se posent des problèmes de saisie : premièrement, le temps nécessaire à la saisie, certains recueils pouvant être constitués de centaines de poèmes, et sachant qu'il faut créer un lien entre l'auteur et chaque œuvre ; deuxièmement, comment procéder si ce recueil de cent poèmes est réédité cinq ans plus tard : faut-il renouveler la saisie ou peut-on copier-coller les éditions avec toutes les liaisons ?

De plus, si certains poèmes ont une « vie » autonome et sont publiés seuls dans un premier temps, puis dans un recueil par la suite ou une autre compilation, d'autres poèmes n'ont pas été pensés comme des entités indépendantes et n'ont de sens que dans l'économie d'un recueil pensé par un auteur, c'est-à-dire à partir d'un choix stylistique, thématique ou autre, puis ordonné par l'auteur pour en faire une œuvre porteuse de sens dans son ensemble. Ou encore : qu'en est-il d'une anthologie réalisée par un éditeur ? L'éditeur n'est l'auteur d'aucun poème, mais le concepteur d'un assemblage d'œuvres. C'est par exemple le cas de *Kontadennoù an amzer-vremañ* édité par Al Liamm en 1960 et composé d'œuvres d'auteurs différents comme *Fin ar bed* de Roparz Hemon, *Un danvez-den* d'Abeozen, etc.

Outre la question informatique, précisément la modélisation des données, qui est très contraignante, reste la question de la pertinence ou non d'apporter une réponse figée et automatique à cette question : l'une des approches empêche-t-elle l'autre et ne peut-on pas les combiner de manière systématique ? Ne peut-on pas ainsi considérer chaque poème comme une œuvre et le recueil également comme une œuvre qui aurait

un lien avec les œuvres qui le composent ? Cette vision formelle ne fait-elle pas fi de l'analyse littéraire elle-même ? Par exemple, que faire du cas du *Farvel göapaer*¹⁵ de Kerenveyer composé d'un opéra-comique, de contes et d'épigrammes ? Que faire du cas du *Barzaz-Breiz*¹⁶ de La Villemarqué ? Peut-on dire que le *Barzaz-Breiz* est une œuvre composée des x chants qui sont eux-mêmes des œuvres et qui le composent ? Ce serait le prendre pour un recueil de chants populaires, définition qui nous semble restrictive et faussée pour comprendre l'œuvre. Ce serait méconnaître la puissance de l'assemblage et de la monumentalisation des chants, ignorer l'ensemble des arguments, notes, introduction et épilogue qui soudent les éléments hétérogènes du volume, oublier les « redressements » opérés par l'auteur sur les chants¹⁷.

Ce raisonnement sur la partie et le tout doit-il donc se faire au cas par cas, mais ce choix ne serait alors pas sans conséquence sur le traitement statistique à partir de la base de données, ou bien doit-il être systématisé pour gagner en exhaustivité, au risque de provoquer des mésinterprétations des œuvres ? Nous nous trouvons là au cœur de l'articulation entre le traitement quantitatif et le traitement qualitatif de la littérature.

Œuvre – Édition vs Expression – Manifestation

Le problème de la partie et du tout se pose en sens inverse, mais débouche sur d'autres réponses, beaucoup plus simples, et ouvre sur d'autres interrogations, plus conceptuelles. Par exemple, lorsqu'une œuvre est parue en feuilletons dans une revue, comme *Itron Varia Garmez* paru en épisodes dans les numéros 5, 6 et 7 de *Galv*, ou lorsqu'une œuvre est éditée en deux tomes comme *C'hoariva*¹⁸ de Roparz Hemon, quelle est l'œuvre ? Cette question trouve facilement réponse : c'est la totalité du texte, de l'expression, qui forme l'œuvre et ce n'est pas la contrainte du format et du nombre de pages qui impose l'endroit de la fin d'une œuvre. Pour pousser le raisonnement plus loin, ce n'est

15. François-Nicolas Pascal DE KERENVEYER, *Ar farvel goapaer. Le bouffon moqueur*, édition et traduction par Ronan Calvez, Brest, CRBC, 2005.

16. Théodore HERSART DE LA VILLEMARQUÉ, *Barzaz-Breiz*, Paris, Didier, 1867 [1839].

17. Voir Nelly BLANCHARD, *Barzaz-Breiz. Une fiction pour s'inventer*, Rennes, PUR, coll. « Interférences », 2006.

18. Roparz HEMON, *C'hoariva*, Brest, Al Liamm, 1969.

pas parce qu'un texte est écrit sur deux cahiers manuscrits, et que la dernière ligne de la dernière page du premier cahier se termine par une entame de phrase qui se poursuit sur la première ligne du deuxième cahier qu'il faut séparer arbitrairement les deux cahiers comme deux œuvres. Il va de soi que cela constitue une seule œuvre.

Cette question de la partie et du tout, ainsi que d'autres questions importantes pour notre domaine, comme la graphie utilisée pour telle ou telle édition ou comme l'édition tardive de textes médiévaux ou plus récents (les mystères médiévaux, *Ar Farvel Gööpaër*¹⁹, *Histor eur famil eus Breiz Izel*²⁰, *Kastel Ker Iann Koatanskour*²¹, etc.), nous ont amenés à distinguer une œuvre d'une édition. Concernant la graphie utilisée pour le manuscrit ou les éditions imprimées, sachant que l'orthographe a évolué au fil du temps et que trois orthographes sont actuellement employées, l'exemple de l'œuvre *Ar Roc'h toull*²² de Jakez Kerrien illustre la différence entre une œuvre et une édition : *Ar Roc'h toull* a aussi été publiée en breton universitaire sous le titre *Ar Rob toull*²³, titre dans lequel on voit que la graphie est bien associée à l'édition. Il s'agit bien de la même œuvre même si le titre physique diffère, mais de deux éditions différentes. Il en est de même pour *Eur galedenn a zen*²⁴ de Yeun ar Gow qui a été édité par la revue *Gwalarn* en 1939, c'est-à-dire avant la création de l'orthographe *peurunvan*²⁵ deux ans plus tard mais réédité sous le titre *Ur galedenn a zen*²⁶ sous forme d'ouvrage en 1973. Au niveau de la modélisation des données, il nous faut d'ailleurs trouver le moyen de signaler qu'il s'agit de la même œuvre alors que les titres diffèrent. Au fil de nos réflexions, nous avons ainsi distingué une œuvre comme œuvre de la pensée, d'une édition comme matérialisation d'une œuvre. Et nous avons pour le moment conservé ces termes, bien que nous ne soyons pas au bout de nos réflexions sur ce sujet.

19. F.-N. Pascal DE KERENVEYER, *op. cit.*

20. Hervé BUREL, *Histor eur famil eus Breiz-Izel. Histoire d'une famille de Basse-Bretagne*, édition et traduction par Nelly Blanchard, Morlaix, Skol Vreizh, 2011.

21. Yves-Marie Gabriel LAOUËNAN, *Kastel Ker Iann Koatanskour. Le château de Kerjean-Coatanskour*, édition et traduction par Yves Le Berre, Brest, CRBC, 2004.

22. Jakez KERRIEN, *Ar Roc'h toull*, Brest, Gwalarn, 1926.

23. Jakez KERRIEN, *Ar Rob toull*, Brest, Emgleo Breiz, 1957.

24. Yeun AR GOW, « Eur galedenn a zen », *Gwalarn*, n° 126, 1939.

25. Il s'agit de la graphie surunifiée.

26. Yeun AR GOW, *Ur galedenn a zen*, Brest, Al Liamm, 1973.

En effet, pour aller plus loin dans la réflexion, nous nous sommes en partie appuyés sur les travaux de la Fédération internationale des associations de bibliothécaires et des bibliothèques (IFLA) qui a développé un modèle de catalogage des notices bibliographiques, réfléchi à des conventions internationales dans ce domaine et émis une modélisation conceptuelle des informations sur les objets fondamentaux pertinents dont relève le domaine²⁷. PRELIB ne vise pas complètement les mêmes objets, mais un grand nombre les recoupe et nous pensons qu’il est intéressant de s’inspirer du modèle élaboré par l’IFLA et appelé FRBR.

En ce qui concerne les œuvres, le FRBR les définit comme des produits d’une activité intellectuelle ou artistique et déploie ce concept en quatre entités distinctes comme le montre le schéma ci-dessous (fig. 3).

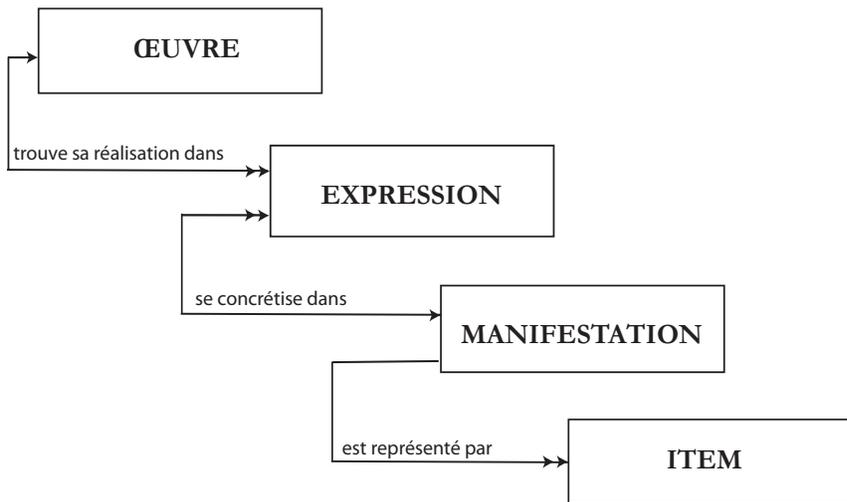


Fig. 3 : Entités du groupe 1 et relations fondamentales (FRBR).

Le contenu intellectuel ou artistique se trouve dans :

- l’œuvre : la création intellectuelle ou artistique, entité abstraite ;
- l’expression : la réalisation intellectuelle ou artistique de cette

27. Voir *Fonctionnalités requises des notices bibliographiques. Rapport final*, Paris, Bnf, 2012.

œuvre sous la forme d'une notation alphanumérique²⁸, musicale ou chorégraphique, ou sous une forme sonore, visuelle objectale cinétique, etc., ou bien sous l'aspect d'une combinaison de ces formes. La lecture à voix haute d'un texte ou sa traduction dans une autre langue constituent des expressions de la même œuvre.

Et la forme matérielle de ce contenu (le contenant en quelque sorte) se trouve dans :

- la manifestation : la matérialisation de l'une des expressions de l'œuvre comme l'ensemble d'exemplaires présentant les mêmes caractéristiques de fond et de forme, issu d'un même processus de matérialisation ou les manuscrits ;
- l'item : un exemplaire isolé d'une manifestation pouvant le plus souvent être un seul objet, parfois plusieurs (par exemple, l'exemplaire de *Ar Roc'h toull* publié par Gwalarn en 1926 et conservé à la bibliothèque du CRBC à la cote C-00121-00).

Pour PRELIB, les items ne sont pas des objets pertinents comme ils peuvent l'être pour une bibliothèque qui veut pouvoir repérer un ouvrage par une cote ou décrire un exemplaire détérioré ou ayant une reliure spécifique, etc. De même, l'œuvre comme création reste présente dans l'absolu comme point de départ de l'acte littéraire, aux contours flous, comme quelque chose de sous-jacent aux expressions, mais l'accès à cette création reste dans la très grande majorité des cas inaccessible. Le concept est surtout intéressant dans le lien qui peut exister entre une œuvre et une œuvre critique sur cette œuvre et qui ne porte pas sur telle ou telle expression mais sur la démarche de création, ou encore sur les liens indirects entre des expressions parentes comme des traductions en différentes langues d'une même œuvre. Ce sont donc surtout les « expressions » et les « manifestations » qui sont les concepts dont nous avons le plus besoin. Ce que nous avons appelé spontanément « œuvre » et « édition » lors de la construction de notre base se trouvent être en fait, si l'on veut suivre les normes FRBR, des « expressions » et des « manifestations ». La question qui reste en suspens pour nous est celle de l'intérêt à changer de vocabulaire pour emprunter

28. La notion d'expression n'englobe toutefois pas les aspects de typographie et de mise en page.

un vocabulaire peu précis pour la plupart des gens qui auront accès au site public mais normé par des instances internationales, ou à nous contenter de nous inspirer de la force des concepts manipulés par ces instances tout en conservant nos termes d'« œuvre » et d'« édition²⁹ ».

La question de ce qu'est une œuvre est bien évidemment au cœur de la base PRELIB sur la littérature de langue bretonne, mais, contrairement au souci des bibliothécaires, ce concept n'est pas le seul à occuper cette place centrale dans le projet car l'objectif de la base est d'appréhender les liens entre les différentes personnes agissant dans ce domaine. Cet aspect bidimensionnel « œuvre/liens interpersonnel » complexifie l'approche que nous proposons et ouvre d'autres types de questions.

Qu'est-ce qu'une relation ?

Auteur ou personne ?

L'intérêt de ce type de bases n'est pas seulement de stocker des données mais surtout de les interroger et si ce n'est pas l'avenir que nous aimerions connaître en interrogeant cet oracle des temps modernes, le langage de requête SQL³⁰ de la base de données relationnelles sur laquelle repose le projet PRELIB permet par exemple de connaître le nom et le nombre de femmes ayant écrit en breton³¹ ou, plus précisément, le nombre de femmes nées dans le Morbihan ayant publié des poèmes en breton après-guerre. D'autres exemples possibles de questions à grande échelle : à quelle période est-ce que le genre théâtral est le plus représenté ? Qui sont les contributeurs de la revue *Brud/Brud Nevez* ? Quelles étaient les professions exercées par les auteurs de langue bretonne ? Quelles sont les périodes où la production est intense ou, au contraire, celles où elle baisse, etc. ? Ou, à partir de requêtes beaucoup

29. On pourrait parler d'œuvre (= expression) et d'édition (= manifestation) sur le site web en mettant à disposition les données dans le SPARQL Endpoint en utilisant FRBR.

30. SQL : Structured Query Language.

31. Cf. Nelly BLANCHARD et Mannaig THOMAS, « Entre sous-représentation et sur-investissement. Les femmes dans la littérature de langue bretonne », dans Arlette Gautier et Yvonne Guichard-Claudic (dir.), *Bretonnes ?*, Rennes, PUR, 2016, p. 35-50.

plus précises connaître le nombre d'œuvres publiées par les individus ayant eu des responsabilités au sein de l'URB.

Afin de faciliter ces requêtes, les méthodes de modélisation de bases de données incitent à regrouper les informations ayant des caractéristiques communes. Ainsi, on aurait naturellement tendance à penser que l'on ne doit pas mélanger les choux et les carottes et à créer des tables pour chaque type d'information : une table pour les « auteurs », une autre pour les « parents », encore une autre pour les « relations », etc. Cependant, ces auteurs, parents et relations partagent plus de ressemblances que de différences. Ce sont tous des individus qui ont un nom, un prénom, des parents, un lieu et une date de naissance et qui peuvent être à la fois parent, relation d'autrui et auteur d'un texte. Nous avons donc tout intérêt à regrouper tous les individus dans la même table de la base de données, quel que soit leur statut (la table a pour titre « Personne »). Mais comment distinguer les auteurs des autres individus ? Pour cela, nul besoin de disposer d'une case à cocher « auteur » sur le formulaire de saisie des individus. La base de données permettant de saisir une œuvre et ses auteurs, il suffit d'interroger la base pour connaître les « personnes » ayant écrit au moins une œuvre. De même, les œuvres étant liées à la table « langue » (voir le schéma simplifié de la base, fig. 2) il est possible d'obtenir la liste des auteurs ayant écrit au moins une œuvre en breton³².

Le statut d'auteur n'est donc pas une donnée explicite mais déductible à partir d'informations saisies par ailleurs.

Comment attester et qualifier des liens entre deux personnes ?

Bien que très utile pour les recherches mentionnées précédemment, le recensement des données sert, avant tout, à poursuivre l'objectif principal de notre base depuis l'origine, à savoir : comprendre les trajectoires sociales et littéraires des auteurs de langue bretonne en mettant en évidence des liens qui les unissent à des personnes ou à des

32. La requête SQL suivante permet de récupérer la liste des auteurs ayant écrit en breton :
 SELECT DISTINCT nom, prenom FROM Personne LEFT JOIN Personne_écrit_Œuvre ON Personne.id = Personne_écrit_oeuvre.Personne_id LEFT JOIN Œuvre_has_Langue ON Personne_écrit_oeuvre .Œuvre_id = Œuvre_has_Langue.Œuvre_id WHERE Œuvre_has_Langue.Langue_id = 3

collectifs afin de voir dans quelle mesure ces relations ont pu avoir un effet sur une entrée dans le monde de l'écriture en breton, sur l'adhésion à un groupe, un parti, une association, sur le choix d'un genre ou d'une graphie en particulier, sur la volonté d'alimenter une revue... On cite très souvent la trajectoire des trois jeunes pensionnaires de Fontarabie : Kerrien, Riou et Drezen³³, et le rôle qu'a pu jouer le Père Müller dans leur volonté d'écrire des ouvrages en langue bretonne mais on peut penser également à la fréquentation du petit séminaire de Saint-Anne d'Auray par François Cadic³⁴ qui y rencontre les abbés Buléon, Le Goff et Duparc ; à Keineg qui participe à un stage de breton, sur les conseils de Loeiz Roparz, lieu où il rencontre, pour la première fois, des membres de ce qui allait devenir l'UDB³⁵...

Le but de PRELIB est donc de recenser des données tangibles, des faits, dans l'idéal précisément datés, qui attestent de liens entre des personnes ou entre une personne et un collectif. Il peut s'agir, par exemple d'une correspondance entre deux personnes, de la participation au bureau d'une association ou d'un parti politique, d'un enseignement donné ou reçu, d'une collaboration littéraire sous forme d'une édition, d'illustrations, d'une préface, etc.

Les questions que nous nous posons sont donc :

- quel(s) « événement(s)³⁶ » nous permet(tent) d'attester d'une relation ?
- quel(s) « événements(s) » nous permet(tent) de qualifier la nature de cette relation ?

Dans les dernières évolutions de la base, nous avons en effet décidé que les preuves de relations interpersonnelles entre des individus devront, comme pour les autres éléments de la base, prendre la forme d'« événements », et c'est à partir de ces derniers que l'existence

33. Cf. Fañch MORVANNOU, « Les novices du Père Wilfrid Müller », dans J. Balcou, Y. Le Gallo et L. Le Guillou (dir.), *Histoire littéraire et culturelle de la Bretagne*, Paris-Genève, Champion-Slatkine, 1987, volume 3, p. 216-232.

34. Cf. Fañch POSTIC, *François Cadic (1864-1929). Un collecteur vannetais « recteur » des Bretons de Paris*, Brest, CRBC et alii, 2012.

35. Mannaig THOMAS, « La poésie engagée-enragée des écrivains de l'UDB (1967-1974) », dans Romain Pasquier et Tudi Kernalegenn (dir.), *L'Union démocratique bretonne. un parti autonomiste dans un état unitaire*, Rennes, PUR, 2014, p. 187-202.

36. Terme employé dans l'ontologie CIDOC-CRM.

d'une relation concrète pourra être déduite. Il y a, par exemple, dans le fonds Hélias conservé au CRBC, une correspondance entre Pierre-Jakez Hélias et Auguste Dupouy datée du début des années 1950³⁷ : ceci permet d'attester d'une relation entre ces deux personnes entre 1951 et 1952. Cette approche par les faits trouve néanmoins une partie de ses limites dans la description des relations humaines : en effet, Hélias et Dupouy ont pu se rencontrer par hasard et discuter longuement dans une rue de Quimper en 1955, cet événement a pourtant toutes les chances de ne jamais apparaître dans la base de données. Par cet exemple aussi imaginaire qu'anecdotique, nous tenons à montrer que nous sommes conscients des limites du traitement informatique de données humaines. Nous pouvons donc affirmer qu'Hélias et Dupouy se connaissaient au moins de manière épistolaire pendant une période précise de leur existence, à charge aux chercheurs de creuser ensuite cette relation si elle leur paraît pertinente pour mieux comprendre l'un ou l'autre des deux auteurs.

L'approche par les événements pose aussi la question de la nature des relations qui peuvent unir deux personnes. Au début de nos réflexions sur la base de données, dans le formulaire « Relations », nous avons choisi de définir la qualité de la relation en question. Très vite il nous est apparu que la qualification de la relation était très difficile à établir, pour ne pas dire illusoire à l'échelle d'une vie. Ainsi, grâce à *l'Anthologie de la littérature bretonne au XX^e siècle* de Francis Favereau, qui est l'une de nos sources principales pour cette période, nous avons commencé à établir une partie³⁸ du réseau des relations de Ronan Huon (cf. fig. 4). Nous apprenons par exemple aux pages 91 et 110 que Yann-Bêr Piriou et Pierre Le Bihan sont des « ami(s) de Ronan Huon » : si des sources ou des archives peuvent permettre d'attester d'une relation entre ces personnes, quels événements peuvent permettre d'en attester la nature ? En d'autres termes, y a-t-il des événements qui permettent de déduire, au-delà de la simple connaissance, qu'une relation d'amitié a existé entre ces personnes et, si oui, sur quelle(s) période(s) ? Penser les relations en termes d'événement, nous fait donc gagner en préci-

37. Voir cotes PJH4C127, PJH4C129, PJH4C130.

38. Il serait illusoire bien sûr de vouloir ici être exhaustif.

sion historique en nous permettant d'en donner des preuves et de les dater mais l'une des limites de cette méthode est de risquer de mettre toutes les relations sur le même plan. Comme si, par exemple, tous les correspondants de Ronan Huon sur une certaine durée entretenaient une relation identique avec lui. Le risque de se baser exclusivement sur des faits est donc de passer à côté de certaines affinités personnelles qui peuvent pourtant jouer un rôle primordial y compris d'un point de vue littéraire (fig. 4).

La correspondance comme preuve de relation

En plus des préfaces/dédicaces, des collaborations à des revues ou à des maisons d'édition, des adhésions au bureau de certaines associations, partis politiques ou collectifs divers, la correspondance nous a paru être, dès le début, un élément primordial dans la prise en compte de relations interpersonnelles. La correspondance éditée constitue une partie des sources que nous allons utiliser, par exemple Luzel/Renan³⁹, Duval/Le Gouic⁴⁰, Souvestre/Turquety⁴¹, etc. Nous disposons également, dans les archives de la bibliothèque du CRBC⁴², de fonds de correspondance importants en lien direct avec le domaine privilégié de PRELIB : par exemple, la correspondance issue des fonds Armand Keravel, Francis Even, Yves Berthou, Tanguy Malmanche, Pierre-Jakez Hélias, Yves Le Febvre, Loeiz Herrieu, etc. Ces données inventoriées (expéditeur, destinataire, date, langue, sujet et cote d'archivage) n'ont pas encore été versées à la base car un certain nombre de questions restent posées. Comme pour les œuvres, l'approche par la partie ou le tout se pose. En effet, nous sommes confrontés à la question suivante : faut-il considérer que chaque lettre échangée entre

39. François-Marie LUZEL et Ernest RENAN, *François-Marie Luzel, Ernest Renan : correspondance (1858-1892)*, édition établie par Françoise Morvan, Rennes, PUR – Terre de brume, 1995.

40. Anjela DUVAL, *Lettres à Gérard Le Gouic. Ha bevet Breizh !*, Brest, Berlobi-Emgleo Breiz, 2013.

41. Émile SOUVESTRE et Édouard TURQUÉTY, *Lettres d'Emile Souvestre à Edouard Turquety : 1826-1852*, édition établie par David Steel, Rennes, PUR, 2012.

42. Voir la liste complète des fonds d'archives déposés à la bibliothèque du CRBC : <https://www.univ-brest.fr/crbc/menu/Bibliothèque+Yves+Le+Gallo+%28UMS3554%29/Fonds+d%27archives>

Relations interpersonnelles			
Relation	Qualité de la relation	Date de début de la relation	Date de fin de la relation
Pierre Le Roux (1874-1975) [PPN:026974738]	enseignant
François Falc'hun (1909-1991) [PPN:026857391]	enseignant
Pol Gourrierec	ami
Le Meliner Hélène	épouse
Gwenael Huon	fils
Tudual Huon [PPN:026928604]	fils
Hervé Huon	fils
Riwal Huon	fils
Jean Yves Charles (1902-1985) [PPN:161021123]	connaissance
René Autret (1914-1984) [PPN:11966416X]	ami
Geneviève Charbonnier du Sireuil (1909-1998) [PPN:057059357]	ami
Louis Paul Nemo (1900-1978) [PPN:026918528]	connaissance
Yves Ollivier (1923-2004) [PPN:027052184]	connaissance
Pierre Denis (1921-2011) [PPN:026823888]	connaissance
Jean Raymond François Plette (1921-1971) [PPN:031221955]	connaissance
Louis Ropars (1921-2007) [PPN:115045988]	connaissance
Jean François Marie Elès (1896-1963) [PPN:026849852]	connaissance
Marcel Le Clerc (1912-1984) [PPN:026950855]	connaissance
Louis Le Floc'h (1909-1986) [PPN:026894106]	connaissance
Yann Gerven (1946-None) [PPN:02958647X]	connaissance
Jean-Pierre Piriou (1937-None) [PPN:027075036]	ami
Erwan Evenou (1940-None) [PPN:030308240]	connaissance
Xavier de Langlais (1906-1975) [PPN:026964457]	connaissance	1937	...
Robert Merle	enseignant	1944	...

Personnes dont Ronan Huon est une des relations			
Qualité de Ronan Huon	Personne	Date de début de la relation	Date de fin de la relation
ami	Pierre Le Bihan (1924-2006) [PPN:02860704X]
ami	Pierre Pennec (1921-1991)
connaissance	Charles Tremel (1885-1965) [PPN:027082849]
connaissance	Joseph Abasq (1923-2006) [WD:Q980750]

Fig. 4 : Une partie du réseau de relations de Ronan Huon au 23 novembre 2016 (non exhaustif).

deux personnes atteste d'une relation ou faut-il considérer l'ensemble de cette correspondance dans sa globalité ? Dans la première approche chaque lettre serait prise comme une œuvre à part entière, ce qui a l'avantage d'une grande précision et qui, quand les lettres sont datées, permet d'évaluer au plus près l'intensité d'une relation en fonction des périodes et de la relier à une éventuelle édition des œuvres. Cette manière de faire revêt, en revanche, au moins deux inconvénients majeurs : premièrement, une lettre adressée à un destinataire précis et à lui seul n'est pas à proprement parler une « œuvre », ce qui pose un problème de définition et de conception de ce qu'est une œuvre littéraire ; deuxièmement, d'un point de vue très pratique, les correspondances représentent une masse de données énorme dont nous disposerons de manière très détaillée pour certains auteurs mais pas du tout pour d'autres (parce que nous n'aurons pas accès aux archives, parce qu'elles ne sont pas inventoriées, parce qu'elles ont disparu...). La dimension très pratique de l'ampleur du travail de saisie – à titre d'exemple, il y a plus de 2500 lettres dans chacun des fonds Herrieu, Even ou Hélias sans parler du fonds Keravel... nous fait enfin également hésiter à lancer ce chantier.

Pour le moment, nous pensons plutôt considérer la correspondance, prise globalement, comme une preuve de relation interpersonnelle. Ceci aura pour avantage de cibler plus précisément ce que nous recherchons à condition de pouvoir mesurer l'importance de ces relations. En effet, dire qu'il y a une relation épistolaire entre X et Y ne suffit pas, il faudra pouvoir circonscrire cette relation dans le temps mais également en nombre de lettres. Or, le risque de mise à plat des spécificités dû au traitement informatisé est encore présent dans ce cas : si nous pouvons en effet distinguer une correspondance assidue sur 30 ans, d'une unique carte de vœux annuelle sur la même période en précisant le nombre de lettres échangées par les correspondants, la question de la nature de ces échanges reste posée : amicaux, intimes, de travail, autre... De nouveau, la base pourra indiquer qu'il y a une relation mais ne pourra se substituer au travail du chercheur qui devra qualifier ces échanges et donc le type de relation évoluant dans le temps entre deux personnes.

Les cas particuliers des relations familiales et des prix littéraires

Un autre type de relations, en apparence plus simple, est également saisi dans la base, il s'agit des relations familiales. La question de la saisie exhaustive de l'ensemble des relations familiales des auteurs n'est évidemment pas l'objet de PRELIB. Nous nous basons donc, pour le moment, sur les sources choisies comme référence en saisissant les relations familiales qui ont un rapport avec les domaines littéraires, culturels ou politiques ou qui permettent d'en savoir plus sur l'origine sociale et donc sur la socialisation primaire des personnes concernées.

Pour la saisie des relations familiales, nous sommes également confrontés à une grande diversité de liens possibles du type « l'oncle de sa femme » ou encore le cas du « mari de la nièce de La Villemarqué » c'est-à-dire Henri-Marie de L'Estang du Rusquec... Tous ces éléments nous orientent vers une simplification des différents types de relations en faisant, dans la mesure du possible, de grandes catégories du type relation familiale, de connaissance, de collaboration, de transmission (pour le cas des enseignants-élèves, directeurs de thèse-doctorants...).

D'un point de vue plus général, que ce soit pour les relations ou d'autres données, nous constatons qu'au fil de l'avancée du travail sur PRELIB, nous avons tendance, de plus en plus, à penser la base différemment en cherchant à simplifier les formulaires de saisie : nous avons ainsi distingué les études des professions exercées par les personnes ou encore les collaborations à des émissions de radio ou de télévision en breton, les collaborations à des maisons d'édition... Il n'est pas impossible qu'à l'avenir toutes ces catégories fusionnent en une seule nommée « activités » et, de nouveau, c'est par déduction que les informations apparaîtront.

Il y a enfin un dernier type de relations à mentionner, il s'agit de liens propres au champ littéraire qui s'établissent, par exemple, entre un auteur et une instance qui lui décerne un prix littéraire ou encore entre un auteur et une collectivité qui donne son nom à une rue, une place, une école... L'implémentation des prix se fait de deux manières complémentaires : la saisie exhaustive, d'une part, de l'ensemble des prix décernés à des œuvres ou à des auteurs de langue bretonne (prix Xavier de Langlais, *Prizioù* (France 3), prix Imram, prix Sten Kidna, *priz ar yaouankiz*, prix de la nouvelle en breton de la ville de Carhaix

mais aussi collier de l'Hermine quand il concerne des acteurs du champ littéraire breton, etc.) ; et d'autre part, les prix ou des distinctions dont tous les lauréats ne seront pas saisis de manière exhaustive : des récompenses pour des œuvres en langue française par exemple ou encore l'attribution de la légion d'honneur à un auteur.

En ce qui concerne l'attribution du nom d'un auteur de langue bretonne à un lieu, il s'agira de se baser sur des sites officiels recensant ce type d'information afin de tenter d'atteindre la plus grande exhaustivité possible dans ce domaine.

En somme, le but recherché dans la saisie des données relationnelles est de faire apparaître des liens entre des personnes et des groupes mais également de montrer l'importance relative des uns et des autres dans le fonctionnement du champ littéraire de langue bretonne : cette importance peut se caractériser par le nombre d'œuvres publiées mais également par la position de centralité occupée par un auteur. Si beaucoup d'auteurs dédicent leur œuvre à une personne X, cela lui attribue nécessairement une position particulière dans le champ. Le but est ensuite de rendre visible ces données sous diverses modélisations afin, dans certains cas, de les faire émerger mais aussi et surtout de les interpréter et de les analyser. Nous espérons que l'ouverture de la base de données au public (avec la possibilité pour les utilisateurs d'ajouter ou de corriger des informations) et les nombreuses perspectives d'exploitation qu'elle ouvre permettront de nombreux échanges et projets pour l'avenir.

Rappel des sites

Description du projet :

https://www.univ-brest.fr/crbc/menu/Projets_de_recherche/Le_point_de_vue_peripherique___Programme_de_recherche_en_litterature_de_langue_bretonne

Consultation de la base : <http://mshb.huma-num.fr/prelib/>

Hébergeur : <http://www.huma-num.fr>