

# L'évolution terminologique des sciences de l'information : Étude sur corpus diachronique d'un nouveau domaine spécialisé

*Terminological Evolution in the Information Sciences: A Diachronic Corpus Study of a New Specialized Field*

18 December 2018.

**Philippe Millot**

✉ <http://preo.ube.fr/textesetcontextes/index.php?id=2318>

Le texte seul, hors citations, est utilisable sous [Licence CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.

Philippe Millot, « L'évolution terminologique des sciences de l'information : Étude sur corpus diachronique d'un nouveau domaine spécialisé », *Textes et contextes* [], 13-2 | 2018, 18 December 2018 and connection on 12 December 2025.

Copyright : Le texte seul, hors citations, est utilisable sous [Licence CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.. URL : <http://preo.ube.fr/textesetcontextes/index.php?id=2318>

PREO

# L'évolution terminologique des sciences de l'information : Étude sur corpus diachronique d'un nouveau domaine spécialisé

*Terminological Evolution in the Information Sciences: A Diachronic Corpus Study of a New Specialized Field*

## **Textes et contextes**

18 December 2018.

**13-2 | 2018**

**Approches diachroniques des discours et cultures spécialisés**

**Philippe Millot**

☞ <http://preo.ube.fr/textesetcontextes/index.php?id=2318>

Le texte seul, hors citations, est utilisable sous [Licence CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>). Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.

---

1. Introduction
2. Cadre théorique
  - 2.1 Sciences et métiers de l'Information : L'hypothèse d'un nouveau domaine spécialisé
  - 2.2. Histoire des domaines spécialisés et approches diachroniques de l'anglais de spécialité
  - 2.3. Les corpus diachroniques en anglais de spécialité
3. Méthode
  - 3.1. Corpus diachronique d'anglais des sciences de l'Information
  - 3.2. Méthode d'analyse
  - 3.3. Bilan méthodologique
4. Résultats
  - 4.1. Évolution terminologique des sciences de l'information
  - 4.2. Évolutions morphologiques : Multiplication des termes-valises et des abréviations
  - 4.3. Évolutions sémantiques : Nouvelles définitions et nouvelles métaphores
  - 4.4. Évolutions phraséologiques
5. Conclusion

# 1. Introduction

- 1 Cette exploration de l'évolution terminologique des sciences de l'information s'inscrit dans le cadre d'une étude plus large qui vise à caractériser une langue, l'anglais, dans un domaine spécialisé particulier, les sciences et les métiers de l'information. Lorsque l'on observe le corpus relativement considérable des travaux en anglais de spécialité en France comme à l'international, il apparaît rapidement que l'anglais des sciences et des métiers de l'information est un domaine de la connaissance largement méconnu des anglicistes de spécialité. Les travaux portant plus particulièrement sur l'anglais utilisé par les professionnels de l'information, qu'il s'agisse des bibliothécaires, des documentalistes, des archivistes, et autres *records managers* sont quasi-inexistants. La seule étude identifiable dans ce domaine est sans doute celle de Georges Fournier qui, en 2012, consacre un article entier à Dan Starer qui exerçait le métier de documentaliste indépendant pour les écrivains américains à la recherche d'informations attestées (Fournier 2012). En dehors de ce cas très spécifique, aucune étude n'aborde vraiment la question de l'anglais utilisé dans le domaine de l'information.
- 2 Pourtant, les métiers de l'Information et les différentes sciences qu'ils mobilisent ont connu une évolution fulgurante depuis la fin du XIX<sup>e</sup> siècle, un moment de l'histoire souvent qualifié par les historiens du domaine « d'explosion documentaire ». À cette période en effet, la forte augmentation de la production scientifique et de la documentation professionnelle s'accompagne d'un développement des méthodes de classification, à la fois dans le monde anglophone et dans le monde francophone. En 1876, Melvil Dewey publie une classification décimale encore en usage aujourd'hui et connue sous l'acronyme DDC (*Dewey Decimal Classification*). Quelques décennies plus tard, en 1934, Paul Otlet, généralement considéré comme le père de la documentation, publie son *Traité de Documentation* qui, selon Leroux et al., « préfigure, dans une certaine mesure, l'invention d'internet un demi-siècle à l'avance » (2010 : 18). C'est également au tournant du XX<sup>e</sup> siècle que naissent les premières grandes institutions et sociétés

savantes du domaine comme l'*American Library Association* (1876), la *Library Association* de Londres (1877) et la *School of Library Economics* (1887, université de Columbia à New York), l'*Association of Special Libraries and Information Bureaux* (1924) qui deviendra par la suite l'*Association for Information Management* et l'*International Federation for Information and Documentation* fondée en 1895 par Paul Otlet. Depuis cette période, la notion d'information n'a cessé d'évoluer. À la fin des années 1940, elle est encore largement centrée sur la matérialité du document mais le développement de l'informatique, dès les années 1950, bouleverse les techniques de recherche d'information pour l'entraîner vers l'immatérialité du *big data* et autres *clouds*.

- 3 Au cours de cette histoire, que nous résumons très brièvement, de nouveaux concepts et par là-même de nouveaux termes, sont apparus et forment aujourd'hui un univers conceptuel particulièrement riche. Au-delà des cas célèbres décrivant l'immatérialité de l'information, l'expansion terminologique couvre un ensemble relativement hétéroclite de branches de la connaissance allant des sciences des bibliothèques à la gestion des connaissances en entreprise en passant par la gestion des archives. Les sciences de l'information sont, à l'évidence, multiples et une question se pose concernant la possibilité de les envisager comme un domaine spécialisé à part entière, c'est-à-dire un domaine observable du point de vue d'une langue de spécialité. La question revêt d'autant plus d'importance que pour les anglicistes de l'enseignement supérieur ou des filières telles que les D.U.T Info-Com, les licences et les masters en sciences de l'information et de la documentation ou encore l'École nationale supérieure des sciences de l'information et des bibliothèques, intègrent depuis longtemps des formations en anglais. Ces formations auraient tout à gagner d'une meilleure connaissance de cette langue de spécialité dont nous tentons ici d'esquisser quelques contours.
- 4 L'étude qui suit est une première exploration de l'évolution terminologique des sciences de l'information entre 1945 et 2015. Nous commençons par conceptualiser ces sciences en domaine spécialisé, observable du point de vue de l'anglistique de spécialité. Nous poursuivons en présentant un corpus diachronique en sciences de l'information conçu pour l'observation de l'évolution terminologique. Nos résultats montrent qu'au-delà du caractère spectaculaire de l'évolution, liée à l'abondance des nouveaux termes, des phénomènes plus pro-

fonds, d'ordre sémantique, sont à l'œuvre. Il s'avère en effet qu'en dépit du maintien du terme, le concept d'information a profondément changé depuis 1945.

## 2. Cadre théorique

### 2.1 Sciences et métiers de l'Information : L'hypothèse d'un nouveau domaine spécialisé

- 5 Il est possible de concevoir l'Information<sup>1</sup> comme un domaine spécialisé à part entière, c'est-à-dire comme un objet observable et légitime du point de vue de l'anglistique de spécialité. À l'instar d'autres domaines tels que l'économie, la finance, le droit ou la médecine, l'Information présente les caractéristiques centrales des domaines spécialisés tels que Van der Yeught (2016) les envisage, c'est-à-dire des concepts ou des ensembles de concepts qui traversent le temps et les communautés spécialisées et dont les origines se trouvent chez des individus particuliers, animés d'une intention particulière :

Dans les domaines spécialisés, les croyances et les désirs d'individus particuliers se métamorphosent en buts abstraits, lesquels buts transcendent les communautés spécialisées déterminées par le temps et l'espace qui les ont produites. Le but ultime du domaine spécialisé de la médecine est de soigner les malades et de préserver la santé, qu'il s'agisse de la médecine pratiquée en Grèce antique, en Chine traditionnelle, à l'époque de la Renaissance en Europe ou de l'Amérique moderne. (Van der Yeught 2016 : 50)<sup>2</sup>

- 6 Les historiens des sciences de l'Information tels que William Aspray, Boyd Rayward ou Michael Buckland nous permettent de retracer l'Information dans ses origines singulières, notamment le fait qu'elle soit issue de l'intention de quelques pionniers considérés aujourd'hui comme les figures de proue du domaine, en particulier Jason Faradayne, Vennevar Bush, Paul Otlet et Gerald Salton. Leur histoire nous permet également d'observer que l'Information a largement dépassé la communauté relativement restreinte des bibliothécaires pour atteindre d'autres sphères professionnelles (des archivistes aux « cour-

tiers en connaissances » en passant par les *records managers*) et scientifiques (étude des systèmes d'information, travaux sur la recherche d'information, la netométrie, la cybermétrie, etc.) et, au fond, la société dans son ensemble à travers le concept de « société de l'information ».

- 7 Outre une histoire partagée d'où émergent certaines grandes figures du domaine, les sciences de l'Information présentent une autre caractéristique des domaines spécialisés, celle de l'existence d'un savoir encyclopédique et, par là-même, d'une terminologie acceptée par les membres des communautés qu'elle représente. Ce savoir est véhiculé par de nombreuses sociétés professionnelles telles que l'*International Federation of Library Associations* pour les bibliothécaires, la *Society of American Archivists* pour les archivistes, l'*Association for Information and Image Management* pour les *records managers*, ou encore l'*Association for Information Science & Technology* pour les chercheurs. En outre, la plupart de ces organisations possèdent leurs propres organes de diffusion, qu'il s'agisse de revues de recherche universitaire (*International Journal of Information Science and Technology*, *Records Management Journal*, *Journal of Documentation*), de revues à orientation plus professionnelle (*The American Archivist*, *Journal of Information Management*, *Journal of Enterprise Information Management*) ou des magazines de vulgarisation (*Wired*, *Information Management Magazine*, *Document Strategy*). Elles possèdent également un ensemble relativement important de normes portées par des organismes nationaux (Afnor pour la France, ISO pour l'international, CEN pour l'Europe) ou par l'*International Federation of Library Association*. Ces institutions définissent les bonnes pratiques liées à la gestion informationnelle et documentaire comme la norme ISO 15489 qui porte sur la gestion des documents d'activités ou *records management*.
- 8 Une autre caractéristique qui pourrait faire de l'Information un domaine spécialisé légitime est le fait que la communauté s'est dotée d'institutions de formations, héritières des écoles de bibliothécaires, dont la première dans le monde anglophone fut fondée en 1877 par Melvil Dewey, au sein de l'université de Columbia aux États-Unis. Depuis cette date, des programmes d'études spécifiques se sont développés dans le cadre de *Bachelors* ou de *Masters* en sciences ou en gestion de l'Information. Certaines organisations telles que l'*Associa-*

tion for Information and Image Management, proposent également des formations tout au long de la vie permettant aux professionnels du domaine d'obtenir une certification comme la certification CIP (Certified Information Professional).

## 2.2. Histoire des domaines spécialisés et approches diachroniques de l'anglais de spécialité

- 9 L'approche historique des domaines spécialisés forme un aspect désormais relativement incontournable pour l'étude des langues de spécialités : l'éclairage porté sur un domaine particulier permet de mieux saisir les usages dont certains sont particulièrement opaques en l'absence de regard historique. Il faut par exemple étudier l'histoire des sciences de l'Information et celle de la formation des différents corps de métiers pour comprendre que le documentation anglophone tire son substrat théorique de courants issus des *library studies*, contrairement au monde francophone où la notion de documentation bénéficie très tôt d'une forte conceptualisation notamment grâce au Belge Paul Otlet. Il n'y a donc rien d'étonnant à ce que le documentaliste anglophone ne se considère pas vraiment comme un *documentalist* mais plutôt comme un *special librarian*, un *specialised librarian* ou encore un *scientific librarian*, le terme *documentalist* étant utilisé de manière tout à fait marginale, généralement dans des cas très spécifiques, comme les documentalistes indépendants (Fournier 2012).
- 10 Outre le caractère éclairant de l'approche historique pour rendre compte des usages de la langue à l'intérieur des domaines spécialisés (Van der Yeught 2009), les études diachroniques de l'anglais de spécialité permettent d'éclairer le domaine dans son ensemble (Banks 2011). Ce type d'études a fait l'objet d'une revue récente et approfondie par Banks (2016) pour qui l'essentiel des travaux porte sur le discours scientifique (Salager-Meyer 1999), qu'il s'agisse du discours médical (Dury 2005, Rowley-Jolivet, 2010), du discours de la biologie (Magnet 2001) ou de celui de la géologie (Château 2010). Parmi les genres de discours étudiés d'un point de vue diachronique, l'article scientifique occupe une place prépondérante. Son archivage, dès son apparition dans les premières *Philosophical Transactions* que Banks a très largement analysées, a rendu possible l'observation des textes à

l'échelle de plusieurs siècles sous des angles relativement divers, qu'il s'agisse de l'approche sociolinguistique (étude des textes dans leur situation historique), l'approche discursive (étude des caractéristiques de genre et de registre), l'approche terminologique voire l'approche iconographique à travers la comparaison des images intégrées aux textes scientifiques. Au-delà des textes scientifiques, Banks (2016) observe que les travaux qui portent sur l'étude diachronique d'autres variétés spécialisées sont en réalité assez rares, ces études se concentrant sur les textes religieux (Taavitsainen 1994), la terminologie juridique (Wagner 2003) et le discours de l'alpinisme (Wozniak 2015).

## 2.3. Les corpus diachroniques en anglais de spécialité

11 Comme nous venons de le voir, l'approche historique des domaines spécialisés occupe une place de choix dans cette branche de l'anglaisque. Mais qu'en est-il des corpus diachroniques spécifiquement conçus pour représenter l'évolution d'une langue de spécialité en tant que telle ? Ce type de représentation est, en réalité, très rare car les corpus spécialisés sont, pour la très grande majorité, conçus comme un échantillonnage synchronique visant à comparer des domaines, des genres ou des registres quelle que soit l'époque. Nous faisons donc ici la différence entre les *corpus historiques spécialisés*, qui visent la représentation d'une variété spécialisée sur une période donnée, et les *corpus diachroniques spécialisés*, qui cherchent à représenter l'évolution d'une variété, sur la base d'une structure organisée par période (Gries et Hilpert : 2008).

12 Comme Dury (2004) le suggère, cette seconde catégorie remplit des fonctions porteuses pour l'étude des variétés spécialisées. Elles nous renseignent tout d'abord sur l'évolution des usages, en particulier les termes, même si la période étudiée est relativement courte :

Les corpus diachroniques, même s'ils couvrent une période relativement courte (...), peuvent fournir des exemples de certains types de structures, certaines façon d'utiliser les termes ou encore apporter un éclairage sur les caractéristiques contextuelles d'une période donnée. (Dury 2004 : 613)<sup>3</sup>

- 13 Les corpus diachroniques spécialisés permettent également d'enrichir les définitions, en nous renseignant sur l'apparition ou la disparition de certaines significations. Ils permettent enfin l'observation de processus terminologiques tels que la terminologisation (l'acquisition du statut de terme pour un mot issu de la langue générale) et la déterminologisation (la perte de ce statut et le passage dans la langue générale).
- 14 Les corpus diachroniques spécialisés que nous avons pu identifier relèvent de l'anglais scientifique, qu'il s'agisse du domaine de l'énergie et de l'écologie (Dury 2004), du domaine médical (Dury 2008a, 2008b, 2013) ou du domaine scientifique dans sa pluridisciplinarité (Degaetano-Ortlieb *et al.* 2012). Ce type de corpus peut être de taille relativement grande comme le corpus SciTex (Degaetano-Ortlieb *et al.* 2012) qui comprend 34 millions de mots ou de taille plus réduite comme le corpus bilingue anglais/ français CIBLSP (Dury 2004) comprenant 1,5 millions de mots dont 700.000 pour la partie anglaise. Ces deux corpus traitent de périodes issues du XX<sup>e</sup> siècle échantillonnées à l'échelle de quelques décennies.
- 15 L'observation des évolutions peut prendre des orientations radicalement différentes. Pour Dury (2013), l'étude diachronique de l'anglais de l'écologie porte sur le lexique, la terminologie et les figures de style telles que la métonymie, ce qui permet à l'auteur d'identifier les représentations et les paradigmes d'une culture spécialisée émergente. Degaetano-Ortlieb *et al.* (2012) abordent la diachronie dans le cadre d'une étude quantitative de registres à partir d'ensembles de traits lexico-grammaticaux considérés comme des dimensions, dans la lignée des dimensions de registre que Biber avait conçues pour l'étude de la langue générale (Biber *et al.* 1998). Ces derniers admettent toutefois que les changements grammaticaux sont moins perceptibles que les changements terminologiques, les premiers étant beaucoup plus lents et diffus que les seconds.
- 16 La méthodologie employée pour la constitution de corpus diachroniques spécialisés correspond à celle de tout corpus spécialisé, c'est-à-dire qu'elle repose sur la question de la représentativité et du dosage des différentes parties qui les composent. Dury (2004) attire toutefois notre attention sur des difficultés spécifiques comme le traitement des textes anciens du domaine, parfois peu accessibles et

difficiles à numériser en raison de leur état. À ces difficultés s'ajoute le problème de la datation de l'émergence d'un domaine. Cette recherche du point de départ, dont l'étude diachronique des termes rend compte de manière efficace, peut toutefois contribuer à l'épistémologie du domaine en tant que tel.

### 3. Méthode

#### 3.1. Corpus diachronique d'anglais des sciences de l'Information

- 17 Notre étude de l'évolution terminologique en sciences de l'Information se fonde principalement sur un corpus de résumés de recherche issus de deux revues phare du domaine, le *Journal of Documentation* et le *Journal of Information Science*. La période couvre 70 ans d'histoire de cette science relativement récente, les premiers résumés datant de 1945, date de la première publication du *Journal of Documentation*, et les derniers de 2015 (tableau 1). Ce corpus témoigne du changement de paradigme que les scientifiques (et les professionnels) de l'information ont connu depuis la seconde moitié du XX<sup>e</sup> siècle, que les experts du domaine caractérisent de manière très générale comme le passage d'une information solide, structurée sur un support facilement identifiable (le livre, le périodique, l'image, etc.) à une information fluide, circulant sur les réseaux informatiques et les réseaux sociaux.

**Tableau 1. Structure du corpus de résumés d'articles de recherche en sciences de l'information**

Période	Résumés (n)	Mots (n)	%
1945-1954	121	19500	12
1979-1988	301	47000	28
2006-2015	506	100000	60
<b>Total</b>	<b>928</b>	<b>166500</b>	<b>100</b>

- 18 Par sa taille relativement modeste et sa visée précise, notre corpus peut être considéré comme un petit corpus spécialisé. Les 928 résumés qui le composent sont structurés en trois parties, correspondant à trois décennies clés du domaine. La première (1945-1954) correspond à la naissance de la discipline des « sciences de l'information » en tant que telles (Shapiro 1995, Bawden et Robinson 2015) tant dans le monde anglophone que dans le monde francophone. C'est d'ailleurs à cette époque de Guerre Froide que certains grands projets documentaires nationaux voient le jour sous l'impulsion d'institutions dédiées à la recherche scientifique et militaire telles que la célèbre Defense Advanced Research Projects Agency (DARPA) aux États-Unis et l'Office for Scientific and Technical Information (OSTI) en Grande-Bretagne. La deuxième partie (1979-1988) correspond au développement de l'informatique qui ouvre la voie à de nouvelles méthodes de recherche documentaire et conduit à la montée en puissance et la structuration du concept d'Information. C'est à cette époque également que paraît le premier numéro du *Journal of Information Science*. Enfin, la troisième décennie (2006-2015) vise à représenter la généra-

lisation d'Internet et l'émergence des réseaux sociaux qui apportent de nouveaux bouleversements aux pratiques informationnelles et au concept d'information lui-même.

### 3.2. Méthode d'analyse

- 19 Le domaine étudié étant relativement nouveau au sein des études anglaises, nous nous sommes livrés à un premier défrichage. La méthode retenue pour l'étude du corpus repose principalement sur une approche quantitative à l'aide du logiciel Antconc (Anthony 2011) que nous avons utilisé dans une double perspective, *corpus-driven* et *corpus-based*. Dans un premier temps, nous avons utilisé la méthode inductive pour dresser des listes de termes les plus fréquents pour chaque période, l'objectif étant d'identifier les grandes tendances terminologiques, c'est-à-dire les sujets spécialisés les plus récurrents dans chaque période. Toutefois, ce type de liste nous renseigne peu sur la spécificité terminologique d'une période par rapport à une autre. En effet, un terme peut être peu fréquent dans une décennie mais peut être parallèlement spécifique à cette dernière et, par là-même, nous renseigner sur les tendances scientifiques nouvelles. Ce type de trait a été identifié à partir du concept de spécificité lexicale ou *keyness* (Bondi et Scott 2010) que les listes de mots-clés obtenues grâce au logiciel Antconc permettent de définir. Pour les besoins de cette étude, nous avons opté pour une définition étroite, technique pour ainsi dire, de la notion de mot-clé à partir de la définition de Scott : « Un mot-clé [...] est un mot (ou un phrasème) dont la fréquence est inhabituelle dans un texte ou ensemble de textes donnés » (Scott 2010 : 149)<sup>4</sup>.
- 20 L'obtention de ces listes suppose, d'une part, de définir des sous-corpus cibles, en l'occurrence les deux sous-corpus représentant les deux dernières décennies étudiées, et d'autre part, des corpus de référence, en l'occurrence la ou les décennies précédant la décennie cible. L'identification de la spécificité lexicale propre à chaque décennie est récapitulée dans le tableau 2.

**Tableau 2. Corpus cibles et corpus de référence dans l'étude diachronique de la spécificité lexicale (keyness)**

<b>Corpus cibles</b>	<b>Corpus de référence</b>
1979-1988	1945-1954
2006-2015	1979-1988 et 1945-1954

21 Cependant, comme Grabowski (2015) le rappelle dans une étude relativement similaire, la liste générée automatiquement par le logiciel ne saurait être considérée en tant que telle car les items contenus dans les listes n'éclairent pas nécessairement la problématique. Par exemple, les listes de mots-clés pour cette étude présentent un nombre relativement important de mots académiques tels que *study*, *focus*, ou encore *article* dans les décennies 1979-1988 et 2006-2015, ce qui tend à montrer que le discours académique en sciences de l'Information, à l'instar de nombreux autres domaines, s'est normalisé au cours du XX<sup>e</sup> siècle, notamment en direction du format IMRAD. Dans le cadre de cette étude donc, les listes brutes, générées automatiquement par le logiciel, ont été traitées de manière à ne retenir que les éléments pertinents pour la problématique de cet article, centrée sur l'évolution de la terminologie du domaine des sciences de l'Information. Notre approche quantitative inductive comprend également une étude des acronymes que nous considérons comme des marqueurs de la terminologisation du domaine. Ces marqueurs ont été identifiés au moyen d'un étiquetage systématique de chaque sous-corpus à l'aide de la balise `<abbr type="acronym">` issue de la *Text Encoding Initiative*. Une fois les trois sous-corpus étiquetés, nous avons pu procéder à la comparaison des schémas de fréquence.

- 22 L'identification des nouveaux termes (mots, phrasèmes et acronymes) nous a alors permis d'étudier leur comportement phraséologique à partir d'une étude inductive (*corpus-based*). Pour obtenir ce type d'information, nous avons eu recours au concordancier d'Antconc et, plus particulièrement, aux outils « *clusters* » et « *collocates* » qui permettent d'extraire les schémas collocationnels des termes précédemment identifiés.
- 23 Enfin, l'approche outillée a été complétée par une approche qualitative, essentiellement centrée sur l'étude des définitions des termes communs aux trois décennies afin d'observer d'éventuelles évolutions sémantiques. Cette partie de l'étude a reposé sur la consultation des différentes éditions d'un même glossaire faisant référence chez les spécialistes du domaine, le *Harrods' Librarians' Glossary and Reference Book* (Harrod 1959, Harrod 1977, Prytherch 1990, Prytherch 1995, Prytherch 2016). Au départ conçu pour le champ relativement restreint de la science des bibliothèques (l'édition de 1959 comprend environ 3000 entrées, centrées sur le livre), son champ d'application a progressivement évolué pour embrasser le champ des sciences de l'information dans son ensemble (la dernière édition, datant de 2016, comprend près de 10200 termes).

### 3.3. Bilan méthodologique

- 24 En résumé, nous avons constitué un petit corpus spécialisé, structuré de telle manière qu'il permet la représentation de l'évolution du domaine des sciences de l'Information depuis 1945. Il convient toutefois de souligner la limite d'un tel échantillon qui représente une période relativement courte à partir de revues en nombre tout aussi limité. Il n'en demeure pas moins que les décennies traitées sont considérées par les experts eux-mêmes comme des périodes cruciales couvrant l'apparition de l'ordinateur et d'internet. Ces phénomènes ont clairement bouleversé les pratiques informationnelles et ont conduit à l'expansion d'un domaine centré, à l'origine, sur le livre imprimé. En outre, le choix des résumés d'articles issus de deux revues phares du domaine nous permet d'obtenir un concentré des tendances scientifiques depuis 1945. À partir de ces éléments, une étude diachronique des sciences de l'Information nous paraît donc envisageable. Nous en présentons les premiers résultats dans la section suivante.

## 4. Résultats

### 4.1. Évolution terminologique des sciences de l'information

- 25 Notre étude montre que la terminologie de l'Information a connu une très forte expansion tant sur le plan conceptuel, de nouvelles techniques générant de nouvelles significations, que sur le plan formel, l'émergence de nouveaux phénomènes informationnels conduisant à la production de nouvelles formes. Une simple comparaison du nombre d'entrées dans l'un des plus célèbres glossaires anglophones du domaine, le *Harrod's Librarians' Glossary and Reference Book*, permet de voir qu'en l'espace de quelques décennies, le nombre de termes a plus que triplé. À la fin des années 1950, la deuxième édition du glossaire (Harrod 1959) comprenait environ 3000 entrées pour atteindre le chiffre de 10250 dans la dernière édition disponible (Prytherch 2016) dont 1700 sont entièrement nouvelles par rapport à la neuvième édition et près de 3000 ont dû être entièrement révisées pour représenter fidèlement la réalité terminologique du domaine en ce début du XXI<sup>e</sup> siècle. Cette expansion terminologique repose tout d'abord sur le maintien des activités classiques, fondées sur le livre, qui occupent encore aujourd'hui une place centrale dans les métiers de l'Information. Les professionnels de l'Information ont toujours pour mission principale de classer, de cataloguer et d'archiver les documents, quel que soit le contexte institutionnel (une bibliothèque ou une grande entreprise). Toutefois, la numérisation des activités liées à la recherche d'information a conduit, dès les années 1950 et les débuts de l'ordinateur, à de nouveaux supports (microfiches, disquettes, disques durs, puis les clouds) et de nouvelles méthodes de recherche d'information (de l'index d'ouvrages aux moteurs de recherche sur Internet).
- 26 Plus particulièrement, l'analyse du corpus montre que les sciences de l'information ont connu une évolution conceptuelle très marquée : elle reflète nettement l'évolution du contexte dans lequel le domaine s'est développé au cours des trois décennies étudiées. Pendant la décennie 1945-1954 (tableau 3), le domaine est essentiellement centré sur le monde du livre et celui des bibliothèques. Des termes tels que

« *book* », « *library* » ou encore « *bibliography* » présentent un fort degré de spécificité par rapport aux autres corpus où ces termes sont bien présents mais ils apparaissent avec un relativement faible degré de fréquence. La recherche d'information, l'un des objets principaux des sciences de l'Information, s'effectue essentiellement par le biais de catalogues, de bibliographies et de microfiches. La situation terminologique des deux décennies suivantes est sensiblement différente (tableaux 4 et 5) : elle témoigne de la révolution conceptuelle du domaine dans laquelle la documentation et ses documents traditionnels, tels que le livre cèdent leur place à l'information et à la donnée (« *data* ») accessible sur ordinateur. La recherche d'information se complexifie avec l'émergence de nouveaux concepts (« *data mining* », « *tagging* », « *metadata* ») et un nouveau langage documentaire (« *XML* »), permettant à leur tour le développement de nouveaux concepts tels que le web sémantique. Outre les concepts proprement dits, cette comparaison des mots-clés de chaque décennie illustre un tournant social allant du bibliothécaire comme acteur principal à l'utilisateur lambda (« *user* ») dont les comportements numériques font désormais l'objet de nouveaux travaux par les chercheurs du domaine.

**Tableau 3. Les mots-clés de la décennie 1945-1954**

Mots-clés	Degré de spécificité
Documentation	471.311
ASLIB	139.446
Libraries	119.67
Books	110.861
Periodicals	101.271
Catalogue	71.218
Bibliographical	56.395
Bibliographies	56.173
Special	51.518
Cards	49.345
Volumes	44.472
Notation	43.103
Microfilm	42.906
List	42.717
Scientific	37.543
Librarians	35.896
Classification	32.18
Librettos	32.18

**Tableau 4. Les mots-clés de la décennie 1979-1988**

Mots-clés	Degré de spécificité
Retrieval	86.296
Systems	85.474
Data	62.623
System	60.079
Online	48.497
Computer	45.644
User	38.264
Database	37.086
Management	37.037
Technology	37.037
Information	35.918

Tableau 5. Les mots-clés de la décennie 2006-2015

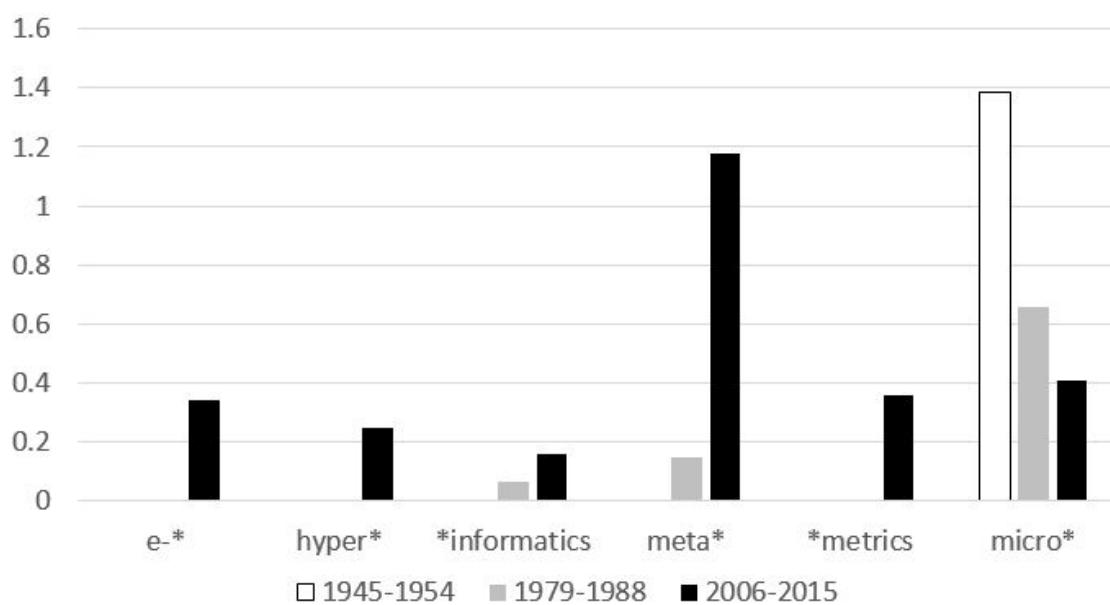
Mots-clés	Degré de spécificité
knowledge	162.368
web	128.485
social	102.882
sharing	73.038
algorithm	65.592
ontology	52.347
metadata	49.417
KM	44.095
XML	44.095
mining	41.054
tagging	38.773

## 4.2. Évolutions morphologiques : Multiplication des termes-valises et des abréviations

- 27 Les termes-valises que nous avons rencontrés au cours de notre analyse s'inscrivent dans la droite lignée des observations précédentes portant sur d'autres domaines spécialisés tels que le commerce en ligne (Humbley 2008) et, plus récemment, l'économie (Resche 2015). Ils consistent en une association de deux termes, éventuellement tronqués, qui constituent alors un nouveau terme employé de manière régulière par les experts du domaine. Comme la figure 1 le montre, la plupart de ces termes apparaissent dans les années 2000 (**e-book**, **hyperlink**, **bibliometrics**). Ils reflètent que les sciences de l'Information sont marquées par le maintien des objets fondamentaux de

la discipline (le livre, l'information, les données, etc.) et par l'introduction de nouvelles techniques venant s'appliquer à ces objets (de nouvelles techniques de mesure, de nouveaux formats de données), de nouveaux environnements technologiques (l'ordinateur, le réseau), de nouvelles typologies (*data* versus **metadata**).

**Figure 1. Préfixes et suffixes formant la base de termes valises en sciences de l'information : Variations de la fréquence entre 1945 et 2015 (par millier de mots)**



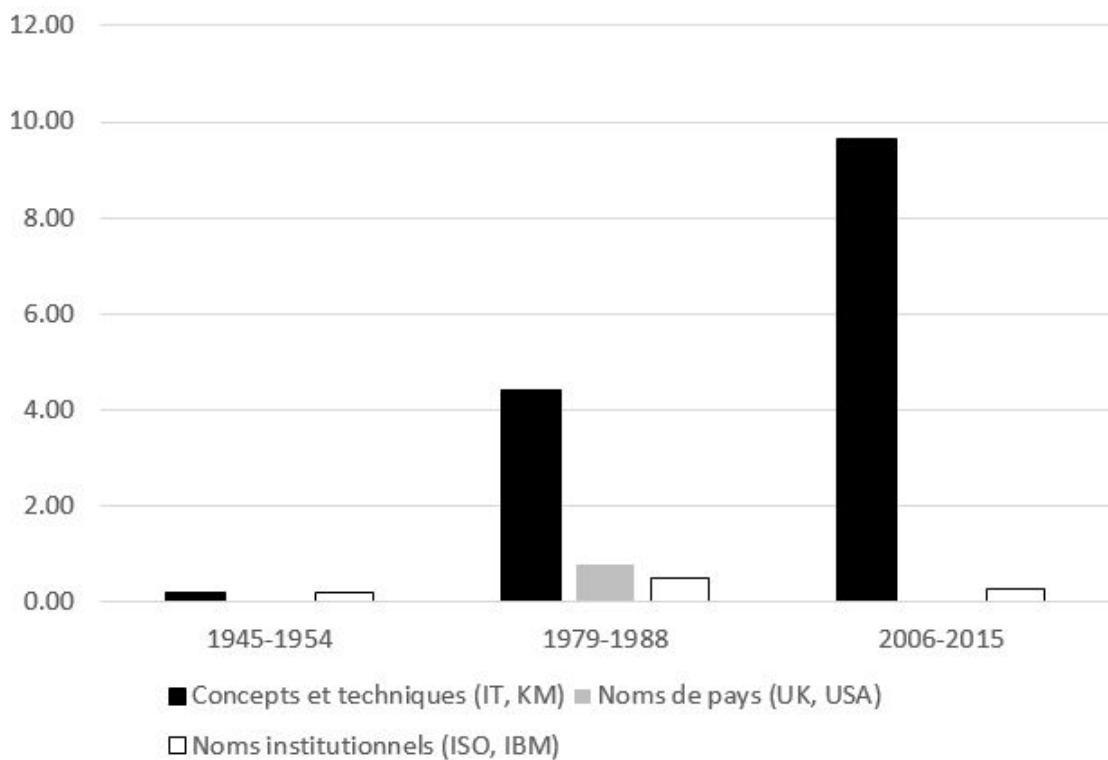
- 28 L'analyse des termes-valises permet également de montrer qu'un pan entier des sciences de l'Information concerne le service documentaire aux autres domaines, comme l'illustrent les cas de *chemoinformatics*<sup>5</sup>, *bioinformatics* et *geoinformatics*. D'autres exemples de formes rencontrées sont présentées dans le tableau 6.

**Tableau 6. Termes-valises en sciences de l'Information : Structures et formes rencontrées**

Structures	Formes rencontrées
<i>e+suffixe</i>	<i>email, e-sources, e-commerce , e-resources, e-consultations, e-books, e-textbooks, e-journals, e-documents, e-government, e-prints, e-business, e-library, e-health, e-form, e-newspaper, e-learning</i>
<i>Hyper + suffixe</i>	<i>hyperlink, hyper-media, hyper-book</i>
Domaine + informatics	chemoinformatics, bioinformatics, geoinformatics
Préfixe + metrics	webometrics, cybermetrics, scientometrics
Meta + suffixe	metadata, metacat, metamodel, meta-classifier, meta-evaluation, meta-attributes

- 29 Cette évolution s'accompagne d'une nette augmentation de la fréquence des abréviations, une catégorie très large rassemblant ici les acronymes et autres « alphabétismes<sup>6</sup> » (Algeo 1991 : 11). Parmi les trois catégories sémantiques que nous avons pu identifier (« concepts et techniques », « noms d'institutions » et « noms de pays »), la catégorie « concepts et techniques » est celle qui rassemble des termes dont les plus fréquents aujourd'hui sont IR (« Information Retrieval »), XML (« Extensible Markup Language ») et KM (« Knowledge Management »). Il faut toutefois noter que la forte augmentation du nombre d'occurrences, à partir de la deuxième décennie (cf. figure 2), ne garantit pas la survie des termes pour autant.

**Figure 2. Évolution de la fréquence des acronymes de 1945 à 2015.**



30 Le maintien des abréviations d'une décennie à l'autre est un phénomène plutôt rare, à l'échelle de quelques cas tels que ISBN (« International Standard Book Number »), IT (« Information Technology ») ou encore ISO (« International Organization for Standardization »). La plupart des cas observés sont en réalité des hapax qui, ici, sont des abréviations qui résultent d'une convention ponctuelle dans la communauté des spécialistes.

### **4.3. Évolutions sémantiques : Nouvelles définitions et nouvelles métaphores**

31 Notre étude comparative de trois périodes clés de l'histoire des sciences de l'information, consolidée par une étude épistémologique, fondée sur le dictionnaire spécialisé Harrod's Librarian's Glossary, nous a permis de dégager trois grandes phases du développement sémantique du domaine. La première phase, datant de la fin des années 1970 concerne le développement du concept d'Information. C'est en effet à cette époque que l'un des pionniers du domaine, Jason Farradane, ancre le concept dans champ de la communication humaine et le

rend indissociable de la notion de « *record* », le document structuré, portant la trace, écrite ou orale, d'une activité humaine (Farradane 1979). À cette première phase de conceptualisation succède une deuxième phase d'évolution épistémologique au cours de laquelle le concept d'Information devient véritablement opérationnel, même si le problème de la définition du concept reste toujours aussi prégnant. Les articles du glossaire Harrod en sont l'illustration. Entre 1977 et 2005, la définition de l'Information passe d'une conception philosophique pour ainsi dire à une définition plus technique, centrée sur le traitement des données.

Information : L'ingrédient essentiel à tout système de contrôle (Harrod 1977 : 418)

Information : Un assemblage de données sous forme intelligible et pouvant être communiquée. Il inclut divers éléments allant du Contenu, quelle que soit la structure véhiculant ces contenus (écrite ou imprimée, stockée dans des bases de données électroniques, données recueillies sur Internet, etc.), aux connaissances personnelles que possèdent les membres d'une organisation. (Prytherch 2016 : 349)<sup>7</sup>

- 32 La structuration du concept conduit alors à une troisième phase de fort développement du champ d'application de l'Information, comme en témoigne l'augmentation relativement fulgurante du nombre d'entrées du dictionnaire à base du terme (2 entrées en 1962, 34 en 1990 et 51 en 2005). Dans la période la plus récente, le concept d'Information porte sur des personnes (« *information broker* », « *information consultant* »), des activités (« *information industry* », « *information audit* », « *information engineering* », « *information mapping* »), et des espaces plus ou moins institutionnalisés, dédiés à l'information (« *information centre* », « *information superhighway* », « *information department* »). Le concept sert également de point d'appui à de nouvelles catégories centrales à la société de l'information tels que « *information literacy* », « *information management* » ou encore « *information retrieval* ». L'observation de ce phénomène de dictionnarisation (Van der Yeught 2012) du terme « *information* » est similaire à celui que connaissent d'autres termes centraux tels que « *data* », « *document* » et, plus récemment « *knowledge* » (cf. tableau 7). À me-

sure qu'ils entrent dans le dictionnaire et qu'ils s'y développent à travers la multiplication du nombre d'entrées, leur définition s'ajuste aux mutations technologiques pour devenir de nouveaux termes à part entière. La définition de ces termes dans le Harrod's Glossary, tout comme dans les encyclopédies auxquelles nous avons eu accès, est en effet indissociable des problématiques techniques liées à la recherche d'information et à l'optimisation des systèmes dans lesquels cette information se trouve partagée. Par exemple, le terme « *knowledge* », et sa mise en œuvre organisationnelle, le « *knowledge management* », sont définis comme l'optimisation des connaissances de telle sorte qu'elles rendent potentiellement service aux institutions (une entreprise ou une grande organisation) qui les mobilisent (Feather et Sturges 2003).

**Tableau 7. Processus de dictionnarisation de quelques termes centraux en sciences de l'information : Comparaison de nombre d'entrées dans quatre éditions du *Harrod's Glossary***

	<b>1959</b>	<b>1977</b>	<b>1990</b>	<b>2005</b>
<b>Data</b>	0	8	12	18
<b>Document</b>	1	11	9	11
<b>Information</b>	2	31	33	42
<b>Knowledge</b>	1	1	4	7

33 Parallèlement, l'expansion sémantique et conceptuelle du domaine produit les nouvelles métaphores de l'Information en réseau. Cette

expansion apparaît dans le corpus de manière tout à fait saillante dès la deuxième décennie étudiée, époque à partir de laquelle le domaine de l'Information bénéficie des progrès de l'informatique et de ses modèles mathématiques. Il s'ensuit que la plupart des métaphores constatées empruntent à ces deux domaines, même si certaines établissent de nouvelles correspondances avec des objets ou des actions concrètes. L'observation du corpus montre que les métaphores en sciences de l'Information remplissent trois grandes fonctions représentatives. La première consiste à représenter l'information dans sa structure et son étendue à travers des termes métaphoriques tels que « *cloud* », « *conceptual map* », « *tree structure* », « *world wide web* », « *data warehouses* », « *knowledge pyramids* » ou encore « *data streams* ». La deuxième catégorie représente l'information dans son immensité à travers des termes tels que « *digital realm* », « *big data* » ou « *crowd sourcing* ». Enfin, la troisième catégorie représente les processus de recherche d'information, qu'il s'agisse de « *data mining* », « *pattern mining* », « *nearest-neighbour searching* ». Ces métaphores témoignent clairement du bouleversement conceptuel qu'a connu l'Information entre 1945 à 2016 en passant d'un état solide et matériel (le livre, le document en format papier) à l'état fluide et immatériel (la donnée et ses multiples bases). À ce bouleversement correspond des nouvelles logiques de recherche d'information qui, sur la période étudiée, passent de la logique claire de la classification hiérarchique telle que la classification de Dewey à la logique floue et approximative (« *fuzzy logic* », « *nearest-neighbour* ») de la recherche d'information par algorithmes.

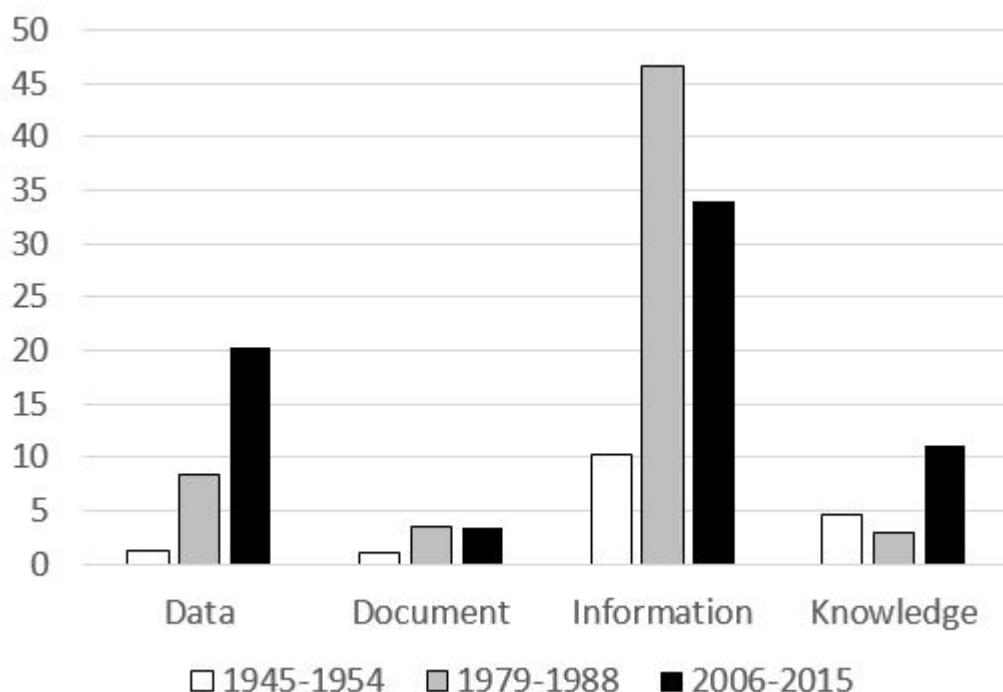
## 4.4. Évolutions phraséologiques

34

L'étude de l'évolution phraséologique du domaine porte ici sur quatre termes transversaux aux décennies étudiées (« *data* », « *document* », « *information* » et « *knowledge* ») dont nous avons exploré l'environnement phraséologique immédiat, c'est-à-dire dans un co-texte réduit à l'échelle de quelques mots à droite comme à gauche. Dans cette partie de l'étude, nous souhaitons mettre en évidence l'expansion phraséologique de termes certes traditionnels dans le domaine mais qui, progressivement, se trouvent associés à de nouvelles problématiques, voire de nouveaux sous-domaines. Cette expansion apparaît tout d'abord de manière générale lorsque nous comparons les

indices de variété lexicale (ratio type/token) des groupes nominaux comprenant ces termes (cf. figure 3).

**Figure 3. Comparaison diachronique de la variété lexicale des clusters de trois termes centraux en sciences de l'Information (ratio type/token)**



- 35 La figure 4 montre l'augmentation sensible du nombre de colocats du mot « *data* ». Alors qu'entre 1945 et 1954, le nombre d'associations de ce terme se limite à quelques notions (« *specialist* », « *empirical* », « *numerical* » et « *scattered* »), ce nombre augmente sensiblement pour former dans la dernière décennie une phraséologie radicalement différente. L'évolution phraséologique témoigne d'une montée en puissance d'un terme progressivement associé à de nouvelles techniques et de nouveaux langages (« *data processing* », « *data-driven* », « *XML data* ») et de nouveaux types (« *raw data* », « *web data* », « *authorship data* », « *spatial data* », « *RDF data* »).

**Figure 4. Évolution phraséologique du terme « data » entre 1945 et 2015.**

1945-1954	1979-1988	2006-2015
Data specialist, empirical data, numerical data, scattered data	Bibliographic data, data bases, raw data, data processing, data communication, data management, data protection, trial data, data analysis, data element, data input, data language, data model	Data sources, data structure, data streams, big data, XML data, data mining, spatial data, data uncertainty, open data, data-driven, bookmarking data, RDF data, data quality, tag data, real, web data, attribute data, authorship data, bibliographic data, biomedical data, data collection

36 Comme les autres cas étudiés le montrent, cette montée en puissance se présente de manière tout à fait spectaculaire avec le terme « *information* » dont la variété lexicale atteint son apogée entre 1979 et 1988 pour ensuite diminuer dans la dernière décennie, époque au cours de laquelle le terme « *knowledge* », lui, gagne du terrain à travers le développement du « *knowledge management* » (la gestion des connaissances), considéré aujourd’hui comme un sous-domaine à part entière des sciences de l’Information.

## 5. Conclusion

37 Cet article avait pour objectif de documenter l’évolution terminologique du domaine des sciences de l’information à travers le prisme de deux revues phare du domaine, tout en présentant les sciences de l’Information comme un nouveau domaine spécialisé observable du point de vue de l’anglistique de spécialité.

38 Il faut tout d’abord nuancer la portée d’une telle recherche qui repose sur un corpus relativement restreint et circonscrit à la dimension scientifique de l’Information. Ici, l’évolution terminologique est envisagée dans le cadre précis du discours scientifique et il conviendrait d’observer le phénomène dans un cadre plus large, incluant le vaste champ des métiers de l’Information (*information managers*, archivistes, documentalistes en institution publique ou privée, etc.) et un corpus de langue plus générale afin de voir si certains termes, à l’ins-

tar de *data* et d'*algorithm*, obtiennent déjà une forme de consécration sociale, pour reprendre l'expression de Cusin-Berche (1998).

- 39 Sur le plan méthodologique, nous avons montré que l'étude diachronique d'un domaine spécialisé s'avère très efficace lorsque celle-ci se fonde directement sur les grandes phases de son développement scientifique. L'approche par mots-clés du corpus a, par exemple, permis d'identifier les premières tendances évolutives, qui tendances ont ensuite été approfondies grâce à une approche qualitative. Cette dernière, fondée sur la comparaison des définitions et l'étude de l'histoire du domaine nous a par ailleurs permis de limiter le biais lié à la petite taille du corpus et à l'échelle, tout aussi réduite, de la diachronie. Il convient toutefois d'ajouter que nous n'avons pas traité la quantité relativement importante d'*hapax* car nous souhaitions nous livrer à une première caractérisation de l'évolution du domaine à travers les termes les plus saillants d'un point de vue statistique. Comme Resche (2013) l'a déjà observé dans ses travaux sur l'économie, certains termes, en dépit de leur faible fréquence dans un corpus donné, et parce qu'ils peuvent relever de sujets fondamentaux ou dénoter de courants de pensée, peuvent potentiellement contribuer au compte rendu d'une langue de spécialité.
- 40 L'évolution de l'Information, et la prolifération de termes qui l'accompagne, pourrait jeter un doute sur la cohérence d'un domaine relativement émergent et dont les préoccupations aujourd'hui ont déjà bien changé depuis les débats scientifiques liés à la classification des documents imprimés dans les années 1940. Cependant, notre observation montre bien qu'à l'échelle synchronique de chaque période étudiée, l'Information présente une homogénéité terminologique relativement forte, essentiellement tournée vers les techniques de classification. S'il est vrai que ces techniques ont radicalement changé depuis 1945, notre étude met en évidence qu'elles s'inscrivent dans une forme d'intentionnalité collective très générale, qui traverse les époques et que les spécialistes de ce domaine appellent « l'organisation des connaissances ». Cette cohérence dans l'évolution, ajoutée aux éléments que nous avons abordés dans cet article, nous mènent donc à la conclusion que l'Information formerait un domaine spécialisé, que des études ultérieures pourront explorer et élargir à d'autres textes et contextes.

### Remerciements

Nous remercions Christian Cote, Maître de conférences HDR en sciences de l'information à l'université de Lyon, pour ses conseils et ses suggestions lors de la rédaction de cet article.

---

ALGEO, John, *Fifty Years among the New Words : A Dictionary of Neologisms, 1941-1991*, Cambridge : Cambridge University Press, 1991.

ALONSO-ALMEIDA, Francisco / Sandra ARRERO-MORALES, « Introduction : Diachronic English for Specific Purposes », in : *Revista de Lenguas para Fines Específicos*, xvii, 2011, p.113-20.

ANTHONY, Laurence, *Antconc version 3.2.4w*. Waseda University, 2011.

BANKS, David, « Diachronic ESP: at the interface of linguistics and cultural studies », in : *ASp*, LXI, 2002, p. 55-70.

BANKS, David, « La référence à autrui dans l'article de recherche scientifique, 1700-1980 », in BANKS, David, Ed. *Aspects diachroniques du texte de spécialité*, Paris : L'Harmattan, 2010, p. 7-20.

BANKS, David, « The place of diachronic studies in LSP », in : *Moderne Sprachen*, LV/2, 2011, p.77-189.

BANKS, David, « Diachronic aspects of ESP », in : *ASp*, LXVIII, 2016, p. 97-112.

BAWDEN, David / Lyn ROBINSON, *Introduction to Information Science*, Londres : Facet Publishing, 2015.

BIBER, Douglas / Susan CONRAD, Randi REPPEN, *Corpus linguistics: Investigating Language Structure and Use*, Cambridge : Cambridge University Press, 1998.

BONDI, Marina / Mike Scott, *Keyness in Texts*, Berne : John Benjamins Publishing, 2010.

CHÂTEAU, Carmela, « Suivre la dérive de 'continents' dans un corpus diachronique d'anglais géologique », in BANKS, D, Ed. *Aspects diachroniques du texte de spécialité*, Paris : L'Harmattan, 2010, p. 21-34.

CUSIN-BERCHE, Françoise, *Le management par les mots: Étude sociolinguistique de la néologie*, Paris : Éditions L'Harmattan, 1998.

DURY, Pascaline, « Building a Bilingual Diachronic Corpus of Ecology: The Long Road to Completion », in : *ICAME Journal*, XXVIII, 2004, p.5-16.

DURY, Pascaline, « Terminology and Specialized Translation: the Relevance of the Diachronic Approach », in : *LSP & Professional Communication*, v/1, 2005, p. 31-41.

DURY, Pascaline. «The Rise of Carbon Neutral and Compensation Carbone: A Diachronic Investigation into the Migration of Vocabulary from the Language of Ecology to Newspaper Language and Vice Versa», in : *Terminology*, XIV/2, 2008a, p.230-248.

DURY, Pascaline, «Les noms du pétrole : Une approche diachronique de la métonymie onomastique», in : *Lexis, E-Journal in English Lexicology*, 2008b, <

<http://journals.openedition.org/lexis/749>.

DURY, Pascaline, « Que montre l'étude de la variation d'une terminologie dans le temps. Quelques pistes de réflexion appliquées au domaine médical », in : *Debate terminológico*, ix, 2013, p.2-10.

FARRADANE, Jason, « The Nature of Information » in : *Information Scientist*, 1/1, 1979, p.13-17.

FEATHER, John/ Paul STURGES *International Encyclopedia of Information and Library Science*. Londres et New York : Routledge, 2003.

FOURNIER ,Georges, « Documenting Fiction ».in : *InMedia: the French Journal of Media and Media Representations in the English-Speaking World*, x, 2012, <<http://journals.openedition.org/inmedia/146>>.

GRABOWSKI, Lukasz. « Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description » in : *English for Specific Purposes*, xxxviii, 2015, p.23-33.

GRIES, Stephan. Th./ Martin HILPERT, « The Identification of Stages in Diachronic Data: Variability-Based Neighbor Clustering », in : *Corpora*, 1, 2008, p. 59-81.

HARROD, Leonard Montague, *The Librarian's Glossary : Terms Used in Librarianship and the Book Crafts*, Londres : Grafton & Co, 1959.

HARROD, Leonard Montague. *The Librarian's Glossary of Terms Used in Librarianship, Documentation and the Book Crafts and Reference Book*, Londres : Andre Deutsch Limited, 1977.

HUMBLEY, John, « Issues of Distance and Proximity in Neologisms, as Instanced

in E-Commerce », in : *ASp*, LIII-LIV, 2008, p.25-38.

MAGNET, Anne, « Diachronic Analysis of the Visuals in the Research Paper: a Corpus-Based Study of the Strategies and Semiotics of Visual Representation in Nutrition Biochemistry », in : *LSP & Professional Communication*, 1/1, 2001, p.55-77.

PRYTHON, Ray, *Harrod's Librarians' Glossary of Terms Used in Librarianship, Documentation and the Book Crafts and Reference Book*, Hants : Gower, 1990.

PRYTHON, Ray, *Harrod's librarians' Glossary and Reference Book: 9,000 Terms Used in Information Management, Library Science, Publishing, the Book Trades and Archive Management*, Hants : Gower, 1995.

PRYTHON, Ray, *Harrod's Librarians' Glossary and Reference Book: A Directory of over 10,200 Terms, Organizations, Projects and Acronyms in the Areas of Information Management, Library Science, Publishing and Archive Management*. New York : Routledge, 2016.

RESCHE, Catherine, *Economic Terms and Beyond : Capitalising on the Wealth of Notions*. Berne : Peter Lang, 2013.

RESCHE, Catherine, « Réflexions sur les termes-valises en anglais de l'économie, de la gestion et de la finance », in : *ASp*, LXVIII, 2015, p. 67-85.

ROWLEY-JOLIVET, Elizabeth, « The Evolution of Medical Imagery in the 19th Century: The Lancet 1823-1905 », in BANKS, David, Ed. *Aspects diachroniques du texte de spécialité*, Paris : L'Harmattan, 2010, 53-74.

SALAGER-MEYER, Françoise, « Referential Behavior in Scientific Writing: a Diachronic study (1810-1995) », in : *English for Specific Purposes*, xviii/3, 1999, p. 279-305.

SCOTT, Mike, « What can Corpus Software Do? », in O'KEEFFE, Anne/ Michael McCARTHY, Eds. *The Routledge Handbook of Corpus Linguistics*, Abingdon : Routledge, 2010, p. 136-151.

SHAPIRO, Fred. R. « Coinage of the Term Information Science », in : *Journal of the Association for Information Science and Technology*, xxxvi/5, 1995, p.384-385.

TAAVITSAINEN, Irma « Subjectivity as a Text-Type Marker in Historical Stylistics », in : *Language and literature*, iii/ 3, 1994, p. 197-212.

VAN DER YEUGHT, Michel, *Une histoire de Wall Street*. Paris : Eska, 2009.

VAN DER YEUGHT, Michel, « A Proposal to Establish Epistemological Foundations for the Study of Specialised Languages », in : *ASp LXIX*, 2016, p.41-63.

WAGNER, Anne. 2003, « Origins and use of English legal terms through history », in : *LSP & Professional Communication*, iii/ 2, 2003, p. 92-106.

WOZNIAK, Séverine, « Étude des domaines spécialisés et de leurs discours en diachronie : Le cas des genres spécialisés de la littérature d'alpinisme aux États-Unis », in : *ASp*, LXVII, 2015, p. 81-99.

---

1 Pour éviter les ambiguïtés, nous utiliserons le terme Information, avec un I majuscule, en référence au domaine question.

2 In specialised domains, the beliefs and desires of singular men turn into abstract purposes which transcend the specialised communities determined by time and space that have generated them. The final purpose of the specialised domain of medicine is to cure sick men and preserve health, whether in ancient Greece, traditional China, Renaissance Europe or modern America (Van der Yeught 2016 : 50)

3 Diachronic corpora, even when they cover a short period of time (...), can provide examples of certain kinds of structures and term use or contexts characteristic of a certain period. (Dury 2004 : 613)

4 « A key word [...] is a word (or word cluster) which is found to occur with unusual frequency in a given text or set of texts » (Scott 2010 : 149).

5 Le sens de ce terme est relativement éloigné de l'informatique au sens où nous l'entendons en français. Il concerne les processus et les méthodes liées au traitement de l'information (stockage, archivage, etc.).

6 Algeo (1991 : 9) distingue les « alphabétismes », qui sont de simples abréviations constituées de lettres de l'alphabet (TV, VIP), des acronymes, qui eux, se prononcent tels des mots (ZIP pour « Zone Improvement Plan »).

7 Information : The essential ingredient of any control system (Harrod 1977 : 418)

Information : An assemblage of data in a comprehensible form capable of communication. This may range from Content in any format – written or printed on paper, stored in electronic databases, collected on the Internet, etc. to the personal knowledge of the staff of an organization. (Prytherch 2016 : 349)

---

## Français

Dans cet article, nous présentons une première caractérisation de l'évolution des sciences de l'information issues du monde anglophone, dans la lignée des travaux récents en anglais de spécialité. La caractérisation se fonde sur un corpus diachronique du domaine en question : un ensemble informatisé et structuré de résumés de recherche dans les revues phares du domaine : le *Journal of Information Science* et le *Journal of Documentation*. La diachronie est ici relativement courte, les premiers résumés datant de 1945 et les derniers de 2015. La méthode de caractérisation repose tout d'abord sur une approche quantitative fondée sur le concept de mot-clé que nous utilisons dans une perspective diachronique. Elle repose ensuite sur une analyse qualitative à travers l'étude de l'histoire du domaine et l'évolution des définitions de termes ayant survécu à l'évolution. L'étude nous permet de confirmer l'hypothèse selon laquelle les sciences de l'information, en dépit de leur éclatement conceptuel lié à la nature des champs tout à fait divers qu'elles couvrent (la science des bibliothèques, la documentation, la gestion des connaissances en entreprise, etc.), peuvent être conçues comme un véritable domaine spécialisé tel qu'il est défini par les anglicistes de spécialité. Ce domaine tirerait sa cohérence d'une histoire commune, marquée par une vision globalement partagée du concept d'organisation des connaissances, et par une finalité liée aux valeurs de partage de l'information et à l'innovation technologique.

## English

This article seeks to characterise the evolution of English terms in the field of Library and Information Science (LIS) in its recent history (1945–2015). Our characterisation is based on a diachronic corpus which consists of a set of abstracts from two major journals in LIS, namely the *Journal of Information Science* and the *Journal of Documentation*. Our quantitative approach is mostly based on the diachronic identification of keywords. This approach is then complemented by a qualitative study of the terms which have survived

the evolution. Our results show that although LIS clearly emerges as a multi-disciplinary field of study – including very diverse subjects such as library science, documentation, and business knowledge management – LIS may also be considered as a coherent specialised domain in its own right. Our claim is based on the observation that the many different sub-fields of LIS have shared a common core of values across their history such as a marked interest in knowledge organization, information dissemination, and technological progress.

---

**Mots-clés**

sciences de l'information, diachronie courte, anglais de spécialité, corpus, terminologie

---

**Philippe Millot**

MCF, Université de Lyon 3, 6 Cours Albert Thomas, 69008 Lyon – philippe.millot [at] univ-lyon3.fr