

h e g

Haute école de gestion
Genève

Pour de nouveaux instruments d'évaluation des publications scientifiques : état des lieux théorique et scénarios applicatifs

Travail de Master réalisé en vue de l'obtention du Master HES

par :

Cynthia A. Germond

Directeur du travail de Master :

Dr. Patrick Ruch

Genève, le 15 août 2020

Haute École de Gestion de Genève (HEG-GE)

Filière Sciences de l'Information

Hes·SO  **GENÈVE**
Haute Ecole Spécialisée
de Suisse occidentale

Déclaration

Ce mémoire de recherche a été réalisé dans le cadre de l'examen final en vue de l'obtention du Master of Science en Sciences de l'Information de la Haute école de gestion de Genève. Il a été remis pour analyse au logiciel de détection de plagiat [URKUND](#).

L'utilisation des conclusions et recommandations formulées dans le présent mémoire de recherche, sans préjuger de leur valeur, n'engage ni la responsabilité du directeur du travail, ni celle du juré ou de la Haute école de gestion de Genève.

« J'atteste avoir réalisé seule le présent travail, sans avoir utilisé d'autres sources que celles citées dans la bibliographie. »

Fait à Genève, le 15 août 2020

Cynthia A. Germond

Note de l'auteur : afin de faciliter la lecture du présent document, nous avons employé le masculin comme genre neutre.

Remerciements

Nous remercions chaleureusement chacune des personnes qui a contribué à la réalisation de la présente recherche.

Le Dr. Patrick Ruch, de nous avoir permis de travailler sur ce passionnant sujet qui nous tient tant à cœur ainsi que de sa disponibilité, de son écoute et de ses conseils avisés, tout au long du projet.

Notre expert, le Dr. Micaela Crespo Quesada, d'avoir accepté de nous accompagner dans cet ultime projet dans le cadre du Master IS.

Les participants à notre enquête, que nous ne pouvons majoritairement pas nommer du fait qu'ils ont souhaité rester anonymes, pour le précieux temps accordé et les éléments portés à notre connaissance qui nous ont permis de pousser notre analyse plus loin. Aussi remercions-nous le domaine Santé de la HES-SO, la Faculté de biologie et de médecine de l'Université de Lausanne, la Faculté de médecine de l'Université de Genève, le domaine Économie et Services de la HES-SO, le Graduate Institute de Genève, La Faculté des lettres de l'Université de Lausanne, la Faculté des lettres de l'Université de Genève et tout particulièrement son Doyen, le Professeur Jan Blanc. Nous remercions également chacun de nos interlocuteurs au sein des facultés contactées, que ces dernières aient ou non pu répondre à notre sollicitation, pour le temps accordé.

Nous exprimons toute notre reconnaissance à nos relecteurs pour leurs lectures critiques et leurs conseils éclairés. Toute erreur qui subsisterait relève de notre seule responsabilité.

Notre conjoint, notre famille et nos amis, pour leur indéfectibles soutien et patience au cours de ces trois dernières années.

Résumé

Les chercheurs sont évalués dès le XVII^e siècle. C'est à partir du XVIII^e siècle que leurs publications passent par un premier filtre de l'évaluation ; les pairs examinent quelles sont ces œuvres de l'esprit qui seront publiées dans les revues savantes. L'évaluation qualitative des publications domine jusqu'au second tiers du XX^e siècle. Les études bibliométriques se développent au tournant des années 1920-1930. Progressivement, l'utilité des mesures quantitatives en la matière s'affirme. Le développement exponentiel de la recherche scientifique, l'intervention des États par le biais de politiques scientifiques et l'avènement de l'informatique vont soutenir le développement de mesures et outils d'évaluation quantitative des publications.

Les chercheurs ne sont pas évalués sur la seule base de leurs publications. Au sein des institutions académiques, l'évaluation de leurs travaux peut faire partie des processus de nomination ou de promotion académique. Elle peut aussi être intégrée aux décisions d'octroi de fonds de recherche, au sein des hautes écoles ou par les bailleurs de fonds.

Au vu de l'actuelle remise en question de la pertinence de certains indicateurs bibliométriques et du contexte de révision des systèmes d'évaluation des chercheurs dans le cadre du déploiement de l'Open Science, le présent travail vise à comprendre le développement historique et la mise en œuvre de l'évaluation des publications scientifiques. Il a également pour objectif de procéder à un état de l'art des indicateurs bibliométriques existants puis des alternatives qui se dessinent. Afin de pouvoir compléter cette analyse, nous procédons à une étude comparative de l'évaluation des publications scientifiques des chercheurs telle qu'elle est menée au sein de certaines institutions académiques. Notre dernier objectif est de formuler, sur la base de nos observations et analyses, des recommandations aux fins de l'évaluation des publications dans le cadre académique.

Mots clefs : publication scientifique, évaluation, évaluation quantitative, scientométrie, bibliométrie, indicateur bibliométrique, citations, *h*-index, *g*-index, facteur d'impact, Open Science, altmetrics

Table des matières

Déclaration	i
Remerciements	ii
Résumé	iii
Liste des tableaux	viii
Liste des figures	viii
Liste des abréviations et acronymes	ix
1. Introduction	1
2. Méthodologie	5
2.1 Type de recherche	5
2.2 Périmètre de la recherche	6
2.3 Question de recherche	6
2.4 Collecte des données	7
2.4.1 Étapes de la collecte.....	7
2.4.1.1 Méthodologie appliquée aux deux premiers objectifs	7
2.4.1.2 Méthodologie de l'enquête.....	8
2.4.1.2.1 Population et échantillonnage pour l'enquête.....	9
2.4.1.2.2 Instruments de collecte des données.....	10
2.5 Qualité de la recherche	11
3. Les fondements historiques de la bibliométrie	12
3.1 L'évolution historique de la science	12
3.2 De la crédibilité à la légitimité scientifique	15
3.3 Historique de l'évaluation des publications scientifiques	19
3.3.1 De l'évaluation des revues scientifiques au concept de « demi-vie » ...	20
3.3.2 De l'indexation aux bases de données informatisées.....	21
3.3.3 De l'essor parallèle de la bibliométrie et des politiques scientifiques....	22
3.3.4 De la bibliométrie comme outil d'évaluation du chercheur.....	23
3.3.5 La culture du "Publish or Perish"	24
4. État de l'art dans le domaine de la bibliométrie appliquée à la production scientifique	28
4.1 L'évaluation des publications scientifiques à l'aune des métriques	28
4.2 Des éléments à prendre en considération en matière d'évaluation bibliométrique	31
4.2.1 L'évaluation bibliométrique : au niveau des disciplines ou des domaines scientifiques ?.....	31
4.2.2 Des habitudes de publication et de citation entre les domaines scientifiques.....	35
4.2.2.1 Les habitudes de publication selon les domaines scientifiques.....	36
4.2.2.2 Les habitudes de citation selon les domaines scientifiques	38
4.2.3 La qualité d'auteur	40

4.2.4	Les biais préexistants	42
4.3	Les lois « piliers » de la bibliométrie.....	44
4.3.1	La loi de Lotka	45
4.3.2	La loi de Bradford	45
4.3.3	La loi de Zipf.....	46
4.4	Les indicateurs bibliométriques	47
4.4.1	Introduction aux indicateurs	48
4.4.1.1	Les caractéristiques des indicateurs	48
4.4.1.2	La validité des indicateurs selon Gingras	49
4.4.1.2.1	L'adéquation entre l'indicateur et ce qu'il doit mesurer	49
4.4.1.2.2	Sensibilité de l'indicateur à l'inertie intrinsèque de l'objet mesuré.....	49
4.4.1.2.3	L'homogénéité des dimensions de l'indicateur	49
4.4.1.3	Du bon usage des indicateurs bibliométriques	50
4.4.2	Les indicateurs bibliométriques classiques	51
4.4.2.1	Le nombre de publications.....	52
4.4.2.2	Le nombre de citations des publications	53
4.4.2.3	Les indicateurs adaptés aux pratiques de publication et de citation des domaines scientifiques.....	57
4.4.3	L'indice de Hirsch et ses variantes.....	59
4.4.3.1	Le <i>h</i> -index	59
4.4.3.2	Les variantes du <i>h</i> -index.....	62
4.4.4	Les indicateurs basés sur les revues	67
4.4.4.1	Le facteur d'impact.....	67
4.4.4.2	Aperçu d'autres indicateurs basés sur les revues.....	70
4.5	Les critiques formulées à l'encontre des indicateurs bibliométriques classiques.....	72
4.5.1	Les critiques sur le fond.....	72
4.5.2	Les indicateurs dont la pertinence n'est pas largement reconnue.....	78
4.5.2.1	Le facteur d'impact des revues scientifiques	78
4.5.2.2	La validité du <i>h</i> -index.....	80
4.6	Les nouvelles tendances en matière de mesure et d'évaluation.....	81
4.6.1	Le développement de nouvelles mesures de la production scientifique sur le Web	81
4.6.2	La refonte de l'évaluation dans le cadre de l'Open Science	85
5.	Étude des pratiques en matière d'évaluation des publications scientifiques au sein des institutions académiques	90
5.1	L'évaluation des publications au sein des facultés et domaines échantillonnés	90
5.2	Autres systèmes d'évaluation des publications.....	96
6.	Discussion et recommandations.....	104
7.	Conclusion	117
	Bibliographie.....	120
	Ouvrages de référence.....	134
	Annexe 1 : Quadrants de la recherche scientifique (Stokes 1997)	135

Annexe 2 : Évolution du nombre de revues scientifiques depuis 1665 (Price 1975)	136
Annexe 3 : Facteurs contribuant à l'irreproductibilité de la recherche (Baker 2016)	137
Annexe 4 : Du Pré-print au Post-print.....	138
Annexe 5 : Les disciplines scientifiques sous différentes perspectives (Krishnan 2009).....	139
Annexe 6 : Répartition des langues de la communication scientifique dans les sciences naturelles : 1980-1996 (Ammon 2006).....	140
Annexe 7 : Répartition des langues de la communication scientifique dans les SHS : 1974-1995 (Ammon 2006)	141
Annexe 8 : Nombre moyen de citations par document et par domaine : 2003-2012 (OCDE et SCImago Research Group 2016)	142
Annexe 9 : Nombre moyen de références par article et par domaine : 1970-1980 (Moed et al. 1985).....	143
Annexe 10 : Taux de citation à court terme dans certains domaines (sciences dures) (Moed et al. 1985)	144
Annexe 11 : Évolution de la moyenne des parts de citations dans les langues allemande, française et anglaise dans les revues d'histoire et de chimie : 1920-1990 (Ammon 1998)	145
Annexe 12 : Langues des publications indexées dans le WoS	146
Annexe 13 : Graphique illustrant la loi de l'inverse du carré appliquée aux citations par auteur (Lotka 1926)	147
Annexe 14 : Biais de publication (Staub 2020)	148
Annexe 15 : Nombre moyen de citations par publication, deux ans après leur parution, livres inclus ou exclus : 1970-1978 (Moed et al. 1985)	149
Annexe 16 : Système de classification des citations (Moravcsik et Murugesan)	150
Annexe 17 : Système de classification des citations (Chubin et Moitra)	151
Annexe 18 : Les quatre contextes de citation de White	152
Annexe 19 : Indicateur de Hirsch-Banks pour les composés et sujets de recherche	153
Annexe 20 : Le <i>h</i>-index affiné par Schreiber (2007), exemples concrets	154
Annexe 21 : Les <i>h</i>-index versus <i>g</i>-index de Leo Egghe et Henry Small ..	155
Annexe 22 : Les <i>h</i>-index versus <i>g</i>-index et les valeurs de <i>g/h</i> des chercheurs décorés de la médaille de Price	156
Annexe 23 : le <i>g</i>-index affiné par Schreiber (2008) et comparé au <i>h</i>-index affiné	157
Annexe 24 : Définition du <i>h</i>-index et de certaines de ses variantes par Bornmann, Mutz et Daniel (2008)	158

Annexe 25 : Avantages de certains indices par rapport au <i>h</i>-index (Bornmann et al. 2011)	159
Annexe 26 : Tableau des avantages et inconvénients des principaux indicateurs bibliométriques relevés dans la présente recherche	160
Annexe 27 : L'obsession du FI (Hicks et al. 2015)	168
Annexe 28 : Situer les <i>webometrics</i> et <i>altmetrics</i>	169
Annexe 29 : L'OS et l'ouverture du processus de recherche	170
Annexe 30 : Publier en OA.....	171
Annexe 31 : Cycle de vie des données de la recherche (UNIL – UNIRIS)	172
Annexe 32 : Guide d'entretien et questionnaire de base	173
Annexe 33 : Synthèse comparative des données concernant l'évaluation des publications au sein des institutions académiques échantillonnées.....	176
Annexe 34 : SEP 2015-2021 – Catégories et 4 degrés d'évaluation (KNAW, NWO et VSNU 2014)	185
Annexe 35 : SEP 2015-2021 – <i>Output</i> indicators (KNAW, NWO et VSNU 2014)	186
Annexe 36 : 6 critères d'évaluation de la qualité (KNAW 2013).....	187
Annexe 37 : Critères et indicateurs de la qualité dans le domaine des humanités (KNAW 2011).....	188
Annexe 38 : Grille d'évaluation CERES (2020).....	189
Annexe 39 : Exemples d'indicateurs pour le modèle PQRST de Ioannidis et Khoury (2014)	190

Liste des tableaux

Tableau 1 : Nombre de projets approuvés et budgets alloués par groupes de domaines par le FNS en 2019.....	10
Tableau 2 : Matrice servant à définir les facteurs d'impact d'une revue scientifique	68
Tableau 3 : Premières préconisations de swissuniversities en termes d'évaluation dans le cadre de l'OS	87
Tableau 4 : Base de la grille d'évaluation.....	116

Liste des figures

Figure 1 : Étapes de la recherche et de la collecte des données.....	7
Figure 2 : Courbe schématique du nombre de citations en fonction du nombre d'articles (Hirsch 2005).....	60
Figure 3 : Vue des profils des chercheurs selon la KNAW (2010).....	101

Liste des abréviations et acronymes

A&HCI	Arts and Humanities Citation Index
ALLEA	All European Academies : The European Federation of Academies of Sciences and Humanities
ALM	Article-level metrics
BOAI	Budapest Open Access Initiative
CERES	Research School for International Development (University of Utrecht)
CIUS	Conseil international des unions scientifiques
CRUS	Conférence des recteurs des universités suisses
CWTS	Centre for Science and Technology Studies (Leiden University)
DG-RTD	Directorate-General for Research and Innovation (European Commission)
DMP	Data Management Plan
DOI	Digital Object Identifier
ENSSIB	École nationale supérieure des sciences de l'information et des bibliothèques
EPF	Écoles polytechniques fédérales
FAIR	Findable, Accessible, Interoperable, Reusable
FBM	Faculté de biologie et de médecine de l'Université de Lausanne (UNIL)
FCS	Field Citation Score
FI	Facteur d'impact
FID	Facteur d'impact diachronique
FNS	Fonds national suisse de la recherche
HES-SO	Haute École Spécialisée de Suisse occidentale
HUG	Hôpitaux Universitaires de Genève

ISI	Institute for Scientific Information
ISSI	International Society for Scientometrics and Infometrics
JCR	Journal Citation Reports
JCS	Journal Citation Score
JIF	Journal Impact Factor
KNAW	The Royal Netherlands Academy of Arts and Science
LDA	Loi fédérale du 9 octobre 1992 sur le droit d'auteur et les droits voisins (LDA ; RS 231.1)
LEHE	Loi fédérale sur l'encouragement des hautes écoles et la coordination dans le domaine suisse des hautes écoles (LEHE ; RS 414.20)
LIBER	Ligue des Bibliothèques européennes de recherche
LIS	Library and Information Science
NIH	National Institute of Health
NSF	National Science Foundation
NISO	National Information Standards Organization
NWO	The Netherlands Organisation for Scientific Research
OA	Open Access
OCDE	Organisation de coopération et de développement économique
ORCID	Open Researcher and Contributor Identifier
ORD	Open Research Data
OS	Open Science
PLOS	Public Library of Science
POP	Publish or perish
REF	Research Excellence Framework
SCI	Scientific Citation Index
SEP	<i>Standard Evaluation Protocol 2015-2021 (KNAW, NWO et VSNU 2014)</i>

SHS	Sciences humaines et sociales
SJR	SCImago Journal Rank
SSCI	Social Science Citation Index
UNESCO	Organisation des Nations unies pour l'éducation, la science et la culture
UNIGE	Université de Genève
UNIL	Université de Lausanne
VSNU	The Association of Universities in the Netherlands
WoK	Web of Knowledge
WoS	Web of Science

1. Introduction

Les activités de recherche sont évaluées selon trois mesures distinctes : les inputs, les outputs et l'impact. Alors que les indicateurs d'*input* reposent sur les ressources qui permettent la recherche scientifique, les indicateurs d'*output* servent à mesurer la connaissance résultant des inputs (Sugimoto et Larivière 2018). Les publications scientifiques en font sans conteste partie. Les mesures d'*impact*, concept omniprésent en matière d'évaluation de la recherche (swissuniversities 2018a), se concentrent quant à elles sur les effets et retombées de la recherche sur la communauté scientifique et plus globalement sur la société. La bibliométrie est utilisée aux fins d'évaluation de l'impact des publications scientifiques.

Les « procédures d'évaluation et d'assurance qualité » de la recherche vont de pair avec « l'histoire de la valorisation de la science et de la pratique scientifique » (swissuniversities 2018a, p. 8). Les chercheurs sont évalués depuis le milieu du XVII^e siècle (Gingras 2014a). Au XVIII^e siècle, l'évaluation de leurs publications sert plus l'étude de leur possible diffusion que l'analyse de leur qualité (swissuniversities 2018a). L'ascendance des pouvoirs publics sur les hautes écoles dès le XIX^e siècle s'accompagne de nouvelles évaluations qui concernent alors l'enseignement (swissuniversities 2018a). À la fin du XIX^e siècle, l'activité et la recherche scientifique se professionnalisent (Withley 2000, cité dans Krishnan 2009). Dès les années 1920, le domaine des sciences de l'information a recours aux statistiques et aux mathématiques (Le Coadic 2005). Si les origines de la scientométrie remontent à ces mêmes années, c'est au milieu du XX^e siècle qu'elle connaît un véritable tournant grâce à la théorisation de Derek J. de Solla Price (Price). Gingras (2014a) précise qu'il faut attendre les années 1970 pour voir la bibliométrie se développer. L'avènement de l'informatique permet le passage de l'indexation aux bases de données, soit au traitement automatisé de données toujours plus nombreuses et ce de manière toujours plus élaborée.

L'évaluation des publications scientifiques peut être menée par l'institution académique au sein de laquelle évolue le chercheur dans le cadre des procédures de nomination, de promotion académique ou d'attribution de fonds de recherche. Une évaluation peut également être effectuée par les bailleurs de fonds aux fins de participation à la recherche. Le Fonds national suisse de la recherche scientifique (FNS 2020d) évalue la compétence scientifique des chercheurs en ayant recours à la lecture de leurs curriculums scientifiques.

L'évaluation des publications scientifiques se fait tant par le biais d'instruments qualitatifs que quantitatifs. Elle ne peut être parfaitement appréhendée sans aborder l'esprit du "Publish or Perish" (POP) (cf. chapitre 3.3.5). Le monde académique est aujourd'hui indissociable de ce qui est désormais considéré comme une « culture » : "[a]cademia has become a POP world" (Mossa 2018, p. 2). Sugimoto et Larivière (2018) perçoivent l'évaluation, soit la mesure de la recherche, comme la promotion de cette culture du POP. Price (1986)¹ dénonce quant à lui les biais engendrés par le seul recours au nombre de publications dans des revues considérées comme prestigieuses en termes d'évaluation. Si l'évaluation qualitative de certaines publications scientifiques passe dans un premier temps par le « filtre » du *peer-review*, plusieurs métriques permettent l'analyse quantitative de la science en tant que système de communication : scientométrie, bibliométrie ou encore infométrie (Mingers et Leydesdorff 2015 ; Yang et Yuan 2017 ; Sugimoto et Larivière 2018) (cf. chapitre 4.1). « Le nombre étant présent dans les systèmes d'information, il est possible de mesurer » (Le Coadic 2005, p. 16). Si nous disposons désormais de mesures et outils de calcul performants, il faut rappeler qu'« [il] n'y a pas de mesure sans erreurs » (Le Coadic 2005, p. 17). Gingras (2014a) relève les critiques des chercheurs vis-à-vis de la place et de l'usage de la bibliométrie en tant que mesure considérée comme « objective » de la qualité et de l'impact de leurs travaux scientifiques. swissuniversities (2018a) précise que l'avènement de l'Open Science (OS), plus particulièrement de l'Open Access (OA), renforce la remise en question des actuels instruments d'évaluation de la recherche avant de préciser que :

« [...] dans le cadre de processus de candidature, [l]es chercheurs sont confrontés au fait que des critères partiellement contradictoires sont utilisés pour prendre des décisions ayant un impact durable sur la carrière scientifique future. »

(swissuniversities 2018a, p. 6)

Une « approche ascendante » est préconisée en matière d'évaluation : elle doit être « reconnue[...] comme judicieuse[...] et efficace[...] par les acteurs au sein des hautes écoles » (swissuniversities 2018a, p. 4). Une approche unique ne saurait toutefois satisfaire chacun des domaines d'études académiques du fait des différentes approches en matière de recherche, de pratiques de publications et de citations (cf. chapitres 4.2.1 et 4.2.2). En vertu de l'art. 20 de la Constitution fédérale de la Confédération suisse du 18 avril 1999 (Cst.) et de l'art. 5 de la Loi fédérale du 30 septembre 2011 sur l'encouragement des hautes écoles et la coordination dans le domaine suisse des hautes écoles (LEHE), les hautes écoles suisses bénéficient de la liberté académique. La Conférence des recteurs des universités suisses (CRUS) rappelait qu'"each

¹ La première édition de *Little science, Big science... and beyond* a été publiée en 1963.

university is responsible for setting its own strategy according to its mission” et que dès lors “[...] it is best that its member universities themselves determine the body of objective quality criteria that most appropriately fit the deliverables emanating from these strategies” (CRUS 2008, cité dans swissuniversities 2018a, p. 7). Les instruments d'évaluation doivent permettre une évaluation appropriée en fonction des domaines de recherche « [...] tout en étant adaptées aux besoins des différentes institutions » (swissuniversities 2018a, p. 7). La CRUS de préciser toutefois que :

“However, no university shall abstain from committing itself to a body of objective quality criteria for its self-chosen deliverables or from communicating them broadly.”
(CRUS 2008, cité dans swissuniversities 2018a, p. 7)

Nous pouvons ici effectuer une première remarque. Sans nullement remettre en question l'autonomie des hautes écoles suisses, relevons le contraste émanant du fait que l'on dénonce le caractère parfois contradictoire des critères d'évaluation tout en estimant que chaque haute école doit pouvoir adapter ces critères à ses propres missions. N'est-ce pas antinomique ? Qu'entendons-nous par une évaluation objective ? Pouvons-nous comparer des évaluations effectuées par diverses institutions alors que les critères sont établis par chaque institution académique ? C'est dans ce contexte et esprit de remise en question des modes d'évaluation des publications qu'« un groupe de rédacteurs en chef et éditeurs de revues savantes » a élaboré la *Déclaration de San Francisco sur l'évaluation de la recherche (Déclaration de Dora)* (2012 §1), depuis lors déjà signée par 2'009 organisations et 16'294 particuliers². De plus, dénonçant l'utilisation non initialement prévue du facteur d'impact des revues scientifiques (FI) (cf. chapitres 4.4.4.1 et 4.5.2.1) ainsi que d'importantes limitations de son application dans l'évaluation de la recherche, la *Déclaration de Dora* (2012) fournit un certain nombre de recommandations pour une évaluation plus pertinente des résultats de la recherche. Trois ans plus tard, Hicks et al. (2015) dénoncent les trop fréquentes mauvaises utilisations et interprétations des métriques et dressent une liste de dix principes destinés à soutenir l'évaluation de la recherche dans *The Leiden Manifesto for research metrics (Manifeste de Leiden)*. Cette même année sort le rapport *The Metric Tide* (Wilsdon et al. 2015), lequel a pour finalité de mettre un terme à l'utilisation parfois brutale des métriques qui peut avoir des conséquences tragiques³. Il s'agit de déterminer la manière d'utiliser de manière responsable et objective le « pouvoir des métriques » :

² Chiffres en date du 11 août 2020 issus de <https://sfdora.org/signers/> [consulté le 11 août 2020].

³ Dans la préface du rapport, James Wildson fait référence à la révision de l'utilisation des métriques de performance du *Imperial College* suite au suicide du Prof. Stefan Grimm à l'automne 2014 (Wilsdon et al. 2015).

*“Metrics hold real power: they are constitutive of values, identities and livelihoods.”
(Wilsdon et al. 2015, p. iii)*

L'évaluation de la recherche et des chercheurs ne se limite pas à la seule évaluation des publications scientifiques. La présente recherche se concentre sur cette dernière et plus particulièrement sur les outils et instruments d'évaluation quantitatifs servant à cette fin. Nous nous concentrons sur les publications scientifiques, fruits de la recherche fondamentale. Nous souhaitons comprendre la manière dont s'est développée l'évaluation de ces publications scientifiques ainsi que l'élaboration des métriques aujourd'hui appliquées ; subséquemment pourquoi, et dans quels cas, leur utilisation est contestée. Paradoxalement, si nombre de chercheurs expriment leur insatisfaction quant aux indicateurs utilisés pour évaluer la qualité ou l'impact de leurs publications, ils sont nombreux à citer :

*« le « facteur d'impact » des revues dans lesquelles ils publient et à afficher leur « indice h » comme gage de leur propre valeur et réussite, sans toujours indiquer la source de ces chiffres, ni même comprendre tout à fait la valeur réelle de tels indicateurs. »
(Gingras 2014a, p. 10)*

La présente recherche comprend quatre objectifs distincts. Il s'agit dans un premier temps de revenir sur l'historique de l'évaluation des publications scientifiques afin de comprendre les bases sur lesquelles reposent l'évaluation telle qu'elle est aujourd'hui mise en œuvre. Nous souhaitons déterminer si ses intentions originelles n'auraient pas été détournées, expliquant par-là même certains des biais aujourd'hui dénoncés. Nous procédons ensuite à un état de l'art des outils et indicateurs utilisés à ces fins. Il s'agit de comprendre leurs fondements ainsi que la manière dont ils devraient être appliqués. Cela nous permettra de définir si « les indicateurs choisis ont [...] bien la signification qu'on leur attribue » (Gingras 2014a, p. 11) et ce que mesure réellement la bibliométrie (Pansu, Dubois et Beauvois 2013). Cet état de l'art intègre en outre les arguments et critiques de l'évaluation des publications scientifiques telle qu'elle est aujourd'hui appliquée et des alternatives qui se dessinent. Il s'agit de saisir le contexte ainsi que l'esprit qui ont abouti à la *Déclaration de Dora* (2012), au *Metric Tide Report* (Wilsdon et al. 2015) ainsi qu'au *Manifeste de Leiden* (Hicks et al. 2015). Dans un troisième temps, nous nous penchons sur les systèmes d'évaluation mis en place dans plusieurs facultés ou domaines d'institutions académiques. Il s'agit de confronter la pratique réelle de l'évaluation des publications au regard de nos précédentes analyses. Notre dernier objectif est de formuler, sur la base de nos précédentes observations et analyses, des recommandations aux fins de l'évaluation des publications dans le cadre académique.

2. Méthodologie

2.1 Type de recherche

Du fait que son point de départ consiste en un état de l'existant, notre recherche se veut descriptive. Nous souhaitons fournir une présentation détaillée de l'évolution ainsi que de la mise en œuvre des mécanismes et outils d'évaluation quantitative des publications scientifiques. Nous avons également pour objectif d'explorer la manière dont elle est concrètement mise en œuvre au sein des institutions académiques. Sa nature est précisément qualitative du fait qu'elle se base principalement sur des données de type narratif (Fortin et Gagnon 2016). Nous basant sur l'existant pour formuler des recommandations, notre processus de recherche est clairement inductif (Fortin et Gagnon 2016).

Du point de vue méthodologique, nous privilégions l'étude de cas multiples. Il s'agit non seulement de décrire, mais également de comprendre les actuels modes d'évaluation des publications scientifiques et la manière dont ils sont aujourd'hui appliqués.

« L'étude de cas est une approche de recherche empirique qui consiste à enquêter sur un phénomène, [...], sélectionné de façon non aléatoire, afin d'en tirer une description précise et une interprétation qui dépasse ses bornes. Le cas étudié est donc bien délimité, mais forme un sous-système dont l'analyse permet de mieux comprendre un système plus large. Ainsi, on s'intéressera aux composantes qui forment le cas, y compris son contexte immédiat [...] et ses différentes dimensions. »
(Roy 2009, pp. 206-207)

Du fait que nous cherchons à déterminer si une solution à une problématique pratique – la remise en question de modes d'évaluation des publications scientifiques tels qu'ils sont aujourd'hui appliqués – est possible (Fortin et Gagnon 2016), la dernière partie de notre recherche se veut également appliquée. Nous avons recours à un raisonnement déductif du fait que nous avons pour objectif de dégager de notre état de l'art des propositions d'ajustement ou de modification du système actuel, soit élaborer une théorie hypothéticodéductive (Fortin et Gagnon 2016).

Notre choix de l'étude de cas est conforté par le fait que celle-ci :

« [...] est souvent utilisée en recherche appliquée, notamment en évaluation de programmes gouvernementaux, où elle est utile pour prendre la mesure de l'efficacité – ou des limites – d'un programme. La méthode facilite en outre l'identification de pratiques administratives exemplaires qui peuvent servir de leçons d'avenir pour l'ensemble du programme évalué. Un exemple typique est l'analyse d'un point de service en particulier d'un programme plus large. »
(Roy 2009, p. 202)

Nous veillons à assurer l'aspect qualitatif inhérent à toute recherche scientifique et prenons soin d'identifier les éventuels biais qui pourraient surgir au fil de notre recherche (cf. chapitre 2.5).

2.2 Périmètre de la recherche

Notre périmètre de recherche s'étend à l'échelle internationale. Nous nous concentrons sur l'évaluation quantitative des publications scientifiques fruits de la recherche fondamentale de chercheurs évoluant dans les institutions académiques, écartant par là-même la recherche appliquée. Là où Price (1975)⁴ distingue la science de la technologie (cf. chapitre 3.1), nous qualifions aujourd'hui la recherche de fondamentale ou appliquée. Alors que « les projets de recherche appliquée vis[ent] une exploitation commerciale » (FNS [2020b] §1), la recherche fondamentale « permet de recueillir de nouvelles connaissances sur l'être humain et la nature » et est considérée comme « impartiale », ne poursuivant « aucun objectif commercial » (FNS [2020a] §2). « L'encouragement de la recherche fondamentale constitue l'activité clé du FNS » (FNS [2020b] §1). Le FNS spécifie toutefois que celle-ci « pose les bases d'innovations techniques et sociales, et garantit ainsi la prospérité durable de la société » (FNS [2020a] §2). La distinction entre recherche fondamentale et appliquée n'est pas toujours évidente : « [à] la fois indissociables et interdépendantes, la recherche fondamentale et la recherche appliquée sont les deux faces d'une même médaille. » (Conseil international des unions scientifiques (CIUS) 2004, cité dans Organisation des Nations unies pour l'éducation, la science et la culture (UNESCO) 2016, p. 9). Le modèle du quadrant proposé par Stokes (1997) permet de distinguer la recherche fondamentale pure (le quadrant de Bohr) de la recherche fondamentale inspirée par une finalité pratique (le quadrant de Pasteur) (cf. Annexe 1).

2.3 Question de recherche

Les objectifs annoncés en clôture de notre introduction justifient notre question de recherche. Au vu du développement historique et de la mise en œuvre de l'évaluation des publications scientifiques, de l'état de l'art des outils et indicateurs bibliométriques existants, de la situation des critiques formulées à leur encontre et des nouvelles tendances en matière de mesure et d'évaluation, puis de l'étude de pratiques en matière d'évaluation des publications au niveau académique, quelles recommandations pouvons-nous formuler aux fins de l'évaluation des publications dans le cadre académique ?

⁴ La première édition de *Science since Babylon* a été publiée en 1961.

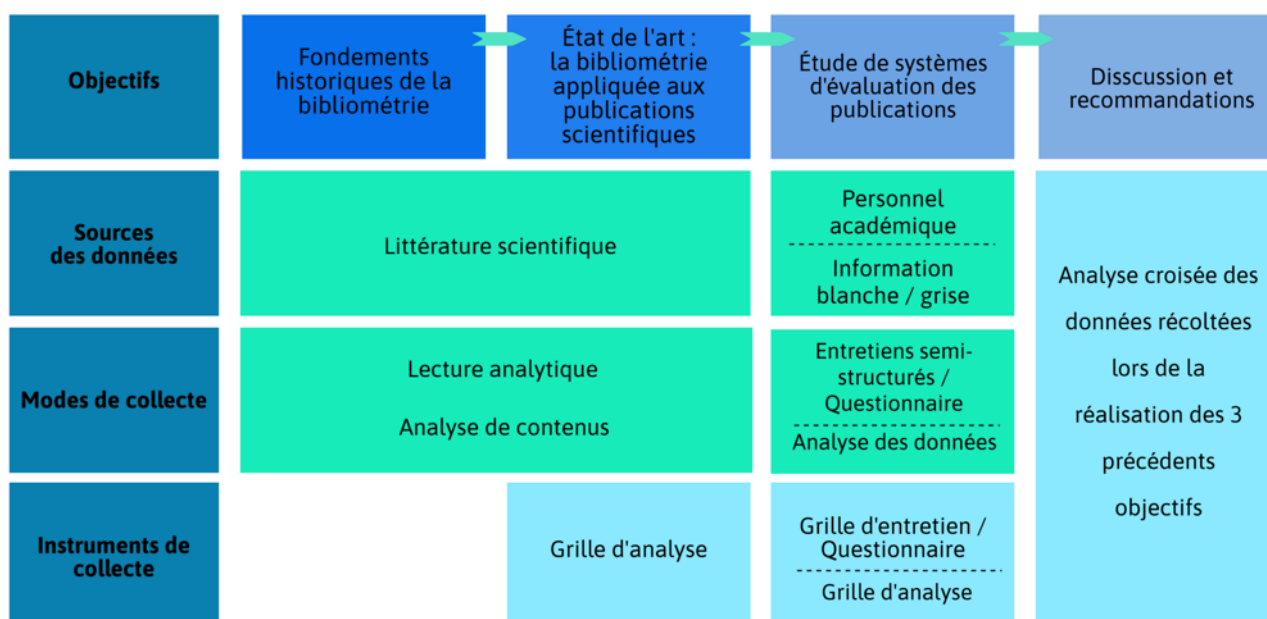
2.4 Collecte des données

Notre périmètre de recherche ainsi que son champ d'application s'étendent à l'échelle internationale. Au vu des objectifs de la présente recherche, nous avons sélectionné deux sources de données distinctes. Les sources documentaires et plus spécifiquement la littérature scientifique permettent de répondre à nos deux premiers objectifs en en dégageant les aspects théoriques de la bibliométrie appliquée aux publications scientifiques. Les unités en charge de l'évaluation des chercheurs et de leurs publications au sein des institutions académiques échantillonnées permettent quant à elles de répondre aux aspects pratiques de la recherche. Pour la réalisation de cet objectif spécifique, nous avons mené une enquête auprès de ces unités.

2.4.1 Étapes de la collecte

Chacun de nos objectifs constitue une étape de la recherche, dès lors de la collecte des données.

Figure 1 : Étapes de la recherche et de la collecte des données



2.4.1.1 Méthodologie appliquée aux deux premiers objectifs

Ayant pour ambition de réaliser une analyse aussi approfondie et exhaustive que possible, nous avons réalisé notre recherche documentaire sans limiter les champs géographiques et temporels. Il nous a semblé pertinent de ne pas la restreindre géographiquement en vue de déterminer si des courants de pensée peuvent, ou non, être observés. Du point de vue temporel, nous ne nous sommes pas fixé d'autre limite que celle engendrée par la faisabilité de la recherche dans le temps imparti. Nous avons eu à cœur de pouvoir remonter aux sources de l'évaluation bibliométrique afin de pouvoir

déterminer ses fondements et intentions originels. Aussi, en matière de recherche documentaire, notre seule limitation a été celle des langues que nous maîtrisons. Notre premier objectif est le fruit de la synthèse post-analyse des sources documentaires que nous avons consultées. L'analyse documentaire orientée vers la réalisation de notre second objectif nous a permis de réaliser un tableau des avantages et inconvénients relevés pour les indicateurs bibliométriques les plus courants (cf. Annexes 26).

2.4.1.2 Méthodologie de l'enquête

Rappelant que notre périmètre de recherche ainsi que son champ d'application s'étendent à l'échelle internationale, nous avons, pour des raisons de faisabilité, dû procéder à un échantillonnage en vue de la réalisation de notre troisième objectif : l'analyse comparative de systèmes d'évaluation mis en place par des hautes écoles au niveau de certaines de leurs facultés ou domaines (cf. chapitre 5).

Dans le cadre de la présente recherche, nous entendons les hautes écoles telles que définies par l'art. 2 al. 2 de la LEHE soit : « a. les hautes écoles universitaires, à savoir les universités cantonales et les écoles polytechniques fédérales (EPF) ; b. les hautes écoles spécialisées et les hautes écoles pédagogiques ». Au niveau helvétique, nous nous basons sur l'ensemble des « [h]autes écoles suisses reconnues ou accréditées » référencées par swissuniversities ([s.d.]). Nous transposons ce regroupement des universités, EPF et hautes écoles spécialisées pour nos recherches et analyses à l'échelle internationale. Nous utilisons le terme d'« institution académique » comme synonyme de « haute école ». C'est en ces lieux de savoir qu'est menée la recherche académique.

Les différentes sphères disciplinaires se regroupent au sein de facultés. C'est au milieu du XIII^e siècle que le vocable de "faculty" (du latin *facultas*) a pris la signification qu'il revêt aujourd'hui encore soit "the administrative subdivision of a university organising the teaching of a discipline" (Vauchez, Dobson et Lapidge 2000, p. 526). Wallace (2015) précise que le terme revêt deux significations : alors même qu'il peut signifier « le personnel enseignant et de recherche » évoluant dans une faculté, nous l'entendons ici dans le sens de la « subdivision organisationnelle elle-même » laquelle est « généralement gérée par un doyen »⁵. Une faculté peut compter plusieurs départements. Dans certaines hautes écoles, à l'instar de la HES-SO ([2020a]), nous parlons non de facultés, mais de domaines, puis des filières qui y sont rattachées.

⁵ <https://www.oxfordreference.com/view/10.1093/acref/9780199679393.001.0001/acref-9780199679393-e-356?rsk=75TWur&result=361> [consulté le 06 mars 2020] [accès par abonnement].

2.4.1.2.1 Population et échantillonnage pour l'enquête

La population cible comprend les institutions académiques reconnues et accréditées en tant que telles à l'échelle internationale, ainsi que leurs différentes facultés ou domaines et les chercheurs qui y officient. Du fait de contraintes géographiques et linguistiques, l'intégralité de la population cible ne peut être considérée comme raisonnablement accessible. S'il ne nous est pas possible d'affirmer que certaines institutions académiques présenteraient plus d'intérêt que d'autres et que chacune d'entre elle saurait être représentative de la problématique de la présente recherche, il existe un nombre conséquent de hautes écoles en suisse au sens de la LEHE. Ce nombre devient considérable à l'échelle internationale. Chaque institution académique regroupe plusieurs domaines de recherche. Nous observons des différences entre les disciplines et domaines scientifiques en matière de pratiques de publications et de citations (cf. chapitre 4.2.2). Procéder à une analyse exhaustive dans le délai imparti est irréalisable. Le nombre de cas d'étude doit dès lors, en sus des contraintes géographiques et linguistiques précitées, être restreint pour des raisons de faisabilité.

Ces considérations motivent un échantillonnage de type non-probabiliste, par choix raisonné. Au niveau territorial, nous avons décidé de circonscrire notre analyse aux pratiques d'évaluation des publications des institutions académiques romandes. Le nombre de domaines et de disciplines scientifiques varie fortement selon les systèmes de classification. "[...] [N]o classification system can be seen as the "truth", as different systems may serve different purposes [...]" (Wang et Schneider 2020, p. 241). Nous avons décidé de nous appuyer sur la liste des « [d]omaines de recherche et disciplines du FNS » qui comprend trois groupes de domaines : i. Sciences humaines et sociales (SHS) ; ii. Mathématiques, sciences naturelles et de l'ingénieur ; iii. Biologie et médecine. Chacun regroupe des domaines spécifiques et leurs disciplines respectives (FNS 2016). Notons que les SHS intègrent les sciences du droit et les sciences économiques.

Toujours en raison des différences des pratiques en matière de publication et de citation entre les disciplines scientifiques, nous avons souhaité observer les pratiques d'évaluation dans deux domaines au moins. Nous souhaitons en outre pouvoir comparer la manière dont les évaluations des publications sont menées dans des universités puis dans des hautes écoles spécialisées afin de voir si des différences notoires sont observables. Aussi nous fallait-il choisir deux disciplines au moins qui sont présentes dans ces différentes hautes écoles. Les *SHS* et la *Biologie et médecine* sont les deux groupes de domaines pour lesquels le FNS a accepté le plus grand nombre de projet en 2019 (cf. tableau 1).

Tableau 1 : Nombre de projets approuvés et budgets alloués par groupes de domaines par le FNS en 2019

Groupe de domaines (FNS)	Nombre de projets (%)	Budget (Mio. de Chf)
Sciences humaines et sociales	1'251 (36,32%)	268
Mathématiques, sciences naturelles et de l'ingénieur	1'009 (29,29%)	354
Biologie et médecine	1'184 (34,37%)	378
Totaux	3'444	1'000

(adapté de FNS [2020c])

Dans le groupe « Biologie et médecine », plus de 53% des projets concernent la médecine (pour un budget à hauteur 49,47% des budgets alloués à ce groupe de domaine) (FNS [2020c]). Concernant les SHS, notons que le niveau de granularité des données concernant les projets acceptés et budgets alloués par le FNS ([2020c]) est bien plus élevé. Si nous ne pouvons pas observer avec précision le nombre de projets accordés par discipline en 2019 dans les données du FNS [2020c], nous avons aléatoirement choisi le domaine des sciences économiques (qui cohabite avec les sciences du droit dans les données du FNS [2020c]). Nous trouvons des facultés ou domaines d'économie dans les universités et les hautes écoles spécialisées suisses. Nous avons aléatoirement choisi d'approcher les facultés des lettres. Quand bien même leurs disciplines ne se trouvent être dans le champ de recherche des hautes écoles spécialisées suisses, nous avons souhaité pouvoir comparer les modes d'évaluation des publications dans plusieurs domaines des SHS. Notre échantillon comprend la Faculté de biologie et de médecine de l'universités de Lausanne (UNIL), la Faculté de médecine de l'université de Genève (UNIGE), la section de médecine de l'Université de Fribourg (UNIFR), les Facultés des lettres de l'UNIL et de l'UNIGE, la Faculté des lettres et des Sciences humaines de l'UNIFR, les différents départements du Graduate Institute de Genève ainsi que les domaines Économie et Services puis Santé de la Haute École Spécialisée de Suisse occidentale (HES-SO).

2.4.1.2.2 Instruments de collecte des données

Notre collecte de données a été réalisée sous deux angles. Au vu du peu d'information blanche disponible quant aux processus d'évaluation des publications des chercheurs, nous avons dans un premier temps opté pour une récolte de données via des entretiens

semi-structurés. La situation sanitaire du moment (COVID-19) ayant une incidence directe sur la disponibilité de certains de nos interlocuteurs, nous avons également créé un questionnaire. Les données récoltées par le biais des entretiens et questionnaires ainsi que par la documentation transmise nous a permis d'élaborer une grille d'analyse (cf. Annexe 33) qui a nourri notre analyse comparative (cf. chapitre 5).

2.5 Qualité de la recherche

Nous avons veillé à assurer l'aspect qualitatif inhérent à toute recherche scientifique tout au long de notre étude. La rédaction d'un devis de recherche nous a permis de précisément délimiter notre problématique et de réfléchir à la méthodologie la plus pertinente en vue de la réalisation de nos objectifs. Tout au long de la recherche, nous avons veillé à adapter notre approche et notre méthodologie lorsque cela s'avérait nécessaire. Conscients que des biais sont inhérents à toute recherche, nous avons pris soin de les relever au fil de sa réalisation. La recherche puis l'analyse documentaire constituent une part importante du présent travail. Nous nous limitons à la bibliométrie en tant que mesure. Nous occultons donc les indicateurs [dits] relationnels (Rostaing 1997) , lesquels concernent la mesure des relations et de la dynamique de la science que nous considérons comme trop éloignés du cœur de notre sujet (cf. chapitre 4.1). Pour des raisons de faisabilité, nous ne pouvons détailler les fondements mathématiques et statistiques des lois piliers de la bibliométrie (cf. chapitre 4.3) et des indicateurs bibliométriques (cf. chapitre 4.4). Nous en énonçons les principes. Nous relevons ce que nous considérons être les principales distinctions entre les domaines scientifiques en matière de publications et de citations. Si elles sont corrélées au domaine scientifique (cf. chapitre 4.2.2), elles reposent également sur d'autres facteurs (culturels par exemple). Une étude exhaustive de ces contrastes devrait faire l'objet d'une recherche distincte. L'aspect interdisciplinaire ne fait pas partie de notre analyse des pratiques d'évaluation des publications au sein des facultés ou domaines des hautes écoles échantillonnées. Enfin, au vu de la situation sanitaire actuelle, nous n'avons pas pu mener notre enquête telle que nous l'aurions fait en temps « normal ». Nous n'avons pas pu rencontrer directement nos interlocuteurs. Nous n'avons dès lors pas pu développer autant que nous le souhaitions initialement l'étude pratique de l'évaluation des publications scientifiques au sein des facultés et domaines échantillonnés.

3. Les fondements historiques de la bibliométrie

3.1 L'évolution historique de la science

Comprendre la science d'aujourd'hui nécessite de « remonter dans le continuum de son histoire » et de repérer ses « moments charnières » qui ne consistent pas forcément en découvertes clefs, mais plutôt aux instants où il aura fallu élaborer de nouveaux modes de pensées, ou introduire des éléments inédits dans la réflexion scientifique (Price 1975, p. 4). Price (1975) explique l'évolution de la science et son état « actuel » par ce qu'il présente comme une succession de « crises ». La première est celle qui a permis à notre civilisation de devenir scientifique, la distinguant ainsi des autres cultures. Dans un second temps, Price (1975) considère le passage de la sphère de la « pensée pure » à la technologie scientifique. Le développement de la science poursuit son cours pour épouser les contours de la science moderne au temps de la Renaissance. La transition entre les « théories classiques du XIX^e siècle et la multiplication explosive des découvertes du XX^e siècle » est qualifiée de « brutale » (Price 1975, p. XV). En s'appuyant sur ses raisonnements et analyses, Price (1975) introduit le développement à venir de la science comme « une future économie interne [...] » qui lui apparaît déjà comme différente (Price 1975, p. XV).

S'ils sont des découvertes scientifiques qui n'ont pas encore eu lieu, rien ne nous permet d'écarter la possibilité que des connaissances existantes ne soient pas déjà portées à la lumière du jour, dès lors intégrées dans les actuelles recherches scientifiques avance Price (1975). Pourquoi donc certaines connaissances scientifiques restent dormantes ? Dans le domaine de la bibliométrie, ce phénomène est appelé « Belle au bois dormant » (van Raan 2004, cité dans Gingras 2014a).

“Galileo, rather than breaking with the past, is perhaps more accurately to be regarded as rounding out a process of refinement of mechanics which had matured during the Middle Ages but had lain dormant for a century.” (Price 1975, p. 94)

Le chercheur se consacre à la recherche, soit “the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions [...]” (Stevenson et Lindberg 2011⁶). Là où nous caractérisons aujourd'hui la recherche de fondamentale ou d'appliquée (cf. chapitre 2.2), Price (1975) distingue la science de la technologie. Lorsque le produit du travail de recherche est considéré comme du savoir et doit être publié pour pouvoir être revendiqué, il s'agit de science. Le chercheur voudra être le premier à publier les résultats de sa recherche de manière à en asseoir la

⁶ https://www.oxfordreference.com/view/10.1093/acref/9780195392883.001.0001/m_en_us1284427?rskey=JSA76S&result=71061 [consulté le 06 mars 2020] [accès par abonnement].

paternité et assurer ses droits en matière de propriété intellectuelle : “[t]he more openly you publish [...] the more secure your claim that the property is exclusively yours” (Price 1975, pp. 122-123). Si, en revanche, le fruit du travail est une « chose » qui peut être achetée ou vendue, il s’agit de technologie. L’invention devra être protégée par un brevet qui lui permettra de rester, un certain temps, dans le domaine privé.

Après avoir rappelé le nécessaire apport constitué par la pratique des citations de publications scientifiques antérieures, Price (1975) évoque la dense structure ressortant de l’analyse des modèles de citation :

“Scientific papers are assembled by a process rather like knitting or the way in which pieces of a jigsaw puzzle are held together by inter locking with their neighbors. Each scientific paper seems to build onto about a dozen previous papers.”
(Price 1975, p. 125)

Précisant que le monde de l’érudition fonctionne ainsi, tel un puzzle, Price (1975) estime que le domaine scientifique connaît un taux de croissance plus rapide du fait du rythme soutenu des publications. “Science has a trick of being eternally very young and new” (Price 1975, p. 126). Dans les années septante, Price (1975) relate déjà l’accélération de la vitesse du temps qui impacte directement les habitudes des chercheurs. Si leurs prédécesseurs avaient pour habitude de se référer aux livres pour accroître leurs connaissances, les chercheurs se sont progressivement tournés vers la lecture d’articles. Le temps gagnant encore en vitesse, Price (1975) parle de ces désormais réseaux de pairs partageant leurs connaissances lors de réunions plus ou moins formelles et néanmoins exclusives, constituant autant de micro-sociétés, qu’il nomme les collègues invisibles⁷. Pour Price (1975), si ces derniers peuvent également être considérés comme l’une des conséquences de l’augmentation exponentielle du volume de la littérature scientifique, le développement du travail de recherche en équipe les nourrit aussi, permettant des avancées notoires dans les domaines de recherche concernés. Price (1975) estime que l’appartenance à un groupe ainsi que les échanges permis lors de réunions parfois informelles sont bénéfiques aux avancées scientifiques

⁷ Originellement parlant, le « Collège invisible » est la réunion d’érudits dont Robert Boyle, Benjamin Worsley et Lady Ranelagh, qui a probablement engendré la *Royal Society*. Webster (1974) relate trois références directes à l’expression « Collège invisible » dans des correspondances de Robert Boyle. Toujours selon Webster (1974), le lien entre le « Collège invisible » et la création de la *Royal Society* aurait été mis par écrit par Thomas Birch, l’éditeur de Robert Boyle, lorsqu’il rédigea les lignes suivantes : “The Invisible College . . . probably refers to that assembly of learned and curious gentlemen, who [...] applied themselves to experimental inquiries, and the study of nature, which was then called the new philosophy, and at length gave birth to the Royal Society.” (Webster 1974, p. 21). Le « Collège invisible » fut actif entre 1645 - ou 1646 et 1660, année de la fondation de la *Royal Society*. Cette relation « originelle » entre le « Collège invisible » et la *Royal Society* est également reprise par Price (1975).

sans manquer de dénoncer leur nature par trop sibylline et le manque de transparence quant aux processus de recrutement et d'attribution de fonds au sein de ces collègues.

“Are these “power groups” dangerously exclusive ? Probably not, but in many ways they may turn out to be not wholly pleasant necessities of the scientific life in its new state of saturation.”
(Price 1975, p. 168)

Price (1975) considère qu'il n'y a pas d'information qui devrait être qualifiée d'inutile et que celle-ci devrait pouvoir être à portée de main : le fruit de ce qu'il nomme des « combinaisons improbables » entre science et technologie permettent en effet d'obtenir des résultats que d'autres ne sauraient trouver (Price 1975, p. 128). À propos de « l'accumulation des contributions » scientifiques où « chaque chercheur ajoute ses briques selon une séquence ordonnée », il observe un « édifice intellectuel construit par l'habileté et l'artifice » lequel « repose sur des fondations primitives » et « s'étend jusqu'aux limites supérieures du front de recherche de la connaissance en pleine croissance » (Price 1975, p. 162). Price (1975) voit dans les revues scientifiques l'une des innovations les plus distinctes de la révolution scientifique et ne manque pas d'évoquer leur rapide croissance en termes de nombre (cf. Annexe 2).

Avec la croissance exponentielle du nombre de revues scientifiques, les chercheurs n'ont plus la possibilité de lire chacun des articles qui pourrait les intéresser, ni même, comme le précise Price (1975), d'être au courant de chacun des travaux pouvant être pertinents pour leur domaine d'étude. Un nouveau dispositif est alors mis en place : les *abstracts*, qui à leur tour deviennent trop nombreux pour pouvoir être exhaustivement lus :

“This is, of course, the reason why during the last decade scientists have been concerned about the need for abstracts of abstracts, calling this an “information problem” which seems to require some process of electronic sorting of abstracts as a means of coping with the rising flood of literature.”
(Price 1975, p. 167)

Observant que le nombre d'auteurs par publication augmente et que la circulation des prépublications peut empiéter sur la publication papier, Price (1975) estime que l'avènement de l'informatique peut apporter des solutions en termes de sélection et de tri des articles. La croissance exponentielle du nombre d'articles scientifiques va de pair avec la croissance du nombre de revues scientifiques et il serait, selon Price (1975), plus intéressant de disposer d'une liste des articles plus complète et sélective que toute liste des revues, du fait qu'elle comprendrait les articles qui sauraient être qualifiés de scientifiques et pourraient être occultés dans du matériel non scientifique. À ce stade de son raisonnement, Price (1975) en arrive à la relation entre le nombre d'articles, le nombre et la qualité des scientifiques et les recherches qu'ils produisent. La relation

entre un scientifique et ses articles est aisée (un chercheur a écrit 1 article, 2 articles...). Citant la loi de Lotka sur la productivité (1926) (cf. chapitre 4.3.1), il estime que “[t]he distribution of productivity among scientists has not changed much over the whole three hundred years for which papers have been produced” (Price 1975, p. 175).

Price (1975) considère que la pratique de la science, dès lors de la recherche, se développe elle aussi de manière exponentielle. Précision est toutefois apportée : la « loi de croissance « normale » » ne l’est pas en ce sens que rien ne peut croître à l’infini : la croissance exponentielle finit par atteindre une limite et voit son élan ralentir, raison pour laquelle il faut parler de courbe logistique (Price 1986, p. 18) . Price (1975) explique la notion de cycle de croissance avec ses périodes de stagnation et de déflation, cycle qui permet d’expliquer la nécessaire évaluation des chercheurs :

“[...] the good scientist will be increasingly in demand and in power, since it must become ever more apparent that it is he who holds the purse strings of civilization in the era we have entered.” (Price 1975, p. 191)

Price (1975) relève l’intérêt croissant pour cette « science de la science », l’augmentation du nombre d’études quantitatives basées sur les revues, articles, citations, auteurs et une première réunion du collège invisible de ces personnes qui les étudient. Souhaitant que la scientométrie devienne un domaine scientifique et de recherche à part entière, Price (1975) souligne l’importance du profil transdisciplinaire de ses chercheurs qui devraient bénéficier d’une formation en sciences, humanités, histoire et philosophie de la science pour pouvoir répondre à la complexité du domaine scientifique.

3.2 De la crédibilité à la légitimité scientifique

L’Académie française (2019) rappelle que la crédibilité caractérise « ce qui mérite d’être cru ; ce qui rend [...] digne d’être cru[e] »⁸, et que la légitimité qualifie non seulement « ce qui est fondé en droit [...] ou conforme à l’équité, à la raison, aux règles établies, à la tradition », mais se dit également pour énoncer « [...] l’adhésion [...] dont peut se réclamer [un homme] et qui fonde [...] son autorité »⁹. Elles s’expriment par la reconnaissance de la communauté, « seul bien dont peut jouir individuellement le scientifique [...] » (Vinck 1995, cité dans Pignard 1999, p. 11). Selon Bourdieu (1975, p. 95 et 1976, p. 91) cette « espèce particulière de capital social » qu’est l’autorité scientifique est définie :

« [...] comme capacité technique et comme pouvoir social, ou si l’on préfère, le monopole de la compétence scientifique, entendue au sens de la capacité de

⁸ <https://www.dictionnaire-academie.fr/article/A9C4860> [consulté le 06 mai 2020].

⁹ <https://www.dictionnaire-academie.fr/article/A9L0538> [consulté le 06 mai 2020].

parler et d'agir légitimement (c'est-à-dire de manière autorisée et avec autorité) en matière de science, ce qui est socialement reconnu à un agent déterminé. »
(Bourdieu 1975, p. 92)

La réputation scientifique intéresse la sociologie des sciences¹⁰ qui l'assimile « [...] à une sorte de dot symbolique dont on peut mesurer l'importance relative au moyen d'indicateurs quantifiables », à l'instar des publications et du nombre de citations des publications d'un chercheur (Ragouet 2000, p. 319) (cf. chapitre 4.4.2). Dans la culture POP, le mérite est mesuré par la quantité et la qualité des publications (Moosa 2018) (cf. chapitre 3.3.5). Les publications scientifiques constituent donc un fondement important de l'octroi de la reconnaissance scientifiques. Elles participent activement à l'établissement de la réputation du chercheur. Basant son modèle sur la notion de capital, Bourdieu estime que :

« [...] les connaissances sont des ressources que le scientifique échange sur une sorte de marché contre du crédit scientifique qu'il peut ensuite réinvestir pour produire de nouvelles connaissances et gagner encore plus de crédit. »
(Bourdieu 1976, cité dans Pignard, p. 11)

Cette accumulation de capital permet au scientifique « de se faire un nom », un nom propre » (Bourdieu 1975, p. 99). La reconnaissance peut être en lien direct avec le degré de spécialisation du chercheur (Ziman 1987, cité dans Ragouet 2000). Elle peut également être considérée comme « une dimension structurante des champs scientifiques » (Ragouet 2000, p. 338), champs qui peuvent être caractérisés par les degrés d'interdépendances fonctionnelles et stratégiques entre le chercheur et « l'incertitude de la tâche scientifique »¹¹ (Whitley 1984, cité dans Ragouet 2000, p. 337). La relation d'interdépendance stratégique, laquelle a trait à la nécessaire persuasion des pairs quand « la pertinence et [...] importance » des sujets traités, est « très liée à l'impératif de la reconnaissance » (Ragouet 2000, pp. 337-338). La relation d'interdépendance fonctionnelle appelle quant à elle les chercheurs à exploiter les pratiques ou résultats de leurs confrères afin de « construire des énoncés qui seront reconnus comme pertinents et utiles » et « n'est pas complètement déconnectée des considérations sur leur notoriété » (Ragouet 2000, pp. 337-338). Dans cette course à

¹⁰ Bourdieu (1976) reconnaît toutefois que la tâche n'est pas aisée pour les sociologues de la science : ils sont eux aussi membres du champ scientifique. Ils doivent donc, s'ils souhaitent être objectifs, « prendre pour enjeu les [...] stratégies de [leurs] adversaires scientifiques [...] et le jeu en tant que tel, qui commande aussi ses propres stratégies, menaçant de gouverner souterrainement sa sociologie, et sa sociologie de la sociologie » (Bourdieu 1976, p. 104).

¹¹ Cette dernière est définie comme ayant « trait à la tâche elle-même, à l'ambiguïté des résultats et des interprétations possibles des phénomènes, à la pluralité des méthodes, à l'instabilité des phénomènes posés et aux buts de l'organisation, aux priorités de la discipline en termes de pertinence des objets » (Legault 1988, cité dans Ragouet 2000, p. 337).

l'« excellence », « créer une certaine uniformité de pratique [...] favoriserait l'appréciation du niveau atteint » (Couzinet 2000, p. 253).

Si le « groupe » attribuant la reconnaissance sociale au chercheur « tend toujours davantage à se réduire à l'ensemble des savants [...] à mesure que s'accroît [...] l'autonomie du champ » (Bourdieu 1975, p. 92), à propos de cette « lutte pour l'autorité scientifique », « seuls des savants engagés dans le même jeu ont les moyens de s'approprier symboliquement l'œuvre scientifique et d'en évaluer les mérites » (Bourdieu 1976, p. 91). Cette « lutte » est tant scientifique que politique du fait « qu'elle oppose des producteurs qui tendent à n'avoir d'autres clients que leurs concurrents » (Bourdieu 1975, p. 95). Whitley (2000, cité dans Krishnan 2009) perçoit les disciplines académiques comme unités de la production et de la validation intellectuelle (cf. chapitre 4.2.1).

Merton (1968) précise que si les récompenses dans le domaine scientifique sont avant tout basées sur la reconnaissance, cette dernière est échelonnée en fonction du niveau d'accomplissement sur la base de l'appréciation des pairs. La consécration par des tiers est indissociable de la reconnaissance :

“Both the self-image and the public image of scientists are largely shaped by the communally validating testimony of significant others that they have variously lived up to the exacting institutional requirements of their role.” (Merton 1968, p. 56)

Si la qualité prime sur la quantité en termes de reconnaissance (Cole et Cole 1967, cités dans Merton 1968), il existerait toutefois une « norme de désintéressement » qui pousserait les scientifiques à nier toute appétence de l'influence ainsi que de l'autorité (Storer 1966, cité dans Merton 1968). Le système de récompense ancré sur la renommée scientifique nécessite cependant un effort constant de la part du chercheur en vue de démontrer la pérennité de son potentiel, dès lors d'asseoir sa notoriété rappelle Merton (1968). Bourdieu (1975) considère que l'appréciation des capacités des scientifiques est fortement influencée par la position hiérarchique du chercheur. La reconnaissance scientifique, envisagée comme un processus cumulatif, peut aboutir à ce que Merton (1968) désigne comme le « *Matthew effect* du système de récompenses »¹² ou principe des avantages cumulatifs, soit :

“[...] the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark.”
(Merton 1968, p. 58)

¹² Merton (1968) constate que le “*Matthew effect*” (Merton 1968) peut également être observé dans l'allocation de ressources ou encore dans la communication scientifique (cf. chapitre 4.4.2.1).

Merton (1968) souligne toutefois le possible effet rétroactif de la renommée lorsque les précédentes publications d'un chercheur sont reconsidérées au moment où celui-ci publie un travail reconnu comme significatif. Utilisant le nombre de publications comme mesure de la productivité scientifique, Crane (1965, cité dans Merton 1968) observe que les scientifiques très productifs émanant de grandes universités obtiennent plus régulièrement la reconnaissance que des scientifiques tout aussi productifs dans des universités moins réputées. La relation entre fertilité scientifique et renommée est, selon Ragouet (2000), confirmée par plusieurs sociologues des sciences. Merton (1968) attire l'attention sur ce qu'il nomme le « phénomène du « 41^e fauteuil » » entendant par-là que s'il n'y a pas plus de quarante académiciens, ce n'est pour autant qu'ils ne sont des personnalités qui mériteraient d'intégrer les rangs des immortels¹³. Aussi déclare-t-il que cet « artefact du nombre fixe » peut être contrebalancé par la remise d'autres distinctions, sans que celles-ci ne véhiculent pour autant le même prestige (Merton 1968, p. 57).

Ragouet (2000) étudie quant à lui la réputation du chercheur à l'aune de ce qu'il nomme l'« arène » de consécration, soit cet « espace au sein duquel se nouent les relations d'échange, de coopération et de compétition pour l'obtention de trophées » (Ragouet 2000, p. 329). Distinguant les arènes académiques, politiques et médiatiques lesquelles ont leurs propres *desiderata* et représentent autant d'espaces i. « d'interaction » où se nouent les relations qui sont au cœur de la réputation ; ii. « de socialisation professionnelle et de contrôle du travail » où l'évaluation se base tant sur la pratique professionnelle que sur les représentations qu'en ont les chercheurs et enfin iii. « de rétribution », qu'elles soient matérielles ou symboliques (Ragouet 2000, pp. 329-330). Si la réputation ne vient pas uniquement des pairs, soit de « l'arène académique », et sans pour autant remettre en doute leur autorité en la matière, la progression de la réputation est régie par « une logique d'accumulation », laquelle est complétée par ce « processus social de diversification des sources de la reconnaissance » (Ragouet 2000, p. 328). Aussi, Ragouet (2000) réfute-t-il le principe des avantages cumulatifs estimant que la reconnaissance scientifique doit être considérée comme un processus i. « séquentiel », du fait que la notoriété n'est ni immédiate, ni garantie ; ii. « arborisé » du fait qu'il faut considérer la notoriété à l'aune des différentes sphères de renommées (les « arènes ») et enfin iii. « réversible » car elle ne peut jamais être considérée comme définitivement acquise ; ce qui l'amène à parler de « carrière réputationnelle » (Ragouet

¹³ Si dans toute organisation il y a des occupants de ce 41^e fauteuil, Merton (1968) considère que l'Histoire s'occupe de renverser des jugements qui auraient été altérés par la « myopie de la contemporanéité » (Merton 1968, p. 57).

2000, p.331). Cette réversibilité va à l'encontre du "ratchet effect" de Duesenberry (1964, cité dans Merton 1968) lequel pourrait, selon Merton (1968), s'appliquer au système de récompenses scientifiques : si un scientifique peut connaître une dépréciation relative de sa renommée au moment où il se fait devancer par quelque confrère, son niveau de reconnaissance acquise ne s'atrophiera pas pour autant.

« C'est bien parce que la définition de l'enjeu de la lutte est un enjeu de lutte [...] que l'on se heurte sans cesse aux antinomies de la légitimité » (Bourdieu 1976, p. 92). Relevons la diversité des termes utilisés pour signifier cette reconnaissance du statut du chercheur gagnant progressivement la reconnaissance de ses pairs, parfois également de la société : crédit, crédibilité, reconnaissance, réputation, médiatisation... Autant de termes signalant au scientifique qu'il assied sa position. Si la reconnaissance des pairs confère un statut au chercheur, l'absence de distinction ne signifie pour autant qu'il ne dispose pas de la légitimité scientifique. Nous avons relevé un certain nombre de faits pouvant avoir une incidence sur le nombre de citations des publications. Ce sont ces éléments qu'il faut retenir pour notre réflexion.

3.3 Historique de l'évaluation des publications scientifiques

1665 est l'année de publication des deux premières revues savantes : le *Journal des Sçavans* et les *Philosophical Transactions of the Royal Society* (Gingras 2014a et Rousseau, Egghe et Guns 2018). Le second conçoit déjà la relecture des contenus par des membres de la société (Hall 2002, cité dans Gingras 2014a).

"[...] the motivating purpose of [those two scholarly journals] was not the publishing of new scientific papers so much as the monitoring and digesting of the learned publications and letters that now were too much for one man to cope with in his daily reading and correspondence."
(Price 1986, p. 13)

Les revues se spécialisent au XIX^e siècle. L'évaluation par les pairs, mode d'évaluation des chercheurs par excellence, se pratique de manière relativement informelle jusqu'au tournant du XX^e siècle (Gingras 2014a). Aux prémices de l'évaluation des chercheurs et de leurs publications se trouvent trois phénomènes distincts. Dans un premier temps, avec le développement exponentiel de la science, la gestion des collections de revues scientifiques devient, dès les années 1920, impérative pour les bibliothèques (Gingras 2014a). L'avènement de l'informatique permet enfin le développement de l'indexation à grande échelle. Dans le courant des années 1970, en plein choc pétrolier, les États pensent et édictent les premières politiques scientifiques. L'élaboration de mesures d'évaluation de la recherche et du développement paraît évidente.

3.3.1 De l'évaluation des revues scientifiques au concept de « demi-vie »

Le nombre de revues scientifiques ne cesse d'augmenter. Les bibliothèques, ayant un budget à respecter, doivent procéder à une sélection.

« C'est dans ce contexte de gestion de collection de revues que l'analyse systématique non plus des publications elles-mêmes, mais des références qu'elles contiennent (les citations), est apparue. » (Gingras 2014a, p. 16)

Elle s'opère non sur « la base [dite *ndlr*] subjective de l'intérêt de certains chercheurs » (Gingras 2014a, p. 17), mais sur la base de l'analyse, par domaine, des revues les plus citées sur une période de temps donnée (Gingras 2014a). Une année après la formulation de la loi de Lotka (1926) (cf. chapitre 4.3.1), paraît dans *Science* une étude menée par Gross et Gross, tous deux chimistes, portant sur l'adéquation des revues présentes dans les bibliothèques universitaires spécialisées dans ce domaine. Se basant sur le volume le plus récent de *The Journal of American Chemical Society* (1926) qui regroupe toutes les références du domaine et considéré comme le plus représentatif de la chimie dans le pays, Gross et Gross (1927) comptent le nombre de références, par revue et par intervalles de quatre années pour la période 1871 à 1925. Ayant relevé de plus de 3'600 citations dans 247 revues, ils effectuent un classement présentant le nombre de revues en fonction du nombre de citations. Il ressort de leur analyse que 60,55% d'entre elles auront été citées moins de trois fois et que seuls 10,1% de ces revues auront été citées entre dix et vingt fois. Gross et Gross (1927) relèvent en outre le fait que le classement, par nombre de citations des revues, de la période 1871 à 1925 ne correspond pas au même classement des années 1916 à 1925. L'emploi des revues scientifiques évolue donc dans le temps. La fréquence de citation de nombreuses revues scientifiques peut aujourd'hui être observée dans le *Journal Citation Report* (JCR) dont Rousseau, Egghe et Guns (2018) rappellent que la première édition est publiée en 1976 puis disponible en ligne à compter de 1997.

Le concept de « demi-vie » a été élaboré pour formaliser le principe de l'évolution de l'utilisation dans le temps. Apparu dans les années 1960, il s'applique tant à la littérature scientifique qu'aux nombre de citations d'un article scientifique. Concernant la première, il permet de définir « la durée de vie utile de la littérature dans un domaine donné » (Gingras 2014a, p. 17). Les chercheurs en histoire, par rapport à leurs confrères évoluant dans le domaine de la chimie, citent régulièrement des articles publiés il y a plus de vingt ans exemplifie Gingras (2014a). Lorsqu'il est appliqué aux citations d'une publication, le concept de « demi-vie » donne « le nombre d'années qui doivent s'écouler avant que soit atteinte la moitié du nombre total de citations reçues » (Gingras 2014a, p. 18).

Gross et Gross (1927) se doutaient-ils que leur raisonnement allait avoir une telle résonance sur la structure et la pratique de la communication scientifique ? L'analyse des citations est désormais omniprésente dans l'évaluation de la pratique scientifique, non seulement au niveau des revues, mais également des chercheurs et de leurs publications (Bergstrom 2007) (cf. chapitres 4.4.2.2, 4.4.2.3 et 4.4.4).

3.3.2 De l'indexation aux bases de données informatisées

S'il n'est désormais plus possible pour le chercheur de lire l'intégralité des publications scientifiques dans son domaine, le recours aux *abstracts* ne lui permet pas pour autant de rester informé de tous les résultats de la recherche. Un système d'indexation devient impératif. Eugene Garfield, le père fondateur de l'*Institute for Scientific Information* (ISI) en 1959, présente un projet de base de données informatisée sur la base des informations qu'il a obtenues quant au fonctionnement de l'*Index Shepard* qui est appliqué au domaine juridique (Gingras 2014a). Il s'agit non seulement de « faciliter la recherche bibliographique », mais également de pouvoir « construire rapidement une bibliographie pertinente » (Gingras 2014a, p. 19). En outre, Gingras (2014a) relève un autre avantage d'une telle base de données à savoir permettre au chercheur de découvrir lesquels de ses confrères s'intéressent au même domaine que lui quand bien même ils ne publieraient pas dans les mêmes revues. En 1961, six ans après la présentation de son projet dans *Science*, Garfield obtient le soutien financier nécessaire pour étudier la faisabilité d'un tel index de la part des *National Science Foundation* (NSF) et *National Institute of Health* (NIH) (Gingras 2014a). Devenu opérationnel, le prototype est dévoilé puis administré par l'ISI en 1963 sous le nom de *Scientific Citation Index* (SCI). Garfield (1972, cité dans Organisation de coopération et de développement économique (OCDE) 1997) considère qu'il n'est pas envisageable d'intégrer de manière exhaustive toutes les revues existantes, du fait de l'absence d'une définition de ce qu'est une revue. Les revues indexées sont sélectionnées en fonction du taux de citation qu'elles démontrent (Garfield 1979[b], cité dans Rostaing 1996). Ce désormais nouvel outil « de recherche bibliographique » n'est alors encore en « lien direct avec l'évaluation de la recherche, thème qui n'est pas encore à l'ordre du jour » (Gingras 2014a, p. 20). Garfield (1968, cité dans OCDE 1997) y voit initialement le moyen de permettre aux chercheurs de trouver les articles pertinents. Un an plus tard, Garfield (1969) rappelle l'intérêt de la communauté scientifique dans l'indexation des citations en matière de recherche et de diffusion. Par la suite, Garfield (1979) reconnaît que la croissance ainsi que la complexification de la science nécessitent l'identification ainsi que l'évaluation de ses contributeurs les plus fructueux. Cet argument conforte l'avis de Durand-Barthez (2009) pour lequel le SCI sert déjà l'évaluation des auteurs et de ceux qui y voient une

« [...] nouvelle technique d'évaluation de l'activité scientifique, fondée sur l'étude des citations que se distribuent les auteurs » (Rostaing 1996, p. 14).

Dès 1973, l'ISI publie le *Social Science Citation Index* (SSCI), suivi du *Arts and Humanities Citation Index* (A&HCI) en 1978 (Rousseau, Egghe et Guns 2018). Disponibles exclusivement sous forme de publications papiers jusqu'aux années 1980, puis également sur CD-ROM, les données électroniques doivent être demandées à l'ISI (Gingras 2014a). C'est dans ces mêmes années septante que la NSF demande à Francis Narin et à sa firme Computer Horizons de traiter les données contenues dans ces bases de données en vue de produire des indicateurs :

« [...] normalisation des noms des pays, vérification minutieuse des orthographes des auteurs, traitement des articles publiés par plusieurs auteurs, sélection des types de documents retenus, classements de ces documents par spécialité, discipline ou domaine. »
(Rostaing 1997, p. 16)

En 1992, l'ISI est acheté par Thomson Scientific & Healthcare¹⁴ et prend le nom de *Thomson ISI* (Rousseau, Egghe et Guns 2018). Les SCI, SSCI et A&HCI, initialement regroupées dans le *Web of Knowledge* Thomson Reuters, font désormais parties de la *Web of Science* (WoS) *Core Collection* (Rousseau, Egghe et Guns 2018). La consultation de ces sources spécifiques de données bibliométriques est onéreuse (Gingras 2014b). Leur analyse est encore le fait d'experts en analyses (Hicks et al. 2015). Le WoS devient plus largement accessible dès 2002, année où Thomson Reuters lance une plateforme internet intégrée (Hicks et al. 2015)¹⁵. En 2004, Elsevier annonce le lancement de sa propre base de données scientifiques, Scopus, année même du lancement de la version *beta* du moteur de recherche Google Scholar (Hicks et al. 2015). Des outils permettant de mesurer la productivité et l'impact de la recherche sur la base des données (InCites pour le WoS et SciVal pour Scopus) et des logiciels destinés à mesurer les profils de citation (Publish or Perish¹⁶) sont développés (Hicks et al. 2015).

3.3.3 De l'essor parallèle de la bibliométrie et des politiques scientifiques

La bibliométrie a été plusieurs décennies durant le fait d'une communauté relativement restreinte de chercheurs précise Gingras (2014a), qui ne manque pas de rappeler que l'intérêt croissant des États pour le domaine scientifique au tournant des années 1970

¹⁴ Désormais intégré à *Thomson Reuters Corporation*.

¹⁵ En 2016, Thomson Reuters vend ses activités de propriété intellectuelle et de science à la Onex Corporation et Baring Private Equity Asia (Rousseau, Egghe et Guns 2018) laquelle créé Clarivate Analytics qui détient et gère désormais le WoS ou encore le JCR.

¹⁶ <https://harzing.com/resources/publish-or-perish> [consulté le 15 avril 2020]. Le logiciel permet aujourd'hui d'interroger les bases de données ou moteurs de recherche que sont Crossref, Google Scholar, Google Scholar Profile, Microsoft Academic, Scopus ainsi que le WoS.

va lui permettre de devenir un domaine de recherche reconnu en tant que tel. Les réflexions des politiques quant aux moyens permettant de stimuler la science nécessitent l'élaboration d'outils permettant d'évaluer la recherche et l'innovation. Ces derniers peuvent être conçus aux niveaux nationaux¹⁷ ou internationaux¹⁸ précise Gingras (2014a). C'est dans ce contexte que « la NSF demande à Francis Narin et à sa firme Computer Horizons d'étudier la possibilité d'utiliser la bibliométrie à des fins évaluatives » (Gingras 2014a, p. 22). Il en résulte un conséquent ouvrage publié en 1975 sous le titre de "Evaluative bibliometrics : the use of publication and citation analysis in the evaluation of scientific activity", une année après une conférence qui réunit les pères de la bibliométrie¹⁹ autour du thème de la métrique et des indicateurs de la science (Gingras 2014a). La désormais notoriété du domaine de la bibliométrie et plus globalement de la scientométrie s'exprime également, selon Gingras (2014a), par la naissance de la revue *Scientometrics* dont le premier numéro paraît en 1978 ou encore avec la création de l'*International Society for Scientometrics and Infometrics* (ISSI) en 1993.

3.3.4 De la bibliométrie comme outil d'évaluation du chercheur

Lotka (1926) (cf. chapitre 4.3.1) est souvent cité comme étant le pionnier de la bibliométrie. Mais Gingras (2014a) précise que des chercheurs venant des domaines de la psychologie et de l'anatomie comparée ont déjà eu recours à l'analyse des publications au début du XX^e siècle ; la différence notoire avec les travaux de Lotka étant que ceux-ci y avaient recours pour retracer, sinon suivre, l'évolution de la croissance de leur discipline. Si les études bibliométriques se développent au tournant des années 1920-1930 et confirment progressivement l'utilité des mesures quantitatives, il faut attendre les années 1980 pour appliquer des outils bibliométriques à l'évaluation des chercheurs (Gingras 2014a).

À compter des années 1990, alors que la subjectivité de la revue par les pairs (*peer-review*) est discutée, d'autres « méthodes de gestion fondées sur la construction d'indicateurs » telles que le "knowledge management"²⁰ ou le "benchmarking" voient dans la bibliométrie l'opportunité de mettre en place une évaluation considérée comme

¹⁷ À l'instar du *Science Indicators* (dès 1972) et futur *Science and Engineering Indicators* (à partir de 1987) aux États-Unis.

¹⁸ L'Organisation de coopération et de développement économique (OCDE) publiera la première édition du *Manuel Frascati* en 1962.

¹⁹ À l'instar de Derek J. de Solla Price, Eugene Garfield, Henry Small, Robert K. Morton...

²⁰ Dans une perspective de knowledge management, la gestion des disciplines académiques présentes dans une institution académique permet d'optimiser son organisation et d'asseoir sa réputation (Krishnan 2009).

plus objective (Gingras 2014a, p. 25). Les États y voient des mesures objectives des performances et du retour sur investissement (Gingras 2014b). La corrélation entre les développements scientifique et économique d'un État est du reste confirmée par le lien direct établi entre le nombre de publications et le produit intérieur brut (Hart et Sommerfeld 1998, cités dans Gingras 2014a).

À l'aube du XXI^e siècle, la majorité des travaux bibliométriques se mène au niveau d'un laboratoire ou d'une université en raison de la méfiance suscitée par le recours à des mesures bibliométriques pour l'évaluation individuelle :

« laquelle semble surtout le fait de gestionnaires de la recherche et des chercheurs qui, depuis que l'accès aux bases de données est devenu très facile [...], se sont mis à faire [...] ce qu'il faut bien appeler de la bibliométrie sauvage, sans aucune rigueur méthodologique [...]. » (Gingras 2014a, p. 25)

Pour Moosa (2018), la désormais inévitable mesure de la performance de la recherche a donné naissance puis nourri la culture POP (cf. chapitre 3.3.5). Nous analyserons puis discuterons du bon usage de la bibliométrie en matière d'évaluation des publications scientifiques dans le chapitre 6.

3.3.5 La culture du “Publish or Perish”

La publication est indissociable de la pratique et de la recherche scientifique. Elle permet, par exemple, d'asseoir la paternité d'une recherche, de rendre compte des fruits du travail scientifique, de les partager avec le monde scientifique et d'une manière plus globale avec la société, de renseigner quant à la manière dont sont utilisés les fonds attribués pour la recherche. Benjamin Franklin avait pour devise “to study, to finish, to publish” (Franklin [s.d.], cité par Sojka et Mayland 1993, cités dans Moosa 2018). Il s'agit de saisir les effets du POP sur la production scientifique, corrélativement ses répercussions en matière de bibliométrie.

Il n'existe pas de consensus quant à l'origine de l'expression de POP (Sojka et Mayland 1993 puis Garfield 1996 et Shapiro 1998, cités dans Moosa 2018). Après avoir été utilisée dans un cadre non universitaire dans les années 1930 (Coolidge et Lord 1932, cités dans Moosa 2018), elle est citée dans le contexte académique peu avant 1940 (Moosa 2018). Il faut entendre par POP cette « pression mise sur les universitaires pour publier rapidement et en continu dans des revues scientifiques », condition qui gouverne leur carrière scientifique : du recrutement aux promotions, parfois même à la conservation d'un poste (Moosa 2018, p. 1). La destinée des universitaires est donc, selon cette doctrine, exclusivement corrélée à une publication scientifique prospère et efficace (De Rond et Miller 2005, cités dans Moosa 2018). Les chercheurs ne s'y conformant pas verraient leur carrière entravée. Une étude menée en 2013 relève quatre

manifestations de la culture POP au sein des institutions académiques : i. Un nombre défini de publications en vue d'une possible avancée dans la carrière académique ii. Les articles revus par les pairs sont désormais considérés comme étant la principale contribution à la carrière scientifique en termes de statut, de potentiel commercial et de rémunération au mérite iii. Un taux de publication élevé augmente les opportunités d'avancement et d'obtention de fonds iv. Du fait qu'un poste de chercheur consiste à produire de la recherche, celle-ci n'est pas considérée comme utile si elle ne rentre pas dans le domaine public (Plume 2013, cité dans Moosa 2018, p. 2). Dans la culture POP, ce sont essentiellement les publications d'articles qui sont prises en considération, au détriment des autres formes de publication scientifique, prônant ainsi, selon Moosa (2018) certains champs disciplinaires (cf. chapitre 4.2.2.1).

La culture POP a eu pour conséquences une augmentation du nombre de publications scientifiques ; près de deux millions en 2012 (Sarewitz 2016). L'importance croissante accordée au nombre de publications des chercheurs a plusieurs effets, positifs comme négatifs ; elle permet d'identifier et distinguer les scientifiques sur la base non du favoritisme, mais du mérite (Van Dalen et Henckens 2012, cités dans Moosa 2018) ; elle a encouragé la "*strategy to leave scientific shadow*" (Moed et al. 1985, p. 189) observée dans les années 1970-1980 : une augmentation de la productivité du fait que le chercheur décide de ne pas simplement rédiger un rapport interne pour rendre compte des fruits de sa recherche, mais bien de les partager avec la communauté. La pression exercée sur les scientifiques ne va pas uniquement dans le sens d'une hyper-publication. Selon Becher et Trowler (2001, cités dans Krishnan 2009), elle aide également à rendre le travail des chercheurs plus pertinent pour la communauté scientifique et pour la société en général.

Un des revers de la médaille consiste en la "*multiplication strategy*" (Moed et al. 1985, p. 189) à savoir le fait de disséminer les fruits de sa recherche dans plusieurs publications au lieu d'une ; l'objectif est d'augmenter le nombre de publications reportées à l'administration de son institution académique (Moosa 2018)²¹. La culture POP peut directement porter atteinte à la qualité des résultats de la recherche. "Mainstream scientific leaders increasingly accept that large bodies of published research are unreliable" (Sarewitz 2016 §5). Moosa (2018) précise que l'évaluation porte sur la quantité des publications et la qualité des revues, non sur la reproductibilité des résultats. Baker (2016) relate une enquête menée par *Nature* auprès de plus de 1'500 chercheurs :

²¹ Moosa (2018) admet toutefois que la pratique ne peut être considérée négativement lorsque plusieurs groupes de recherche travaillant sur un projet commun produisent diverses publications lesquelles permettent de préciser son contexte global.

plus de 70% de ceux qui auraient tenté de reproduire leur expérimentation auraient échoué, reconnaissant par là-même une « crise de la reproductibilité ». Baker (2016) constate que la pression à la publication et la communication sélective seraient les principaux responsables de cette « crise » auxquels s'ajoutent, par exemple, de mauvaises méthodologies et gestions des données (cf. Annexe 3). Une fois publiés, des analyses et résultats erronés peuvent à nouveau être repris ou cités, se voyant ainsi largement diffusés. Sarewitz (2016) estime que cela s'observe dans divers domaines de recherche. La problématique est connue dans le domaine de la recherche contre cancer où certaines lignées de cellules cancéreuses utilisées pour la recherche s'avèrent être contaminées :

“The average biomedical research paper gets cited between 10 and 20 times in 5 years, and as many as one-third of all cell lines used in research are thought to be contaminated, so the arithmetic is easy enough to do : by one estimate, 10,000 published papers a year cite work based on contaminated cancer cell lines. Metastasis has spread to the cancer literature.”

(Sarewitz 2016 §6)

Certains estiment que la subordination de récompenses financières à la productivité en termes de publication pousserait certains chercheurs à l'inconduite (Fanelli, Costas et Larivière 2015, cités dans Moosa 2018).

Nous estimons que le développement de l'Open Research Data (ORD) pourrait être une réponse, du moins partielle, à cette « crise ». Price (1986) estimait que la croissance exponentielle de la science amènerait à un état de saturation puis à quelque chose de foncièrement différent, du fait que la saturation n'est pas forcément synonyme de mort (cf. chapitre 3.1). La « *Nouvelle science* » ou « *Saturation stable* » doit être planifiée en conséquence, sans quoi il faudra parler de « sénilité » (Price 1986, p. 29). Le POP tendrait aujourd'hui à nous rapprocher de cet état de « sénilité », du fait que la science « se [noie] dans le bruit de sa propre productivité croissante » Sarewitz (2016 §11). Pour contrer cela, Sarewitz (2016) plaide en faveur d'une publication plus sélective, dès lors plus qualitative. Pour Moosa (2018), la culture POP détourne les scientifiques de recherches qui pourraient être bénéfiques à la société du fait que pour subsister ils doivent suivre les règles du POP et s'y faire leur nom. Enfin, lorsque la culture POP domine, les activités des scientifiques qui ne sont pas directement liées à la recherche, à l'instar de l'enseignement et des services rendus à la communauté, ne rentrent pas dans le spectre du mérite (Moosa 2018). Si la promotion et la sensibilisation de l'éthique de la recherche peuvent en partie contrebalancer les effets pervers du POP, nous sommes d'avis que le cadre posé par l'OS renforcera l'intégrité de la recherche scientifique. La rédaction d'un Data Management Plan (DMP) devenant pratique

courante, les chercheurs doivent désormais penser à la gestion puis à la réutilisation de leurs données de recherche avant de les archiver et de les rendre accessibles (ORD). Les principes FAIR²² ne s'appliquent pas exclusivement aux sciences « dures ». En outre, l'OA permet d'augmenter le lectorat des publications scientifiques. L'OS a donc des implications concrètes sur tout le cycle de la recherche. Nous estimons que la qualité de la recherche y gagnera. Si nous n'avons pas encore assez de recul pour pouvoir l'affirmer, nous voyons dans l'OS une possible tendance post-POP. L'OS concerne aujourd'hui particulièrement la recherche financée par des fonds publics. Nous estimons toutefois que son esprit aura des répercussions sur les fondements méthodologiques et comportementaux des chercheurs et pourrait amener à un nouveau paradigme. Nous revenons sur l'OS et plus particulièrement sur la refonte souhaitée du système d'évaluation dans le chapitre 4.6.2.

²² FAIR : **F**indable, **A**ccessible, **I**nteroperable, **R**eusable.

4. État de l'art dans le domaine de la bibliométrie appliquée à la production scientifique

Dans le présent chapitre, nous expliquons les métriques pouvant s'appliquer à la production scientifique. Après avoir réfléchi au niveau, domaine ou discipline scientifique, auquel devraient s'appliquer les mesures bibliométriques, nous introduisons les principales distinctions en termes de publications et de citations en fonction des domaines de la recherche. Nous présentons les trois lois considérées comme les piliers de la bibliométrie et introduisons les indicateurs bibliométriques les plus courants. Nous avançons aussi les principales critiques formulées à leur rencontre, que ce soit sur le fonds ou sur la forme. Nous concluons le présent chapitre avec les nouvelles tendances en matière de mesure et d'évaluation de la production scientifique.

Wilsdon et al. (2017) précisent que l'objet des évaluations bibliométrique peut être une publication, un chercheur, un groupe de recherche, une institution ou encore un pays. Nous distinguons ce que nous nommons les études des évaluations bibliométriques. Les premières sont menées par des chercheurs, organismes de recherche ou institutions (trans-) étatiques aux fins du développement des techniques et outils bibliométriques ainsi que des connaissances du fonctionnement du monde scientifique. Elles sont le plus souvent réalisées au niveau macro. Les secondes sont menées par les institutions académiques aux fins de l'évaluation des chercheurs qui officient en leur sein, soit au niveau micro. Elles se basent généralement sur des données institutionnelles. Les publications entrant dans leur champ d'analyse sont définies par les hautes écoles. Certaines institutions académiques ont toutefois recours à des indicateurs bibliométriques issus de bases de données scientifiques internationales (cf. chapitre 4.2.4).

4.1 L'évaluation des publications scientifiques à l'aune des métriques

L'évaluation des publications des auteurs scientifiques est réalisée aux niveaux qualitatifs et quantitatifs. Si la mesure de la qualité du contenu s'opère différemment selon les disciplines (Bibliothèque de Sorbonne Université 2017a), l'examen par les pairs ou *peer-review* en est l'expression la plus commune pour certains types de publications (articles et actes de conférences en particulier) (cf. Annexe 4). Celui-ci se fait en amont de la publication, lorsque celle-ci est corrélée à cet examen. L'éthique veut que cette évaluation soit impartiale (All European Academies (ALLEA²³) 2018, Académies suisses

²³ L'ALLEA est la Fédération européenne des académies nationales des sciences et des humanités.

des sciences 2008). En outre, nombre de politiques d'évaluation de la recherche en appellent désormais à un examen éclairé des pairs (Colwell et al. 2012 puis KNAW 2010 et Phillips 2012, cités dans Wouters 2014), signifiant par là-même le rôle prépondérant qu'ils jouent en matière d'évaluation de la qualité (Wouters 2014), mais pas uniquement au moment de la soumission d'une publication. Le recours à des « preuves bibliométriques » permet de limiter les biais de l'évaluation par les pairs estime Wouters (2014). Comment combiner de l'information quantitative et qualitative ? Wouters (2014) estime que la réponse est laissée à la discrétion des pairs, quand bien même le concept d'évaluation éclairée est aujourd'hui insuffisamment développé. Du point de vue quantitatif, l'évaluation de ces publications se fait par le biais de métriques. Elle s'opère en aval de la publication.

La métrique désigne la science de la mesure appliquée à un domaine spécifique (Stevenson et Lindberg 2011²⁴). Estimant que le XX^e siècle peut être considéré comme le siècle des sciences métriques, Dutta (2014) rappelle l'origine latine du terme et le définit puis le situe comme :

“a measure for something or any mean of deriving quantitative measurement or approximation. [...] This word envelops number of subject domains. One of the most well-known uses of this word is found as unit of measurement in various subjects and LIS²⁵ also. Any system of measurement invariably involves one crux subject from methodological viewpoint, which is nothing but statistics.”

(Dutta 2014, [p. 2])

En matière d'évaluation des publications, plusieurs métriques doivent être mentionnées. Le terme d'infométrie est apparu à la fin des années 1970 (Dutta 2014 ; Rousseau, Egghe et Guns 2018). Englobant tant la scientométrie que la bibliométrie, le champ de l'infométrie est vaste, raison pour laquelle elle fut dans un premier temps définie comme : “ [...] the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists” (Tague-Sutcliffe 1992 puis Ingwersen et Björneborn 2004, cités dans Rousseau, Egghe et Guns 2018, p. 3). Du fait qu'il n'est plus nécessaire de rappeler son étendue, elle peut désormais être définie comme étant « l'étude des aspects quantitatifs de l'information sous toutes ses formes et pour tous les groupes sociaux » (Rousseau, Egghe et Guns 2018, p. 3).

Le terme « наукометрия » (« *naukometriya* » soit scientométrie) a été utilisé pour la première fois utilisé en 1969 par Nalimov et Mulchenko (Dutta 2014 ; Rousseau, Egghe et Guns 2018). La scientométrie ou « science de la science » (Price 1986, p. XX) :

²⁴ https://www.oxfordreference.com/view/10.1093/acref/9780195392883.001.0001/m_en_us1267545?rskey=dX218J&result=7 [consulté le 9 mars 2020] [accès par abonnement].

²⁵ LIS = *Library and Information Science* soit bibliothéconomie et sciences de l'information.

« [...] porte sur la mesure quantitative de l'ensemble des activités scientifiques, toutes disciplines confondues. Ses données incluent les montants investis pour la recherche-développement (R&D), la formation du personnel scientifique et la production d'articles et de brevets. »
(Gingras 2014a, p. 15)

Pour Sugimoto et Larivière (2018), la scientométrie regroupe les modes de collecte et d'analyse quantitative des données de la science. La production et le contenu des publications scientifiques en font sans conteste partie.

L'application de méthodes statistiques au domaine de la publication remonte à la fin du XIX^e siècle (Sengupta 1992, cité dans Dutta 2014). Le terme de bibliométrie avait déjà été utilisé par Paul Otlet dans son célèbre *Traité de documentation* (1934, cité dans Dutta 2014 et Rousseau, Egghe et Guns 2018) qui entendait par-là la mesure de tout aspect lié aux livres et documents. Avant l'avènement des bases de données informatiques, la bibliométrie consistait en une analyse manuelle limitée à certains domaines scientifiques (Gingras 2014a). Au vu de la croissance exponentielle de la recherche (Price 1975 et 1986), l'évaluation dans le domaine des sciences de l'information par le biais du recours aux mathématiques et statistiques apparaît comme de plus en plus évident et nécessaire. Il s'exprime sous divers termes à l'instar de "statistical bibliography" dans les années 1920 (Hulme 1923, cité dans Rousseau, Egghe et Guns 2018) ou encore de "librametrics" ou "librametry" dès la fin des années 1940 (Ranganathan 1948, cité dans Dutta 2014 et Rousseau, Egghe et Guns 2018). C'est vingt ans plus tard que le terme de "bibliometrics" est redéfini par Pritchard (1969, cité dans Garfield, Malin et Small 1983 ; Dutta 2014 et Rousseau, Egghe et Guns 2018) comme étant :

"the application of mathematics and statistical methods to books and other media of communication. It is the metrology of the information transfer process; its purpose is analysis and control of the process. In short: bibliometrics is the scientific study of recorded discourse."

(Pritchard 1969, cité dans Rousseau, Egghe et Guns 2018, p. 2)

Le suffixe « métrie » « renvoie aussi bien à la mesure qu'à la métrique » (Rostaing 1997, p. 22). Dans le présent chapitre, après être revenus sur « la modélisation des distributions des éléments bibliométriques » (Rostaing 1997, p. 23) (cf. chapitre 4.3), nous nous concentrons sur la bibliométrie en tant que mesure quantitative, soit le recours à des indicateurs permettant d'apprécier l'activité scientifique (cf. chapitre 4.4). Le « concept de métrique est plus spécifique aux indicateurs [dits] relationnels » (Rostaing 1997, p. 22) lesquels permettent d'apprécier « les liens et les interactions entre différents acteurs des systèmes des S[ciences]-T[echnologie] [...]» (OCDE 1997, p. 26). La bibliométrie en tant que mesure, « [...] sous-ensemble de la scientométrie [...] se limite à l'analyse des *publications* et de leurs propriétés » (Gingras 2014a, p. 16). Cette

« méthode de recherche [...] consiste à utiliser les publications scientifiques et leurs citations comme indicateurs de la production scientifique et de ses usages » (Gingras 2014a, p. 9). Nous pouvons en outre distinguer la bibliométrie descriptive de la bibliométrie évaluative :

“Descriptive [bibliometrics] concern geographic distribution or temporal distribution of productivity count [whereas] [e]valuative bibliometrics includes references or citations that are known as literature usage count.” (Dutta 2014, p. [4])

Là où Gingras (2014a) et Rousseau, Egghe et Guns (2018) précisent que les termes de scientométrie et bibliométrie sont devenus interchangeables, « tant l’usage des données bibliométriques dans l’étude de la dynamique des sciences est désormais routinier » (Gingras 2014a, p. 15), Yang et Yuan (2017) se réfèrent à la scientométrie, à la bibliométrie enfin à l’infométrie comme aux trois métriques (“*Three metrics*”), leurs noms étant désormais utilisés de manière interchangeable. Toutefois, si “[t]he three metrics refer to component fields related to the study of the dynamics of disciplines as reflected in the production of their literature” (Hood et Wilson 2001, cités dans Yang et Yuan 2017, p. 1508), “the terms differ in their discipline attribute” (Yang et Yuan 2017, p. 1508). Dans le cadre de la présente recherche, nous recourons au terme de bibliométrie pour désigner l’ensemble des mesures quantitatives permettant d’évaluer la publication scientifique d’un chercheur.

4.2 Des éléments à prendre en considération en matière d’évaluation bibliométrique

Ils sont des éléments dont il faut être conscient en termes d’études et d’évaluations bibliométriques. Dans le présent chapitre, nous cherchons à déterminer si celles-ci doivent être réalisées au niveau des disciplines ou des domaines scientifiques. Il s’agit également de saisir quelles sont les principales différences en termes d’habitudes de publication et de citation entre les champs disciplinaires. En outre, ils sont de possibles sources de biais préexistants. Après les avoir relevés, nous formulons un certain nombre de recommandations en vue d’asseoir la pertinence des études et évaluations bibliométriques.

4.2.1 L’évaluation bibliométrique : au niveau des disciplines ou des domaines scientifiques ?

Du fait de son développement exponentiel, la science est aujourd’hui subdivisée en divers domaines (Price 1975). Les domaines d’étude et de recherche académiques, également nommés champs scientifiques (Bourdieu 1976) ou disciplinaires, sont constitués de nombreuses disciplines. Ils peuvent être fortement corrélés aux modes et habitudes de publication des chercheurs (cf. chapitre 4.2.2). L’Académie française

(2019) définit la discipline comme : « [b]ranche de la connaissance, domaine d'activité, matière d'enseignement et d'étude »²⁶. Heilbron (2003) précise que :

“Discipline refers to a demarcated body of knowledge identified as a separate science[...] with its own methods, problems, and practitioners. [...] Members of different disciplines compete among each other for funding and academic turf. To exist disciplines require considerable political, economic, social, and cultural resources.”
(Heilbron 2003 §1²⁷)

Il s'agit de déterminer quelle est l'unité d'analyse la plus pertinente en matière de bibliométrie : la discipline ou le domaine disciplinaire. Il semble opportun de saisir dans un premier temps ce qu'est une discipline académique. Aristote considérait déjà une division de la connaissance, distinguant la pensée « pure » de l'observation de la nature (Thomson Klein 2005, cité dans Krishnan 2009). Price (1975) estime toutefois qu'avant le XVI^e siècle, alors que la polymathie prédominait, la science était considérée comme un tout. Thomson Klein (1990, cité dans Krishnan 2009) fait remonter au Moyen Âge tardif l'institution de disciplines scientifiques distinctes²⁸. Aujourd'hui, du fait de son développement, proportionnel au nombre de chercheurs et d'écrits scientifiques, la science est subdivisée en divers domaines (Price 1975), soit autant de champs d'études ou disciplines. Kuhn considère cette “succession of scientific revolutions that from time to time fundamentally reorganize scientific fields or disciplines” (Kuhn 1962, cité dans Krishnan 2009, p. 15). À compter du XIX^e siècle, la disciplinarisation permet de répondre à un besoin social : former les spécialistes qui accompagneront le développement de l'industrialisation et permettront l'avancée de la technologie (Moran 2002, cité dans Krishnan 2009). Le contour des disciplines n'étant pas figé, leur nombre évolue dans le temps²⁹. Les disciplines se regroupent autour de ce qui peut être perçu comme une communauté partageant des caractéristiques communes (Pignard 1999) ; Krishnan (2009) considérant les disciplines comme autant de communautés de la connaissance. Le partage d'une approche et d'une méthodologie par les chercheurs d'une discipline assoient sa crédibilité du fait de sa propre rigueur scientifique (Bridges 2005, cité dans Krishnan 2009). Les disciplines académiques sont souvent identifiées par les sujets enseignés au sein des institutions académiques, sans pour autant que chacun d'entre eux soit considéré comme une discipline à part entière, ni que chaque discipline soit

²⁶ <https://www.dictionnaire-academie.fr/article/A9D2636> [consulté le 06 mars 2020].

²⁷ <https://www.oxfordreference.com/view/10.1093/acref/9780195112290.001.0001/acref-9780195112290-e-0190?rskey=JnlHws&result=181> [consulté le 09 mars 2020] [accès par abonnement].

²⁸ N.B. Vauchez Dobson et Lapidge (2000) parlent d'une subdivision des disciplines au sein de facultés dès la fin du XIII^e siècle (cf. chapitre 2.4.1.2).

²⁹ Pour les néo-positivistes, le nombre de disciplines ne devrait pas connaître de forte variabilité du fait de leur vision d'une science unifiée et basée sur une même rationalité scientifique universelle et sur un langage scientifique universel (Friedman 1999, cité dans Krishnan 2009).

représentée par un département ou une faculté qui lui soit propre précise Krishnan (2009) : les disciplines émergentes peuvent dans un premier temps être considérées comme centres de recherches, le temps qu'elles développent leurs propres théorisations et méthodologies ; la reconnaissance d'une nouvelle discipline est généralement corrélée à la création d'une chaire professorale. Si un certain nombre de caractéristiques permettent de déterminer si un sujet peut être considéré comme une discipline³⁰, toutes ne comprennent pas pour autant chacune de ces caractéristiques précise Krishnan (2009). Les disciplines académiques les mieux implantées comprennent les principales caractéristiques d'autres professions à savoir que cette communauté de professionnels dispose d'une autonomie collégiale, d'un ensemble distinct de connaissances et compétences ou encore d'une éthique professionnelle qui leur est propre précise encore Krishnan (2009). Si certains définissent donc les disciplines selon leurs caractéristiques, d'autres, à l'instar de Biglan (1973, cité dans Krishnan 2009) les distinguent selon leurs dimensions « épistémologiques et culturelles » : les disciplines dites « dures ou paradigmatiques » versus les disciplines « molles ou pré-paradigmatiques » soit la séparation entre les sciences dites naturelles des SHS, les disciplines « pures ou principalement théoriques » par rapport aux disciplines « appliquées » qui permettent de distinguer les mathématiques de l'ingénierie ou encore les disciplines qui impliquent ou non des systèmes vivants à l'instar de la biologie puis de l'histoire (Biglan 1973, cité dans Krishnan 2009, p. 11).

Quand bien même le chercheur évolue dans une discipline spécifique, il peut travailler sur des recherches trans- ou inter-disciplinaires. L'interdisciplinarité est, dans sa définition la plus globale, considérée comme "any form of dialogue or interaction between two or more disciplines" (Moran 2001, cité dans Krishnan 2009 p. 6). Les chercheurs la percevraient comme un « moyen bienvenu d'hybridation des savoir » tandis que les administrateurs y verraient « un outil pour favoriser les échanges [...] et ainsi rentabiliser des disciplines toujours perçues comme trop coûteuses » (Boelaert et al. 2015, p. 20). Une nouvelle discipline peut naître d'un projet interdisciplinaire du moment qu'elle apporte des éléments et perspectives inédits (Krishnan 2009). La notion de progrès scientifique est indissociable de l'existence d'une discipline laquelle doit constamment développer de nouvelles matières et méthodes, perdant par-là même ses cohérence et

³⁰ Ces caractéristiques sont : un objet de recherche particulier – sachant que celui-ci peut être partagé avec d'autres disciplines ; un corpus de connaissances spécialisées se référant à l'objet d'étude ; des théories et concepts qui sont propres à la discipline ; le recours à une terminologie ou à un langage technique spécifiques adaptés à l'objet de recherche ; le développement de méthodes de recherche spécifiques ou encore sa manifestation institutionnelle, soit son déploiement au sein d'une institution académique et la connexion avec des associations professionnelles) (Krishnan 2009, p. 9).

démarcation, dès lors son adéquation selon Krishnan (2009). Aussi est-ce la méthodologie empruntée et non le domaine de recherche qui permettrait d'identifier les disciplines (Grieb 1974, cité dans Krishnan 2009). S'il est difficile de définir avec minutie ce qui constitue une discipline, relevons que certains sociologues des sciences estiment que cette notion ne serait pas pertinente, du fait que le réseau des chercheurs ne se bornent à leur champ disciplinaire et de la labilité des structures de ces réseaux (Ragouet 2000). Krishnan (2009), sans manquer de préciser que la division de la science en diverses disciplines implique que notre connaissance du monde est divisée en un nombre conséquent de branches différentes, mène une analyse comparative intéressante de la « discipline » sous les perspectives philosophiques, anthropologiques, sociologiques, historiques et managériales³¹ laquelle permet de rendre compte qu'il y a autant de visions et interprétations de ce qu'est une discipline et de son existence selon le paradigme d'approche et l'intellection de l'analyste (cf. Annexe 5). La problématique d'une définition unifiée réside dans le fait que :

“Disciplinary and interdisciplinary are intrinsically connected to the problem of the correspondence or non-correspondence of knowledge to an objective reality and the problem of the unity or disunity of all knowledge.” (Krishnan 2009, p. 13)

Krishnan (2009) conclut que si certains considèrent que l'existence des disciplines en tant qu'organisation, académique et des connaissances, est vouée à disparaître et que nous tendrions en direction d'une science post-disciplinaire³², il est fort à parier que l'organisation de la science en disciplines va persister : si celles-ci vont continuer à naître ou à disparaître et qu'elles doivent conserver leur identité distincte tout en étant plus ouvertes au travail de leurs consœurs, la solution résiderait en :

“[...] embracing interdisciplinary while keeping and nurturing disciplines as the ultimate reference points. Only such a mature and self-conscious science will be worthwhile pursuing and deserves future.” (Krishnan 2009, p.51)

Du fait de la labilité du terme de discipline, il semble indispensable d'élargir le rayon de l'analyse aux domaines scientifiques ou champs disciplinaires en vue

³¹ Krishnan (2009) intègre également à son analyse la vision de la discipline du point de vue de l'éducation, vision que nous n'intégrons dans notre analyse estimant, sans nullement remettre en question son intérêt, qu'elle ne nourrit cette dernière.

³² Si la disciplinarisation constituait une organisation effective au XIX^e siècle, certains la considère aujourd'hui inefficace et dépassée du fait de la duplication des efforts entre les départements académiques et du fait que cette organisation ne correspond plus ni à la réalité intellectuelle, ni aux actuelles tendances de production et de gestion de la connaissance (Krishnan 2009) ; un nouveau mode de production de la connaissance serait en cours de développement et s'articulerait autour de thèmes plus globaux à l'instar des études environnementales ou les études de genre lesquelles intègrent les aspects de plusieurs disciplines sans pour autant être considérées en tant que telles (Gibbons et al. 1994, cités dans Krishnan 2009).

de saisir les distinctions dans les modes de publications et de citations. Cela implique que les indicateurs bibliométriques doivent être adaptés non à la discipline, mais aux champs disciplinaires. Historiquement du reste, les indicateurs bibliométriques reposent sur un système de classification par champs disciplinaire (Braun et Glänzel 1990 et Moed et al. 1995 puis Waltman et al. 2011, cités dans Waltman et van Eck 2013). Il est nécessaire de tenir compte des distinctions inter-domaines en matière d'évaluation des publications, ce d'autant plus que "[v]alues of metrics differ by field and hence their interpretation must be adapted to the corresponding field or even subfield" (Smolinsky et Lercher 2012, cités dans Rousseau, Egghe et Guns 2018, p. 256). swissuniversities (2018a) confirme que les pratiques d'évaluation appliquées à certains domaines de recherche ne sont pas applicables à d'autres domaines du fait de ces différentes pratiques en termes de publications. Hirsch (2005) relève que les valeurs de référence du *h*-index (cf. chapitre 4.4.3) varient selon les champs disciplinaires du fait du nombre moyen de publications des chercheurs du domaine³³, du nombre de chercheurs dans le domaine, du nombre moyen de citations par publication (cf. chapitre 4.2.2.2). Il est de l'avis de Van Raan (2006) que les indicateurs bibliométriques standards tiennent compte de ces distinctions. Au vu de notre analyse, s'il est établi que les indicateurs bibliométriques s'appliquent au niveau du domaine et non des disciplines qui le constituent, nous rejoignons Waltman et van Eck (2013) qui rappellent qu'il ne faut pas perdre de vue le fait que les frontières entre les champs disciplinaires sont floues et qu'un même champ peut être défini selon différents niveaux de granularité, raisons pour lesquelles ils ne peuvent être considérés comme des entités homogènes.

4.2.2 Des habitudes de publication et de citation entre les domaines scientifiques

Ayant déterminé quelle est l'unité de mesure la plus utilisée en matière de bibliométrie, nous relevons ici les principales distinctions en termes d'habitudes de publication et de citation selon les domaines de la science. Nous ne pouvons pas procéder à une présentation détaillée et exhaustive de ces distinctions par domaine. Elles pourraient faire l'objet d'une étude spécifique. Nous relevons dès lors les principales distinctions

³³ "Scientists working in nonmainstream areas will not achieve the same very high h values as the top echelon of those working in highly topical areas". (Hirsch 2005, p. 16571).

dont il faut tenir compte en matière d'évaluation des publications et illustrent certaines de leurs applications selon les domaines scientifiques.

4.2.2.1 Les habitudes de publication selon les domaines scientifiques

Les habitudes de publication sont corrélées au champ disciplinaire dans lequel évolue le chercheur. Outre l'aspect méthodologique et les différentes conceptions de la notion de qualité (swissuniversities 2018a), les principales distinctions concernent le type de support servant la communication ainsi que la langue de publication. Aussi, en reprenant la distinction entre les sciences dites « dures » ou « molles » :

“[...] natural scientists would be more focused on producing journal articles [...] In contrast, [...] practitioners [of “soft” sciences] would be more focused on [...] publishing monographs [...].” (Biglan 1973, cité dans Krishnan 2009, p. 11)

Hicks et al. (2015) relèvent de même que les historiens et chercheurs du domaine des sciences sociales publient plus de monographies que d'articles tandis que les informaticiens revendiquent l'intégration des actes de conférences. Ingwersen et al. (2001) précisent en effet que les chercheurs évoluant dans le domaine de l'informatique et des technologies de l'information demandent à ce que les résultats de leurs recherches présentés lors des conférences soient intégrés dans l'évaluation de leurs publications. « En littérature et études culturelles, les livres dominent [...] [alors que] le recours aux articles en ligne [...] [est] très courant en santé et psychologie sociale » (Sanna et Hanni 2003, cités dans Couzinet 2018). Selon Pignard (1999), les sciences dures seraient en avance en matière de publications électroniques. swissuniversities (2018a) relève les pratiques de publications dans le domaine du droit, la distinction peu aisée entre les travaux relevant de la recherche de ceux relevant de la pratique ainsi que de la question de la langue.

Les domaines scientifiques ont recours à une langue de travail dominante, sans pour autant que tous les chercheurs rédigent dans celle-ci. La langue « de référence » peut varier dans le temps. Selon Hamel (2007), au tournant du XX^e siècle, le français, l'allemand et l'anglais tiennent une place égale dans le domaine scientifique, sans toutefois être utilisés de la même manière selon les disciplines : la médecine, la biologie et la chimie s'expriment principalement en allemand et dans des revues allemandes, le français étant utilisé dans les domaines du droit et des sciences politiques alors que l'anglais tient une place prépondérante dans le domaine de l'économie politique et de la géologie.

“Throughout the course of the 20th century, however, this balance was lost, not because of intrinsic dynamics in the field of science itself, but due to socio-economic and political factors.”
(Hamel 2007, p. 56)

Gross et Gross (1927) relèvent que, dans le domaine de la chimie, plus de 50% des citations concernent des publications en langue allemande. Si la chimie tend aujourd’hui encore à avoir une distribution linguistique plus large (Hamel 2007), l’anglais prédomine dans les sciences naturelles (Ammon 2006) lequel “is more prominent in the ‘pure’ or theoretical than in the applied sciences” (Skudlik 1990, cité dans Ammon 2006, p. 2) (cf. Annexe 6). Dans les SHS, si le recours à l’anglais prédomine (Ammon 2006) la proportion des autres langues est bien plus grande que dans les sciences naturelles, en particulier pour ce qui est du français et de l’allemand (Hamel 2007) (cf. Annexe 7). Ammon (2006) précise que le lectorat auquel s’adresse la publication ainsi que ses compétences linguistiques peut influencer la langue dans laquelle elle est rédigée : dans le domaine des sciences appliquées et SHS, les thèmes sont souvent intra-national et s’adressent à leurs lectorats respectifs. L’utilisation correcte de la langue aux fins des recherches en SHS nécessite un niveau de maîtrise qui n’est pas évident à atteindre dans une langue qui n’est pas sa langue maternelle. Concernant les sciences naturelles, Ammon (2006) relève que leur intérêt peut-être plus universel. Le lectorat se trouvant au-delà des frontières nationales, l’anglais est donc la langue de choix. Le vocabulaire technique des sciences naturelles est aussi plus formalisé, ce qui facilite l’usage d’une langue « étrangère ».

Les bases de données scientifiques archivent aujourd’hui majoritairement des revues et publications en langue anglaise facilitant ainsi leur visibilité, au détriment d’articles publiés dans une autre langue (cf. chapitre 4.2.4). Ceci peut donc, en fonction des domaines de recherche et de leurs méthodologies, avoir un impact direct sur les citations (cf. chapitre 4.2.2.2). Notons que dans le cadre de l’évaluation des publications au niveau d’une institution académique, celle-ci peut [aussi] être réalisée sur la base des publications déposées dans l’archive institutionnelle (cf. chapitre 5). Toutefois, en matière d’évaluation bibliométrique reposant sur des bases de données « externes », il faut impérativement être conscient de leurs limites géographiques, temporelles, linguistiques et disciplinaires et tenir compte des spécificités du domaine dans lequel évolue le chercheur. Du fait que “[o]ld and new fields may differ in growth rates, degree of interdisciplinarity and resources that are needed as inputs”, la production de chercheurs évoluant dans différents domaines n’est en effet pas comparable (Rousseau, Egghe et Guns 2018, p. 256).

4.2.2.2 Les habitudes de citation selon les domaines scientifiques

Le recours à des publications antérieures soit la « construction documentaire personnelle » du chercheur lui permet « d'obtenir des résultats nouveaux qui, rendus publics à leur tour, pourront être réinvestis dans d'autres travaux » (Couzinet 2000, p. 257). Les habitudes de citations varient entre les disciplines scientifiques (Sanna & Hanni 2003, cités dans Couzinet 2016) ; plus globalement selon le domaine de recherche (Costas et al. 2009 ; Hicks et al. 2015 ; Rousseau, Egghe et Guns 2018). Les pratiques en matière de citation diffèrent tant au niveau quantitatif que temporel (Rousseau, Egghe et Guns 2018) entraînant ainsi une variation considérable du taux de citation entre les différents domaines de recherche (Hicks et al. 2015). Aussi, Garfield (1979[b], cité dans Moed et al. 1985) qualifie d'inadéquante la comparaison du nombre de citations entre les domaines du fait que le « potentiel de citation » peut considérablement varier entre eux et que telle comparaison n'illustrerait pas réellement des différences en termes d'impact des publications des différents domaines ; aussi préconise-t-il le recours au taux de citation moyen par publication par domaine (cf. chapitre 4.4.2.3).

Du point de vue quantitatif, le nombre de citations dans une publication diffère selon les domaines scientifiques (Moed et al. 1985 ; Waltman et al. 2011). L'OCDE et le SCImago Research Group (2016) précisent que le nombre de citations pouvant être considéré comme important varie d'un domaine à un autre (Cf. Annexe 8). En outre, "[t]here is considerable variation in the skewness of citation distributions even within a given subfield" (Hirsch 2005, p. 16571) (cf. chapitre 4.4.3). Relevons encore le fait que « [...] le nombre moyen de citations par auteur dépend fortement de la taille des disciplines et des spécialités [...] » (Gingras 2014a, p. 33). Dans leur étude, Moed et al. (1985) se penchent sur le nombre moyen de citations dans les articles publiés par les treize disciplines représentées au sein des facultés de mathématiques et sciences naturelles de l'université de Leiden³⁴ : ce nombre diffère effectivement entre elles (cf. Annexe 9). Une étude se penchant sur la production d'articles dans les disciplines que sont « l'écologie, les mathématiques appliquées, la sociologie et l'économie » arrive au constat que les articles disponibles en OA sont plus régulièrement cités que les articles

³⁴ Moed et al. (1985) n'ont pas procédé à une sélection aléatoire de leur échantillon – d'où un possible « biais de Leiden » – ni testé la distribution normale de la population estimant que "in case of small sample size even quite severe departures from normality seem to make little differences in the conclusions reached" (Hayes 1973, cité dans Moed et al. 1985, p. 182). Pour mener leur étude, Moed et al. (1985) se basent sur les données du JCR de 1979. Ils ont constitué des groupes de revues de la manière suivante : des disciplines publiant régulièrement dans les mêmes revues sont considérées comme un domaine ayant un groupe spécifique de revues – les disciplines publiant dans des revues hétérogènes sont considérées de manière distincte.

qui sont publiés dans des revues disponibles sous format papier relevant toutefois une différence entre ces disciplines : « [I]es articles de sociologie sont les plus cités alors qu'ils sont moins nombreux en circulation en version numérique que dans les autres disciplines [précitées] » (Norris, Oppenheim et Rowland 2008, cités dans Couzinet 2018, p. 255). Garfield (1979) qualifie donc d'inadéquante la comparaison des taux de citation entre les domaines scientifiques, du fait que le potentiel de citation varie entre eux. En outre, le potentiel de citation n'est pas exclusivement corrélé au nombre de chercheurs dans le domaine précise-t-il :

“It also seems to have much to do with the ratio of publishing authors to total research population, the distribution of published papers over the population, and the distribution of references over the existing literature.” (Garfield 1979, p. 366)

Du point de vue temporel, nous observons que les citations des publications des chercheurs ne sont réparties de la même manière dans le temps selon la discipline ou le domaine scientifique dans lequel ils évoluent. Les variations caractéristiques des citations comprennent i. la rapidité à laquelle une publication sera citée ii. le temps nécessaire pour que le taux de citation atteigne son point maximal ou encore iii. le temps durant lequel la publication sera citée (Garfield 1979[b], cité dans Moed et al. 1985, p. 178). Citons les SHS dont les citations s'inscrivent dans un temps plus long par rapport aux domaines des sciences de la technique et de la médecine (Couzinet 2018). L'étude précitée de Moed et al. (1985) distingue les disciplines selon que leur taux de citation à court terme soit élevé, moyen ou bas (cf. Annexe 10). Gingras (2014a) précise toutefois que si à court terme, soit sur deux ans, la proportion des articles non cités dans le domaine des SHS est supérieure à celle des articles émanant des sciences biomédicales, cet écart se réduit à moyen terme.

Précisons en outre que la dimension linguistique doit être considérée en matière d'habitudes de citations : langue des publications scientifiques ainsi que maîtrise des langues par les scientifiques. Ammon (2006) estime que les anglophones ne tiennent compte des résultats de la recherche publiés dans d'autres langues que la leur que de manière très restreinte tout en précisant que les non-anglophones tiendraient compte des publications anglophones ainsi que de celles dans leur propre langue. L'évolution des pratiques de citation selon les domaines au regard de la langue de publication ne sont toutefois pas comparables. Ammon (1998) compare l'évolution de la moyenne des parts de citations dans les langues allemandes, françaises et anglaises dans des revues d'histoire et de chimie entre 1920 et 1990 (cf. Annexe 11). Dans le domaine de la chimie, tant la progression de la langue anglaise que la décroissance des langues allemande puis française sont linéaires. Les citations en anglais prennent le pas à la fin des années

1940. Pour ce qui est des revues d'histoire, l'évolution est tout sauf linéaire. C'est au tournant des années 1970 que les citations en anglais distanceraient celles des deux autres langues. Notons que la croissance du nombre de publications en anglais explique en partie l'augmentation du nombre de références à ces dernières. La non-linéarité de l'évolution temporelle des citations dans les trois langues dans les revues d'histoire est liée i. au fait que le domaine ne connaît pas une même progression en langue anglaise que d'autres domaines et ii. ces articles connaissent un temps de citation plus long en langue nationale que celui observé dans les sciences naturelles. Ainsi donc :

« [...] le nombre absolu de citations signifie généralement peu de chose en lui-même ou en comparaison avec d'autres disciplines, [...] il doit toujours être rapporté aux pratiques de la discipline du chercheur à une période donnée. »
(Gingras 2014a, p. 33)

S'il n'est donc pas pertinent de comparer les taux de citation (cf. chapitre 4.4.2.2) ainsi que le *h*-index (cf. chapitre 4.4.3.1) entre champs disciplinaires et qu'ils sont des indicateurs adaptés aux pratiques de citation selon les domaines, Hicks et al. (2015) en appellent à l'utilisation d'indicateurs de normalisation (cf. chapitre 4.4.2.3). Il reste à déterminer s'il est pertinent et recommandé d'y avoir recours.

4.2.3 La qualité d'auteur

Les publications scientifiques sont rédigées par un ou plusieurs auteurs.

« L'auteur est la personne qui, par son travail scientifique personnel, a fourni une contribution essentielle à la planification, à la réalisation, à l'évaluation ou au contrôle du travail de recherche. » (Académies suisses des sciences 2008, p. 18)

Nous observons une augmentation du nombre d'auteurs par publications³⁵, ce qui peut avoir une incidence directe sur certains indicateurs bibliométriques. Ce fait participe aux critères qui permettent de définir qui doit être considéré comme qu'auteur. Les principes en matière d'intégrité de la recherche offrent des indications quant à la gestion de la qualité d'auteur. Il est évident que si le nom d'un auteur ne figure pas sur la publication ou n'est pas intégré dans les métadonnées des bases de données, ladite publication ne saurait être portée à son actif. Le fait est que la « collectivisation de la recherche » et le nombre d'auteurs par article n'a cessé de croître au XX^e siècle (Gingras 2014a, p. 40). Il s'agit de savoir qui est considéré comme auteur et de qualifier la qualité de l'auteur. L'*International Committee of Medical Journal Editors* (ICMJE) (2020) dresse une liste de quatre conditions cumulatives permettant d'être légitimement considéré comme auteur ; le chercheur doit :

³⁵ Lindsey et Brown (1977, cités dans Garfield 1979) relèvent que la proportion des publications rédigées par plusieurs chercheurs varie selon les domaines scientifiques.

« [i.] Avoir contribué de manière substantielle à l'œuvre sinon avoir travaillé à l'acquisition, analyse ou interprétation des données puis [ii.] Avoir rédigé l'œuvre sinon avoir procédé à une revue critique du cœur du contenu intellectuel puis [iii.] Avoir approuvé la version finale destinée à la publication puis [iv.] Accepter d'assurer la responsabilité de tous les aspects de l'œuvre tout en s'assurant que l'exactitude et l'intégrité du travail soient convenablement examinées et résolues. (ICMJE 2020, §2)

Comme le rappellent les Académies suisses des sciences (2013), s'ils sont plusieurs, les co-auteurs doivent être mentionnés selon les pratiques reconnues en matière d'élaboration des listes d'auteurs. L'ordre de citation ainsi que la notion de « qualité d'auteur » dans le cas de publications qui reflètent le fruit du travail d'un groupe de recherche ne sont pas toujours évidentes et peuvent amener à la discorde (Académies suisses des sciences 2013). Les principes en matière d'intégrité de la recherche renseignent sur la gestion de la qualité d'auteur. S'y trouvent, pour ne citer que ces exemples, la question de la responsabilité de la publication (Académies suisses des sciences 2008 et ALLEA 2018) ou la nécessaire entente sur l'ordre de citation (ALLEA 2018). Rousseau, Egghe et Guns (2018) confirment l'accroissement des imbroglios connexes à l'augmentation du nombre de co-auteurs : ils peuvent potentiellement rendre certains auteurs invisibles (nombre d'auteurs limité par la revue scientifique, recours à un nom de groupe...) ou encore soulever des problèmes éthiques (recours à des prête-plumes, paternité à titre honorifique...) (Moosa 2018).

Rousseau, Egghe et Guns (2018) rappellent qu'un chercheur qui publie en son nom propre sera considéré comme plus productif que ceux qui collaborent à la production du même nombre de publications sur une même période donnée. À l'opposé, Garfield (1979) stipule que la collaboration entre chercheurs augmente la productivité, ce qui peut influencer le nombre total de citations. Comment alors évaluer les publications signées de plusieurs plumes ? Pour Rousseau, Egghe et Guns (2018), les méthodes de comptage peuvent, sur la base de deux critères, être réparties en quatre groupes. Le premier critère est basé sur la décision d'accorder, ou non, crédit à chacun des auteurs. Le second consiste à attribuer un nombre entier naturel à chaque crédit ou de le fractionner de manière à ce leur addition s'élève à un crédit total de 1. Il existe plusieurs modes de comptage fractionné précisent enfin Rousseau, Egghe et Guns (2018) : i. attribuer à chaque auteur $1/x$ crédits, x étant le nombre d'auteurs ou ii. attribuer des crédits en fonction d'une progression dite harmonique qui permet d'octroyer un crédit équivalent aux auteurs considérés comme ayant contribué de manière significative à la publication. L'OCDE (1997) précise qu'une méthode analogue est utilisée lorsqu'une publication est rédigée par des chercheurs basés dans différents pays : i. chaque pays

se voit attribuer un crédit complet (soit 1) ; ii. une fraction du crédit est octroyée à chaque auteur en vue d'atteindre le 100%.

« Chaque manière de compter possède [...] sa propre logique, mais encore faut-il que le fractionnement soit totalement maîtrisé par les bibliométriciens et compris par les utilisateurs. »
(OCDE 1997, p. 23)

Le comptage entier serait la meilleure des solutions selon Van Raan et Tijssen (1990, cités dans OCDE 1997). Au moment d'analyser les données, si les modes de calculs diffèrent, des manipulations seront nécessaires, augmentant par-là même le risque d'erreurs.

La question de la prise en considération de la paternité des publications est centrale. Elle a des répercussions directes sur le nombre de citations comptabilisé par auteur. Rousseau, Egghe et Guns (2018) citent les deux exemples suivants : i. sur la base du comptage normalisé, un écrit, rédigé par quatre auteurs et cité dix fois, verra attribué 2,5 citations à chacun d'entre eux. Toutefois ii. si seul le premier auteur est crédité, il se verra attribuer les dix citations. Les répercussions s'observent aussi dans les citations de la publication en question. Selon les pratiques, ce ne sont généralement que le premier ou les deux premiers auteurs qui se voient cités, ce qui a des répercussions directes sur le nombre de citations des autres co-auteurs. Comme disait Price (1986), à l'ère de la « Big Science »³⁶ et de la recherche collaborative, il devient plus difficile d'évaluer individuellement le chercheur (Price 1986, p. 36).

4.2.4 Les biais préexistants

Les biais que nous relevons ici sont à considérer de manière globale. S'il faut en tenir compte en matière d'étude bibliométrique, ils n'ont pas toujours une incidence sur l'évaluation du chercheur, selon la procédure d'évaluation de son institution académique et de la source des données qui sous-tendent l'évaluation. De nombreux biais sont liés aux bases de données scientifiques, qu'elles soient internationales (WoS, Scopus...) ou nationales (HAL Archives...). Aucune de ces bases de données n'est exhaustive. Ce ne sont au demeurant pas leur objectif. Il faut être conscient du fait qu'elles ne couvrent ni l'intégralité des revues, ni l'intégralité des publications scientifiques. Elles sont orientées vers certains domaines et disciplines scientifiques spécifiques. Rousseau, Egghe et Guns (2018) précisent que les variations du nombre de références ont une incidence directe sur le nombre de citations par domaine. Qui plus est :

³⁶ Price (1986) parle de « Big Science » versus « Small Science », la première se distinguant de la seconde en termes du nombre important de personnes qui se dédient à la science, dès lors à la recherche, ainsi que des fonds financiers injectés dans le domaine le transformant en un segment majeur de l'économie.

“[...] this reinforces the previous point: the number of citations in a poorly covered field may in reality be much higher than what one can deduce from one single database.”
(Rousseau, Egghe et Guns 2018, p. 115)

Enfin, le nombre et la qualification des domaines et disciplines varient d'une base de données à une autre. La couverture temporelle peut, au sein d'une même base de données, varier entre les domaines et disciplines. Les types de publication insérés ne sont pas toujours les mêmes, tout comme les schémas de métadonnées ou le vocabulaire contrôlé. Basant leur étude sur la SCI, Moed et al. (1985) relatent l'incorporation de nombreuses nouvelles revues scientifiques dans la base de données, le transfert de revues du SCI au nouvel SSCI (1973) ainsi que l'intégration de publications désormais plus uniquement « limitées » aux articles entre 1970 et 1980³⁷.

Un autre biais inhérent aux bases de données est celui de la langue. Les bases de données internationales sont biaisées en faveur de l'anglais (Ammon 2006 ; Hicks et al. 2015 ; Rousseau, Egghe et Guns 2018) (cf. Annexe 12), ce d'autant plus dans les domaines où cette langue ne « domine » pas la communication scientifique (cf. chapitre 4.2.2.1). En outre, en compilant les références dans les bases de données, les rédacteurs peuvent faire quelques erreurs qui peuvent être répercutées dans les citations, corrélativement dans leur analyse (Rousseau, Egghe et Guns 2018).

Les différents biais introduits ici ne sont pas à prendre comme une critique des bases de données scientifiques. Celles-ci s'avèrent être d'extraordinaires et puissants outils en matière de recherche et de communication scientifique. En outre :

« En tenant compte des biais inhérents aux différentes bases de données, une analyse de la production scientifique dans le temps, à l'échelle d'un pays, d'une région ou d'une institution, fournit en effet des informations importantes impossibles à obtenir autrement. »
(Gingras 2014a, p. 37)

Mentionnant la possible absence d'unification dans le nom des chercheurs et de leurs homonymes, Costas et al. (2009) soulignent l'importance de contrôler le nom ainsi que l'affiliation des auteurs en matière d'analyse bibliométrique. Schreiber (2007), qui précise que les noms des chercheurs peuvent également changer au gré des retranscriptions ou lors du changement de leur état civil, souligne qu'une solution est de contrôler l'affiliation des auteurs, tout en admettant la complexité du travail lorsque les chercheurs évoluent au sein de plusieurs institutions successives et de la possible absence d'unification dans la dénomination de ces dernières dans les publications. Le recours systématisé aux identifiants uniques et pérennes à l'instar de l'*Open Researcher and*

³⁷ Moed et al. (1985) précisent qu'à compter de 1977, le SCI intègre également ce qu'il regroupe sous la dénomination de « livre » : actes de conférences, collections thématiques de documents, monographies...

Contributor Identifier (ORCID) pour les auteurs et des *Digital Object Identifier* (DOI) pour les publications et jeux de données archivés nous paraît être la solution la plus à même d'endiguer les incertitudes quant aux attributions : de la paternité comme de l'œuvre. Le *Metric Tide* (Wilsdon et al. 2015) en appelle à l'interopérabilité des bases de données et au recours d'identifiants uniques, pour les auteurs ainsi que pour les travaux de recherche, en vue d'accroître la robustesse des indicateurs.

Il s'agit d'être conscient de ces biais en vue d'assurer la pertinence des études bibliométriques. D'où l'importance de clairement identifier la base de données la plus pertinente pour chaque étude, de préciser le domaine ou la discipline de son champ, la manière dont est entendue, si elle est considérée, la trans- ou interdisciplinarité. La définition de chacun des termes de l'évaluation est importante du fait qu'elle a une incidence directe sur les résultats de l'analyse. Du point de vue temporel, la période sur laquelle se penche l'analyse doit être spécifiée, tant du point de vue des publications que des citations, tout comme la date à laquelle les données sont récoltées. Les types de publications prises en considération doivent être précisées. Concernant l'attribution des publications et citations à son (ses) auteur(s), Garfield (1979) préconise deux solutions selon la dimension de l'analyse : i. si le nombre d'auteurs concernés est restreint, les données de la base de données peuvent être croisées avec l'examen des revues dans lesquelles la publication a été publiée puis citée ; ii. si en revanche l'analyse porte sur un nombre plus conséquent d'auteurs, les évaluateurs devraient toujours se référer, en sus de la base de données, à la bibliographie exhaustive des chercheurs. Du fait de la nécessaire prudence en matière d'utilisation des indicateurs bibliométriques, l'OCDE (1997) préconise de comparer les résultats obtenus dans plusieurs bases de données. Les jeux de données utilisés aux fins d'études ou d'évaluations devraient être archivés de manière pérenne en vue de (ré-)utilisations futures.

La même prudence est requise lors de la citation d'indicateurs issus de ces bases de données. En matière de bibliométrie, comme dans tout domaine recourant à des données et usant de techniques mathématiques et statistiques, la méthodologie doit être mûrement réfléchie et communiquée de manière transparente. La même vigilance prévaut aux fins d'une lecture critique des études et des évaluations bibliométriques.

4.3 Les lois « piliers » de la bibliométrie

Les lois de Lotka, Bradford et Zipf sont les piliers de la bibliométrie (Bailón-Moreno et al. 2005). Expliquant des phénomènes distincts, elles présentent toutefois un certain nombre de ressemblances (Rostaing 1996). “[T]hey express the empirical relation between sources and the items they deliver” (De Bellis 2014, p. 37). De Bellis (2014)

relève le fait que leur dénominateur commun est le constat d'une disparité dans le modèle qu'elles analysent : la majeure partie des publications scientifiques ayant un impact dans un domaine scientifique sont le fait de quelques chercheurs (loi de Lotka) ; la majorité des articles significatifs dans un champ disciplinaire sont publiés dans quelques revues (loi de Bradford) et enfin un nombre relativement restreint de « groupes » de mots récurrents domine les comportements linguistiques individuels dans la communication scientifique (loi de Zipf).

4.3.1 La loi de Lotka

Basant son étude sur le *decennial index of chemical abstracts*³⁸ et l'*index d'Auerbach's Geschichtstafeln der Physik*, Lotka (1926) souhaite savoir s'il est possible de déterminer la manière dont des chercheurs de différents acabit contribuent au progrès de la science. Soit la distribution de la productivité scientifique des chercheurs (Gingras 2014a). Lotka (1926) observe que seul un nombre restreint d'auteurs publie un nombre conséquent d'articles à savoir que ladite productivité se limite à un certain nombre de scientifiques : le nombre de chercheurs publiant n contributions scientifiques est égal à $1/n^2$ et est dès lors inversement proportionnel au carré de n (cf. Annexe 13). Si des études menées depuis lors ont confirmé cette loi, d'autres la remettent en question. Ainsi, Price (1986) relève que le nombre de chercheurs les plus prolifiques en termes de publication s'approcherait plus de l'inverse du cube et non du carré³⁹. Certains estiment que « [l]a présentation générale et la régularité de la distribution [...] ne sont plus remises en cause [et que] seule la formulation mathématique reste sujet de discussion » (Rostaing 1997, p. 42). La loi de Pareto ou loi des 80/20 est une extension de la loi de Lotka au domaine de l'économie et de la démographie (Bailón-Moreno et al. 2005).

4.3.2 La loi de Bradford

Les revues contiennent des articles pouvant intéresser plus d'un domaine scientifique. Partant de l'observation que seul un tiers des articles publiés sont présents dans les journaux d'abstract et d'indexation, Bradford (1934) souhaite découvrir s'il est un moyen de sélectionner les revues les plus pertinentes⁴⁰ en cherchant à définir quel en est leur "nucleus" (noyau). Il analyse la manière dont les publications spécifiques à un domaine sont réparties dans les revues qui se voient classées en trois groupes : i. celles qui publient plus de quatre articles ; ii. celles qui publient entre un et quatre articles et iii. celles qui en publient un, si ce n'est moins, par an. Les revues de ce dernier groupe sont

³⁸ Lotka (1926) a sélectionné les chercheurs dont le nom commence par A ou B.

³⁹ En analysant les réseaux de citation, il trouve toutefois une loi de distribution très similaire à celle de la loi de Lotka (Price 1965).

⁴⁰ Dans sa recherche, Bradford (1934) se concentre sur la géophysique appliquée et la lubrification.

les plus nombreuses et de nature très variée. Bradford (1934) s'étonne du fait que certaines d'entre elles, pourtant orientées vers ses domaines de recherche, ne publient pas plus d'articles sur le sujet. Il exprime sa loi de distribution des articles traitant d'un sujet spécifique dans les revues de la manière suivante :

"[...] if scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several groups or zones containing the same number of articles as the nucleus, when the numbers of periodicals in the nucleus and succeeding zones will be as 1 : n : n²..."

(Bradford 1985, p. 178, publication originale Bradford 1934)

Au terme de son étude, Bradford (1934) arrive au constat qu'un changement radical est nécessaire : la littérature scientifique ne doit plus être résumée et indexée par sujet, mais par source.

En vue du lancement du SCI, Garfield cherche à déterminer le nombre de revues scientifiques qui permettront de recueillir le nombre optimal d'information scientifique. L'OCDE (1997) relate les étapes suivies par Garfield. Dans un premier temps, il détermine, sur la base d'une loi elle-même fondée sur celle de Bradford, que ce sont entre cinq cent et mille revues qui doivent être consultées pour obtenir 95% « de la littérature « significative » dans un domaine donné » (OCDE 1997, p. 18). Dans un second temps, du fait que la loi de Bradford se réfère à un domaine scientifique, Garfield cherche à déterminer le nombre de revues à sélectionner en vue de couvrir plusieurs domaines. Il associe une loi de concentration qu'il a développée à la loi de dispersion de Bradford. Puis établit qu'il ne faut pas multiplier le nombre de revues par le nombre de domaines qu'il souhaite couvrir : du fait que nombre de « disciplines se recoupent, le cœur de la littérature pour l'ensemble de ces disciplines peut également être couvert par environ 500 à 1'000 revues » (OCDE 1997, p. 18). Il s'agit alors pour Garfield de repérer ces revues. Il se réfère au nombre de citations :

« Premièrement, on compte le nombre de fois qu'un article est cité dans une revue donnée. Ensuite, on calcule le facteur d'impact [...] en divisant le nombre de citations par le nombre d'articles contenus dans cette revue. Cela permet d'éliminer tout avantage lié à aux dimensions de la revue, et rend la citation proportionnelle au nombre d'articles. »

(OCDE 1997, p. 18)

Nous traitons du fonctionnement du nombre de citations dans le chapitre 4.4.2.2 puis du facteur d'impact dans le chapitre 4.4.4.

4.3.3 La loi de Zipf

George Kingsely Zipf dénombre les occurrences des près de trente mille mots se trouvant dans Ulysses de James Joyce avant de les classer par ordre décroissant de fréquence f et leur attribuer un rang r (Rostaing 1996). Multipliant f par r , Zipf observe

que le résultat obtenu est toujours proche d'une constante C , soit $f(r).r = C$. La fréquence est inversement proportionnelle au rang du mot. Constatant que l'être humain préfère avoir recours à des mots ordinaires plutôt que singuliers, Zipf nomme en 1949⁴¹ sa loi comme étant celle « du principe du moindre effort » (Rostaing 1997 ; Bailón-Moreno et al. 2005). Si cette loi de puissance ou “*power law*” est aujourd'hui connue sous le nom de loi de Zipf, elle avait déjà été relevée plus tôt dans le XX^e siècle par le sténographe Jean-Baptiste Estoup puis par le physicien nucléaire Edward Uhler Condon (Bailón-Moreno et al. 2005 ; Rousseau, Egghe et Guns 2018). En outre :

“The regularities he observed do not follow from a universal property of the human mind, but they are rather a consequence of the laws of probability. Zipf curves are merely expressing a consequence of regarding a message source as a stochastic process.”
(Rousseau 2002, p. 14)

Ce qui fut originellement formulé par Benoit Mandelbrot selon Rousseau (2002) et peut être considéré comme une généralisation de la Loi de Zipf : Mandelbrot intègre à sa fonction un paramètre supplémentaire qui s'applique dès lors plus facilement aux données observées (Rousseau, Egghe et Guns 2018). Plus récemment, il a été démontré que la notoriété des pages internet peut être décrite selon la loi de Zipf (Aida, Takahashi et Abe 1998, cités dans Rousseau 2002). Si les lois de puissances que sont les lois de Lotka et de Zipf sont considérées comme similaires :

“[...] starting with ‘Zipf’ leads to ‘Lotka’. Starting with ‘Lotka’ does not in general lead to ‘Zipf’ but rather to a generalization due to Mandelbrot.”
(Rousseau 1990, cité dans Rousseau 2002, p. 16)

4.4 Les indicateurs bibliométriques

Comme précédemment énoncé (cf. chapitre 4.1), nous avons recours au terme de bibliométrie pour signaler l'ensemble des mesures quantitatives qui permettent d'évaluer la publication scientifique d'un chercheur. Chacune de ces mesures se base sur des indicateurs qui permettent l'analyse et l'interprétation des données. Les informations relatives aux publications scientifiques sont désormais disponibles sous forme de données numériques dans des bases de données ; elles peuvent dès lors être communiquées et interprétées. Un indicateur est une expression mathématique qui, appliqué à des données numériques, permet d'obtenir une valeur indicatrice (“*indicator value*”) (Rousseau, Egghe et Guns 2018, p. 201). Autrement dit :

⁴¹ Selon Rousseau, Egghe et Guns (2018), Zipf aurait déjà fait référence à « sa » loi en 1935 dans “The psychobiology of language”. En outre, selon Rousseau (2002), Zipf ne serait pas l'auteur de « sa » loi qui lui aurait été suggérée par un ami mystérieux.

“[...] an indicator is a variable that can be measured and that aims to at faithfully representing a giver concept referring to the property of an object that one wants to measure.”
(Lazarsfeld 1958, cité dans Gingras 2014b, p. 112)

Il doit être utilisé « selon un objectif précis et dans un contexte particulier », « [i]l n'est pas un objectif, il apporte un élément éclairant un objectif [...] [et] permet la conversion de données [...] en informations concises à interpréter » (Bibliothèque Sorbonne Université 2017b⁴²). Pour pouvoir être considérés comme fiables, les indicateurs doivent regrouper un certain nombre de caractéristiques. Ils doivent en outre être utilisés de manière adéquate et appropriée. Les indicateurs bibliométriques connaissent un développement exponentiel faisant qu'il est aujourd'hui “[...] all but impossible [...] to count today's metrics” (Van Noorden 2010, cité dans Wouters 2014, p. 51). À défaut de pouvoir être exhaustifs, nous présentons ici les principaux indicateurs bibliométriques.

4.4.1 Introduction aux indicateurs

Il est impératif de clairement exprimer ce que mesure un indicateur, de démontrer qu'il mesure concrètement ce qu'il est censé mesurer ainsi que sa validité (Gingras 2014b ; Rousseau, Egghe et Guns 2018). En effet :

“[The] evaluation fever has resulted in rampant multiplication and misuses of faulty indicators.”
(Gingras 2014b, p. 111)

Rousseau, Egghe et Guns (2018) attirent l'attention sur le fait que les indicateurs sont trop souvent utilisés comme variables de substitution (“*proxy variable*”) :

“The journal impact factor (JIF), for instance, is often used, be it inappropriately, as a proxy for journal quality; similarly, the number of downloads is taken as a proxy for readership.”
(Rousseau, Egghe et Guns 2018, p. 201-202)

Il ne faut pas perdre de vue que chaque indicateur a ses avantages et ses limites ; ils doivent être considérés comme complémentaires (OCDE 1997).

4.4.1.1 Les caractéristiques des indicateurs

Un indicateur doit être juste, à savoir exempt de biais (“*said to be fair*”), utile (fournit des informations essentielles) et fiable (soit reproductible) précisent Rousseau, Egghe et Guns (2018, p. 202). D'autres caractéristiques, parfois regroupées sous le terme de SMART, précisent encore qu'un indicateur doit être significatif, mesurable, acceptable, robuste et temporellement défini (Bibliothèque de Sorbonne Université 2017b ; École nationale supérieure des sciences de l'information et des bibliothèques (ENSSIB) 2013).

⁴² <https://paris-sorbonne.libguides.com/c.php?g=640699&p=4488054> [consulté le 17 avril 2020].

4.4.1.2 La validité des indicateurs selon Gingras

Pour Gingras (2014b), trois critères permettent d'évaluer la validité d'un indicateur.

4.4.1.2.1 L'adéquation entre l'indicateur et ce qu'il doit mesurer

Les propriétés de l'indicateur doivent correspondre aux caractéristiques inhérentes à l'objet que l'on souhaite mesurer (Gingras 2014b) : si certains indicateurs sont plus évidents à appréhender – à l'instar des indicateurs de productivité – la mesure de concepts abstraits tels que la qualité, la visibilité ou l'impact est plus sibylline.

Gingras (2014b) stipule que le lien entre le concept et l'indicateur doit être testé : le nombre calculé de citations reçues par une publication peut être utilisé comme indicateur de la qualité, de la visibilité ou de l'impact - qui sont alors synonymes du fait qu'ils se réfèrent au même indicateur. Énonçant que le lien entre la qualité ou la visibilité et l'indicateur que sont les citations doit être démontré par relation entre les citations et les mesures indépendantes de la qualité déjà acceptées comme telles, Gingras (2014b) estime que les études sociologiques et bibliométriques menées ces cinquante dernières années ont démontré la corrélation entre le nombre de citations d'un chercheur et sa renommée laquelle peut être évaluée par les nominations, prix et récompenses. Si l'adéquation des indicateurs que sont le nombre de publications et de citations est désormais largement reconnu dans les domaines des sciences naturelles, Gingras (2014b) estime qu'ils ne peuvent pas pour autant être automatiquement appliqués aux domaines des sciences humaines et sociales du fait des différences entre les pratiques de publications et de citations (cf. chapitre 4.2.2).

4.4.1.2.2 Sensibilité de l'indicateur à l'inertie intrinsèque de l'objet mesuré

Rappelant que par inertie il faut entendre la capacité à résister au changement et d'impacter la célérité à laquelle ce changement peut survenir, Gingras (2014b) qualifie de bon indicateur celui qui évolue de manière cohérente avec l'inertie de l'objet mesuré. Aussi, pour Gingras (2014b), la fréquence à laquelle sont effectués les classements et évaluations doit être adaptée à l'objet d'étude.

4.4.1.2.3 L'homogénéité des dimensions de l'indicateur

L'homogénéité dans la composition de l'indicateur est primordiale pour Gingras (2014a) qui distingue les indices composés (tel que l'indice des prix à la consommation) des indices hétérogènes ; ces derniers combinent arbitrairement deux ou plusieurs indicateurs dont les mesures sont différentes, à l'instar du *h*-index (cf. chapitre 4.4.3).

“Combining different indicators into a single number is like transforming a multidimensional space into a zero-dimension point, thus losing nearly all the information contained in the different axes.”
(Gingras 2014b, p. 116)

Gingras (2014b) plaide donc en faveur du maintien des indicateurs pris individuellement et leur représentation sur un diagramme de Kiviat afin qu'ils rendent compte de chacune des différentes composantes du concept évalué.

4.4.1.3 Du bon usage des indicateurs bibliométriques

Selon Gingras (2014a), les « malaises » qui entourent l'usage de la bibliométrie en tant qu'outil d'évaluation des chercheurs proviennent :

« [...] d'une part, du flou entourant les différents concepts et indicateurs utilisés et, d'autre part, de désaccords sur l'échelle à laquelle ils peuvent s'avérer valides et utiles. »
(Gingras 2014a, p. 11)

Si les évaluations font désormais partie intégrante du domaine de la recherche et que leur bien-fondé ne doit pas être remis en question, « [i]l est [...] essentiel de critiquer rigoureusement les indicateurs mal construits, et dont l'utilisation peut engendrer des effets pervers » (Gingras 2014a, p. 12). Une réflexion sur leurs « fondements épistémologiques » doit précéder leur critique (Gingras 2014b). Il s'agit de « bien comprendre les propriétés spécifiques des indicateurs bibliométriques » ainsi que « leurs champs d'application » (Gingras 2014a, p. 11), soit de s'assurer qu'ils révèlent bien ce qu'ils sont censés indiquer ; à défaut de quoi ils doivent être considérés comme non pertinents (Gingras 2014b). Un chiffre isolé ne peut ni exprimer la multidimensionnalité de la recherche, ni servir l'évaluation de la qualité ou de l'impact :

“The very existence (and persistence) of such biased indicators and rankings seems to be a consequence of the unwritten rule that any number beats to number.”
(Gingras 2014b, p. 111)

En outre, la transparence des indicateurs utilisés en matière d'évaluation est capitale. Il est important que les chercheurs connaissent la signification ainsi que l'exacte formulation mathématique des indicateurs qui sont utilisés pour les évaluer ainsi que les manières dont ils peuvent être détournés : il importe qu'ils deviennent “metric-wise” (Rousseau 2015 et Rousseau et 2017, cités dans Rousseau, Egghe et Guns 2018, p. 1)

Le Coadic (2005) rappelle trois possibles sources d'erreurs en matière de mesure :

« 1. Les erreurs provoquées par les outils bibliographiques classiques [...] [soit] les erreurs de mesure dans le paratexte [qui sont] de l'ordre de 10 à 20%. 2. « Les erreurs provoquées par les outils informatiques [...] 3. Les erreurs provoquées par les méthodes d'observation. »
(Le Coadic 2005, p. 17)

L'évaluation doit donc être de qualité. Pour ce faire :

« [...] il apparaît, qu'idéalement, les politiques et pratiques d'évaluations devraient s'intégrer dans le concept qualité de l'institution. Ceci permettrait d'articuler de façon cohérente évaluation et qualité. »
(swissuniversities 2018a, p. 16)

Nous listons les différents indicateurs bibliométriques, précisons ce qu'ils définissent ainsi que la manière dont ils doivent être appliqués, sans manquer de signaler leurs limites. Ces dernières ne peuvent à elles seules constituer des arguments contre leur utilisation. C'est bien la mauvaise utilisation des indicateurs considérés comme pertinents, la malheureuse interprétation des résultats et l'utilisation d'indicateurs non pertinents qui doivent être écartés. Il existe des limites inhérentes à tout indicateur.

“Although fairness is a natural requirement, it is in practice probably impossible that an indicator is fair to everyone and in all circumstances.”

(Rousseau, Egghe et Guns 2018, p. 202)

4.4.2 Les indicateurs bibliométriques classiques

La bibliométrie repose sur deux postulats rappelle Rostaing (1997) : i. l'écrit scientifique, qui permet au chercheur de rendre compte de son activité de recherche et de convaincre la communauté de la pertinence de son travail ; ii. le chercheur confronte sans cesse son savoir et sa pensée avec ceux de ses confrères et fait référence à leurs travaux en vue d'asseoir son point de vue. Les publications et les citations sont les deux indicateurs classiques.

La bibliométrie consiste en « [...] l'analyse des publications et leurs propriétés » (Gingras 2014a, p. 15). Aussi, comme le précise Rostaing (1997), les données quantitatives sur lesquelles reposent la bibliométrie se basent sur le calcul statistique : des publications ainsi que des éléments qui en sont extraits. Le Coadic (2005) qualifie de paratexte « [...] le texte qui enveloppe le texte [de la publication] » à l'instar du « [...] titre [de la publication], le nom de (ou des) auteur(e), [...], les références... » (Le Coadic 2005, p. 17).

« La littérature scientifique se prête [...] à une quantification en ce sens que chaque [publication] contenant ces éléments bibliographiques paratextuels [...] permet de constituer un ensemble de variables textuelles, personnelles et temporelles sur lesquelles on peut effectuer des mesures et des analyses quantitatives puis qualitatives [...]. »
(Le Coadic 2005, p. 17)

Ces indicateurs que sont le nombre de publications et de citations des chercheurs permettent la comparaison entre les chercheurs d'un domaine scientifique (cf. chapitre 4.2.2). En vue d'une interprétation correcte des indicateurs, Garfield (1979) souligne que l'évaluation doit toujours être réalisée par – du moins en étroite consultation avec – des personnes accoutumées au domaine scientifique concerné. Relevons qu'au fil du temps, de nouveaux indicateurs se référant aux nombres de citations et de publications se développent. Hirsch (2005) liste trois autres indicateurs régulièrement utilisés en vue de l'évaluation de la production scientifique du chercheur : i. Le nombre de citations par publication ii. le nombre de publications considérées comme importantes – précisant

qu'il est établi comme étant le nombre de publications présentant plus de C citations et enfin iii. le nombre de citations de chacune des P publications les plus citées. Les avantages et inconvénients relevés pour les principaux indicateurs bibliométriques présentés dans le chapitre 4.4 sont synthétisés dans l'Annexe 26.

4.4.2.1 Le nombre de publications

Les publications scientifiques, soit la communication écrite des fruits de la recherche, sont considérées comme l'un de ses *outputs* (produits). Elles sont protégées en tant qu'œuvres par le droit d'auteur si elles sont le résultat d'une création de l'esprit et revêtent un caractère individuel (art. 2 al. 1 et 2 de la *Loi fédérale du 9 octobre 1992 sur le droit d'auteur et les droits voisins* (LDA)).

Le dénombrement du nombre de publication « [...] fournit une première mesure, simplifiée et approximative, de la quantité de travail fournit par un chercheur [...] »
(OCDE 1997, p. 26)

Considéré individuellement, ce nombre est un « indicateur bibliométrique grossier » (OCDE 1997, p. 26). Comparé à celui d'autres chercheurs, il indique une « [...] mesure plus significative de la puissance relative des sujets étudiés » (OCDE 1997, p. 26). Une fois divisé par le nombre de scientifiques – ou par le budget alloué – il peut renseigner quant à « la « productivité » des travaux en question » (OCDE 1997, p. 26).

Les écrits qui sont regroupés sous le terme de publication scientifique varient selon les sources ou les domaines scientifiques (cf. chapitre 5). En vue de leur analyse bibliométrique, les publications doivent néanmoins être disponibles sous la forme de bases de données (Gingras 2014a). Dans la présente recherche, nous regroupons sous le terme de publications scientifiques les communications scientifiques, fruits de recherches fondamentales, publiées par une institution académique ou un éditeur scientifique.

Price (1986) relève le fait qu'en vue d'obtenir la reconnaissance, un scientifique doit faire preuve de ténacité et de persévérance, ce qui implique un rythme soutenu de publication. En matière d'évaluation, il dénonce toutefois le seul recours au nombre de publications dans des revues considérées comme prestigieuses, conscient de ses effets sur la culture du POP (cf. chapitre 3.3.5). Toutefois, Price (1986) considère qu'il s'agit là d'un point de départ qui doit être affiné : si la capacité scientifique ne peut pas directement être mesurée, les « propriétés de sa distribution présumée » peuvent être déduites (Price 1986, p. 36). Relevons néanmoins le possible « biais de publication » relevé par Staub (2020) qui apparaît lorsque certains chercheurs décident de ne pas publier une étude du moment que cette dernière n'a pas abouti à des résultats qu'il

considère comme positifs. L'amplitude du biais de publication peut être évalué en effectuant le lien entre efficacité et précision au sein d'une recherche selon (Staub 2000) (cf. Annexe 14). « Une mesure efficace contre un biais de publication est l'enregistrement prospectif d'études prévues dans des registres d'études » précise Staub (2000 p. 18), ces registres permettant d'assurer le respect des analyses programmées ainsi que la publication future.

Pour Merton (1968), le *Matthew effect* (cf. chapitre 3.2) s'observe également en matière de communication scientifique : la visibilité de la publication d'un chercheur dont la réputation est établie est plus grande que celle de la visibilité d'un jeune chercheur. En outre, avec la croissance exponentielle du nombre de publications scientifiques, le *Matthew effect* augmente en « fréquence et intensité » (Merton 1968, p. 56). Corrélativement, les publications des chercheurs de renom sont susceptibles d'entrer « rapidement » et « largement » dans le réseau de communication scientifique, dès lors d'accélérer son développement (Merton 1968, p. 57). Aussi Merton (1968) de conclure que l'innovation doit être communiquée de manière effective – c'est ce que l'on nomme la « contribution à la science », soit l'apport au « fonds commun de la connaissance » (Merton 1968, p. 56). “For the development of science, only work that is effectively perceived and utilized by other scientists, then and there, matters.” (Merton 1968, pp. 56-57). La question de savoir si un article est plus cité en raison de son contenu ou de son auteur doit-elle se poser au vu des arguments de Merton (1968) ? Les citations des publications permettent de découvrir quels sont les travaux des chercheurs qui, dans un domaine et à un moment donnés, sont utilisés par leurs confrères. Savoir comment le chercheur dresse la liste de ses références est une question complexe qui, à notre avis, doit être analysée sous divers angles. La discipline ou le domaine scientifique (cf. chapitre 4.2.2), la méthodologie employée, la finalité de la recherche sont autant de pistes de réponses tangibles. Viennent ensuite les motivations personnelles et la psychologie du chercheur qui ne sont pas aisément supputables.

Pour l'OCDE (1997), cette mesure, prise individuellement, ne rend pas compte de la qualité des travaux. En outre, elle ne représente pas la quantité de travail fourni ; si une publication a été rédigée par plusieurs chercheurs, eux seuls connaissent la contribution réelle de chacun des auteurs (OCDE 1997).

4.4.2.2 Le nombre de citations des publications

Le lancement du SCI en 1963 (cf. chapitre 3.3.2) explique l'augmentation du nombre de travaux destinés à étudier le rapport entre le nombre de citations et d'autres « variables caractérisant les chercheurs » en vue d'« identifier les déterminants de la productivité et

de l'impact de la recherche » (Gingras 2014a, p. 35). Dès ses débuts, l'évaluation de la performance des chercheurs par le recours au nombre de citations soulève craintes et critiques. Garfield (1979) reconnaît que ceux qui ont recours à cet indicateur aux fins d'évaluation doivent : i. être des connaisseurs du domaine scientifique en question, ii. conscients des subtilités et des limites de ces données, iii. observer la méthodologie pertinente (cf. chapitre 4.2.4) et iv. être capables d'interpréter les résultats. Souvent, ce sont les termes d'impact ou de qualité qui sont utilisés pour qualifier le taux de citation. Garfield (1979) préfère le terme d'utilité, en ce sens que cet indicateur démontre l'intérêt des autres chercheurs pour un sujet publié, aussi se rapproche-t-il d'une évaluation et de la reconnaissance par les pairs.

Nous avons déjà relevé les différentes pratiques en la matière au regard des domaines scientifiques (cf. chapitre 4.2.2.2). Il s'agit maintenant de préciser la manière dont fonctionnent les citations. Pour Price (1980, cité dans Moed et al. 1985), le nombre moyen de citations dans une publication est déterminé par le volume de la littérature scientifique disponible dans le domaine. L'émergence de la culture POP (cf. chapitre 3.3.5) dans les années 1970-1980 a fait augmenter le volume de la littérature scientifique sur la base de laquelle les chercheurs peuvent s'appuyer pour étayer leurs recherches. Moed et al. (1985) observent une augmentation du nombre de citations dans ces mêmes années : les données reposent sur les articles publiés ou sur les nouvelles catégories de publications intégrées dans le SCI (cf. chapitre 3.3.2) (cf. Annexe 15). Le nombre de citations par article, « [...] dépend du nombre de références que les articles contiennent ; plus ce nombre est élevé, plus les chances d'être cité augmentent » (Gingras 2014a, p. 33). Au vu de la croissance exponentielle de la science et du nombre de ses praticiens, Gingras (2014a) rappelle en outre que le nombre de références par publication augmente tout comme le nombre moyen de citations par parution. Élargissant leur période d'analyse à la décennie 1970-1980, Moed et al. (1985) constatent qu'en se basant sur les citations des articles publiés les deux précédentes années, leur nombre augmente proportionnellement au nombre de publications ; alors qu'en étudiant les citations des articles publiés les trois précédentes années, le nombre de citations augmente plus que le nombre de publications. Il faut néanmoins être conscient d'un certain nombre de facteurs pouvant influencer les résultats précisent Moed et al. (1985) à l'instar de l'évolution du nombre de chercheurs, des modifications dans les bases de données (cf. chapitre 4.2.4) ainsi que les habitudes de publication et de citation selon les domaines scientifiques (cf. chapitre 4.2.2).

Basant son argumentation sur des études ayant démontré que la sélection d'articles pouvait se faire sur la base de leurs auteurs, Merton (1968) considère que les chercheurs

ont tendance à prendre comme repère la réputation scientifique de leurs confrères (cf. chapitre 3.2) en vue de sélectionner les écrits qu'ils considèrent comme les plus significatifs. Les collègues invisibles (cf. chapitre 3.1) et les traditions en termes de réseautage entre les chercheurs émanant de différents domaines scientifiques peuvent aussi avoir une influence sur leurs habitudes de citation :

"[...] natural scientists [...] would enjoy a greater degree of social connectedness in their specialist fields. [...] practitioners [of "soft" sciences] would be more focused on teaching and [...] fare more loosely connected."

(Biglan 1973, cité dans Krishnan 2009, p. 11)

Selon Moed et al. (1985), sans pour autant pouvoir quantifier cet argument, l'importance croissante du nombre de citations en matière d'évaluation de la performance scientifique pourrait inciter certains scientifiques à se référer plus régulièrement aux publications de leur réseau. Schreiber (2007) serait curieux de connaître le *h*-index des membres de ces réseaux une fois les autocitations écartées.

Les articles scientifiques forment un réseau qui est relié par les citations, soit autant de témoignages individuels des auteurs quant aux publications de leurs pairs qu'ils considèrent comme pertinentes pour leur propre recherche précise (Bergstrom 2007). L'objectivité de la mesure est le principal argument des défenseurs de cet indicateur dénué de toute interprétation personnelle (Brown et Gardner 1985 puis Beattie et Goodacre 2006, cités dans Moosa 2018). Un certain nombre d'arguments cherche à discréditer l'indicateur des citations, qui est toujours calculé sur la base d'un ensemble défini de revues (Ingwersen et al. 2001), soit celles intégrées dans la base de données sur laquelle se base l'analyse. L'un d'entre eux a trait aux citations dites négatives. Or, le taux de citation ne sert pas à compter le nombre de fois où l'on considère qu'un chercheur a raison énonce Garfield (1979) qui le considère comme "[...] very general measure of the level of contribution an individual makes to the practice of science" (Garfield 1979, p. 362). White (2001, cité dans Rousseau, Egghe et Guns 2018) souligne qu'avant de citer une publication, l'auteur évalue si cette référence appuie ou nourrit son analyse et estime si elle a suffisamment de poids pour étayer son raisonnement. Une autre critique relevée par Garfield (1979) émane de ceux qui considèrent que toutes les publications ne devraient pas être comptabilisées sous cet indicateur – à l'instar des ouvrages méthodologiques. Garfield (1979) de rappeler que les types de publications cités varient selon les domaines scientifiques et que certains d'entre eux sont précisément orientés vers la méthodologie, à l'instar de la chimie analytique. À ceux qui critiquent le fait que l'indicateur ne tient pas compte du prestige de la revue dans laquelle est parue la publication, il précise : "[t]hough it is easy to speculate on the effects that

journal prestige may have upon citation counts, it does not seem to be a very important factor” (Garfield 1979, p. 365).

La principale critique envers cet indicateur est l'autocitation qui permettrait, en théorie, de manipuler son taux de citation. Se basant sur le SCI, Garfield et Sher (1963) avancent un taux d'autocitation de 8% pour les premiers auteurs. Sans définir le sens sous lequel il entend l'autocitation, Schreiber (2008) avance un taux de d'autocitations entre 15% et 25% selon les études dans le domaine de la physique. Garfield (1979) qualifie la pratique d'ordinaire et de raisonnable estimant que moins de 10% des citations sont concernées si l'on se réfère à l'autocitation comme étant la citation d'une œuvre dont le chercheur est le principal auteur. Ce taux augmente si le terme est défini de manière à inclure les citations des publications dont le chercheur était à une autre position dans la liste des auteurs précise Garfield (1979). Il faut préciser que le terme d'autocitation n'est pas toujours défini de la même manière selon les évaluateurs (Garfield 1979 ; Moed et al. 1985). Ceci a pour effet de modifier le nombre de publications prises en considération dans l'analyse bibliométrique. Moed et al. (1985) font référence à ce qu'ils nomment le « cas de la biochimie » qui étudie des groupes de recherche de l'université de Leiden (cf. chapitre 4.2.2.2) et où deux évaluations ont été menées. La première à l'interne, la seconde par des enquêteurs sur les comités nationaux en biochimie. L'évaluation menée à l'interne ne considère pas comme autocitation la citation effectuée par un auteur principal qui n'était pas membre du personnel de recherche, mais aurait publié avec des membres des groupes de recherche de l'université de Leiden. Les enquêteurs sur les comités nationaux en biochimie entendent eux par autocitation la citation d'une publication dont un des auteurs cités apparaît comme auteur dans celle-ci. Costas et al. (2009) parlent de citations externes pour signaler l'exclusion des autocitations qu'ils définissent comme étant la citation effectuée par un des auteurs de la publication citée.

Afin de déterminer si cet indicateur est pertinent, il est important de comprendre pourquoi les chercheurs ont recours aux citations. Comme le relèvent Rousseau, Egghe et Guns (2018), il existe de justes et honnêtes fondements en matière de citation tout comme des pratiques qui ne peuvent pas être qualifiées de bienséantes. Si elles « [...] fournissent la lignée historique du savoir et sont le reflet d'une dette intellectuelle », les intentions derrière les citations peuvent être d'autres natures : « citation-récompense », « citation-politique », « citation-alibi », « citation-persuasion » ... (Le Coadic 2005, p. 19). Des chercheurs se sont penchés sur la classification des schémas de citation (Moravcsik et Murugesan 1975 puis Chubin et Moitra 1975, cités dans Rousseau, Egghe et Guns 2018) (cf. Annexes 16 et 17). D'autres ont cherché à déterminer le contexte dans lequel une publication est citée (White 2001, cité dans Rousseau, Egghe et Guns 2018) (cf.

Annexe 18). Nous regrettons de ne pas pouvoir détailler ces différentes visions lesquelles devraient faire l'objet d'une recherche spécifique. Comprendre le fonctionnement des citations revêt une importance cruciale en matière de bibliométrie : “[s]olving this puzzle is of eminent practical relevance for the evaluation practice of the near futur” précise Wouters (2014, p. 61). Or, quand bien même étudié depuis les années 1960, il n'a pas encore trouvé de réponse.

Avec le développement exponentiel de la science, Garfield (1979) estime qu'il devient indispensable d'être en mesure d'identifier les personnalités et groupes qui contribuent en plus grande partie à la science. S'il reconnaît que l'indicateur de citation des publications est avant tout un outil d'évaluation qui doit être manié avec précaution par des personnes avisées, il estime que :

“[...] citation measures have been demonstrated to be a valid form of peer judgment that introduces a useful element of objectivity into the evaluation process [...]”
(Garfield 1979, p. 373)

Les citations jouent un rôle primordial, bien que fortement critiqué, en matière de bibliométrie. Les citations des publications sont intégrées dans le *h*-index (Hirsch 2005) (cf. chapitre 4.4.3) et servent à mesurer le FI (cf. chapitre 4.4.4) :

“[...] both indicators have been naturalized as instantiations of quality irrespective of the methodological critiques by professional scientometricians, and even their inventors, Hirsch and Garfield, have not been able to limit the range of applications of the indicators.”
(Wouters 2014, p. 58)

Le nombre de citations peut être calculé i. par publications ; ii. par groupe de publications – à l'instar de celles considérées comme les plus importantes ; iii. pour l'ensemble des publications. Les autocitations peuvent, ou non, être prises en compte. L'université de Laval (2020) préconise de ne pas comparer le nombre de citations entre différents types de publications, citant l'exemple des actes de conférences qui sont plus rapidement cités que des articles après leur publication, mais dont les citations se tarissent ensuite plus rapidement. Rappelons l'importance de toujours préciser la manière dont l'indicateur est calculé et de définir les termes qui sont utilisés pour le mesurer.

4.4.2.3 Les indicateurs adaptés aux pratiques de publication et de citation des domaines scientifiques

Nous avons identifié les principales distinctions en termes d'habitudes de publication et de citations entre les domaines scientifiques (cf. chapitre 4.2.2). Nous estimons qu'il n'est dès lors pas pertinent de comparer les indicateurs précités entre domaines. Ils sont toutefois des indicateurs qui permettent l'évaluation des publications au sein d'un même domaine scientifique.

Pour Garfield (1979[b], cité dans Moed et al. 1985), les variations caractéristiques des citations comprennent i. la rapidité à laquelle une publication sera citée ii. le temps nécessaire pour que le taux de citation atteigne son apogée puis iii. le temps durant lequel la publication sera citée. Aussi estime-t-il que le nombre moyen de citations par publication par domaine est l'indicateur le plus précis du potentiel de citation au sein d'un domaine. Le nombre moyen de citations par publication varie bien évidemment pour des publications qui n'auraient pas été faites la même année (Waltman et al. 2011).

Basant son étude sur des groupes de chercheurs évoluant dans les domaines de la chimie et du génie chimique, Van Raan (2006) énumère puis teste un certain nombre d'indicateurs qu'il présente comme avancés : ils prennent en compte les habitudes de citation et de publication selon les domaines (cf. chapitre 4.2.2). Costas et al. (2009) reprennent ces mêmes indicateurs pour voir si les constats de Van Raan (2006) sont observables au niveau individuel du chercheur. Il ressort de ces deux études que la taille du domaine a une influence non négligeable sur la valeur des indicateurs bibliométriques. Les indicateurs sur lesquels Van Raan (2006) et Costas et al. (2009) basent leurs études comprennent, outre le nombre de publications P et le nombre de citations C : le nombre moyen de citation par publication (CPP), le *Journal Citation Score* (JCS) et le JCS moyen soit pondéré (JCS_m) (cf. chapitre 4.4.4.2), le *Field Citation Score* (FCS) sans les autocitations, la moyenne pondérée du score de citation spécifique au domaine (FCS_m) qui rend compte de la densité de citation et est plus pertinent lorsque différentes disciplines sont regroupées au sein d'un domaine, les ratios CPP/JCS_m puis CPP/FCS_m . Ce dernier indicateur est décrit comme étant le *crown indicator* par Van Raan (2006) et Costas et al. (2009) en ce sens qu'il permet d'observer où se situe la performance de l'unité d'analyse, respectivement le groupe de recherche ou le chercheur, par rapport à l'impact standard international dans le domaine⁴³. Notons qu'il est parfois, à tort selon Rousseau, Egghe et Guns (2018), nommé *Leiden crown indicator* ou *CWTS crown indicator* (pour Center for Science and Technology Studies de l'Université de Leiden). Waltman et al. (2011) relèvent que le CPP/FCS_m est évidemment bas pour les publications récentes. Aussi étudient-ils le recours à un autre indicateur, le *mean normalized citation score* ($MNCS$), qui accorde le même poids à toutes les publications, quelle que soit leur année de publication. Waltman et al. (2011) constatent que les publications récentes introduisent du bruit dans le $MNCS$, raison pour laquelle ils préconisent de ne pas en tenir compte dans ces deux indicateurs. Il existerait une

⁴³ Un $CPP/FCS_m < 0,5$ indique une performance nettement inférieure à l'impact standard global. Entre 0,5 et 0,8 elle est considérée comme similaire. Entre 0,8 et 1,2 elle est qualifiée de supérieure et devient bien supérieure au-delà de 1,2.

bonne corrélation entre le CPP/ FCS_m et la revue par les pairs (Rinia, van Leeuwen, van Vuren et van Raan 1998 et 2001, cités dans Costas et al. 2009). Les fenêtres de publication et de citation considérées ainsi que les types de publications intégrées au calcul des indicateurs doivent être les mêmes pour calculer le CPP, le JCS_m et le FCS_m . Ces divers indicateurs permettent de situer un chercheur par rapport à ses confrères dans un domaine donné. Les biais inhérents aux bases de données se répercutent sur ces indicateurs (cf. chapitre 4.2.4). En outre, en vue de comparer les données au sein d'un même domaine, il nous semble pertinent de voir si notre définition du domaine est la même que celle de la base de données sur laquelle nous effectuons notre analyse et le calcul de ces indicateurs. Quant aux modes de calculs en termes d'évaluation :

“In the context of research evaluation an AoR [average of ratios] is the better approach. Yet, we claim that the impact of a field can best be described as an RoA [ratios of average]. We recall that when geometric means are used instead of arithmetic ones, the AoR ver[s]us RoA problem disappears.”

(Rousseau, Egghe et Guns 2018, p. 272)

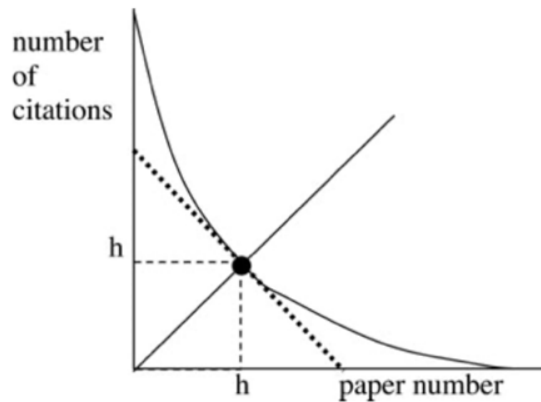
4.4.3 L'indice de Hirsch et ses variantes

4.4.3.1 Le h -index

Les données désormais disponibles grâce à l'archivage des publications et citations des chercheurs permettent de fournir des informations telles que « le nombre (N_p) de publications sur n années, le nombre de citations (N_c^j) de chaque publication (j), les revues dans lesquelles les articles ont été publiés, ... » (Hirsch 2005, p. 16'569). Partant des constats que i. ces données peuvent être évaluées en fonction de différents critères par différentes personnes et que ii. la mesure des fruits de la recherche est nécessaire en termes d'évaluation et de comparaison des chercheurs, Jorge E. Hirsch propose en 2005 d'avoir recours à une mesure unique pour quantifier la production et l'impact des résultats scientifiques d'un chercheur : le h -index. Il doit, selon Hirsch (2005), contrer les inconvénients des indicateurs bibliométriques standards (cf. chapitre 4.4.2). Son indice se base sur le nombre de publications du chercheur ainsi que sur le nombre de ses citations :

“A scientist has index h if h of his or her N_p papers have at least h citations each and the other ($N_p - h$) papers have $\leq h$ citations each.” (Hirsch 2005, p. 16569)

Figure 2 : Courbe schématique du nombre de citations en fonction du nombre d'articles (Hirsch 2005)



(Hirsch 2005, p. 16570)

Le h -index indique le nombre le plus élevé d'articles ayant reçu h ou plus de citations lequel permet d'estimer la visibilité et l'impact général de la contribution cumulative d'un chercheur (Schreiber 2007). Quand bien même le nombre d'articles ou de citations entre deux chercheurs diffère, leur impact scientifique global est considéré comme analogue par Hirsch (2005) du moment que leur h -index est similaire. Le rapprochement du h -index ne peut se faire qu'entre chercheurs d'un même domaine scientifique (Schreiber 2007).

Élevé au carré, le h -index donne une sous-estimation du nombre total de citations pour Van Raan (2006) qui confirme l'avis de Hirsch (2005) que le carré i . sous-estime le nombre total des citations des h articles les plus cités tout en ii . ignorant les articles qui ont moins de h citations. Cela s'expliquerait du fait de la distribution asymétrique des citations (C) sur les publications (P) décrite par la loi « $P(C) = \alpha C^s$, où la pente s et le facteur α pourrait être empiriquement établi à partir des données » (Van Raan 2006, p. 491). Aussi, Hirsch (2005) précise que h^2 donne la limite inférieure de ce nombre, raison pour laquelle il introduit une constante de proportionnalité a :

“The relation between $N_{c,tot}$ and h will depend on the detailed form of the particular distribution (1), and it is useful to define the proportionality constant a as : $N_{c,tot} = ah^2$ ”
(Hirsch 2005, p. 16'569)

Associant un indice m résultant de la pente de h par rapport au nombre d'années (n), soit $h \sim mn$, Hirsch (2005) considère que m pourrait constituer un critère pour comparer des chercheurs qui n'en sont pas au même stade dans leur carrière, du moment qu'ils maintiennent leur niveau de productivité.

“i. A value of $m \approx 1$ (i.e., an h index of 20 after 20years of scientific activity), characterizes a successful scientist. ii. A value of $m \approx 2$ (i.e., an h index of 40 after 20years of scientific activity), characterizes outstanding scientists, likely to be

found only at the top universities or major research laboratories. iii. A value of $m \approx 3$ or higher (i.e., an h index of 60 after 20 years, or 90 after 30 years), characterizes truly unique individuals.”
(Hirsch 2005, p. 16571)

Hirsch (2005) relève plusieurs des limites de son modèle. Dans un premier temps, il précise qu'un chiffre ne constitue qu'une estimation, il ne peut donc pas être considéré seul. En termes d'évaluation, le h -index doit être combiné avec d'autres facteurs afin de représenter toutes les facettes du profil du chercheur. Dans un second temps, Hirsch (2005) rappelle qu'en matière de décisions qui peuvent avoir un impact durable sur une carrière, qu'il peut toujours y avoir des exceptions à la règle. Troisième point et non des moindres sur lequel Hirsch (2005) attire l'attention est le fait que les valeurs h considérées comme références ("*typical*") diffèrent selon les domaines scientifiques du fait notamment du degré de développement et d'avancement des disciplines, du nombre de chercheurs par domaine et des différentes habitudes de publications et de citations (cf. chapitre 4.2.2).

Hirsch (2005) met également en garde contre les possibles mauvaises interprétations de l'index. S'il est convaincu qu'un h -index élevé reflète bien la performance, Hirsch (2005) précise que l'inverse n'est pas toujours avéré. Il faut, en outre, tenir compte de l'importante asymétrie de la distribution des citations :

“for an author with a relatively low h that has a few seminal papers with extraordinarily high citation counts, the h index will not fully reflect that scientist's accomplishments. Conversely, a scientist with a high h achieved mostly through papers with many coauthors would be treated overly kindly by his or her h .”
(Hirsch 2005, p. 16'571)

La valeur du h est donc plus élevée dans les domaines scientifiques où prédomine la collaboration et donc l'écriture à plusieurs. Hirsch (2005) préconise, lorsque le nombre de co-auteurs varie fortement, de normaliser h par un facteur reflétant le nombre moyen de co-auteurs afin de pouvoir réaliser une comparaison individuelle. Concernant la pratique de l'autocitation, conscient du fait qu'elle puisse augmenter la valeur du h , Hirsch (2005) estime toutefois que leur impact est moindre par rapport au nombre total de citations. Il distingue les autocitations dont l'effet doit être considéré comme insignifiant de celles dont l'effet sur h doit être corrigé ; relevant, par là même, la manière dont le chercheur pourrait alors augmenter son h -index :

“First, all self-citations to papers with $<h$ citations are irrelevant, as are the self-citations to papers with many more than h citations. To correct h for self-citations, one would consider the papers with number of citations just $>h$ and count the number of self-citations in each. If a paper with $h + n$ citations has $>n$ self-citations, it would be dropped from the h count, and h would drop by 1. As the other face of this coin, scientists intent in increasing their h index by self-citations would naturally target those papers with citations just $<h$.”
(Hirsch 2005, p. 16571)

Hirsch (2005) précise toutefois que le nombre de corrections ne concernerait que peu de publications, voire aucune. Relevons encore le fait que Hirsch (2005) a construit son indice et formulé ses observations et remarques sur des données issues du WoS. S'il peut être calculé avec les données émanant des diverses bases de données scientifiques aujourd'hui disponibles, il ne peut être cité sans préciser cette source ainsi que le moment où il a été calculé, du fait qu'il évolue logiquement dans le temps. Si nous avons déjà abordé la notion de demi-vie des revues et citations (cf. chapitre 3.3.1), Hirsch (2005) précise que si certains articles de début de carrière peuvent ne pas alimenter le h -index d'un chercheur tout au long de sa carrière, des publications peuvent encore être citées après son départ à la retraite, avant d'affirmer que le h -index ne diminuera toutefois jamais avec le temps. Hirsch (2005) ne manque pas de signaler que le h -index peut également être calculé pour un groupe de chercheurs.

4.4.3.2 Les variantes du h -index

L'indice Hirsch-Banks

Banks (2006) exploite le h -index en vue d'élaborer un indice qui permet de définir quels sont les sujets et composés en vogue dans le domaine de la physique. Il cherche à déterminer s'il est possible d'anticiper si un sujet de recherche aura un écho important auprès des pairs.

“The Hirsch-Banks index is defined in analogy to h as the highest number of papers in a particular field or on a specific compound that received h or more citations.”
(Schreiber 2007, p. 30002 : 1)

Banks (2006) déclare que la relation linéaire entre le h - b et le nombre d'années écoulées depuis la première publication est $h-b \sim mn$ - où m est le gradient. Il teste son approche sur les données du WoS. Banks (2006) parvient ainsi à distinguer les composés i. qui peuvent intéresser les chercheurs du domaine qui est représenté par une communauté relativement restreinte de scientifiques ($0 < m < 0.5$), ii. des composés pouvant être « brûlants » d'actualité et qui concernent une large communauté de chercheurs - s'ils ne présentent pas des caractéristiques qu'il qualifie d'exceptionnelles ($0.5 < m \leq 2$) et enfin iii. les composés dont le rayon d'influence va au-delà du domaine de la physique ($m > 2$) (cf. Annexe 19). Du point de vue des sujets, Banks (2006) distingue ceux pouvant être qualifiés de « brûlants » d'actualité ($m \geq 3$), de ceux qui sont toujours d'actualité, mais pour lesquels un nombre conséquent de recherches a déjà été effectué ($h-b \geq 100$ et $m \geq 3$) des sujets démodés ($h-b \geq 100$ et $m \leq 2$). Ainsi, en 2006, les sujets définis comme étant les plus « tendance » dans le domaine de la physique sont, pour Banks (2006), les nanotubes de carbone, les nanofils et les points quantiques (cf. Annexe 19). L'indice Hirsch-Banks (2006) permettrait donc d'évaluer l'état d'avancée de la recherche dans

un domaine donné sur un sujet ou sur un objet spécifique, et pourrait aussi selon son concepteur, servir à l'attribution de fonds pour la recherche.

Le h-index affiné par Schreiber

Schreiber (2007) souhaite démontrer que contrairement aux dires de Hirsch (2005), les conséquences des autocitations sur le h -index peuvent être importantes, ce d'autant plus pour les jeunes scientifiques dont le h -index est encore modeste. L'identification des autocitations prenant du temps, Schreiber (2007) analyse ses propres publications et exclut toutes les citations qu'il aurait lui-même introduites ou qui émaneraient de ses coauteurs. Son h -index baisse de 18%. Il reproduit l'expérience en se basant sur les publications d'un de ses confrères aînés puis d'un de ses cadets : leurs h -index diminuent respectivement de 13% et 46%. Les difficultés liées à la graphie des noms des chercheurs dans les bases de données (cf. chapitre 4.2.4) se répercutent dans l'analyse des autocitations rappelle Schreiber (2007) qui reconnaît qu'il est impossible de ne pas omettre quelques autocitations. Quand bien même Schreiber (2007) considère certaines autocitations comme étant légitimes (lorsqu'elles servent, par exemple, à ne pas ré-exprimer une idée déjà développée dans une publication antérieure), il existe des pratiques discutables. Ces dernières consistent en les autocitations effectuées lorsque le chercheur se réfère à ses publications antérieures pour des raisons de facilité estime Schreiber (2007) qui précise que le h -index :

"[...] is a single number which can be relatively easily enhanced by specifically citing those papers for which the citation count is closed to but below the critical value h ."
(Schreiber 2007, p. 3002:2)

Schreiber (2007) étudie trois manières de parfaire le h -index. La première consiste à corriger les autocitations effectuées par le chercheur lui-même (h_o) ; la seconde à supprimer les citations effectuées par des co-auteurs (h_c) ; la troisième consistant à déterminer l'autocitation entre plusieurs co-auteurs en étudiant chacune des publications citant (h_s). Leurs effets sur le h -index sont illustrés dans l'Annexe 20.

S'il recommande l'exclusion de tous types d'autocitations, Schreiber (2007) reconnaît que ce sont avant tout les auto-autocitations qui devraient systématiquement être écartées lors de l'évaluation du fait que ce sont celles qui sont le plus aisément manipulables par le principal intéressé. Schreiber (2007) admet l'intérêt de l'étude des citations au sein d'une communauté lorsque l'on cherche à identifier les domaines tendances ; le cas échéant, les autocitations constituent d'excellents témoignages.

Le g-index

Reconnaissant les avantages du h -index, Egghe (2006a) observe que sa robustesse le rend insensible aux articles pas ou peu cités ainsi qu'aux articles qui le seraient beaucoup plus. Sans remettre en question la pertinence de la traîne des articles les moins cités, Egghe (2006a) considère que le h -index devrait être sensible au niveau des publications les plus citées : une fois qu'un article atteint h , les citations ultérieures ne sont pas prises en considération, quel que soit le nombre de citations. Egghe (2006a) estime que la mesure de la qualité globale d'un chercheur – ou d'une revue - devrait intégrer la performance de ses meilleures publications et englober l'intégralité des citations : cela peut être réalisé en modifiant le h -index en g -index lequel conserve les avantages du précédent et corrige l'effet précité. Le g -index est défini par Egghe (2006a) comme étant le nombre le plus élevé de g publications qui ont reçu g^2 ou plus de citations.

“So for all authors or journals, the g -score will be higher than the h -score but, what is interesting in this, the higher the number of citations in the top-class (in other words, the skewer the citation distribution) the higher the g -score will be.”
(Egghe 2006a, p. 8)

Egghe (2006a) teste sa méthode sur ses publications ainsi que sur celles d'Henry Small en classant les publications par ordre décroissant de citations (TC), celle ayant recueilli le plus de citations étant en rang (r) 1, puis en indiquant $\sum TC$ qui désigne les scores cumulatifs jusqu'au rang r (cf. Annexe 21). Il constate que les différences entre les chercheurs sont plus visibles avec le g -index qu'elles ne le sont avec le h -index ce qui, pour Egghe (2006a) permet de faciliter la comparaison entre scientifiques évoluant dans un même domaine. Dans une seconde publication, Egghe (2006b) teste sa méthode sur les auteurs ayant reçu la médaille Price pour leurs travaux. Prenant les données dans le Web of Knowledge (WoK), Egghe (2006b) est conscient de ne pas avoir accès à l'intégralité de leurs publications. Il relève que :

[...] the g -index, as simple as the h -index (a single measure, containing publication and citation elements), contains more comparative information from the raw data than the h -index and resembles more the overall feeling of “visibility” or “life time achievement”.
(Egghe 2006b, p. 143)

Egghe (2006b) souhaite observer l'évolution relative de g/h et constate que : “[...] we see remarkable order changes with respect to the h - or g -orderings.” (Egghe 200b, p. 143) (cf. Annexe 22).

Le g-index affiné par Schreiber

Egghe (2006a et 2006b) ne tient pas compte des autocitations dans le g -index. Aussi Schreiber (2008) souhaite voir comment réagit le « nouvel » index si les autocitations sont supprimées. Se basant sur ses publications et celles de huit autres de ses collègues dans le WoS, il compare les h -index et g -index avec et sans les autocitations et constate que :

“[...] the g index after exclusion of self-citations allows for a better distinction between the impact of the publications of the different authors. Thus it is superior to the original g index, which in turn is superior to the h index.”

(Schreiber 2008, p. 189)

Rappelant que le h -index est vulnérable aux autocitations, Schreiber (2008) estime que le g -index l'y est moins : une modification raisonnable du nombre de citations d'une publication qui est proche de la valeur de g n'a, dans la major partie des cas, aucune incidence sur cet indice. Toutefois, le nombre d'autocitation produit un effet significatif sur le h -index.

À l'instar de ce qu'il a fait avec le h -index (Schreiber 2007), Schreiber (2008) applique trois types de correction des autocitations : la première (g_o) concerne les autocitations directes, soit celles réalisées par l'auteur lui-même ; la seconde (g_c) a trait aux autocitations effectuées par les co-auteurs et la troisième (g_s) qui cherche à déterminer les autocitations de tous les co-auteurs, sachant qu'il ne s'agit pas de les additionner : il arrive que certains des auteurs ont également écrit d'autres articles ensemble ; s'ils citent la première publication, cela compterait comme une autocitation pour chacun des auteurs. Dans ce cas, Schreiber (2008) précise qu'il faut vérifier chaque article cité pour déterminer s'il y a de multiples autocitations des auteurs, comme cela avait été fait pour le h -index. Leurs effets sur le g -index ainsi que sur le h -index sont illustrés dans l'Annexe 23. Schreiber (2008) conclut que les deux indices peuvent être affinés par la suppression des autocitations.

Aperçu d'autres variantes au h-index

Sans pouvoir présenter de manière détaillée chacune des variantes du h -index développées depuis 2005, nous présentons ici brièvement certaines d'entre elles. Vuillemin-Raval (2018) évoque l'existence du h -5 index qui couvre les publications et leurs citations pour une durée de cinq ans ainsi que celle du $i10$ -index de Google Scholar qui donne le nombre de publications qui ont reçu au moins dix citations. Bornmann, Mutz et Daniel (2008) définissent puis examinent plusieurs de ces nouveaux indices en vue de définir lesquels serviraient au mieux l'évaluation des chercheurs (cf. Annexe 24). Le

premier de ces indices étudié par Bornmann, Mutz et Daniel (2008) est le m quotient qui propose de diviser le h -index par le nombre d'années d'activité du chercheur, ce que Hirsch (2005) préconisait déjà. À l'instar du g -index, le $h(2)$ index tient compte des articles les plus cités : “[a] scientist’s $h(2)$ index is defined as the highest natural number such that his $h(2)$ most-cited papers received each at least $[h(2)]^2$ citations” (Kosmulski 2006, cité dans Bornmann, Mutz et Daniel 2008, p. 832). Pour Burrell (2007, cité dans Bornmann, Mutz et Daniel 2008), le h -index cherche à repérer le noyau le plus productif des publications d’un chercheur donné en fonction du nombre de citations reçues. Rousseau (2006, cité dans Bornmann, Mutz et Daniel 2008) parle du noyau de Hirsch. Le a -index n’est autre que le nombre moyen des citations dudit noyau (Bornmann, Mutz et Daniel 2008). Le noyau de Hirsch sert également à calculer le m -index proposé par Bornmann, Mutz et Daniel (2008) qui utilise non la moyenne arithmétique, mais la médiane, comme mesure de la tendance centrale. Jin et al. (2007, cités dans Bornmann, Mutz et Daniel 2008) proposent un r -index calculé par le biais de la racine carrée. Il se base également sur le noyau de Hirsch : il est ainsi très sensible à un nombre restreint de publications obtenant un nombre élevé de citations (Bornmann, Mutz et Daniel 2008). Adapté à partir du a -index, le ar -index tient compte de l’intensité des citations dans le noyau de Hirsch ainsi que de l’âge des citations qui s’y trouvent : pour Jin et al. (2007, cités dans Bornmann, Mutz et Daniel 2008), l’indice doit pouvoir être sensible aux variations de performance dans le temps en vue d’une évaluation pertinente. Jin et al. (2007, cités dans Bornmann, Mutz et Daniel 2008, p. 833) ont calculé divers indices pour Bertram C. Brookes qui a obtenu la médaille Price en 1989 : pour la période 2002-2007, son h -index est resté à 12, son r -index a augmenté de 5% tandis que son ar -index a régressé de 5%. Bornmann, Mutz et Daniel (2008) font également référence au h_w -index proposé en 2008 par Egghe et Rousseau qui serait le h -index pondéré par les citations. Bornmann, Mutz et Daniel (2008) testent ces différents indices sur plus de quatre cent post-doctorants candidats pour l’obtention d’une bourse *Boehringer Ingelheim Fonds* entre 1990 et 1995 et leurs plus de 1’500 articles publiés entre 1986 et 1994 qui ont été cités, selon le SCI, à plus de 60’800 reprises. Leur analyse les amène à distinguer deux types d’indices lesquels sont complémentaires :

“The one type of indices (h index, m quotient, g index and $h(2)$ index) describe the most productive core of the output of a scientist and tell us the number of papers in the core. The other indices (a index, m index, r index, ar index, and h_w index) depict the impact of the papers in the core.” (Bornmann, Mutz et Daniel 2008, p. 836)

Bornmann, Mutz et Daniel (2008) concluent qu’aux fins de l’évaluation, la combinaison de différents indices est recommandable, préconisant le recours à au moins un indice de chacun des deux types précités. Les variantes aux h -index sont plus nombreuses

que celles relevées ici. Certaines d'entre elles se retrouvent dans les trente-sept identifiées par la suite par Bornmann et al. (2011). Elles ont été catégorisées en quatre groupes par Egghe (2010, cité dans Bornmann et al. 2011) selon qu'elles tiennent compte i. des différents taux de citation entre les domaines scientifiques ; ii. de l'autocitation ; iii. des co-auteurs ; iv. de la durée de la carrière ou de l'âge de la publication (cf. Annexe 25).

4.4.4 Les indicateurs basés sur les revues

La spécialisation croissante de la science a engendré une augmentation du nombre de revues spécifiques aux divers domaines qui la composent. Elles cohabitent avec des revues considérées comme multi- ou transdisciplinaires, dont certaines jouissent d'une réputation internationale. Les revues considérées comme étant de notoriété se voient soumettre un nombre considérable de publications. D'après Lawrence (2003), le nombre d'articles soumis à *Nature* a doublé en l'espace de dix ans : il avance le chiffre de 9'000 manuscrits reçus annuellement, dont 95% seraient rejetés. La sélection des articles par les éditeurs peut augmenter l'impact de leurs revues précisent Rousseau, Egghe et Guns (2018) qui y voient une des expressions du *Matthew effect* (cf. chapitres 3.2 et 4.4.2.1). Le produit de la recherche scientifique est considéré sous l'angle de la quantité et de la qualité. Cette dernière est souvent appréciée au regard du statut de la revue ou de l'éditeur (Moosa 2018). Les revues peuvent être classées selon i. le nombre de citations ii. l'opinion ou appréciation iii. le marché iv. la fréquence de téléchargements précise Moosa (2018, p. 76). L'approche basée sur les citations est devenue le principal mode de classification des revues (Moosa 2018 ; Rousseau, Egghe et Guns 2018) :

“In this way, well-known (= heavily cited) scientists publish mainly and preferably in well-known (= heavily cited) journals.” (Rousseau, Egghe et Guns 2018, p. 156)

4.4.4.1 Le facteur d'impact

L'année du lancement du SCI (cf. chapitre 3.3.2), Garfield et Sher (1963) publient un article portant sur les nouveaux facteurs dans l'évaluation des publications scientifiques par le biais de l'indexation des citations. Relatant la diversité des données disponibles dans ledit index, ils introduisent le *journal impact factor*, soit le facteur d'impact (FI) d'une revue :

“One of the most interesting correlation is the “journal impact factor”. In the usual citation count methods, as e.g. Gross and Gross [1927], the importance of a journal is determined by the absolute number of citations to it.”
(Garfield et Sher 1963, p. 200)

Le FI est aujourd'hui considéré comme la mesure de la qualité des revues scientifiques la plus largement reconnue (Moosa et al. 2018). Il sert principalement à la sélection des revues dans les services d'information puis aux fins de l'évaluation de la performance

de la recherche (Ingwersen et al. 2001). Une enquête menée par Buéla-Casal et Zych (2012) auprès de plus de 1'700 chercheurs dans 86 pays révèle que près de 90% des participants le considèrent comme étant important, voire très important. Le FI demeure l'indicateur le plus critiqué en termes d'évaluation des publications (cf. chapitre 4.5.2.1). C'est pour cette raison que nous détaillons ici les différents types d'IF et leurs calculs respectifs en vue de saisir pourquoi il est frappé d'anathème en la matière.

Le FI est calculé au vu du nombre d'articles publiés annuellement dans une revue et du nombre de fois où ceux, pris en compte dans le calcul du FI, sont cités : “[e]ditorials, meeting abstracts, book reviews and similar publications are considered “uncitable.” (Rousseau, Egghe et Guns 2018, p. 157) durant la fenêtre de citation, soit l'intervalle de temps utilisé pour calculer le FI. Ingwersen et al. (2001) de rajouter que le choix de la fenêtre de publication est tout aussi important.

Tableau 2 : Matrice servant à définir les facteurs d'impact d'une revue scientifique

Year	1995	1996	1997	1998	1999	2000	Total
# publications	10	15	20	25	30	35	
# citations received in the year 1995	10						10
# citations received in the year 1996	20	10					30
# citations received in the year 1997	25	15	10				50
# citations received in the year 1998	30	19	12	10			71
# citations received in the year 1999	24	26	18	14	10		92
# citations received in the year 2000	23	28	29	17	13	10	120
Total	132	98	69	41	23	10	

(Ingwersen et al. 2001, p. 525)

À partir de cette matrice, plusieurs FI et mêmes types de FI peuvent être calculés. La principale distinction relève du fait qu'ils sont synchroniques, à savoir que les citations prises en considération pour calculer le FI ont été faites la même année ou diachroniques, lorsque les données utilisées pour calculer le FI reflètent différentes années et donc l'évolution dans le temps (Ingwersen et al. 2001 ; Rousseau, Egghe et Guns 2018).

À l'instar du FI des revues accessibles dans le JCR et annuellement généré par *Clarivate Analytics* (2020a)⁴⁴, le FI synchronique « standard », aussi dit de Garfield [-Sher], est calculé sur la base d'une année de citation pour deux années de publication⁴⁵ (Ingwersen et al. 2001 ; Rousseau, Egghe et Guns 2018).

$$IF_n(Y) = \frac{\sum_{i=1}^n CIT(Y, Y-i)}{\sum_{i=1}^n PUB(Y-i)}.$$

(Ingwersen et al. 2001, p. 525)

En reprenant les données du tableau 2, le FI₂ (soit calculé pour deux ans) de la revue pour l'année 1998 est de : 12 + 19 / 20 + 15 ≈ 0,889. Le FI₃ est quant à lui de : 12 + 19 + 30 / 20 + 15 + 10 ≈ 1,36.

Le FI diachronique (FID) (IMP dans la formule d'Ingwersen et al. 2001 reproduite ci-dessous) permet d'observer son évolution dans le temps.

$$IMP_n(Y) = \frac{\sum_{i=k}^{k+n-1} CIT(Y+i, Y)}{PUB(Y)},$$

(Ingwersen et al. 2001, p. 525)

Ingwersen et al. (2001) précisent deux modes de calcul distincts du FID : i avec les données des années subséquentes ; en reprenant celles du tableau 2, le FID₂ (1998) est de : 14 + 17 / 25 = 1,24 ; sachant que le plus souvent ii. on y inclut l'année de publication soit un FID₂⁰ (1998) de : 10 + 14 / 25 = 0,96. Comme pour le FI, le FID peut être calculé pour plus de deux ans.

Quel est le FI le plus approprié en matière d'évaluation ? Pour Ingwersen et al. (2001), le FI synchronique est plus robuste lorsqu'il s'agit d'évaluer la revue et son impact global ; le FID est par contre plus pertinent quand s'agit d'exprimer son impact actuel :

“Consequently when the article (or the scientist who wrote it) is being evaluated diachronous impact factors are more commonly employed.”
(Ingwersen et al. 2001, p. 526)

En outre précisent Ingwersen et al. (2001), le FID permet une évaluation plus équitable en ce sens que le recoupement se fait entre articles publiés la même année que la publication évaluée (*ndlr* : cet argument est selon nous valable lorsque le FID est calculé

⁴⁴ Depuis son édition de 2007, le JCR fournit désormais également le FI synchrone sur cinq ans (Rousseau, Egghe et Guns 2018).

⁴⁵ Le FI synchrone peut être basé sur plus de deux années ; la méthode de calcul ne change pour autant.

au regard de l'année de publication, cf. supra), différence d'autant plus notoire que la fenêtre de citation est élargie et où, contrairement au FI, le FID se base toujours sur le même groupe d'articles⁴⁶. Enfin, Ingwersen et al. (2001) estiment que seuls les articles pourraient faire l'objet d'une évaluation par le biais du FI alors que le FID permettrait de tenir compte de publications annuelles « ponctuelles » (*one-off publications*), à l'instar des actes de conférences qui représentent une source importante de publication dans certains domaines (cf. chapitre 4.2.2.1).

De par leur facilité d'accès, les IF sont couramment utilisés pour l'évaluation de la recherche précisent Ingwersen et al. (2001) qui ne manquent pas de souligner que leur utilisation malavisée induit des résultats incorrects : les FI des revues sont attribués aux scientifiques ou à leurs articles évalués. Les résultats ainsi obtenus ne peuvent pas être considérés comme valides stipulent Ingwersen et al. (2001) du fait que i. le FI des revues ne correspond pas à l'impact d'un article qui y est publié, tous les articles de la revue ne recevant le même nombre de citations, aussi : “[a]ssigning JCR’s journal impact factor to each article could well result in an unfair assessment of research impact.” (Ingwersen et al. 2001, p. 527) ; ii. ils sont des revues qui sont citées sur de plus longues périodes que d'autres après publication, leur FI est donc moins impacté par une courte période de citation ; iii. il n'est pas pertinent de comparer les FI entre domaines (cf. chapitre 4.2.2.2). Seglen (1992, cité dans Ingwersen et al. 2001) donne l'exemple de la médecine clinique qui cite des articles de la médecine fondamentale, l'inverse n'étant pas forcément observé, d'où un FI moyen plus élevé dans ce second domaine. Pour ces raisons, aux fins d'une évaluation plus objective, le recours au FI et à FID devrait être envisagé “as a measure of the expected impact of a group of articles” (Ingwersen et al. 2001, p. 527). Ingwersen et al. (2001) relèvent en outre la pertinence du recours au FCS (cf. chapitre 4.4.2.3) en vue d'une comparaison intégrée des FID.

4.4.4.2 Aperçu d'autres indicateurs basés sur les revues

Si leur nom peut laisser entendre qu'il s'agit d'autres types d'indicateurs, ils ont en réalité trait aux FI des revues. Citons le SCImago Journal Rank (SJR) qui exprime le nombre de citations reçues l'année n par les documents publiés les trois précédentes années ($n-1$, $n-2$, $n-3$) dans les journaux de la base de données *Scopus* (Elsevier) (SCImago [2020]). Certains font référence au *Journal Citation Score* (JCS) (Van Raan 2006 ; Šember, Utrobičić et Petrak 2010 ; Zweigbibliothek Medizin 2020) qui est un type de FI

⁴⁶ Un autre argument avancé en faveur du FID par Ingwersen et al. (2001) est le fait que d'une année à une autre, le contenu des revues peut varier, citant l'exemple d'une revue qui publie, l'année n , avant tout des actes de conférences, puis l'année $n+1$, des numéros thématiques, rendant ainsi la comparaison entre articles publiées en n puis en $n+1$ d'autant plus invraisemblable.

englobant généralement une période de citation plus longue (Zweighbibliothek Medizin 2020). Le JCS se calcule sur la base des données disponibles dans le WoS, Scopus et/ou Google Scholar. Recours peut être fait au *mean Journal Citation Score* (JCS_m), soit le taux de citation attendu pour les publications effectuées dans les revues dans lesquelles un chercheur ou groupe de recherche publie habituellement. Le ratio JCS_m/FCS_m indique l'impact relatif de la revue dans un domaine scientifique donné (Costas et al. 2009). L'Eigenfactor se fonde quant à lui sur des données provenant du JCR ainsi que sur près de 110'000 références documentaires auxquelles les revues qui y sont représentées font référence – qu'il s'agisse de livres, journaux... – (Bergstrom 2007). Il a pour objectif de permettre une analyse plus sophistiquée des données de citation en calculant le *Eigenfactor Score*, soit la mesure de l'importance totale d'une revue pour la communauté scientifique ainsi que l'*Article InfluenceTM Score*, l'influence moyenne, par article, des publications d'une revue qui s'apparente donc au FI (Bergstrom, West et Wiseman 2008). En 2016, Elsevier lance une nouvelle métrique : *CiteScore* qui donne le nombre de citations reçues par une revue pour les documents qu'elle a publiés les trois années précédentes, divisé par le nombre de documents indexés dans Scopus et publiés ces mêmes trois années (Zijlstra et McCullough 2016). Ces différents FI se basent donc sur des fenêtres de citation qui peuvent différer et sont calculés sur en ayant recours à différentes bases de données.

Bergstrom (2007) et Bergstrom, West et Wiseman (2008) estiment qu'il n'y a pas de manière plus adéquate d'évaluer un scientifique et ses publications que de lire et comprendre ces dernières et d'échanger avec le chercheur quant à son travail, ce qu'ils considèrent comme la référence absolue ou « étalon-or » (Bergstrom, West et Wiseman 2008, p. 11433). Les données de citation sont utiles, mais il s'agit d'en faire bon usage. Considérant que les mesures quantitatives ne sont pas pertinentes en matière d'évaluation individuelle des scientifiques, ils sont des questions auxquelles seule des analyses de grande ampleur peuvent répondre, c'est à ces fins qu'a été développé l'algorithme Eigenfactor (Bergstrom 2007). Aussi intéressant soit-il, nous considérons qu'il est plus pertinent aux fins d'études bibliométriques relationnelles, raison pour laquelle nous ne développerons pas son étude ici. Faut-il sinon encore rappeler que ces différents facteurs d'impact des revues sont calculés sur des bases de données et leurs données dérivées, dans le cas de l'Eigenfactor. Les biais inhérents à ces dernières (cf. chapitre 4.2.4) s'y répercutent dès lors. En outre, corrélativement aux habitudes de citations selon les disciplines (cf. chapitre 4.2.2.2), il n'est pas pertinent d'avoir recours au FI pour comparer l'impact des revues entre elles (Dutta 2014).

4.5 Les critiques formulées à l'encontre des indicateurs bibliométriques classiques

Les voix s'élevant contre le recours à des indicateurs bibliométriques en matière d'évaluation des publications scientifiques sont diverses et nombreuses. Certaines avancent des arguments ayant trait à la manière dont ces indicateurs sont appliqués. D'autres dénoncent la non-pertinence de certains d'entre eux en la matière. Parfois, ce sont l'interprétation ou l'utilisation des résultats qui sont dénoncés. Enfin, ils sont des chercheurs qui critiquent, d'une manière plus globale, le fonctionnement du système de publication scientifique. Dans le présent chapitre, nous revenons sur les critiques formulées à l'encontre de la bibliométrie et plus spécifiquement sur certains de ses indicateurs et des biais qu'ils engendrent. Dans un second temps, nous relevons les principales critiques formulées émises contre les indicateurs que sont le *h*-index et le FI. Considérant que le sujet qu'est le fonctionnement des publications scientifiques est bien plus vaste que le cœur de notre recherche, nous n'avons pas la prétention de procéder à l'état de l'art de ses critiques. Nous ne ferons que rappeler certains des arguments qui auraient une incidence directe sur les évaluations bibliométriques. Rappelons que nous distinguons les études des évaluations bibliométriques (cf. chapitre 4).

4.5.1 Les critiques sur le fond

Dans le présent chapitre, nous ne revenons pas sur les bases de données dont nous avons situé les origines (cf. chapitre 3.3.2) et les biais (cf. chapitre 4.2.4). Nous ne revenons pas non plus sur les métriques de citations (cf. chapitre 4.4.2.2).

La « quantofrénésie » et ses dérivés

Ou « quand les métriques prennent le pas sur l'appréciation qualitative des fruits de la recherche ». Les critiques face aux nombres et mesures sont de deux ordres. D'une part, certains dénoncent la suprématie du chiffre et les dérivés engendrés par une société qui serait avant tout conduite par la performance et l'audit. D'autre part, les critiques se tournent vers la méconnaissance des chiffres, de leur manipulation, de leur signification.

L'une des principales critiques du modèle bibliométrique classique est que l'évaluation ne se résume pas aux seuls comptages et nombres (Couzinet 2018) et son attention exacerbée sur ce qui peut être mesuré, au détriment du reste (Wilsdon et al. 2017). Une évaluation serait désormais guidée par les données plutôt que par l'appréciation (Hicks et al. 2015), ce que Le Coadic (2005, p. 16) nomme la « quantofrénésie » ou « religion du nombre ». Le *Manifeste de Leiden* (Hicks et al. 2015) considère que “[r]esearch evaluations that were once bespoke and performed by peers are now routine and reliant

on metrics” (Hicks et al. 2015, p. 429). Van Raan (2006) insiste sur le nécessaire recours à plusieurs indicateurs en vue de renseigner sur les différents aspects de la productivité ; en ayant recours à une seule mesure, il y voit même un danger : laisser entendre aux décisionnaires stratégiques ou administrateurs que la performance scientifique peut s’exprimer au travers d’une seule note.

Aux yeux de certains, l’omnipotence des métriques nourrit une culture excessive de l’audit (Collini 2016 puis Martin 2016, cités dans Wilsdon et al. 2017). Lawrence (2003) dénonce cette désormais “société de l’audit” dans laquelle les indicateurs de performance deviennent une fin en soi. Cette « philosophie » est, selon Lawrence (2003) désormais également partagée par les chercheurs qui considèrent leurs confrères non sur la base du message scientifique qu’ils véhiculent, mais en fonction du prestige des revues dans lesquelles ils publient. Pour Lawrence (2003), les mesures d’évaluation des chercheurs – nombre d’articles, position dans les listes d’auteurs et le facteur d’impact des revues – sont devenus des objectifs en soi.

La place prépondérante des indicateurs de performance dans la gouvernance bouleverse la dynamique de la production scientifique (Wouters 2014). Ils augmentent la pression mise sur les acteurs et détournent des effets souhaités : répondre aux critères de performance. Ils peuvent générer des effets pervers :

“The research community responds strategically, and this may in turn create unintended effects, either through the mechanism of goal displacement or through more structural changes in research priorities, publication activities, or research capacity and organization.”
(Wouters 2014, p. 55)

Certains estiment que les modes d’évaluation de la recherche, quantitatifs comme qualitatifs, tendent à influencer le comportement des scientifiques s’ils sont corrélés à l’octroi de fonds ou ont un impact sur la réputation des chercheurs (Butler 2007, cité dans Wouters 2014). Un de ces effets peut être d’augmenter la quantité des publications au détriment de leur qualité (Colwell et al. 2012, cités dans Wouters 2014) (cf. chapitre 3.3.5), bien que cela ne soit pas toujours observé (Georghiou et al. 2000, cités dans Wouters 2014). Ce n’est pas le nombre en lui-même qui est remis en question, mais le risque de sa suprématie, de ses effets en termes de gouvernance au sein de certaines institutions et de son incidence sur la pratique de quelques scientifiques. Partant de la nécessaire évaluation de « la qualité et l’impact de la production scientifique » pour les bailleurs de fonds et hautes écoles, cette dernière doit être « mesurée rigoureusement et évaluée avec discernement » (*Déclaration de Dora* 2012 §1). Le *Manifeste de Leiden* (Hicks et al. 2015) souligne les risques causés par la méconnaissance de l’application et de l’interprétation des métriques :

“Metrics have proliferated: usually well intentioned, not always well informed, often ill applied. We risk damaging the system with the very tools designed to improve it [...]”
(Hicks et al. 2015, p. 429)

Pansu, Dubois et Beauvois (2013) insistent sur le fait qu’il faut savoir ce que mesure réellement la bibliométrie. Le *Manifeste de Leiden* (Hicks et al. 2015) exprime parfaitement plusieurs des critiques formulées à l’encontre des indicateurs bibliométriques. Qui souhaite y avoir recours doit savoir ce qu’ils mesurent et comment ils se calculent. Nous considérons que détenir les réponses à ces questions est indispensable, non seulement pour les acteurs qui y ont recours comme outils d’évaluation ou d’étude, mais également pour les chercheurs dont la production scientifique est évaluée avec l’un ou l’autre de ces outils. Ou comme dit Le Coadic (2005) :

« Face à cette religion du nombre, à cette quantofrénésie, on trouve en général des hommes et des femmes toujours aussi désarmés dans leur compréhension des nombres et leur maniement, des hommes et des femmes qui par ailleurs peuvent être fort cultivés mais aussi tout à fait capables de perdre le sens des réalités ou celui des mots quand il s’agit de nombres. » (Le Coadic 2005, p. 16)

Pour Gingras (2014b), l’anarchie qui peut régner dans le domaine de l’évaluation serait également corrélée au lancement de bases de données ou réseaux sociaux gratuits et non contrôlés. Ce ne sont dès lors plus exclusivement des spécialistes en bibliométrie et des pairs avisés qui y font leurs mesures. On ne s’improvise pas analyste ou évaluateur. S’il n’y a pas besoin de compétences particulières pour collecter des données, leur manipulation et leur interprétation requièrent un certain savoir-faire (Rousseau, Egghe et Guns 2018). Mesurer et communiquer des résultats provenant de bases de données dont on ne connaît pas le fonctionnement et générer puis comparer des indicateurs sans en connaître les fondements et les finalités diminue la confiance et ternit l’image de la bibliométrie.

Le besoin de mesures de la qualité

Ces premières critiques relevées signalent également le manque ressenti quant à l’évaluation qualitative. La *Déclaration de Dora* (2012) en appelle à une meilleure évaluation de la qualité des fruits de la recherche. L’évaluation qualitative des publications se fait aujourd’hui quasi exclusivement par les pairs. Le mode d’évaluation traditionnel (Abramo et al. 2011 et Bornmann 2011 puis Moed 2007, cités dans Mingers et Leydesdorff 2015). “Scientific peer review is the evaluation of scientific research findings or proposals for competence, significance and originality, by qualified experts (peers) who research and submit work for publication in the same field.” (Brown 2004, cité dans Rousseau, Egghe et Guns 2018, p. 33). Rousseau, Egghe et Guns (2018)

précisent qu'il ne se limite pas aux revues scientifiques et peut également se déployer lors de l'examen de demandes de subventions ou encore lors des différentes étapes d'une carrière académique (promotion, évaluation des performances...). Les retours des confrères émanant d'autres domaines et plus globalement de la société civile peuvent-ils participer à l'évaluation ? Les webometrics et altmetrics seront-ils un jour considérés comme des indicateurs qualitatifs ? (cf. chapitre 4.6.1). Le temps nous le dira. Malgré les critiques formulées à l'encontre des pairs, nous sommes comme Bourdieu (1976, p. 91) d'avis que : « [...] seuls des savants engagés dans le même jeu ont les moyens de s'approprier symboliquement l'œuvre scientifique et d'en évaluer les mérites ». Notre recherche se concentrant sur les modes d'évaluation quantitatifs, nous ne procédons pas à l'état de l'art de l'évaluation par les pairs. Relevons toutefois certaines des limites qui sont avancées : la revue par les pairs est parfois considérée comme longue et coûteuse (Abramo et D'Angelo 2011, cités dans Mingers et Leydesdorff 2015), sujet à des biais et distorsions (Horrobin 1990 puis Moxham et Anderson 1992, cités dans Mingers et Leydesdorff 2015), globalement opaque (Reale, Barbara et Costantini 2007, cités dans Mingers et Leydesdorff 2015).

La nécessaire définition des termes et des outils de l'évaluation

Une autre des problématiques relevées est l'absence de définitions claires. Chacun des termes de l'étude ou de l'évaluation doit être clairement définis. Tant les évaluateurs que les chercheurs doivent savoir ce qui est évalué. En cas de recours aux indicateurs bibliométriques : quel indicateur est utilisé, quelle est la source des données, quelles sont les fenêtres de citation et de publication. Nous avons déjà abordé les effets de l'absence de définition claire quant à l'autocitation (cf. chapitre 4.4.2.2). Couzinet (2018) cite l'exemple des évaluateurs qui cherchent à déterminer la productivité des chercheurs, sans savoir avec précision ce qui est entendu par là.

Les publications qui sortent du périmètre de l'évaluation

Nous avons déjà traité la question des habitudes de publications et de citations selon les domaines (cf. chapitre 4.2.2). Reconnaisant que les articles revus par les pairs demeurent des produits scientifiques essentiels, la *Déclaration de Dora* (2012) insiste sur le fait que de nouvelles composantes de la production scientifique doivent désormais aussi être évaluées (cf. chapitre 4.6.2). Les données de la recherche en font partie. La *Déclaration de Dora* (2012) estime en outre qu'il nous faut tirer parti du Web, en termes de publication et d'évaluation (cf. chapitre 4.6.1).

Moosa (2018) rappelle que dans certaines facultés les publications qui ne sont pas des articles sortent de la liste des publications considérées comme *outputs* de la recherche. « Un publiant est [...] quelqu'un qui publie sous une certaine forme et pas sous une autre » (Gingras 2014a, p. 40). Dans certains domaines, l'écriture d'un livre serait considérée comme une « perte de temps » (Moosa 2018, p. 34). Dans les domaines de SHS, certains considèrent comme « réducteur et inapproprié » le fait de s'en tenir aux nombres de publications et de citations des articles (Perret et Kadelbach 2017, p. 4). Les livres y constituent une part conséquente de la production des chercheurs et sont considérés comme les pierres angulaires de courants de pensées. Or, estime Gingras (2014a), ne pouvant attribuer le FI aux livres, ce mode de publication est dévalorisé. Cette vision d'un type limité de publications par domaine, écartant de facto les publications qui ne rentrent pas dans ce spectre, est anti-productive. Lawrence (2003) dénonce la pression exercée dans certains domaines de la recherche pour publier dans des revues considérées comme éminentes. Rousseau, Egghe et Guns (2018) avancent que le fait de n'évaluer que les publications qui se font dans un groupe restreint de revues, considérées comme étant les plus réputées, devrait être aboli : ils sont des publications influentes qui se font dans d'autres revues et toutes les revues de renom ne publient pas que d'excellents articles. « [...] les pratiques sont adaptées aux critères, alors que ce devrait être l'inverse » (Gingras 2014a, p. 40). Comprenant les raisons amenant à publier dans une revue où un article aura le plus de probabilités d'être lu, Lawrence (2003) dénonce toutefois la priorité qui leur est accordée allant jusqu'à dire "we turn ourselves into philistines in our own world." (Lawrence 2008, p. 259). Selon Moosa (2018), dans le contexte du POP (cf. chapitre 3.3.5), la pression est telle pour publier dans une revue considérée comme prestigieuse dans son domaine de recherche, qu'elle entrave le développement des connaissances nationales. Prenant l'exemple du domaine de l'économie, Moosa (2018) cite l'exemple de l'Australie où la revue la plus réputée est l'*Economic Record*, toutefois moins cotée que le *Journal of Economic Behaviour and Organization* américain que lui-même considère moins que la revue australienne. Afin de stimuler leur positionnement, des chercheurs australiens seraient plus enclins à travailler sur des problématiques américaines ou internationales, au détriment de travaux pouvant directement intéresser leur pays. (Moosa 2018). Ce culte des revues considérées comme prestigieuse serait-il aussi alimenté par la presse ? Lawrence (2003) cite l'exemple du journal *El Pais* qui relayerait exclusivement les articles scientifiques des chercheurs espagnols qui publient dans *Nature*, *Cell* ou *Science*. Très souvent, les ouvrages de vulgarisation sortent du périmètre de l'évaluation. Couzinet (2018) estime que la communication avec la société dans son ensemble peut être vue comme un retour sur investissement du financement de la

recherche. Toutefois, ces types de publication ne sont pas considérés comme des travaux de recherche : l'ouverture vers la société se ferait donc au détriment de la recherche « pure » (Couzinet 2018). La pertinence sociale de la recherche est parfois intégrée aux critères d'évaluation de la recherche (cf. chapitre 5.2). Cet aspect de la recherche devrait être porté à l'actif de l'évaluation qualitative. Les instruments bibliométriques donnent des pistes quant aux aspects quantitatifs de la production scientifique. Il nous apparaît que nombre des critiques contre la bibliométrie ont trait au fait que celle-ci donne un certain nombre d'indicateurs qui sont dénués de nuance. Ce n'est pas leur objectif. Mais ce manque de profondeur de l'évaluation peut heurter certaines sensibilités. L'analyse qualitative est, à notre avis, du ressort des pairs, que ce soit en amont ou en aval de toute publication. Wouters (2014) souhaite le recours à une évaluation éclairée des pairs qui combine leur avis prépondérant quant à la qualité et le recours à des preuves bibliométriques. La complexité réside dans la combinaison des données qualitatives et quantitatives précise Wouters (2014) qui rappelle que la question est laissée à la discrétion des pairs. Le concept d'évaluation éclairée par les pairs n'étant pas encore défini, Wouters (2014) avertit que nous ne pouvons déjà nous assurer du fait que la combinaison des indicateurs bibliométriques et de la revue par les pairs permettra d'accroître la qualité de l'évaluation, estimant même que dans certaines disciplines : "[...] particular disciplinary biases may be reinforced rather than weakened, especially when subdisciplines have diverging publication or citation cultures" (Wouters 2014, p. 59). Rousseau, Egghe et Guns (2018) préconisent de baser l'évaluation sur une liste des publications fournie par le chercheur.

La démultiplication artificielle du nombre de publications

Selon Lawrence (2003), les chercheurs se doivent de publier le plus rapidement possible. Il s'agit d'éviter le risque de ne pas être publié du fait qu'une recherche similaire a déjà été portée à la connaissance des lecteurs. Les éditeurs peuvent être réticents à l'idée de publier un article portant sur un thème similaire à un article publié par un de leurs concurrents. En cas de recherches similaires, les articles doivent donc être publiés en même temps. Lawrence (2003) dénonce également le découpage d'articles en vue de leur soumission à différentes revues en vue d'augmenter le nombre d'articles à l'actif du chercheur. À cette même fin, Lawrence (2003) relève par ailleurs cette pratique qui consiste à publier un premier article suivi d'un second qui sert à corriger les erreurs qui auraient été présentes dans le premier. Ces usages ont pour effet d'augmenter le nombre d'articles d'un chercheur et ce de manière artificielle. Aussi ont-ils une conséquence directe sur l'évaluation des publications du moment que tous les articles d'un même auteur sur une recherche donnée sont comptabilisés plusieurs fois.

4.5.2 Les indicateurs dont la pertinence n'est pas largement reconnue

4.5.2.1 Le facteur d'impact des revues scientifiques

En réponse aux critiques concernant le calcul du nombre de citations, Garfield (1979) réfléchit déjà à la pertinence de la pondération du nombre de citations en regard du prestige des revues. Il doute que la réputation des revues augmente la chance d'être cité, du fait que les services d'indexation et d'*abstract* permettent d'augmenter la visibilité des articles publiés dans la majeure partie des revues. En outre, estimant que les publications suffisamment citées pour « faire la différence » sont publiées dans un groupe réduit de revues qui bénéficie dès lors d'un haut degré de prestige et de visibilité, Garfield (1979) se demande :

“Is not the prestige of the journal that published the citing work just as important as the one that published the cited work?”
(Garfield 1979, p. 365)

Le FI est sans conteste l'indicateur le plus controversé dans le domaine. Le *Manifeste de Leiden* (Hicks et al. 2015) parle de cette obsession du FI, illustrant l'évolution des mentions dans les publications (cf. Annexe 27). La controverse autour du FI est de deux ordres précisent Wilsdon et al. (2017) : i. la manière dont il est calculé ; ii. la manière dont il est utilisé dans l'évaluation de la recherche. Nous avons précédemment expliqué ce qu'est le FI et les manières dont il est calculé (cf. chapitre 4.4.4). Cet indicateur renseigne quant à l'impact ou à la visibilité d'une revue, calculé en fonction du nombre de citations que reçoivent les articles qu'elle publie. Le FI a été développé en vue de permettre la sélection des revues auxquelles les bibliothèques vont s'abonner, « non pour mesurer la qualité scientifique de la recherche exposée dans un article » (*Déclaration de Dora 2012, § 3*).

La principale critique provient du détournement du FI, à savoir son recours en tant qu'indicateur de l'impact d'un article, alors utilisé comme substitut du nombre de citations d'une publication, ce qui est erroné (Wilsdon et al. 2017). Tous les articles publiés dans une revue ne seront pas cités de manière identique. De plus, “not every paper published in a top journal is high-quality paper” (Moosa 2018, p. 7). Dans leur étude portant sur le domaine de la pédiatrie, Jelacic Kadic et al. (2020) observent que :

“Articles that use observational research methodology, and are published in journals with lower impact factors, can become citation classics.”
(Jelacic Kadic et al. 2020, p. 7)

En matière d'évaluation, Balaban (2012) avise que certaines publications ne seraient pas considérées du moment que la valeur du FI serait inférieure à un certain seuil. Les articles publiés dans des revues à fort FI ayant plus de chance d'être cités, Balaban

(2012) suggère d'inverser la pratique en attribuant plus d'importance, dans l'évaluation, aux citations d'articles publiés dans des revues à faible FI. Certains spécialistes du domaine n'écarteraient pas pour autant le recours aux indicateurs d'impact des revues dans le cadre de l'évaluation des publications récentes du fait que leur taux de citation n'est pas encore conséquent (Wilsdon et al. 2015). La *Déclaration de Dora* (2012) relève quatre faiblesses dans de FI :

« A) les distributions des citations dans les revues sont très asymétriques [...] ; B) les propriétés du facteur d'impact sont propres à chaque domaine : il s'agit d'un agrégat de types d'articles multiples et très divers, avec des articles primaires comme de synthèse [...] ; C) les facteurs d'impact peuvent être manipulés (ou « instrumentalisés ») par une politique éditoriale [...] ; et D) les données utilisées pour calculer les facteurs d'impact ne sont ni transparentes ni ouvertement accessibles au public. »
(*Déclaration de Dora 2012, [p. 1]*)

La relative facilité avec laquelle le FI peut, intentionnellement ou non, être manipulé par le biais de politiques et stratégies éditoriales a été analysée par Reedijk et Moed (2008). S'il leur a été impossible d'évaluer empiriquement la fréquence de ces comportements, Reedijk et Moed (2008) préconisent la prudence en cas de recours au FI. Moosa (2018) d'ajouter que le FI est biaisé : en faveur des revues américaines, en défaveur des revues non-anglophones et des jeunes revues avant de préciser qu'il ne peut être utilisé que pour comparer des revues au sein d'un même domaine.

À propos du recours au FI dans le cadre de l'évaluation de la recherche, Clarivate Analytics (2020a) rappelle qu'il n'est autre qu'une approximation brute du rayonnement de la revue et qu'il devrait être combiné avec d'autres mesures et l'appréciation par les pairs. La *Déclaration de Dora* (2012) demande à ce que les indicateurs basés sur les revues ne servent plus dans le cadre de l'évaluation des publications et chercheurs, particulièrement lorsqu'elle est corrélée à des décisions impactant la carrière scientifique ou l'octroi de fonds (point 1 de la *Déclaration de Dora 2012*). Elle préconise que la recherche soit désormais évaluée au regard de sa « valeur intrinsèque », non de l'aura des revues (*Déclaration de Dora 2012, [p. 2]*). Les éditeurs devraient développer des indicateurs qui permettent d'évaluer un article sur son contenu, non sur la revue (point 7 de la *Déclaration de Dora 2012*). Ils se voient en outre demander d'abandonner le FI comme outil de promotion, sinon du moins y associer d'autres indicateurs relatifs aux revues tels que :

« [...] [FI] sur 5 ans, EigenFactor, [SJR], indice h, temps de traitement éditorial et de publication, etc. [...] qui offrent une vision plus riche de la performance d'une revue. »
(*Déclaration de Dora 2012, point 6*)

Au vu des arguments avancés, ne faudrait-il pas désormais clairement « [...] distingue[r] [...] deux types d'évaluation, celle des auteurs et celle des sources, c'est-à-dire des revues » (Durand-Barthez 2009, p. 128) ?

4.5.2.2 La validité du *h*-index

Rappelant que Hirsch (2005) a développé le *h*-index (cf. chapitre 4.4.3.1) du fait que le seul nombre de publications d'un chercheur ne peut pas être l'unique gage de qualité au vu de la culture POP (cf. chapitre 3.3.5), Gingras (2014b) évalue la validité du *h*-index à la lecture des trois critères de validité des indicateurs (cf. chapitre 4.4.1.2). Le souhait de Hirsch (2005) est de contrebalancer le poids du seul recours au nombre de publications en le combinant avec le nombre de citations. Gingras (2014b) qualifie d'ironique cette forte corrélation avec le nombre de publications qui détermine de par nature le *h*-index qui, dès lors, ne remplit pas son premier critère de validité. Le *h*-index ne remplit pas non plus le critère d'homogénéité proposé par Gingras (2014b) du fait qu'il est une mesure composite du nombre d'articles publiés et du nombre de citations. Il ne peut dès lors pas être considéré comme un indicateur de la quantité et de la qualité.

“Interestingly, its heterogeneous combination of quantity and “quality” confuses the author himself about what his h-index really measures, for the title of Hirsch’s paper was “An Index to Quantify an Individual’s Scientific Research Output” (my italics)”.
(Gingras 2014b, p. 118)

Gingras (2014b) reconnaît que malgré ses failles, le *h*-index est largement relayé par la communauté scientifique et désormais intégré dans nombre de bases de données. Du fait qu'il le considère comme invalide, Gingras (2014b) prévient de son effet pervers en matière d'évaluation :

“In case of the h-index, we have situations where a lower index often hides a better-quality researcher.”
(Gingras 2014b, p. 119)

Toujours selon Gingras (2014b), les failles du *h*-index ne peuvent pas être corrigées du fait qu'il ne présente pas de relation intuitive avec ce qu'il cherche à mesurer ; aussi toute opération qui serait effectuée sur cet indice “only make an invalid indicator even less transparent” (Gingras 2014b, p. 119).

Une autre critique formulée à l'encontre du *h*-index est qu'il est de facto plus élevé pour un chercheur avancé dans sa carrière, quand bien même il n'aurait pas publié de façon continue (Hicks et al. 2015). Il ne permet pas de comparer des chercheurs qui n'en sont pas au même stade de développement de leur carrière (Vuillemin-Raval 2018). En outre, le *Manifeste de Leiden* (Hicks et al. 2015) rappelle que sa valeur varie en fonction de la base de données qui l'a généré. D'où l'importance de toujours préciser la source du *h*-index que l'on cite.

4.6 Les nouvelles tendances en matière de mesure et d'évaluation

En parallèle des métriques et indicateurs bibliométriques classiques, l'avènement de l'informatique et le déploiement de l'OS soutiennent le développement et la mise en œuvre de nouveaux instruments potentiels d'évaluation. L'OS impacte tout le cycle de la recherche, de la conceptualisation du projet à la publication de ses résultats. Il entraîne le passage des pratiques de publications telles que nous les connaissions jusqu'à présent au partage des connaissances et données le plus tôt possible dans le processus de recherche (Wilsdon et al. 2017) (cf. Annexe 29). Si nous ne disposons pas de suffisamment de recul pour pouvoir procéder aux lectures et analyses critiques des nouvelles métriques potentielles que nous présentons ici, nous estimons nécessaire de relever tant les nouveaux supports de publication scientifique que les possibles nouvelles mesures de la production scientifique, en vue de nos recommandations finales.

4.6.1 Le développement de nouvelles mesures de la production scientifique sur le Web

Alors que les métriques de citations représentent une importante mesure de l'impact, la National Information Standards Organization (NISO) (2016) rappelle leurs limites au regard des nouvelles pratiques de communication scientifique sur le Web. La définition des métriques évolue vers :

" [...] indicators based on recorded events of acts (e.g., viewing, reading, saving, diffusing, mentioning, citing, reusing, modifying) related to scholarly documents (e.g., papers, books, blog posts, datasets, code) or scholarly agents (e.g., researchers, universities, funders, journals)."
(Haustein 2016, cité dans Coombs et al. 2018, p. 2)

C'est à la fin des années 1990 que le terme de web(o)metrics est introduit par Almind et Ingwersen (1997, cités par Rousseau, Egghe et Guns 2018). La webométrie est définie comme :

"the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Internet drawing on bibliometric and informetric approaches."
(Rousseau, Egghe et Guns 2018, p. 3)

Elle ne consiste donc pas au recours aux sources sur le Web, mais à l'étude de leur utilisation précisent Rousseau, Egghe et Guns (2018). Le *Web Impact Factor*, introduit par Ingwersen (1998, cité dans Dutta 2014), est l'une des mesures développées à ces fins. Il est calculé au regard du nombre de sites internet qui font référence à d'autres sites. Moins de quinze ans plus tard, Priem et al. (2010) déclarent que des métriques alternatives, les *altmetrics* doivent venir compléter les mesures d'évaluation existantes ;

la revue par les pairs et le recours au nombre de citations étant devenus trop lents pour l'évaluation de la science, des chercheurs et de leurs publications, d'autant plus que le recours au FI est le plus souvent inadéquat en la matière (cf. chapitre 4.5.2.1). Partant de l'observation qu'une partie non négligeable de la communication scientifique se fait sur internet et ne se limite plus aux articles et aux monographies, que les logiciels de référencement bibliographiques tels que Zotero ou Mendeley constituent d'extraordinaires réservoirs de références qui dépasseraient désormais les bases de données scientifiques : "Altmetrics expand our view of what impact looks like, but also of what's making the impact" (Priem et al. 2010⁴⁷). De nouvelles formes de communication se mettent en place en place. Wouters (2014) considère qu'il faut dorénavant tenir compte : i. du partage de la « science brute » (*raw science*) tels que les sets de données et codes ; ii. de nouveaux formats de publication à l'instar des nanopublications puis de la généralisation du recours à l'auto-publication. La NISO (2016) a dressé une liste des résultats de la recherche qui regroupe les canaux traditionnels de publications scientifiques et les nouvelles alternatives⁴⁸. Pour la NISO (2016), tous pourraient entrer dans le champ de l'évaluation et servir à mesurer l'impact de la recherche, estimant qu'il doit désormais être évalué au-delà des voies de publications conventionnelles.

Les *altmetrics* s'intègrent aux *webometrics* (Wilsdon et al. 2017) et sont un sous-champ de l'infométrie (Rousseau, Egghe et Guns 2018) (cf. Annexe 28). Les mesures traditionnelles de la science et de ses activités peuvent désormais être complétées par les données disponibles sur le Web. La NISO (2016) considère que les *altmetrics* donnent un aperçu de l'impact en mesurant la portée, la pertinence sociale ainsi que l'attention d'une communauté spécifique. Les *altmetrics* permettent de capturer l'impact via les métriques que sont le nombre de vues, de téléchargements, de sauvegardes ou enregistrements (Mendeley...) d'échanges, de recommandations (F1000Prime...) et de citations, y compris dans des sources qui ne sont pas considérées comme primaires (Lin et Fenner 2013 ; Taylor 2013, cités dans Rousseau, Egghe et Guns 2018). Ils sont essentiellement basés sur les médias sociaux comme les blogs, Twitter, ResearchGate ou encore Mendeley précisent Wilsdon et al. (2017). Colledge (2017) intègre les *altmetrics* dans les mesures de l'impact social de la production scientifique, au même titre que l'engagement public ou la reconnaissance académique, les distinguant par là-même des indicateurs de publication et de citation. L'impact de la recherche a des retombées qui dépassent le cadre strictement scientifique. Cet aspect intéresse l'Open Science : « [...] l'évaluation des impacts sociétaux a gagné en importance parallèlement

⁴⁷ <http://altmetrics.org/manifesto/> [consulté le 6 juillet 2020].

⁴⁸ <https://sites.google.com/a/niso.org/scholarlyoutputs/> [consulté le 6 juillet 2020].

à celle des qualités de la recherche » (Wilsdon et al. 2017, p. 6) (cf. chapitre 4.6.2). En matière d'OS, les métriques, *altmetrics* incluses, peuvent soutenir l'ouverture à la science et permettre l'évaluation des performances des chercheurs considèrent Wilsdon et al. (2017). Elles doivent être utilisées de manière responsable comme le préconisent la *Déclaration de Dora* (2012), le *Manifeste de Leiden* (Hicks et al. 2015) ou encore le *Metric Tide* (Wilsdon et al. 2015).

Les métriques d'utilisation, qui se situent entre les métriques classiques et les *altmetrics*, se mesurent au nombre de vues ou de téléchargements et permettent d'obtenir des données qui ne se reflètent pas via les citations⁴⁹ ainsi que des données sur les publications non traditionnelles, données et logiciels ouverts inclus (Wilsdon et al. 2017). Wilsdon et al. (2017) estiment que les métriques d'utilisation constituent de potentielles mesures quantitatives de l'influence des chercheurs devant être utilisées de manière responsable et pouvant encore être améliorées.

Haustein et al. (2016, cités dans Wilsdon et al. 2017) ont élaboré un cadre conceptuel dans lequel ils définissent les actes donnant lieu à des événements qui sont enregistrés et peuvent dès lors être mesurés ; les actes sont catégorisés en trois niveaux d'implication du « visiteur » : Accès – Évaluation – Mise en application. L'impact de la recherche est souvent évalué au niveau de la revue. Certains éditeurs indiquent désormais le nombre de téléchargements d'un article (Wilsdon et al. 2017). L'éditeur OA PLOS (*Public Library of Science*) fournit un excellent exemple de l'application de ces nouvelles métriques au niveau des articles. En 2009, il est le premier à fournir des *Article-level metrics* (ALM) qui renseignent les chercheurs quant à l'impact de leur article au regard : de leurs utilisations online, citations et aux *altmetrics*, ce avant même que ce terme ne soit utilisé (Lin et Fenner 2013). Wilsdon et al. (2017) considèrent que les *altmetrics* peuvent soutenir le développement des principes de l'OS tels que la collaboration, le réseautage et bien évidemment le partage. Ils présenteraient en outre, précisent Wilsdon et al. (2017), un potentiel dans l'évaluation de la recherche transdisciplinaire et de l'impact de la recherche sur la société dans son ensemble. S'ils sont souvent considérés pour leur apport quantitatif, les *altmetrics* offriraient la possibilité d'analyses qualitatives selon Wilsdon et al. (2017) qui citent l'exemple de l'analyse du contenu des commentaires effectués ou des profils des utilisateurs.

Sachant que les livres sont sous-représentés dans les bases de données scientifiques et qu'ils constituent un mode de publication apprécié dans des disciplines des SHS

⁴⁹ À l'instar du FI d'utilisation (Kurtz et Bollen 2008, cités dans Wilsdon et al. 2017) et des Libcitations (White et al. 2009).

notamment, relevons le développement de métriques basées sur ces types de publications. Citons l'exemple de *Bookmetrix* développé par Springer Nature et *Almetric* (Digital Science & Research Ltd. 2015), qui permet aux auteurs de découvrir le nombre de citations, mentions en ligne et téléchargements de leurs ouvrages et des chapitres qui les composent (Springer [2020]). Pour White et al. (2009), en sélectionnant les livres pour compléter leurs collections, les bibliothèques créent un instrument qui permet d'évaluer leur impact. Les *Libcitations* sont une autre de ces nouvelles métriques orientées vers les livres. Les catalogues des bibliothèques sont ici considérés comme des « index de citation des bibliothécaires » (White et al. 2009, p. 1084). Les données proviennent de catalogues collectifs, nationaux ou internationaux, et permettent de connaître le nombre de bibliothèques qui possèdent un ouvrage.

Rappelons que nous sommes en faveur du recours généralisé aux identifiants pérennes (cf. chapitre 4.2.4). Ceux-ci permettent d'établir le lien entre l'auteur et ses œuvres, de suivre leur utilisation et d'agréger des données provenant de différentes sources. Ils permettent de même une analyse transversale et interdisciplinaire des champs de recherche de l'auteur. La curation des données nous apparaît comme évidente, à l'instar de celles provenant de bases de données scientifiques. Mesurer la science sur la Toile nous apparaît comme un potentiel complément aux outils bibliométriques classiques. Tout un chacun ne peut pas pour autant s'improviser évaluateur et nous ne pouvons pas faire fi des possibles biais d'interprétation et des polémiques sans fondements. Des réactions et commentaires peuvent être effectués de manière anonyme, ce qui nous paraît être inadéquat en matière d'évaluation objective. Si certaines informations, comme le score Mendeleev, sont moins susceptibles d'être manipulés, le danger de voir manipulés des résultats *altmetrics* est supérieur à celui de la manipulation des citations avertissent Rousseau, Egghe et Guns (2018). Comme pour tout indicateur, il s'agit de savoir comment ils sont construits et ce qu'ils sont censés mesurer. L'audience est-elle toujours synonyme de qualité ? Rousseau, Egghe et Guns (2018) estiment que l'évaluation contemporaine devrait intégrer les nouveaux modes de communication scientifique ainsi que les répercussions sociales de la recherche et penchent pour une approche multi-métriques. S'ils n'ont pas de frontières géographiques, ces outils sont récents et nous n'avons pas encore suffisamment de recul pour pouvoir évaluer leur pertinence. Wilsdon et al. (2017) précisent que les études menées sur les *altmetrics* sont encore essentiellement empiriques. La NISO (2016) précise qu'à l'instar de tout recours à des indicateurs, les *altmetrics* doivent être combinés à des données qualitatives.

4.6.2 La refonte de l'évaluation dans le cadre de l'Open Science

L'OS regroupe l'OA et ses différentes voies de publication (cf. Annexe 30) et l'ORD (cf. Annexe 31). Les chercheurs dont la recherche est financée par des fonds publics doivent, en sus de la publication en OA, archiver leurs données dans un dépôt à but non lucratif remplissant les critères FAIR en vue de leur partage (ORD) (Université de Lausanne [2019]). Des textes fondateurs de l'OA, deux font référence à l'évaluation des futures publications OA. Les déclarations du groupe de travail des institutions et bailleurs de fonds du *Bethesda Statement on Open Access publishing* (2003) précisent, en leur point quatre, que dans le cadre des évaluations pour les nominations, promotions et subventions, les publications en OA seront considérées comme un indicateur des services rendus à la communauté ; les déclarations du groupe de travail des scientifiques et sociétés scientifiques préconisent, en leur cinquième point, que les chercheurs plaident en faveur de changements dans l'évaluation aux fins de promotions ou de titularisations allant dans le sens de la reconnaissance de la valeur de leurs articles individuels, sans tenir compte des revues dans lesquelles ils publient. Les dispositions finales de la *Déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales* (2003) font référence à deux types d'évaluation devant être mis en œuvre dans le cadre de l'OA : l'évaluation au sens du monitoring de l'OA puis la prise en compte des publications OA dans le cadre des évaluations à des fins de promotions. Le Gold OA (cf. Annexe 30) implique de ne plus publier dans les revues classiques, ce qui a des conséquences directes sur l'évaluation et la réputation des chercheurs et prétérite d'autant plus les jeunes chercheurs (swissuniversities 2018b). À l'occasion des dix ans de la parution de la *Budapest Open Access Initiative* (BOAI), la BOAI10 (2012) plaide contre le recours au FI et en appelle à l'élaboration de « métriques alternatives visant la mesure de l'impact et de la qualité, qui soient moins simplistes, plus fiables et complètement ouvertes à l'utilisation et la réutilisation » (BOAI 10 2012, point 1.5).

Au niveau européen, le *Directorate-General for Research and Innovation* de la Commission européenne (DG-RTD) (2016) annonce que le système d'évaluation doit s'adapter à la nouvelle dynamique qu'est l'OS. Le DG-RTS (2016) précise que les chercheurs s'attendent à ce que tous les produits de la recherche soient désormais pris en considération dans leur évaluation. En Suisse, la révision du système d'évaluation des publications fait partie intégrante de la Stratégie nationale suisse sur l'OA (swissuniversities 2017) et de son Plan d'action (swissuniversities 2018b). La Stratégie nationale sur l'OA rappelle que les initiatives en vue du déploiement de l'OA « représentent [...] une possibilité de signaler les failles et les conséquences des

mécanismes actuels d'évaluation et de réputation dans la science » (swissuniversities 2017, p. 1). Parmi ses cinq principes directeurs se trouve le « système de révision de l'évaluation de la qualité » : il rappelle les failles et limites de l'actuel système d'évaluation, fait référence à la *Déclaration de Dora* (2012), en appelle à la prise en considération de l'OA et de l'OS, sans manquer de signaler l'importance de l'adhésion des chercheurs au système d'évaluation à venir (swissuniversities 2017, p. 3). Le troisième point de ses lignes d'action portant sur la coordination et le regroupement des ressources relève qu'« [u]ne attention particulière doit être portée sur la gestion de la qualité [et] l'évaluation par les pairs [...] » (swissuniversities 2017, p. 4).

Dans les mesures proposées dans le Plan d'action de mise en œuvre de la Stratégie nationale sur l'OA se trouvent « la participation à une réforme de l'évaluation de la recherche » qui s'inspirera de *Dora* (2012) et du *Manifeste de Leiden* (Hicks et al. 2015) ainsi que des acquis en la matière de précédents programmes de swissuniversities⁵⁰ (swissuniversities 2018b, p. 3). Le catalogue des mesures annexé au Plan d'action (swissuniversities 2018b) apporte d'intéressantes précisions et propositions en vue de la refonte du système d'évaluation des publications à venir qui s'inscrivent dans la « mise en œuvre de la ligne d'action 4 : Des formes alternatives de publication »⁵¹ (swissuniversities 2018b, p. 9 de l'Annexe 2). Du point de vue de la réforme de l'évaluation, orientée tant vers les chercheurs que leurs projets, swissuniversities (2018b) réitère son souhait de mettre l'accent non sur la revue ou l'éditeur, mais sur la qualité du travail, que l'évaluation soit menée aux fins de nomination, de promotion académique ou de sélection de projets. Les propositions de swissuniversities (2018b) vont plus loin (cf. tableau 3).

⁵⁰ « Mesurer les performances de la recherche (2008- 2011/12) et Performances de la recherche en sciences humaines et sociales (2013-2016) [soit le CUS 2013-2016 P-3] » (swissuniversities 2018b, p. 3).

⁵¹ Celles-ci se divisent en trois pans « soutenir [...] le basculement de journaux vers l'[OA] (4.1) ; [...] le lancement de nouvelles formes de publication (4.2) [...] de réformer l'évaluation de la recherche de manière à tenir compte de l'Open Access (4.3). » (swissuniversities 2018b, p. 9 de l'Annexe 2).

Tableau 3 : Premières préconisations de swissuniversities en termes d'évaluation dans le cadre de l'OS

Do	Don't
Demander aux chercheurs de soumettre un nombre limité de publications	Récompenser les longues listes de publication
Exiger des membres des comités d'évaluation de lire effectivement ces publications	Se concentrer sur les journaux à fort FI
Honorer les efforts pour l'Open Science, en particulier la publication en Open Access (Green et Gold)	Se concentrer sur les maisons d'édition renommées
Donner la préférence à une analyse des contenus. Pour les approches quantitatives, utiliser des métriques modernes et appropriées (et en aucun cas le FI)	

(adapté de swissuniversities 2018b, p. 11 de l'Annexe 2)

La révision du système d'évaluation a déjà été abordée dans le programme P3 de swissuniversities (2018a) : son rapport de conclusion présente les résultats de deux programmes menés dans les hautes écoles nationales entre 2007 et 2016 portant sur la visibilité de la qualité et de l'impact de la recherche dans les domaines des SHS et le développement de nouveaux instruments d'évaluation. L'un des fruits de ces programmes est l'élaboration d'une liste de dix thèses pour une évaluation performante de la recherche posant des bases de discussion au sein des organes de swissuniversities (2018a). Celles-ci en appellent à la prise en considération des distinctions entre les disciplines (thèse 1), des unités de recherche et institutions (thèse 2), préconisent une approche bottom-up dans l'élaboration et l'application des procédures d'évaluation (thèse 3), une définition claire et transparente des finalités et méthodologies (thèse 4), une communication efficiente favorisant la participation de toutes les parties prenantes (thèse 5), un recours aux instruments quantitatifs « [...] dans les cas où des informations qualitatives sont disponibles au même moment et que celles-ci permettent de hiérarchiser et de contextualiser les résultats » (thèse 6), la prise en considération des différentes formes d'impacts sociétaux selon les disciplines (thèse 7), l'adaptation de l'évaluation et de ses instruments aux contextes (thèse 8), le nécessaire recul temporel en matière d'évaluation de l'impact de la recherche (thèse 9) et enfin de mener des évaluations qui répondent aux « exigences de qualité scientifique » et évoluent grâce aux partages de connaissances entre les acteurs engagés (thèse 10) (swissuniversities 2018a, pp. 40-41). La poursuite de l'étude de l'évaluation est au planning de la phase A de la mise en œuvre du plan d'action OA au cours du prochain programme de

swissuniversities : *Open Science I* (2021-2024). La proposition de projet soumise au Conseil des hautes écoles exprime le souhait de :

« [...] réviser la manière dont les scientifiques et les hautes écoles sont évalués au regard de leurs réalisations en matière de recherche, indépendamment du nombre ou du support des publications. » (swissuniversities 2020, p. 6)

Nous constatons que la mise en œuvre progressive de ce nouveau paradigme qu'est l'OS en appelle à une refonte des modes d'évaluation des publications pour les raisons évoquées dans ce travail. Il s'agit dans un premier temps d'intégrer les publications OA à l'évaluation, quelle que soit la voie choisie. Dans ce cadre, l'évaluation sert autant d'incitatif que de valorisation de l'OS. Une autre motivation est d'intégrer à l'évaluation les *outputs* de la recherche qui n'entraient pas jusqu'alors pas dans son champ tels que les jeux de données et, d'une manière plus globale, la communication digitale des fruits de la recherche. Des initiatives sont élaborées en vue de permettre la transition. La ligue des Bibliothèques européennes de recherche (LIBER) a créé un groupe de travail portant sur les "*innovative metrics*" dont l'un des objectifs est de formuler des recommandations sur l'évaluation de la recherche qui prennent en compte les nouveaux canaux de communication digitale ainsi que du cadre posé par l'OS (Coombs et al. 2018). Coombs et al. (2018) proposent de définir les métriques ouvertes comme les indicateurs de la production et de l'impact de données ouvertes. Par données, il faut entendre :

"[...] (bibliographic) metadata, data related to scholarly outputs (e.g., number of published articles), data related to the impact of scholarly works (e.g., number of citations), and qualitative information about engagement with scholarly works and other stakeholders of the research enterprise (e.g., demographic information about users tweeting a scientific article and the tweet containing the DOI to the article)." (Coombs et al. 2018, p. 2)

Les recommandations de Coombs et al. (2018) en matière d'évaluation sont de quatre ordres. Dans un premier temps, les objectifs de l'évaluation doivent être fixés en amont de la sélection des bases de données et métriques pertinentes. Les évaluateurs doivent clairement identifier et définir ce qu'ils cherchent à mesurer. Tant les métriques, que les indicateurs et sources de données doivent correspondre à ce que l'on cherche à mesurer. L'utilisation et l'interprétation des données doivent être effectuées en toute transparence. Les données brutes et agrégées ayant servi à l'évaluation doivent être mises à disposition des autres parties. Enfin, l'évaluation doit tenir compte du contexte des activités de recherche et de publication du chercheur. Dans un second temps, Coombs et al. (2018) préconisent d'inclure dans l'évaluation les différents canaux de communication et de publication savantes des disciplines scientifiques en ayant recours aux différentes bases de données y relatives. Dans un troisième temps, Coombs et al.

(2018) conseillent de baser l'évaluation sur des sources de données objectives, indépendantes et communément admises en vue de fournir des mesures fiables et transparentes. Le recours à des données contrôlées puis harmonisées provenant de différentes sources augmente l'objectivité de l'évaluation estiment Coombs et al. (2018). Tout élément considéré comme subjectif doit être notifié en toute transparence. Enfin, Coombs et al. (2018) conseillent de ne pas avoir recours à des indicateurs composites et de veiller à ne pas confondre les différents aspects du travail et de l'impact scientifique ; les évaluateurs devraient avoir recours à différents indicateurs en vue de permettre une interprétation multidimensionnelle des résultats.

Les nouveaux outils et paradigmes présentés dans le présent chapitre n'en sont pas encore au stade de la maturité. Ils offrent un extraordinaire terrain de réflexion pour l'évaluation de demain. La question de savoir comment les indicateurs classiques et les nouvelles métriques peuvent se combiner pour permettre une évaluation plus objective est au cœur de la réflexion de projets tels que le *Snowball Metrics* ❄️ (Colledge 2017). Son comité de pilotage, regroupant plusieurs hautes écoles anglaises ainsi qu'Elsevier, cherche à déterminer si une norme internationale de mesure de la recherche ne pourrait pas être déduite des innombrables systèmes institutionnels développés au fil du temps en vue de permettre un alignement de l'évaluation.

5. Étude des pratiques en matière d'évaluation des publications scientifiques au sein des institutions académiques

Pour Wouters (2014), la progression de la bibliométrie s'accorde avec une nouvelle tendance de la revue avisée par les pairs. Ces derniers posent le cadre de l'évaluation qui se voit complété d'outils bibliométriques. La bibliométrie devient un outil d'analyse et de gestion ordinaire au sein des institutions académiques (Wouters 2014). Partant de ce postulat, l'objectif du présent chapitre est de présenter certains des systèmes d'évaluation des publications dans le domaine académique. Dans un premier temps, il s'agit de déterminer si, et dans quelle mesure, l'évaluation des publications des chercheurs en Suisse romande intègre la bibliométrie. Il est également question de savoir si les pratiques sont homogènes ou si, au contraire, des tendances divergentes se dessinent. Dans un deuxième temps, nous relevons des procédures et principes d'évaluation observés ou suggérés. Ce chapitre termine par les recommandations aux institutions académiques formulées dans la *Déclaration de Dora* (2012) puis par Anna Hatch, l'actuelle directrice du programme DORA, et sa consœur Ruth Schmidt.

5.1 L'évaluation des publications au sein des facultés et domaines échantillonnés

L'article 27 LEHE stipule que les hautes écoles suisses doivent régulièrement procéder à des contrôles de qualité de l'enseignement, des prestations de service et de la recherche. Les institutions académiques doivent « veill[er] à l'assurance et au développement de la qualité à long terme » (art. 27 LEHE *in fine*). L'« accréditation d'institution » est corrélée à la mise en place d'un « système d'assurance de la qualité » lequel doit s'assurer de la qualité de la recherche, en plus de celui de l'enseignement et des prestations de service de la haute école (art. 30 al. 1 let. a ch. 1 LEHE). Dans la pratique, nous observons que l'évaluation des publications scientifiques s'inscrit dans le cadre d'évaluations qui ne s'y bornent pas. Comme décrit plus haut, ces dernières ont trait, au niveau des institutions académiques, des facultés et des domaines échantillonnés, aux nominations et promotions académiques et/ou à l'octroi de fonds pour la recherche.

Au fil de notre enquête et pour chacun des contextes dans lesquels peut intervenir l'évaluation des publications, nous avons cherché à obtenir les informations suivantes (cf. Annexe 32) :

- Le niveau auquel se fait l'évaluation: chercheur et/ou groupes de recherche ;

- Le cadre dans lequel intervient l'évaluation : nomination, promotion académique et/ou octroi de fonds pour la recherche ;
- Le dépôt d'une publication dans l'archive institutionnelle est-il une condition pour que celle-ci intègre l'évaluation. Dans le cas contraire, d'où provient la liste des publications évaluées ;
- Les types de publications entrant dans le champ de l'évaluation ;
- Les critères ou indicateurs bibliométriques utilisés aux fins de l'évaluation des publications.

L'Annexe 33 présente la synthèse comparative des données concernant l'évaluation des publications au sein des institutions académiques analysées (cf. chapitre 2.4.1.2). Elles proviennent de la littérature grise et blanche en notre possession, et/ou des informations communiquées par les facultés et les domaines échantillonnés ayant pu répondre à notre requête. La granularité des informations varie d'une faculté ou domaine à l'autre. Nous ne prétendons pas détenir l'intégralité des données relatives à l'évaluation des publications en leur sein. Nous avons dû adapter notre mode de collecte à la singularité de la période que nous traversons (COVID-19) (cf. chapitre 2.5). Les données en notre possession sont toutefois suffisantes pour procéder à une première analyse comparative.

Des cinq institutions académiques approchées pour notre enquête, trois ont pu répondre à notre sollicitation : la HES-SO, l'UNIL et l'UNIGE. De tous les domaines et facultés contactés, nous avons un taux de réponse de 50%. La question de savoir si des procédures d'évaluation des publications établies existent bien au sein de chacune des facultés suisses peut se poser. Toutes les hautes écoles échantillonnées ont signé la *Déclaration de Dora* (2012). Il n'est pas exclu que la discrétion de certaines facultés à propos de leurs modes d'évaluation des publications soit corrélée à la transposition en cours de l'esprit de *Dora* (2012) dans leurs procédures. Le Graduate Institute de Genève n'a pas pu répondre à notre sollicitation du fait de la refonte de son processus d'évaluation des publications en cours. Des quatre institutions académiques choisies dans le domaine de la médecine et de la santé, nous disposons d'informations quant à l'évaluation des publications de trois d'entre elles : celles de l'UNIL, de l'UNIGE et de la HES-SO. Relevons le fait que de notre population échantillonnée, tous domaines confondus, la Faculté de biologie et de médecine (FBM) de l'UNIL ([2020a]) est la seule à fournir de l'information blanche quant à son recours à la bibliométrie. Des quatre facultés et domaines échantillonnés dans le champ de l'économie et du management, la HES-SO est la seule institution dont nous avons obtenu des données. Des trois facultés des lettres approchées, celles de l'UNIGE puis de l'UNIL ont répondu.

À la question de savoir dans quel cadre intervient l'évaluation des publications, relevons le fait que concernant la HES-SO, nous ne présentons que les données qui relatives à la prise en considération des publications dans le cadre de l'attribution de fonds pour la recherche. Celle-ci est menée au niveau des domaines en vertu des dispositions d'application des règlements de leurs fonds de recherche et d'impulsion (FRI) respectifs (Rectorat de la HES-SO 2016, 2017a et 2017b). Les écoles des domaines constituant la HES-SO perçoivent un financement de base composé d'une part fixe puis d'une part variable basée sur la performance des écoles en matière de recherche. L'évaluation des publications intervient dans le cadre du calcul et de l'attribution de cette part financière dite variable. Dans le domaine de la Santé, les projets et les publications se voient attribuer des points qui déterminent le montant accordé aux différentes écoles (Conseil de domaine Santé de la HES-SO 2018). Le domaine Économie et Services se base quant à lui sur huit indicateurs (I) ayant chacun leur propre pondération (Conseil de domaine Économie et Services de la HES-SO 2015). Trois indicateurs concernent les publications : I1 (« [...] full papers dans des conférences internationales »), I2 (« [...] articles dans les revues scientifiques ») et I4 (« [...] apparitions médias, publications et présentations publiques, au niveau régional et national [...]») (art. 15 des dispositions d'application du règlement du FRI applicables au domaine, Rectorat de la HES-SO 2017b). Concernant la possible évaluation des publications dans les contextes de nomination et de promotion académique, l'institution dispose de règles communes pour le personnel d'enseignement et de recherche (Convention intercantonale sur la HES-SO du 26 mai 2011). Ces dernières renseignent quant aux qualifications à l'engagement et aux fonctions et missions du personnel de recherche (Comité gouvernemental de la HES-SO 2014), non quant aux modalités d'engagement et de renouvellement des contrats. Ces dernières seraient élaborées au niveau cantonal. Ainsi, le *Règlement interne sur le personnel de la Haute école spécialisée de Suisse occidentale – Genève* stipule en son art. 34 que les décisions relatives à l'engagement, au renouvellement et à la fin des rapports de travail sont prises par la direction générale, sur proposition de la direction des écoles (Conseil de direction de la HES-SO – Genève 2017). Le guide des ressources humaines (RH) relatif à l'engagement préconise toutefois de considérer la qualité, non la quantité des publications, et de tenir compte du profil complet et de l'âge des candidats au moment de l'évaluation de leurs publications (HES-SO [2020b]).

Des facultés et domaines contactés et des informations en notre possession, seule la Faculté de médecine de l'UNIGE procède à l'évaluation des publications des groupes de recherche (recherche fondamentale et recherche clinique). Celle-ci s'intègre au système d'évaluation *MIMOSA*, créé en 1999 (Vos et Monnet 2007), qui attribue des

fonds pour la recherche. Les critères d'évaluation *MIMOSA* sont en cours de refonte. Les indicateurs aujourd'hui utilisés pourraient ne plus l'être dans un proche avenir. L'évaluation des publications des chercheurs est sinon réalisée au niveau individuel. Exception faite de la HES-SO, pour les raisons susmentionnées, elle intègre les processus de nomination et de promotion académique.

À la question de savoir si les publications intégrant l'évaluation doivent avoir été déposées dans l'archive institutionnelle, la FBM de l'UNIL et *MIMOSA* en font une condition. Si nous ne disposons pas de cette information pour le domaine Santé de la HES-SO, le domaine Économie et Services de l'institution requiert l'archivage des publications intégrant les indicateurs I1 et I2 dans ArODES. Les publications et les apparitions dans les médias (I4) doivent quant à elles être transmises à l'Infothèque, la bibliothèque de la Haute École de Gestion de Genève. La Faculté des lettres de l'UNIL corrèlera l'évaluation des publications au dépôt dans SERVAL dès la fin 2020. La Faculté des lettres de l'UNIGE et la Faculté de médecine de l'UNIGE, en termes d'évaluation individuelle des chercheurs (nomination ou promotion), n'exigent pas aujourd'hui le dépôt des publications. Dans ces trois cas, le *curriculum vitae* (CV) du chercheur sert de base à l'évaluation de ses publications.

Concernant les types de publications qui entrent dans le champ de l'évaluation, nous pouvons distinguer deux courants : i. les facultés qui considèrent l'ensemble des publications des chercheurs et des groupes de recherche ; ii. les facultés et domaines qui précisent les types de publications évalués. Qu'elle concerne le chercheur ou les groupes de recherche, la Faculté de médecine de l'UNIGE examine l'intégralité des listes de publications. Pour ce qui est des groupes de recherche, ce sont les publications des trois dernières années qui entrent dans le champ de l'évaluation *MIMOSA* du moment qu'un des membres du groupe de recherche figure dans la liste des auteurs et que la publication mentionne l'UNIGE ou les Hôpitaux Universitaires de Genève (HUG). L'évaluation des publications des groupes de recherche fondamentale et clinique étant corrélée au dépôt des publications dans l'Archive ouverte de l'Université de Genève, une nuance doit être apportée ici. La Directive de la Faculté de médecine de l'UNIGE portant sur l'archive ouverte de l'Université précise les cinq types de publications qui peuvent être déposées ainsi que les critères auxquelles elles doivent répondre (Bureau décanal élargi de la Faculté de médecine de l'UNIGE 2012). Les types de publications concernés sont les articles, les livres et les parties de livres, les contributions à des congrès, les thèses et les mémoires de maîtrise, enfin les rapports et les "*working papers*". Dans les Facultés des lettres de l'UNIGE et de l'UNIL, le dossier du chercheur est considéré dans son intégralité. À l'UNIGE, relevons que les œuvres de vulgarisation et les traductions

entrent dans le champ de l'évaluation. Par traduction, il faut entendre les œuvres traduites par le chercheur ainsi que la traduction de ses publications. Dans le cadre des procédures de nomination dans ces deux facultés des lettres, les publications de certains candidats sont lues par les commissions. À l'UNIGE, les candidats retenus pour le second tour présentent une monographie et deux articles qui seront lus par les membres de la commission de nomination. Parmi les domaines et facultés qui précisent quelles sont les publications qui intègrent l'évaluation, nous observons trois approches distinctes. i. Le Conseil de domaine Santé de la HES-SO (2018) nomme les articles, *peer-reviewed* ou non, les livres et chapitres de livres. ii. La FBM de l'UNIL ([2020d]) précise les types de publications concernés par l'évaluation avant de spécifier quelles sont les publications qui entrent dans leurs définitions ainsi que leur pondération. Cette liste des types de publications précise la nomenclature y relative dans SERVAL. Relevons que la Directive de la Direction [de l'UNIL] 4.6 portant sur le dépôt et la diffusion des publications scientifiques dans SERVAL approuvée le 17 mars de cette année et entrée en vigueur le 1^e juillet 2020 (Direction de l'UNIL 2020) n'a pas encore été transposée au moment où nous rédigeons ces lignes ; la terminologie utilisée dans la liste des types de publication de la FBM de l'UNIL ([2020d]) ne s'accorde pas encore avec les types de publications prévus à l'art. 2 de la nouvelle directive de la Direction de l'université (Direction de l'UNIL 2020). iii. Le Conseil de domaine Économie et Services de la HES-SO (2015) détaille pour chaque indicateur les critères cumulatifs que les publications doivent remplir pour être comptabilisées. Précisons que la désignation I2 « [...] articles dans les revues scientifiques » peut porter à confusion : elle intègre, outre les articles, les livres et chapitres de livres scientifiques ainsi que les thèses de doctorat. Aucune des données en notre possession ne laisse entendre qu'il serait des publications exclues de l'évaluation dans les Facultés des lettres de l'UNIGE et de l'UNIL ainsi qu'à la Faculté de médecine de l'UNIGE. La FBM de l'UNIL précise que des publications de l'année courante ne sont pas intégrées à l'évaluation et que « [s]euls les articles publiés dans des journaux à politique éditoriale de la plate-forme *InCites Journal Citation Report (JCR)* de Thomson Reuters sont pris en compte » (FBM UNIL [2020a] point C. et [2020b] §1). Le Conseil de domaine Santé de la HES-SO (2018) exclut du champ de l'évaluation les présentations dans les congrès qui sont publiées sous forme de résumés ainsi que les travaux de bachelor, master ou doctorat. Le Conseil de domaine Économie et Services de la HES-SO (2015) fournit quant à lui une liste non exhaustive des publications qui sont écartées de I2 (à l'instar des nouvelles éditions remaniées d'un livre) et de I4 (tels que les posters).

Dans les contextes d'évaluation aux fins d'attribution de fonds pour la recherche, l'évaluation des groupes de recherche dans le cadre de *MIMOSA* intègre les publications des trois années précédentes (Faculté de médecine de l'UNIGE [s.d.]). Dans le domaine de la Santé de la HES-SO, l'évaluation porte sur les « [...] publications acceptées parues au cours de l'année précédente » (art. 2.1.1 des Modalités d'attribution de l'enveloppe Ra&D du domaine Santé (Conseil de domaine Santé de la HES-SO 2018)). Pour l'année de financement n , le temps de référence pour les I1, I2 et I4 du domaine Économie et Services couvre la période allant du 1^e octobre $n-2$ au 30 septembre $n-1$ (Rectorat de la HES-SO 2017b). Dans les contextes de nomination et de promotion académique et des informations en notre possession, la période de référence n'est pas précisée. Cela nous paraît raisonnable, du fait que l'évaluation sert alors à apprécier l'ensemble de la carrière du chercheur. La FBM de l'UNIL ([2020b]) précise que l'année civile en cours n'est pas considérée.

Quels sont les indicateurs utilisés pour l'évaluation des publications par ces différents domaines et facultés ? Au niveau individuel des chercheurs et dans le contexte d'octroi de fonds pour la recherche, les domaines de la Santé puis de l'Économie et Services de la HES-SO se réfèrent au nombre de publications répondant aux critères établis. Dans le contexte des nominations et promotions académiques, le nombre de publication est utilisé par la Faculté de médecine de l'UNIGE et les Facultés des lettres de l'UNIGE puis de l'UNIL. Ces trois facultés se réfèrent également au nombre de citations. La Faculté des lettres de l'UNIGE précise que le recours à cet indicateur n'est pas systématique. La FBM (UNIL) décline l'indicateur en quatre mesures : le nombre total de citations (pour l'ensemble de la carrière du chercheur et pour les six dernières années) et le nombre total de citations pour les publications où le chercheur est premier ou dernier auteur (pour l'ensemble de la carrière puis pour les six dernières années). Les deux facultés de médecine sont les seules à avoir recours au h -index. Celle de l'UNIGE demande aux chercheurs de lui fournir leur h -index issu du WoS. La FBM de l'UNIL calcule quant à elle le h -index des chercheurs pour l'intégralité de leur carrière académique puis pour les six dernières années. Précisons que les trois indicateurs utilisés par la Faculté de médecine de l'UNIGE pour l'évaluation des publications de ses chercheurs (nombre de publications, nombre de citations et h -index) sont ensuite comparés avec les données de la base de données des professeurs de la Faculté créée en 2014. La Faculté des lettres de l'UNIGE considère le FI des revues et le prestige des éditeurs. Outre le nombre des publications précité, la Faculté estime la variété des types de publications. Le premier critère d'évaluation est la qualité. Elle est évaluée au regard des habitudes de la discipline concernée par les commissions de nomination et de renouvellement. Des

actuels critères d'évaluation *MIMOSA*, le FI constitue un indicateur central. Nous avons été informés du fait qu'il sera vraisemblablement abandonné dans le futur. Jusqu'à présent, le FI total du groupe de recherche est obtenu en additionnant les IF corrigés du groupe. La correction signifie qu'une pondération est effectuée en fonction du type de publication et selon la place des auteurs. Les premier, second, avant dernier et dernier auteurs peuvent compter sur 100% du taux d'attribution *MIMOSA*. Sinon, celui-ci est calculé en fonction du nombre total d'auteurs de la publication. Le FI moyen du groupe de recherche est calculé en divisant son IF total par le nombre total de publications. Pour les groupes de recherche clinique, *MIMOSA* intègre aujourd'hui encore le *Research Production Unit* (RPU), un dérivé du FI qui tient compte du domaine spécifique des revues scientifiques. Nous présumons que le RPU pourrait aussi être délaissé : suite à la signature de la *Déclaration de Dora* (2012), la FBM de l'UNIL a abandonné le recours au FI et au RPU (FBM de l'UNIL [2020a]). Les fonds pour la recherche attribués via *MIMOSA* se basent en outre sur la totalité de ce que le système nomme les « salaires » perçus par le groupe qui concernent certaines sources de financement spécifiques. Un autre des critères d'évaluation *MIMOSA* est la productivité du groupe de recherche qui est calculée sur la base des IF ou RPU au regard de la totalité des « salaires » du groupe (Faculté de médecine de l'UNIGE [s.d.]). La FBM de l'UNIL ([2020c]) compare les données bibliométriques des chercheurs avec celles des autres chercheurs de médecine clinique et fondamentale par leur illustration sous forme de box plots.

5.2 Autres systèmes d'évaluation des publications

Au fil de notre recherche, nous avons trouvé des systèmes d'évaluation qui incluent les publications scientifiques, ou plus globalement des produits de la recherche. Certains ont un champ d'application national. D'autres se concentrent sur un domaine scientifique, voire une discipline. Certains enfin réfléchissent à un socle commun de l'évaluation, qui serait complété par des mesures spécifiques aux divers domaines de la recherche. S'il ne nous est pas possible de présenter de manière exhaustive tous les systèmes d'évaluation intégrant les publications scientifiques, nous avons choisi d'introduire certains d'entre eux en raison de leur intérêt particulier.

Au Royaume-Uni, le *Research Excellence Framework* (REF) sert à évaluer la qualité de la recherche menée au sein des établissements d'enseignement supérieur. Les critères du REF 2021 ont été élaborés par le *Department for the Economy of Northern Ireland*, le *Higher Education Funding Council for Wales*, *Research England* et le *Scottish Funding Council* (2019). Le REF 2021 proscrit le recours au FI ou à tout système de hiérarchisation des revues dans l'évaluation des produits de la recherche. Il fournit la liste des dix-neuf types de résultats de la recherche qui peuvent être évalués. Cette liste

sert de classification. Les œuvres soumises à l'évaluation doivent pouvoir intégrer l'une ou l'autre des catégories. Sans toutes les nommer, relevons les jeux de données et bases de données de même que les logiciels. La dernière catégorie, nommée « autre », permet d'inclure dans l'évaluation des œuvres et produits de la recherche particuliers à l'instar des projets de curation. Dans tous les cas, les membres du comité d'évaluation doivent pouvoir saisir les dimensions du projet de recherche et apprécier son importance, son originalité et sa rigueur (Department for the Economy of Northern Ireland, le Higher Education Funding Council for Wales, Research England et le Scottish Funding Council 2019, p. 97). Le REF 21 fournit en outre une liste des sphères dans lesquels la recherche peut avoir un impact pour la société : sur le bien-être social, le commerce et l'économie, la politique publique ou le droit... Leur analyse doit être complétée par l'évaluation de l'originalité, de la portée et de la rigueur scientifique de la recherche.

La *Royal Netherlands Academy of Arts and Science* (KNAW), la *Netherlands Organisation for Scientific Research* (NWO) et l'*Association of Universities in the Netherlands* (VSNU) (2014) ont élaboré le *Standard Evaluation Protocol (SEP) 2015-2021*. Les KNAW, NWO et VSNU (2014) précisent que si le SEP 2015-2021 tient compte des différentes facettes de la recherche, il ne prétend pas être exhaustif. Le SEP 2015-2021 fournit les bases pour « une évaluation transparente et équitable de la qualité et de la pertinence de la recherche financée par des fonds publics » (KNAW, NWO et VSNU 2014, p. 4). Les groupes cibles du SEP 2015-2021 sont les chercheurs et les responsables de groupes de recherche, les institutions académiques et leurs directions, le gouvernement ainsi que le secteur privé et la société civile. L'objectif principal du SEP 2015-2021 est d'évaluer la pertinence et la qualité de la recherche pour la société dans son ensemble. S'il a été élaboré aux fins de l'évaluation des groupes de recherche (KNAW, NWO et VSNU 2014), ses critères fournissent des éléments qu'il nous semble intéressant de relever en termes d'évaluation individuelle des chercheurs. Ce protocole d'évaluation concerne les aspects scientifiques, sociétaux et stratégiques de la recherche (KNAW, NWO et VSNU 2014). Le comité d'évaluation tient compte des tendances et du développement de la recherche et de la science dans le domaine dans lequel évoluent les chercheurs ainsi que des objectifs fixés à ces derniers (KNAW, NWO et VSNU 2014). Les trois principaux critères d'évaluation sont : la qualité de la recherche, sa pertinence pour la société ainsi que sa viabilité, soit le fait de disposer des ressources nécessaires pour la poursuite des activités de recherche ; elle se veut quantitative et qualitative et est mesurée sur quatre niveaux (KNAW, NWO et VSNU 2014) (cf. Annexe 34). Trois autres perspectives qualitatives sont prises en compte : les programmes de

doctorat, l'intégrité puis la diversité de la recherche (KNAW, NWO et VSNU 2014). Par diversité, les KNAW, NWO et VSNU (2014) entendent la présence de profils complémentaires qui permettent de rassembler des perspectives et des opinions différentes. Ramenée au niveau individuel, elle pourrait servir de critère à la nomination de nouveaux chercheurs ainsi qu'à la comparaison des *outputs* individuels de la recherche au sein d'une même discipline. La SEP 2015-2021 (KNAW, NWO et VSNU 2014) propose une grille d'évaluation dans laquelle l'axe concerne la qualité, qui distingue la qualité de la recherche de son adéquation sociale. L'ordonnée se décline en trois mesures d'évaluation : i. les produits dont les indicateurs intègrent, par exemple, les publications et les jeux de données ; ii. l'utilisation concrète des produits de la recherche dont les indicateurs peuvent compter les citations ou encore le recours aux jeux de données par des pairs et iii. les signes de reconnaissances manifestes tels que la remise de prix ou l'octroi de fonds pour la recherche (cf. Annexe 35).

La KNAW (2013) avait déjà eu recours à ces critères pour mesurer la qualité de la recherche sociale au regard de ses productions, utilisations et reconnaissances concrètes. Distinguant la qualité scientifique de l'intérêt de la recherche pour la société, la première est conférée par les pairs, la seconde par le public cible de la recherche précise la KNAW (2013) (cf. Annexe 36). Deux ans auparavant, la KNAW (2011, pp. 20-21) définissait les « indicateurs de qualité » dans le contexte de l'évaluation de la recherche comme toute indication réellement déterminable de la qualité scientifique et sociétale autre que le contenu de la production scientifique elle-même. Ils permettent de systématiser et normaliser les procédures d'évaluation précise la KNAW (2011). L'appréciation par les pairs est un des indicateurs de qualité. Ces derniers peuvent ne pas toujours s'accorder. La KNAW (2011, p. 10) introduit ce qu'elle nomme des « indicateurs de qualité externe » qui ne reposent sur l'évaluation par les pairs et fournissent à cette dernière une assise vérifiable : les « indicateurs de résultat et les indicateurs d'estime ». Tout système d'évaluation de la qualité de la recherche devrait, selon la KNAW (2011), comprendre ces deux types d'indicateurs, ne pas reposer sur la seule qualité scientifique et tenir compte de la pertinence de la recherche pour la société. En outre, la KNAW (2011) estime que le système d'évaluation doit être suffisamment souple pour être adapté aux contextes spécifiques et ajusté aux nouvelles données et pratiques numériques. La KNAW (2011) de préciser que le système d'indicateurs doit être suffisamment développé pour tenir compte des distinctions entre les disciplines (cf. chapitre 4.2.2). La KNAW (2011) rappelle que les indicateurs bibliométriques provenant de bases de données scientifiques internationales ne présentent pas la même pertinence en fonction des domaines. Concernant les disciplines émanant des

humanités, la KNAW (2011) préconise la prudence en matière de recours aux indicateurs que sont le nombre de publications et le nombre de citations qui ne s'avèrent pertinents que pour les disciplines et domaines largement couverts par ces bases de données. Aussi propose-t-elle l'élaboration de listes de classement des supports de publications (articles, livres...), sur trois niveaux (A, B et C), régulièrement revues et admises par les chercheurs et leurs pairs. Pour les humanités, la KNAW (2011) propose une évaluation fondée sur la production et la qualité scientifique se voyant chacune attribuer les trois mêmes critères d'évaluation : i. les publications et produits de la recherche ii. l'utilisation et iii. la reconnaissance ; chacun ayant ses propres indicateurs (cf. Annexe 37).

Dans les domaines des SHS, la CERES *Research School for International Development* néerlandaise (CERES) (2020) a élaboré un outil d'évaluation ayant pour objectif de fournir une vue d'ensemble des apports à la recherche et de ses résultats. Aussi, CERES (2020) combine les publications et produits scientifiques, qu'ils soient ou non destinés à un public académique, la supervision de doctorats et les « autres activités » de la recherche (critiques de livres, productions de films scientifiques dont la distribution est effectuée par des professionnels du domaine, acquisition de fonds auprès de bailleurs de fonds...). Les chercheurs se voient attribuer un certain nombre de crédits en fonction de ces différents paramètres. Les revues et éditeurs de livres sont respectivement classifiés en cinq catégories allant de A à E. Les éditeurs de livres sont classés en fonction de la visibilité sur Google Scholar⁵². Les chercheurs se voient en outre attribuer des crédits en fonction de la longueur de la publication. Les revues scientifiques sont classées en fonction de leur renommée, sinon de leur type⁵³. La valeur attribuée aux revues et éditeurs scientifiques est revue annuellement par un comité⁵⁴.

Le nombre de crédits octroyés au chercheur varie selon la catégorie attribuée aux revues et éditeurs de livres. Une pondération est également effectuée selon que le chercheur est le seul auteur de la publication (S) ou que celle-ci ait été rédigée à plusieurs (M). P indique le nombre de thèses de doctorat supervisées avec succès. Le nombre de crédits octroyés au chercheur évolue selon le nombre de superviseurs. Les rapports de

⁵² Catégorie A : plus de 500'000 visites. Catégorie B : entre 50'000 et 500'000 visites. Catégorie C : entre 1'000 et 50'000 visites. Catégorie D : entre 100 et 1'000 visites. Catégorie E : moins de 100 visites ou visibilité incertaine (CERES 2020 : <https://ceres.sites.uu.nl/about-the-valuation-system/> [consulté le 12 février 2020]).

⁵³ Catégorie A : Le tiers des revues les plus citées dans les catégories de l'ISI. Catégorie B : Les autres revues indexées dans l'ISI. Catégorie C : Revues académiques avec comité de lecture. Catégorie D : Autres revues académiques et les revues semi-académiques. E : Revues destinées à un lectorat qui n'est pas issu du monde académique (CERES 2020 : <https://ceres.sites.uu.nl/about-the-valuation-system/> [consulté le 12 février 2020]).

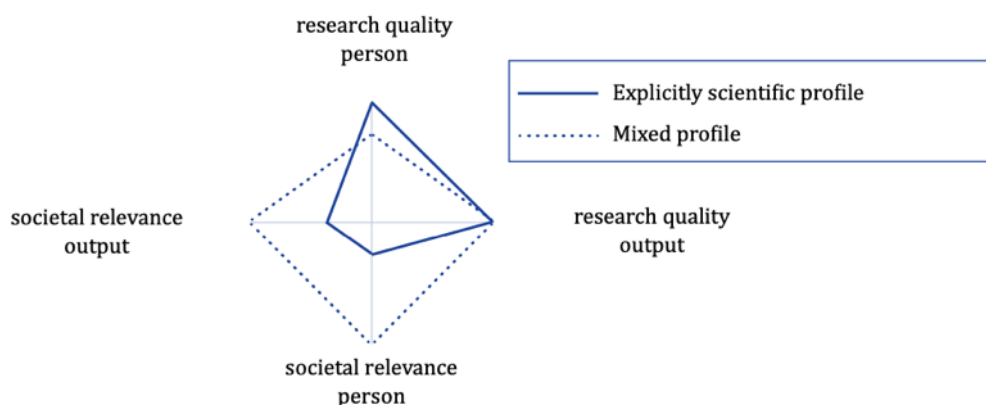
⁵⁴ Comité dont la composition n'est pas communiquée dans CERES (2020).

recherche effectués pour des tiers (R) se voient attribuer des crédits en fonction de leur nombre de pages et du nombre de rédacteurs. Les « autres activités » (O) sont quant à elles créditées en fonction de leur nature. Le système CERES (2020) calcule ainsi la « productivité totale » du chercheur qui comprend l'intégralité des indicateurs. Selon CERES (2020), la « qualité académique » est mesurée au regard des catégories $A + B + C + P$ soit des revues et livres et du nombre de thèses dirigées avec succès. L'« adéquation sociale » est quant à elle être évaluée en fonction de $D + E + R + O$ à savoir les revues intégrant les deux dernières catégories selon CERES (2020) et les livres ayant reçu 1'000 visites ou moins sur Google Scholar. L'Annexe 38 offre le détail des crédits CERES (2020) octroyés aux chercheurs en fonction de leur contributions et apports à la science.

Ne limitant pas son analyse aux disciplines des SHS, la KNAW (2010) estime que l'évaluation des chercheurs devrait certes tenir compte des spécificités des disciplines et des domaines de recherche, mais que les normes d'évaluation ne devraient pas pour autant être élaborées de manière individuelle, au risque d'être considérées comme arbitraires. Aussi propose-t-elle l'élaboration de cadres d'évaluation reposant sur deux composantes : i. des critères de qualité applicables à toutes les disciplines et ii. des indicateurs qualitatifs et quantitatifs spécifiques aux différentes disciplines (KNAW 2010, p. 10). Les ajustements en termes d'évaluation de la qualité propre aux disciplines et domaines trouve son expression dans la mise au point d'indicateurs spécifiques dont la pondération est adaptée pour i. la qualité de la recherche et ii. son adéquation sociale (KNAW 2010). L'importance et la pondération des indicateurs pourrait être adaptée aux finalités de l'évaluation (nomination, promotion académique, octroi de fonds pour la recherche...) précise la KNAW (2010). Pour les disciplines portant sur la conception et l'ingénierie ("*design and engineering*"), la KNAW (2010) considère que les artefacts tels que les produits, processus ou logiciels devraient être pris en considération dans l'évaluation du moment qu'ils engendrent de nouvelles connaissances. Dans le domaine des sciences techniques, la KNAW (2010) estime que la qualité peut être évaluée au regard de i. la qualité la recherche et ii. son adéquation sociale. Les indicateurs de production de la qualité scientifique pourraient être i. les publications scientifiques, ii. la conception d'artefacts, iii. l'impact de la recherche et iv. son potentiel impact qui servirait à l'évaluation dans le cadre de soumission de projets. La pertinence sociale est quant à elle évaluée au regard de i. l'utilisation des résultats de la recherche par des parties prenantes externes (la recherche contribue à résoudre un problème social) ii. ou par des professionnels (la recherche a une utilisation commerciale) iii. l'intérêt de la société civile pour une recherche iv. la contribution à la diffusion des connaissances qui consistent en

des publications professionnelles ou non qualifiées de scientifiques ainsi qu'en la participation à des événements en lien avec les résultats de la recherche. La KNAW (2010) de rappeler l'importance de la prise en considération du profil du chercheur dont les traits peuvent varier d'une discipline à l'autre, voire en leur sein. Aussi propose-t-elle une vue des profils sous la forme d'un diagramme de Kiviati (cf. figure 3).

Figure 3 : Vue des profils des chercheurs selon la KNAW (2010)



(Source : KNAW 2010, p. 34)

La révision des publications par les pairs favorise les contrôles croisés de la qualité, contribuent à diffuser la connaissance scientifique ainsi qu'à la « scientisation » des disciplines (KNAW 2010, p. 12). Le rôle des pairs est considéré comme primordial par la KNAW (2010) qui précise toutefois que la discipline doit être suffisamment homogène, avoir une portée globale et internationale et que le nombre de pairs doit être adéquat. La KNAW (2010) rappelle que l'objectif de l'évaluation de la qualité doit être clairement défini en amont et que l'évaluation doit être considérée à la lumière de cet objectif spécifique.

En réponse aux possibles dérives des systèmes de récompenses et de la culture POP (cf. chapitre 3.3.5), Ioannidis et Khoury (2014) imaginent les bases d'un système d'évaluation dont les critères sont fondés sur le résultat attendu de la recherche scientifique : le PQRST ; P pour productivité, Q pour qualité, R pour reproductibilité, S pour partageable ("*shareable*") et T pour une recherche dont l'influence est translationnelle. La réflexion de Ioannidis et Khoury (2014) portant sur la recherche biomédicale, les auteurs précisent que les indicateurs qu'ils proposent doivent être affinés aux fins de l'évaluation, quel que soit le domaine de la recherche qui aurait recours au PQRST. À propos de la productivité, Ioannidis et Khoury (2014) n'entendent pas récompenser un fort taux de publication ; ils souhaitent la mesure des publications les plus influentes. Citant les universités qui demandent aux chercheurs, dans le cadre

de nomination ou de promotion académique, de leur transmettre ce qu'ils considèrent être leurs meilleures publications, Ioannidis et Khoury (2014) estiment que le processus peut être automatisé en ayant recours aux bases de données. Si cela peut s'avérer vrai dans certains domaines des sciences dites dures, l'automatisation de la mesure ne nous apparaît pas aujourd'hui pertinente dans tous les domaines, en particulier pour les SHS (cf. chapitre 4.2.4). Concernant la qualité, Ioannidis et Khoury (2014) rappellent qu'elle n'est pas toujours corrélée au nombre de citations et sa difficile définition objective avant de préciser que chaque domaine scientifique devrait définir quelles sont les caractéristiques essentielles de la qualité. À propos de la reproductibilité, Ioannidis et Khoury (2014) reconnaissent que si certains types de contrôles peuvent être qualifiés d'aisés, d'autres sont d'une difficulté rédhibitoire, d'où la nécessité de déterminer dans quelles circonstances particulières un contrôle doit être effectué. Ioannidis et Khoury (2014) considèrent le partage comme mesurable ; ils entendent par-là l'évaluation du nombre de publications qui sont accompagnées de données, matériel ou protocoles. Enfin, concernant la mesure du caractère translationnel de la recherche, Ioannidis et Khoury (2014) reconnaissent que c'est à ce jour encore aux évaluateurs de procéder manuellement à la mesure, préconisant de ne la faire que pour les articles les plus cités. Des exemples d'indicateurs envisagés par Ioannidis et Khoury (2014) se trouvent dans l'Annexe 39. Ioannidis et Khoury (2014) relèvent enfin qu'un effort doit être mené en vue d'évaluer la recherche transdisciplinaire.

Les recommandations aux institutions académiques (Dora 2012 ; Hatch et Schmidt 2020)

La *Déclaration de Dora* (2012) formule deux recommandations spécifiques aux institutions académiques. La première a trait à la nécessaire transparence des « critères utilisés dans les décisions de recrutement, de titularisation et de promotion » puis à la sensibilisation des chercheurs quant à la primauté du contenu scientifique de leurs publications sur « les indicateurs de publication ou l'image de marque de la revue dans laquelle il a été publié » (*Déclaration de DORA 2012*, point 4). La seconde enjoint ces institutions à ne pas tenir compte des seules publications scientifiques et d'y incorporer tous les *outputs* des travaux de recherche de leurs chercheurs. L'évaluation doit tenir compte tant de leur valeur que de leur impact. Les hautes écoles devraient « envisager un large éventail de mesures d'impact, y compris des indicateurs qualitatifs sur les retombées des travaux, comme leur influence sur les politiques et les pratiques » (*Déclaration de DORA 2012*, point 5). Une autre manière d'introduire l'adéquation sociale. Au printemps 2020, Hatch et Schmidt publient cinq recommandations destinées aux institutions académiques en vue de l'élaboration puis de la mise en œuvre de

meilleures pratiques d'évaluation. La première de ces indications est la mise en place de normes et structures dans les processus d'évaluation en vue d'asseoir la prise de décision ; Hatch et Schmidt (2020) encouragent les échanges entre les évaluateurs en vue de déterminer les attentes et seuils souhaités pour l'évaluation. La seconde suggestion de Hatch et Schmidt (2020) est d'intégrer les parties prenantes, chercheurs inclus, à la réflexion puis à la définition des critères d'évaluation. Dans un troisième temps, Hatch et Schmidt (2020) soulignent l'importance de l'équité et de la transparence en matière d'évaluation. Elles recommandent ensuite l'élaboration des procédures d'évaluation reposant sur une vue d'ensemble des contributions des chercheurs. La cinquième recommandation de Hatch et Schmidt (2020) est de fixer des objectifs, à court et long termes, aux pratiques et politiques d'évaluation. Ceci aux fins de la mesure de leur pertinence et de permettre leur révision lorsque cela s'avère nécessaire. Ces recommandations nous semblent fondées.

6. Discussion et recommandations

Dans le présent chapitre, nous nous basons sur nos précédentes observations et analyses pour formuler des recommandations en vue de l'évaluation des publications dans le cadre académique. Le modèle que nous suggérons est à considérer comme la base d'une réflexion qui nécessitera d'être approfondie. Nos considérations sont globales. Elles devront faire l'objet d'une adaptation aux différentes disciplines scientifiques ainsi qu'aux procédures internes des institutions académiques. À ce stade de notre raisonnement, nous nous concentrons sur l'évaluation des chercheurs, non des groupes de recherche. Nous n'intégrons pas l'évaluation des chercheurs appartenant à différentes hautes écoles et travaillant sur un même projet de recherche.

Pour pouvoir développer puis tester ce modèle, il nous faudrait disposer d'informations complémentaires telles que le nombre de chercheurs composant chacun des domaines et des facultés des institutions académiques suisses. À ce stade, la base du modèle que nous proposons ne prend pas en compte le temps nécessaire à la réalisation d'une telle évaluation. Nous aurions également besoin de la liste exhaustive de tous les possibles *outputs* de la recherche en vue de mieux déterminer ceux qui devraient participer à l'évaluation. Cette réflexion devrait être menée au niveau des disciplines scientifiques. Il nous faudrait en outre connaître le fonctionnement des diverses procédures de nomination, de promotion académique et d'octroi de fonds au sein de chacune des institutions académiques. Il serait intéressant de connaître la composition des comités œuvrant à ces fins en vue de déterminer si elles reposent sur des bases spécifiques ou partagées. Nous estimons que l'évaluation devrait se baser sur des objectifs communs à toutes les parties concernées : institutions académiques, décanats ou conseils de domaine, facultés ou domaines et chercheurs. Nous ne pouvons pas prétendre connaître chacun de ces objectifs. Ils devraient être définis au niveau des facultés ou des domaines et être reconnus comme tels par les chercheurs. Nous estimons toutefois disposer de suffisamment d'informations pour proposer la base d'un modèle d'évaluation.

Nous considérons et envisageons la bibliométrie comme complément à l'évaluation par les pairs. Nous nous concentrons, dans la présente recherche, sur l'évaluation quantitative des publications. Nous n'avons pas procédé à l'état de l'art de l'évaluation par les pairs. Nous considérons néanmoins qu'ils sont les plus fins connaisseurs du domaine scientifique dans lequel ils évoluent avec leurs confrères. Ils sont les plus à même de connaître l'état et l'évolution de la recherche dans leurs discipline et domaine. Les pairs sont parfois en désaccord. Des antagonismes entre le chercheur et ses pairs

qui reposent sur des considérations professionnelles et/ou personnelles peuvent survenir. Mais cela n'est pas spécifique au domaine scientifique. Il en va de même des collègues ou réseaux. Les divergences d'intérêt et de points de vue sont autant de stimulants qui peuvent produire de bonnes ou malheureuses conséquences selon les personnalités et les dynamiques en présence. Nous sommes en faveur de la présence de plusieurs pairs dans les comités d'évaluation, quelle que soit la finalité de cette dernière (octroi de fonds, nomination, promotion académique et évaluations ponctuelles des chercheurs). Au vu de la progression de la recherche transdisciplinaire, nous estimons important que les comités d'évaluation soient composés de pairs représentant les diverses disciplines concernées par la recherche. L'éthique et la transparence doivent sans conteste former la base conceptuelle de l'évaluation. Le dialogue doit être encouragé et ouvert entre toutes les parties prenantes, de la conception de la procédure d'évaluation à sa mise en œuvre et à son rendu.

Il ressort de notre étude de certains systèmes d'évaluation et des informations en notre possession que deux approches distinctes se dégagent (cf. chapitre 5). La première consiste en ces procédures d'évaluation qui sont concises, mais ne fournissent pas toujours toutes les informations ou les données que nous considérons comme importantes, à l'instar de la définition retenue selon les types de publications, de la question du dépôt, de la place du chercheur dans la liste des auteurs. La seconde comprend les procédures qui sont fournies, mais dont la lecture est absconse. Il est difficile de prétendre qu'une procédure d'évaluation puisse être qualifiée d'exhaustive. Les critères formels doivent être clairement exprimés. Il ne faudrait pas que la rigidité de la procédure entrave l'évaluation ou porte atteinte à l'évaluation qualitative des pairs. Une révision régulière permettrait de la parfaire tout en y intégrant les nouvelles données pertinentes, par exemple l'intégration de nouveaux types d'*outputs*. Elle gagnerait sans conteste à être clairement formulée et facilement compréhensible. Nous qualifions de procédure [d'évaluation] le document formel qui sert de cadre aux évaluations. Nous ne les qualifions pas de « politique » du fait qu'elles s'inscrivent dans le cadre de procédures préexistantes, élaborées aux fins de l'attribution de fonds pour la recherche, de nomination, de promotion académique ou encore des évaluations ponctuelles des membres du corps académique.

Sans nullement remettre en question, ni vouloir porter atteinte à l'autonomie des institutions académiques helvétiques, nous préconisons un certain degré d'alignement entre leurs procédures d'évaluation des *outputs* de la recherche. Nous considérons qu'ils sont des éléments qui devraient être communs à chacune d'entre elles. Observant une augmentation des recherches transdisciplinaires et/ou interinstitutionnelles, il ne serait ni

stimulant, ni équitable, qu'un confrère soit plus ou moins valorisé par son institution pour un travail commun et un effort partagé à un niveau d'engagement similaire. Cet aspect gagnerait à être discuté et arrêté au niveau national en tenant compte des distinctions entre les disciplines et domaines de la recherche scientifiques (cf. chapitre 4.2.2).

Nous avons construit la base de ce modèle d'évaluation en considérant chacune des parties prenantes : la faculté ou le domaine, le comité d'évaluation et le chercheur. L'évaluation ne doit pas amener les chercheurs à modifier leur comportement en vue de « cadrer » avec les critères qui serviront à apprécier leurs activités. Elle doit néanmoins correspondre à la substance même de l'activité scientifique. Aussi, les critères d'évaluation gagneraient à être élaborés en tenant compte de l'évaluation des chercheurs et des *outputs* ainsi que des projets de recherche dans leur globalité. Nous préconisons le recours à une terminologie et à des critères d'emblée identiques entre la documentation et les procédures relatives aux projets de recherche et ceux sur lesquels se fonde l'évaluation. Nous avons cherché à intégrer dans le socle de notre modèle ce nouveau paradigme que nous voyons dans l'OS (cf. chapitre 4.6.2). Lors de l'élaboration de nos recommandations, nous nous sommes efforcés de ne pas nourrir certains des effets néfastes de la culture POP (cf. chapitre 3.3.5).

Nous indiquons quelles sont les parties que nous intégrerions à l'élaboration des procédures d'évaluation. Après avoir précisé nos recommandations quant à leur contenu, nous abordons la question centrale de la définition des *outputs*. Nous traitons ensuite du recours aux indicateurs bibliométriques puis de l'évaluation de la qualité et de l'impact des *outputs*. Avant de présenter la base de notre modèle d'évaluation, nous spécifions le déroulement que pourrait prendre l'évaluation.

Remarques préliminaires

Pansu, Dubois et Beauvois (2013) posent trois questions : qui doit évaluer, à quelle fréquence et sur quelles bases. La réponse dépend des finalités de l'évaluation, des spécificités propres aux divers domaines et disciplines scientifiques ainsi que des procédures des institutions académiques en la matière. À la question de qui doit évaluer les chercheurs, nous estimons que les évaluateurs devraient au moins comprendre un ou plusieurs membres des instances de gouvernance de la faculté/du domaine et des pairs du chercheur. Concernant ces derniers et au regard de la finalité de l'évaluation, la question de la pertinence de la présence de pairs émanant d'autres institutions académiques peut légitimement se poser. La fréquence de l'évaluation est logiquement définie par la finalité de l'évaluation, en fonction des soumissions ou appels à projets, ouvertures de postes ou évaluation internes ponctuelles. Les bases de l'évaluation

doivent de plus être adaptées au domaine voire à la discipline de spécialisation du chercheur et s'inscrire dans le cadre des politiques et procédures de l'institution académique dans laquelle il évolue.

Parties prenantes à l'élaboration et à la mise à jour des procédures d'évaluation

En matière d'élaboration et de mises à jour des procédures, nous envisageons une approche mixte : top-down et bottom-up. Les Décansats, les Conseils de domaine et/ou les Comités Ra&D ainsi que les responsables de la recherche bénéficient d'une vue d'ensemble et des connaissances nécessaires quant aux objectifs académiques et stratégiques de leurs unités de recherche ainsi que des procédures de nomination, de promotion académique et d'octroi de fonds de leur institution. Nous préconisons d'intégrer les chercheurs non seulement aux réflexions portant sur les critères de l'évaluation, mais également à celles ayant pour finalité la définition des mesures ou seuils de l'évaluation et des *outputs* de la recherche qui devraient venir compléter les « traditionnelles » publications en fonction des disciplines scientifiques (jeux de données, logiciels...). Ceci en vue de déterminer de concert les objectifs communs aux parties prenantes qui s'y identifieront alors. Nous intégrerions en outre les bibliothèques académiques et les responsables OS aux groupes de travail œuvrant à l'élaboration et à la mise à jour des procédures d'évaluation. Leurs connaissances et compétences peuvent s'avérer précieuses, particulièrement pour ce qui a trait aux *outputs* de la recherche, des critères et procédures de dépôts et d'archivage, en termes de connaissance des revues et éditeurs scientifiques et de mise en œuvre de l'OS.

Recommandations quant au contenu des procédures

Nous recommandons de fournir aux parties prenantes à l'évaluation un document formel regroupant toutes les informations pertinentes et constituant l'évaluation. Nous entendons par là les informations relatives à :

- La finalité et la fréquence de l'évaluation ;
- Les critères de l'évaluation (qualité, impact...), leurs définitions, mesures et éventuelles pondérations ;
- Les types de publications et autres *outputs* évalués, les conditions à remplir pour qu'ils intègrent l'évaluation (dépôt, affiliation, *openess*...). La pondération attribuée aux *outputs* remplissant les critères doit être précisée ainsi que la période de référence de l'évaluation (année civile, année académique...) ;
- La manière dont est considérée la place de l'auteur lorsqu'une publication a été rédigée par plusieurs chercheurs ;
- La durée sur laquelle se fonde l'évaluation (la carrière du chercheur, les six dernières années...) ;

- La/les sources de données des *outputs* : interne à l'institution (dépôt institutionnel ou autre base de données interne) ; propre au chercheur (son *Curriculum Vitae* (CV)), bases de données scientifiques, brevets ;
- S'il est fait recours à des indicateurs [bibliométriques] : la manière dont ils sont calculés, la source des données et la manière dont ils sont interprétés. Les fenêtres temporelles doivent être spécifiées (fenêtre de publication, fenêtre de citation) ;
- La définition de chacun des termes employés dans le document ;
- La personne ou l'entité auprès de qui les parties prenantes peuvent se tourner en cas de questions, remarques ou litiges ;
- La date et la version du document en vigueur.

La question des définitions est centrale et primordiale. Nous constatons que plusieurs termes peuvent être utilisés à mauvais escient. Citons l'exemple du terme de facteur d'impact qui est parfois utilisé non en référence à l'indicateur qu'est le FI, mais pour signaler la réputation d'une revue scientifique ou d'un éditeur de livres. Ou le terme « article » qui peut être utilisé pour signifier une publication scientifique. Nous considérons que les définitions doivent être élaborées au niveau des disciplines scientifiques, voire des domaines et arrêtées par consensus entre les parties. Les mesures, indicateurs et *outputs* gagnent à être clairement définis dans les procédures d'évaluation. Par mesure, nous entendons les critères sur lesquels les chercheurs sont évalués.

La documentation relative à la procédure d'évaluation devra elle aussi être mise à jour. Celle-ci doit toujours correspondre à la réalité de l'activité et de la communication scientifique. Elle doit pouvoir être adaptée aux retours d'expériences de la pratique concrète de l'évaluation par les parties. La fréquence de ses révisions doit être définie. Nous recommandons aux institutions académiques de toujours dater leurs procédures. La version du document pourrait aussi y figurer. Toute modification dans la procédure doit être traçable.

Les procédures d'évaluation ne doivent ni être conçues, ni être perçues comme un carcan. Elles fournissent un cadre à l'évaluation. Elles ne doivent en aucun cas être un frein à l'appréciation du CV du chercheur dans son intégralité. Nous ne considérons pas l'évaluation comme un grand oral où le chercheur est jugé sur la base de critères fixes. Elle permet de situer le scientifique et son travail par rapport aux objectifs des projets du chercheur, de ceux de la faculté et plus globalement au domaine dans lequel le chercheur évolue. Les évaluations sont autant d'opportunités d'échanger avec les pairs quant aux impacts et qualités perçus par chacune des parties. Ces dernières ne doivent pas perdre de vue le fait que les retombées concrètes de la recherche fondamentale peuvent n'être visibles que des années plus tard.

La définition des *outputs* intégrant l'évaluation

La définition des *outputs*, soit des « produits » de la recherche qui intègrent l'évaluation, est centrale. Elle doit être définie par les parties prenantes selon les pratiques des disciplines et domaines concernés. Une Annexe de la procédure d'évaluation pourrait être réservée aux types d'*outputs* intégrant l'évaluation. Elle détaillerait les types de publications et autres *outputs* en signalant pour chacun d'entre eux :

- La définition retenue pour chaque output ;
- Les critères qu'ils doivent individuellement remplir pour pouvoir intégrer l'évaluation ;
- La date qui fait foi (la date de dépôt, la date de la publication, la date de l'événement...);
- La version qui fait foi (pré-print, post-print, la version « papier » ou la version numérique, une publication dans la langue du chercheur et/ou une traduction...);
- La pondération attribuée à chaque output.

Pour ne citer que l'exemple des articles scientifiques, la procédure d'évaluation devrait spécifier et définir quels sont ceux qui intègrent l'évaluation. Sont-ils tous considérés de la même manière ou sont-ils scindés en différentes classes ? La publication doit-elle avoir été validée par un comité de lecture ? Il y a-t-il un nombre de pages minimum/maximum ? Les articles scientifiques sont-ils distingués des articles parus dans la presse professionnelle ? La revue dans laquelle est publié l'article doit-elle être indexée dans une base de données internationale...

Le cas des traductions devra lui aussi être traité. Par traduction nous entendons la traduction des publications du chercheur ainsi que les traductions de publications effectuées par le chercheur. La seconde engendre un investissement temporel conséquent. Il devrait ainsi intégrer l'évaluation, tout en demeurant un *output* « collectif ». Les traductions de publications du chercheur par un tiers méritent réflexion. Devraient-elles être mises à l'actif de l'impact de sa recherche ? La renommée de l'éditeur qui a publié la traduction doit-elle être prise en considération et de quelle manière ?

À propos de la place du chercheur dans la liste des auteurs, il est essentiel d'apprécier sa réelle contribution s'il n'est pas en première ou seconde position. La CERES (2020) propose deux niveaux : le *single authorship* versus le *multiple authorship*. Aux fins de notre modèle, nous distinguons trois cas de figure : i. le chercheur est le seul auteur de la publication (A) ; ii. le chercheur figure comme premier, second ou dernier auteur (B) ; iii. le chercheur est à une autre place dans la liste des auteurs (C). Sans nullement

souhaiter prêter la recherche et la publication à plusieurs, nous considérons équitable de distinguer les *outputs* nés du seul effort du chercheur de ceux engendrés à plusieurs.

La question du dépôt obligatoire en vue d'être considéré comme un *output* intégré dans l'évaluation est importante. Les directives et politiques qui règlent l'archivage dans les dépôts institutionnels prévoient les types d'*outputs* qui doivent être déposés. Comment traiter les *outputs* qui devraient intégrer l'évaluation, mais qui n'y figurent pas ? Est-ce aux directives et politiques de dépôt d'être amendées, ou à la procédure d'évaluation de nommer les cas où l'évaluation n'est pas corrélée au dépôt en raison des dispositions institutionnelles en la matière ? Cette solution nous semble être la plus appropriée.

Recours aux indicateurs bibliométriques

En cas de recours à des indicateurs bibliométriques, il nous apparaît primordial de toujours préciser les indicateurs utilisés, la raison de leur utilisation, la manière dont ils sont calculés, la source des données qui permet de les générer, les fenêtres temporelles considérées et les valeurs de référence du comité d'évaluation. Ces dernières sont données à titre indicatif. Elles ne constituent pas un critère d'évaluation en soi. Elles pourront nourrir la discussion entre le chercheur et le comité d'évaluation. Si une évaluation ultérieure intervenait pour un même chercheur, les mêmes bases devraient servir de référence en vue de permettre une comparaison objective des données et des commentaires y relatifs. Toute modification, dans la procédure ou dans la source des données (ajouts, suppressions, de types de publications, de revues...), doit être signalée aux parties prenantes ainsi que dans le rapport d'évaluation. Le point quatre du *Manifeste de Leiden* (Hicks et al. 2015) en appelle à une collecte des données aux fins de l'évaluation et à des processus analytiques ouverts, transparents et simples. À propos du recours aux indicateurs bibliométriques :

“Simplicity is a virtue in an indicator because it enhances transparency. But simplistic metrics can distort the record [...]. Evaluators must strive for balance : simple indicators true to the complexity of the research process.”
(Hicks et al. 2015, p. 430)

D'une manière globale, nous considérons que le recours aux indicateurs bibliométriques est plus approprié en matière d'étude bibliométrique. Nous n'écartons pas pour autant l'utilisation de certains de ces indicateurs en matière d'évaluation, lorsque cela s'avère opportun, à savoir qu'ils apportent des informations utiles et une plus-value aux commentaires et à l'analyse des évaluateurs. Le cas échéant, nous sommes en faveur du recours à des indicateurs simples. Nous entendons par là ceux qui fournissent une information, sans être calculés sur plus d'un type de données. Les indicateurs plus

élaborés peuvent quant à eux apporter d'intéressantes matières à réflexions et à analyses aux études bibliométriques.

Nous avons annoncé ne pas souhaiter nourrir certains des effets négatifs de la culture POP. Pour cette raison, nous n'intégrons pas à notre modèle l'indicateur brut qu'est le nombre des publications (cf. chapitre 4.4.2.1). Les comités d'évaluation ont certes besoin de connaître le volume des publications. Nous considérons toutefois que ce chiffre isolé ne suffit pas à constituer un critère d'évaluation. Nous sommes en faveur de l'application de cet indicateur aux différents types de publication. Ceci en vue de savoir si le chercheur a recours à différents canaux de communication. La liste des *outputs* intégrant l'évaluation, types de publications inclus, fait partie intégrante de l'évaluation et devrait être élaborée de concert par les parties prenantes. Des seuils quantitatifs peuvent en outre être définis. La qualité des publications doit néanmoins toujours primer sur la quantité.

Nous considérons le nombre de citations comme un des signaux de l'impact des publications dans les disciplines et les domaines largement couverts dans les bases de données internationales. Nous préconisons l'exclusion des autocitations, terme qui devrait être clairement défini dans la procédure d'évaluation (cf. chapitre 4.4.2.2). À la discrétion de chaque institution de préciser quel pourcentage d'autocitation est admissible. Si nous ne disposons pas aujourd'hui du recul nécessaire pour évaluer leur pertinence et que, comme pour tout indicateur, les parties prenantes doivent être conscientes des limites de la mesure, nous voyons dans les *altmetrics*, les métriques d'utilisation, les *Bookmetrics* et les *Libcitations*, utilisées de manière appropriée, de potentielles futures mesures de l'impact social.

Si nous privilégions les indicateurs que nous qualifions de simples, nous n'écartons pas pour autant *de facto* le recours au *h*-index pour les chercheurs évoluant dans des disciplines et domaines scientifiques largement couverts dans les bases de données scientifiques internationales, ayant plusieurs années de recherche à leur actif et un profil académique en rapport. Les parties doivent connaître les limites de cet indicateur et l'utiliser à bon escient.

Nous ne sommes pas en faveur du recours au FI que nous considérons comme n'étant pas pertinent en matière d'évaluation des chercheurs et de leurs publications. Comme énoncé précédemment, le terme FI est parfois utilisé pour signaler la réputation d'une revue ou d'un éditeur de livres. Nous prêchons en faveur d'une évaluation sur le fond des publications. Il est un élément que nous devons toutefois prendre en considération : dans tout domaine scientifique, des revues et des éditeurs de livres sont considérés

comme étant le *Graal* par les chercheurs et les pairs. Ils possèdent très souvent une stricte politique de sélection et disposent de comités de lecture. Leur réputation est telle qu'ils reçoivent un nombre important de demandes de publications. Le taux d'acceptation est le plus souvent [très] faible. Aussi, si nous écartons *de facto* le recours au FI, nous considérons qu'il est honnête et équitable de tenir compte de la renommée des revues et éditeurs dans l'évaluation, tout comme le fait la CERES (2020). Concernant les éditeurs de livres, contrairement à la CERES (2020), nous ne baserions pas la sélection sur leur [seule] visibilité dans Google Scholar. Les rayonnements, régionaux, nationaux ou internationaux des revues et éditeurs de livres devrait aussi intégrer l'évaluation. Aussi proposons-nous que la réflexion soit ouverte au sein des facultés et des domaines, voire des disciplines et des filières, en vue de déterminer quelle classification est attribuée aux différents revues et éditeurs. Celle-ci pourrait, par exemple, se faire sur une base de trois ou cinq niveaux et devrait être validée par les parties prenantes des procédures d'évaluation. À nouveau, il nous semble pertinent d'associer les bibliothèques académiques à la discussion. La classification devrait régulièrement être revue afin de refléter la perception réelle des chercheurs et de leurs pairs.

Évaluer la qualité et l'impact des *outputs* intégrant l'évaluation

Il ressort de nos précédentes analyses que la mesure de l'impact et de la qualité de la recherche ne se limite désormais plus au cadre strictement scientifique. Nous observons une tendance qui consiste à intégrer la pertinence ou l'adéquation sociale à l'évaluation (cf. chapitre 4.6). Elle est mesurée au regard de l'intérêt pour les fruits de la recherche et de l'utilisation de ses résultats par la société civile. Pour cette raison, nous préférons avoir recours au terme d'impact social. Nous relevons de même une inclination à ne plus attendre uniquement des chercheurs des publications qualifiées de scientifiques. Nous n'avons pas ici pour objectif d'apprécier ces nouvelles tendances. Nous les prenons toutefois en compte dans l'élaboration de nos recommandations. L'approche de la KNAW (2010) (cf. chapitre 5.2) aux fins de la définition des profils de chercheurs nous semble intéressante sous deux angles. Dans un premier temps, les facultés et les domaines peuvent avoir recours au diagramme de Kiviati proposé par la KNAW (2010) en vue de déterminer quel est le profil du chercheur recherché (octroi de fonds, nominations, promotion académique) ou de situer un chercheur par rapport à ses confrères (évaluations ponctuelles). Nous n'entendons pas par-là que tous les scientifiques doivent avoir un profil similaire. Il s'agit bien de pouvoir disposer, lorsque cela s'avère pertinent et nécessaire, de profils scientifiques complémentaires. Ce diagramme peut en outre servir à l'élaboration des critères et des seuils de l'évaluation, toutes finalités confondues, au vu des objectifs scientifiques et stratégiques de la faculté

ou du domaine, en fonction des profils scientifiques de ses chercheurs. Si cette approche peut de prime abord sembler mieux correspondre au domaine des SHS, nous estimons qu'il est intéressant d'élargir la réflexion aux sciences dures. L'esprit de l'OS et le développement de la recherche transdisciplinaire soutiennent cet argument.

Comme le rappelle Furner (2014), nous pouvons distinguer l'analyse des publications de celle de leur utilisation. Le terme de publication doit être élargi aux *outputs* intégrant l'évaluation. Nous considérons que l'analyse des *outputs* permet d'en évaluer la qualité. L'étude de leurs utilisations permet quant à elle d'en apprécier l'impact. Rappelons que la KNAW (2013) considère la qualité sous deux angles : la qualité scientifique et l'intérêt de la recherche pour la société. Il est de notre avis que la qualité scientifique doit primer. Une fois établie, la recherche pourra déployer ses effets dans la société. Les critères de qualité doivent être définis au niveau des disciplines et des domaines. D'une manière globale, nous estimons que les pairs sont les plus à même d'évaluer la qualité scientifique des *outputs*. Aussi important que soit ce travail, rien ne saurait se substituer à la lecture des publications du chercheur et à l'analyse des autres *outputs* de ses recherches antérieures. Il est toutefois un certain nombre de critères qui peuvent soutenir l'évaluation. Le premier est la méthodologie observée par le chercheur. La véracité, l'indépendance et la généralisabilité des résultats constituent d'autres critères de la qualité d'une recherche (Makhlouf-Shabou 2018). Nous ajoutons le critère de l'originalité par lequel nous entendons un regard nouveau apporté à une problématique existante ou l'étude d'une problématique entièrement nouvelle et pertinente pour le domaine scientifique. Dans le domaine des sciences dures, la reproductibilité peut être perçue comme un des fondements de la qualité de la recherche. Le caractère FAIR des données de la recherche doit intégrer les critères d'évaluation.

Nous déclinons l'utilisation dans l'impact scientifique puis social (professionnel et société civile). Nous définissons l'impact comme étant l'utilisation concrète d'un *output* par un tiers. Il comprend, par exemple, les citations effectuées par d'autres chercheurs ou encore l'utilisation, pour d'autres recherches, de données déposées.

Déroulement préconisé de l'évaluation

Nos recommandations quant au déroulement de l'évaluation reposent sur trois phases :

- i. Le comité d'évaluation prend connaissance de l'intégralité du CV du chercheur.
- ii. les *outputs* retenus par les parties prenantes peuvent être analysés et crédités des points qui leurs sont attribués sur la base d'une grille d'évaluation précédemment établie.
- iii. procéder à une sélection des publications ou d'autres *outputs* en vue d'une discussion entre le chercheur et le comité d'évaluation. Comme nous l'avons vu précédemment,

des facultés procèdent déjà de la sorte. Nous proposons d'étendre la procédure à l'ensemble des évaluations. Ainsi, le chercheur et le comité d'évaluation peuvent individuellement opérer un choix dans la liste des *outputs* présents dans le CV du chercheur. Le comité d'évaluation en prend connaissance (lire un article, tester un logiciel...). Les raisons de la sélection sont discutées au moment de l'évaluation. Les parties prenantes peuvent expliquer leur choix et aborder les questions de la qualité et de l'impact.

Présentation du socle de la grille d'évaluation

Nous avons construit une grille d'évaluation (cf. Tableau 4), qui constitue une base et pourra être complétée. La présentation s'inspire du modèle CERES (2020). Celui-ci présente l'avantage d'être clair et de regrouper en un seul et même endroit les *outputs* intégrant l'évaluation. Notre grille diffère de celle du CERES notamment au niveau des *outputs*, de la place du chercheur dans la liste des auteurs ou encore dans l'octroi de points/crédits pour la publication OA et le dépôt des données de la recherche. La grille d'évaluation doit faire partie de la procédure d'évaluation. Elle ne peut être appréhendée sans cette dernière, hors contexte.

Les *outputs* sont signalés à titre indicatifs. Ils devront être adaptés aux listes des *outputs* élaborées de concert entre les parties et les facultés et les domaines concernés. À propos du classement des revues et éditeurs de livres, comme énoncé précédemment, nous estimons que ce sont aux facultés et aux domaines de les élaborer, de concert avec les parties prenantes. Pour illustrer notre grille, nous avons simulé un classement. Nous n'avons volontairement pas indiqué la pondération qui pourrait être attribuée à chacun de ces *outputs* au regard de leurs caractéristiques individuelles. Cette décision est du ressort des facultés et des domaines, voire des parties prenantes. Nous nous sommes concentrés sur certains *outputs*, à titre d'exemple. Nous n'avons volontairement pas intégré la direction de travaux de doctorats. Comme nous avons pu l'observer précédemment, celles-ci intègrent déjà certaines procédures d'évaluation, mais pas toutes (cf. chapitre 5.1). Nous n'avons pas non plus intégré l'organisation de congrès ou colloques, les possibles interactions avec les acteurs de la société civile ou économique ou encore l'obtention de prix et distinctions. Ces activités devraient certes intégrer le cadre de l'évaluation, mais nous les considérons trop éloignées de notre cible : les publications, déjà étendues à d'autres *outputs*. Dans notre grille (cf. Tableau 4), les *outputs* en bleu sont ceux destinés à la communauté scientifique, voire professionnelle. Ils sont distincts de ceux dont le public-cible est la société, signalés en jaune.

Nos lecteurs constateront que nous avons prévu l'octroi de points ou de crédits supplémentaires pour les publications OA ainsi que pour les publications pour lesquelles les données de recherche sont elles aussi mises à disposition (ORD) et remplissent les critères FAIR. Sans revenir sur le fonctionnement de l'OS, signalons que certaines publications et données ne sont pas immédiatement accessibles (délais d'embargo...). Pour ce qui est de l'ouverture des données, celle-ci peut ne pas toujours être possible ou pertinente. L'OS respecte les normes d'éthique et de déontologie du domaine scientifique ainsi que les droits liés à la propriété intellectuelle. S'il existe déjà des dépôts de données, toutes les institutions académiques suisses ne disposent pas déjà de politiques en la matière, ni d'un dépôt reconnu comme institutionnel. Dans notre grille d'évaluation, nous préconisons de signaler N/A ("*not applicable*") dans les cas où le chercheur n'a pas déposé ses données pour des raisons juridiques, éthiques ou pour toute autre raison qui rend la chose impossible ou non indispensable. Le comité d'évaluation sera à même de décider si les raisons et motivations du chercheur sont considérées comme acceptables.

Tableau 4 : Base de la grille d'évaluation

Catégorie revue / éditeur de livre		1			2			3			4			5		
Place du chercheur dans la liste des auteurs		A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
<i>Outputs</i>	Remarques															
Articles	OA : +1 ; ORD FAIR : +1															
Monographies	OA : +1 ; ORD FAIR : +1															
Chapitres de livres	OA : +1 ; ORD FAIR : +1															
Direction / Édition scientifique																
Éditoriaux																
Lettres																

Place du chercheur dans la liste des auteurs

- A Le chercheur est l'unique auteur
- B Le chercheur est le 1er, le 2e ou le dernier auteur [si pas "opportuniste"] dans la liste
- C Le chercheur est à une autre place dans la liste des auteurs

Revue [proposition de "classement" pour illustrer notre modèle] *

- 1 Revues scientifiques avec comité de lecture
- 2 Revues scientifiques sans comité de lecture
- 3 Revues professionnelles / spécialisées avec comité de lecture
- 4 Revues professionnelles / spécialisées sans comité de lecture
- 5 Revues dont le lectorat-cible n'est pas majoritairement issu du monde de la recherche

Éditeurs de livres [proposition de "classement" pour illustrer notre modèle] *

- 1 Éditeurs scientifiques avec comités de lecture
- 2 Éditeurs scientifiques sans comités de lecture
- 3 Éditeurs professionnels / spécialisés avec comité de lecture
- 4 Éditeurs professionnels / spécialisés sans comité de lecture
- 5 Éditeurs dont le lectorat-cible n'est pas majoritairement issu du monde de la recherche

* Tous les éditeurs n'ayant pas la même importance aux yeux de la communauté scientifique et selon les domaines, il conviendrait de distinguer plus finement leur qualité perçue, dès lors de prévoir différents niveaux d'octroi de points/crédits par catégories de revues / éditeurs de livres (p. ex. 1.1 - 1.2 - 1.3 puis 2.1 - 2.2 - 2.3 etc.)

	Niveau international		Niveau national	
	Avec comité de lecture	Sans comité de lecture	Avec comité de lecture	Sans comité de lecture
Actes de colloques, congrès ou conférence				
Communication dans des colloques, congrès ou conférences				
	Niveau international		Niveau national	
Édition d'actes de colloques, congrès ou conférences				
Posters				
Ouvrages de vulgarisation				
Alimentations de carnets de recherche et de blogs				
Émission (radio ou TV) en lien avec le domaine de recherche				
Articles ou interviews dans la presse en lien avec le domaine de recherche				
Participations à des débats en lien avec le domaine de recherche				
Prototypes				
Logiciels				
Brevets / Licences				
Rapports d'expertise				

7. Conclusion

L'évaluation quantitative des publications se base sur des chiffres. Chiffres bruts, à l'instar du nombre de publications, ou chiffres générés par la combinaison d'indicateurs, tel que le nombre de citations par publication. Chiffrer permet de mesurer, donc de comparer. L'évaluation ne consiste pas en un classement des chercheurs. Si la mesure donne une dimension tangible à l'évaluation, cette dernière ne saurait se limiter et se résumer à une suite de nombres. La place et le rôle des pairs sont prépondérants en la matière. Ils sont les plus à même d'évaluer la qualité puis l'impact de la production scientifique des chercheurs évoluant dans la même discipline ou le même domaine scientifique qu'eux.

L'évaluation doit reposer i. sur des objectifs communs aux parties prenantes et ii. sur des critères tangibles et transparents. Les caractéristiques propres aux différentes disciplines et aux domaines scientifiques doivent impérativement être prises en compte. Des mesures bibliométriques ne sauraient jamais être utilisées à des fins de comparaisons inter-disciplinaires. Les bases de données scientifiques présentent un certain nombre de limites en la matière. Aucune d'entre elles ne peut, à ce jour, prétendre être exhaustive. Elles sont d'autant plus importantes dans les disciplines scientifiques qui y sont le moins représentées. L'indicateur qu'est le nombre de publication peut provenir des données disponibles dans les dépôts institutionnels. Les autres indicateurs bibliométriques proviennent sinon de ces bases de données. Nous avons distingué les évaluations des études bibliométriques. Ces dernières portent sur les analyses du fonctionnement et des dynamiques du domaine scientifique. Elles peuvent également servir à l'élaboration de nouveaux outils bibliométriques. Le plus souvent menées au niveau macro, elles se fondent sur les données disponibles dans les bases de données scientifiques.

Nous avons pu déterminer la manière dont les indicateurs bibliométriques classiques ont été élaborés ainsi que leurs finalités, dès lors ce qu'ils signifient et mesurent concrètement. Les limites du recours aux indicateurs bibliométriques en matière d'évaluation ont été relevées. Des voix s'élèvent contre l'utilisation d'indicateurs non pertinents en matière d'évaluation. Le premier d'entre eux est le FI des revues. Les chercheurs et les pairs sont très attachés non au FI, mais à la renommée des revues ou éditeurs de livres qui publient leurs travaux. Il est raisonnable et approprié d'en tenir compte dans l'évaluation. Nous ne souhaitons pas pour autant pas discréditer les indicateurs bibliométriques : ils se révèlent forts utiles en matière d'études bibliométriques. Ils peuvent en outre compléter l'évaluation par les pairs, dans les

disciplines et domaines largement couverts par les bases de données scientifiques. Tant les chercheurs que les comités d'évaluation doivent connaître ce qu'ils mesurent, d'où ils proviennent et sur quelles bases l'interprétation est faite. Il ressort de notre analyse que peu d'indicateurs bibliométriques servent aujourd'hui à l'évaluation des chercheurs en Suisse romande. Il en va de même des autres systèmes d'évaluations présentés. Selon les domaines de recherche, l'accent peut être mis sur certains types de publications. Mais les évaluations des chercheurs ne se limitent jamais à leurs seules publications. Désormais, ces dernières ne constituent plus le seul des *outputs* intégrant l'évaluation. Le dossier ou le CV du chercheur doit être étudié dans son intégralité.

La culture du POP et le désormais ancrage de l'évaluation dans la société ne doivent pas faire perdre de vue la diversité des missions et objectifs des chercheurs. En outre, l'évaluation comporte immanquablement des aspects politiques et économiques qui ne sont pas dénués d'intérêt. L'éthique doit être considérée comme une ouverture à la discussion et ne doit en aucun cas la bloquer. Le respect de ce cadre est nécessaire à un développement sain et fécond de la science. L'OS augmente le contrôle de la qualité de la recherche tout au long du cycle de la recherche. Les encouragements financiers sont un autre type de stimulus positif. La reconnaissance du travail des chercheurs, si elle n'est uniquement symbolique, est favorable au développement de la science.

Notre analyse permet de relever une phase transitoire dans le domaine de l'évaluation qui est liée à plusieurs considérations. Il ressort de nos analyses que l'esprit de la *Déclaration de Dora* (2012), du *Manifeste de Leiden* (Hicks et al. 2015) et du *Metric Tide* (Wilsdon et al. 2015) intègrent progressivement le domaine de l'évaluation. L'avènement de l'OS en appelle à de nouveaux modes de publication et au partage des données qui sous-tendent la recherche dans le respect et le cadre des normes juridiques et éthiques. Le mouvement en appelle à une refonte de l'évaluation des *outputs* de la recherche. L'évaluation ne doit pas pour autant se fonder sur ces seuls critères. Nous ne disposons malheureusement pas encore de suffisamment de recul pour pouvoir parfaitement intégrer ces nouvelles données au cadre futur de l'évaluation. Sans qu'il ne soit exclusivement lié à l'OS, un nouvel aspect de l'évaluation se dessine : la prise en compte de l'impact social de la recherche. De nouvelles métriques pourraient renseigner à ce propos. Nous observons le développement de nouvelles mesures provenant du Web. Si nous n'en sommes encore qu'aux prémices du développement des *webometrics*, des *altmetrics*, des *Bookmetrics* et des *Libcitations*, ce sont autant de nouvelles données qui pourraient un jour intégrer l'évaluation. La vigilance reste toutefois de mise. Nous ne disposons pas de suffisamment de recul pour évaluer aujourd'hui leur pertinence. Nous avons de plus relevé que certaines de ces métriques pourraient être encore plus

manipulables que d'autres indicateurs. En outre, nous rappelons que l'évaluation ne peut pas se passer de la revue qualitative par les pairs et qu'aucun indicateur ne peut s'y substituer. S'il peut être intéressant de mesurer l'impact d'une recherche ou du travail d'un chercheur sur le Web, les données issues de ces nouvelles mesures devront être curées. Elles pourraient un jour servir à mesurer l'impact social de la recherche. L'impact scientifique doit toujours primer.

La précipitation induite par des situations telles que le COVID-19 peut avoir un impact sur la qualité des publications. La difficulté réside à trouver un juste milieu entre l'avancée scientifique et la communication de résultats probants et de qualité. Nous voulons ne pas voir un mélange pervers des effets négatifs de la culture POP et de la médiatisation de résultats scientifiques obtenus sans strict respect de la méthodologie. Nous ne désirons pas non plus en arriver au stade de la « *twitterisation* » ou de l'« *instagramisation* » de la science.

Au terme de ce passionnant voyage au cœur de l'évaluation quantitative des publications et au vu de nos analyses et constats, nous estimons qu'il serait pertinent de mener une enquête au niveau national, au sein des facultés et des domaines des institutions académiques helvétiques. Une telle étude permettrait de recueillir l'ensemble des pratiques et des procédures d'évaluation et de les comparer. Cette enquête pourrait servir de base à une réflexion nationale en vue de l'élaboration de procédures d'évaluation tenant compte des nouveaux paramètres et paradigmes susmentionnés, à l'instar de ce qui se fait pour l'OA. L'évaluation de la recherche transdisciplinaire et interinstitutionnelle devrait sans conteste être intégrée aux discussions. Il nous semble pertinent de réfléchir à un certain degré d'alignement entre les procédures, dans le respect de l'autonomie des institutions académiques helvétiques. Toutes les parties prenantes de l'évaluation devraient être invitées à la table des discussions : institutions académiques, décanats ou conseils de domaine, chercheurs auxquels nous souhaitons associer les bibliothèques académiques.

Bibliographie

ACADÉMIES SUISSES DES SCIENCES, 2008. *L'intégrité dans la recherche scientifique : principes de base et procédures* [en ligne]. Berne : Académies suisses des sciences. [Consulté le 2 avril 2020]. ISBN 3-905870-05-3. Disponible à l'adresse : http://www.akademien-schweiz.ch/fr/dms/D/Projekte-und-Themen/Integritaet/Richtlinien_Directive_F.pdf

ACADÉMIES SUISSES DES SCIENCES, 2013. *Qualité d'auteur des publications scientifiques : analyse et recommandations* [en ligne]. Berne : Commission « Intégrité scientifique » des Académies suisses des sciences. [Consulté le 27 mars 2020]. ISBN 978-3-905870-35-0. Disponible à l'adresse : <http://akademien-schweiz.ch/fr/index/Schwerpunkte/Wissenschaftliche-Integritaet.html>

ALL EUROPEAN ACADEMIES (ALLEA), 2018. *Code de conduite européen pour l'intégrité en recherche* [en ligne]. Édition révisée. Berlin : All European Academies (ALLEA). [Consulté le 30 mars 2020]. Disponible à l'adresse : https://www.allea.org/wp-content/uploads/2018/01/FR_ALLEA_Code_de_conduite_europeen_pour_lintegrite_en_recherche.pdf

AMMON, Ulrich, 1998. *Ist Deutsch noch internationale Wissenschaftssprache ? : Englisch auch für die Lehre an den deutschsprachigen Hochschulen*. Berlin ; New York : De Gruyter. ISBN 978-3-11-016149-6.

AMMON, Ulrich, 2006. Language Planning for International Scientific Communication: An Overview of Questions and Potential Solutions. *Current Issues in Language Planning* [en ligne]. 15 février 2006. Vol. 7, n° 1, pp. 1-30. [Consulté le 15 mai 2020]. Disponible à l'adresse : <https://doi.org/10.2167/cilp088.0>

BAILÓN-MORENO, Rafael, JURADO-ALAMEDA, Encarnación, RUIZ-BAÑOS, Rosario et COURTIAL, Jean-Pierre, 2005. Bibliometric laws: Empirical flaws of fit. *Scientometrics* [en ligne]. 1 avril 2005. Vol. 63, n° 2, pp. 209-229. [Consulté le 1 mai 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-005-0211-5> [accès par abonnement]

BAKER, Monya, 2016. Is there a reproducibility crisis ? *Nature News* [en ligne]. 26 mai 2016. Vol. 533, n° 7604, pp. 452-454. [Consulté le 20 mai 2020]. Disponible à l'adresse : <http://doi.org/10.1038/533452a>

BANKS, Michael G., 2006. An extension of the Hirsch index : Indexing scientific topics and compounds. *Scientometrics* [en ligne]. 1^{er} octobre 2006. Vol. 69, n° 1, pp. 161-168. [Consulté le 12 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-006-0146-5> [accès par abonnement]

BALABAN, Alexandru T., 2012. Positive and negative aspects of citation indices and journal impact factors. *Scientometrics* [en ligne]. 1^{er} août 2012. Vol. 92, n° 2, pp. 241-247. [Consulté le 2 juillet 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-012-0637-5>

BERGSTROM, Carl T., 2007. Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News* [en ligne]. Mai 2007. Vol. 68, n° 5, pp. 314-316. [Consulté le 18 juin 2020]. Disponible à l'adresse : <https://doi.org/10.5860/crln.68.5.7804>

BERGSTROM, Carl T., WEST, Jevin D. et WISEMAN, Marc A., 2008. The Eigenfactor™ Metrics. *Journal of Neuroscience* [en ligne]. 5 novembre 2008. Vol. 28, n° 45, pp. 11433-11434. [Consulté le 18 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1523/JNEUROSCI.0003-08.2008>

Bethesda Statement on Open Access Publishing. [en ligne]. 20 juin 2003. [Consulté le 14 janvier 2020]. Disponible à l'adresse : <http://legacy.earlham.edu/~peters/fos/bethesda.htm>

BIBLIOTHÈQUE DE SORBONNE UNIVERSITÉ, 2017a. BiblioGuide : Infométrie, bibliométrie : Pourquoi ? Comment ? : Contexte et enjeux. *Portail documentaire Sorbonne Université* [en ligne]. 2017. [Consulté le 16 avril 2020]. Disponible à l'adresse : <https://paris-sorbonne.libguides.com/c.php?g=640699&p=4487136>

BIBLIOTHÈQUE DE SORBONNE UNIVERSITÉ, 2017b. BiblioGuide : Infométrie, bibliométrie : Pourquoi ? Comment ? : Familles d'indicateurs. *Portail documentaire Sorbonne Université* [en ligne]. 2017. [Consulté le 17 avril 2020]. Disponible à l'adresse : <https://paris-sorbonne.libguides.com/c.php?g=640699&p=4488054>

BOELAERT, Julien, MARIOT, Nicolas, OLLION, Étienne et PAGIS, Julie, 2015. Les aléas de l'interdisciplinarité : Genèses et l'espace des sciences sociales françaises (1990-2014). *Genèses* [en ligne]. 2015. Vol. 3-4, n° 100-101, pp. 20-49. [Consulté le 15 avril 2020]. Disponible à l'adresse : <https://www.cairn.info/revue-geneses-2015-3-page-20.htm>

BORNMANN, Lutz, MUTZ, Rüdiger et DANIEL, Hans-Dieter, 2008. Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology* [en ligne]. 2008. Vol. 59, n° 5, pp. 830-837. [Consulté le 13 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1002/asi.20806>

BORNMANN, Lutz, MUTZ, Rüdiger, HUG, Sven E. et DANIEL, Hans-Dieter, 2011. A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics* [en ligne]. 1^{er} juillet 2011. Vol. 5, n° 3, pp. 346-359. [Consulté le 14 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1016/j.joi.2011.01.006> [accès par abonnement]

BOURDIEU, Pierre, 1975. La spécificité du champ scientifique et les conditions sociales du progrès de la raison. *Sociologie et sociétés* [en ligne]. 1975. Vol. 7, n° 1, pp. 91-118. [Consulté le 6 mai 2020]. Disponible à l'adresse : <https://doi.org/10.7202/001089ar>

BOURDIEU, Pierre, 1976. Le champ scientifique. *Actes de la Recherche en Sciences Sociales* [en ligne]. Juin 1976. Vol. 2, n° 2-3, pp. 88-104. [Consulté le 23 mars 2020]. Disponible à l'adresse : [10.3406/arss.1976.3454](https://doi.org/10.3406/arss.1976.3454)

BRADFORD, Samuel C., 1934. Sources of information on specific subjects. *Engineering an Illustrated Weekly Journal*. Vol 137, pp. 85–86.

BRADFORD, Samuel C., 1985. Sources of information on specific subjects. *Journal of Information Science* [en ligne]. 1985. Vol. 10, n°4, pp. 176-180. [Consulté le 12 avril 2020]. Disponible à l'adresse : <http://doi.org/10.1177/016555158501000407> [accès par abonnement]

BUELA-CASAL, Gualberto et ZYCH, Izabela, 2012. What do the scientists think about the impact factor ? *Scientometrics* [en ligne]. 1^{er} août 2012. Vol. 92, n° 2, pp. 281-292. [Consulté le 16 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-012-0676-y>

Budapest Open Access Initiative 10 (BOAI10) [en ligne]. 12 septembre 2012. [Consulté le 3 juillet 2020]. Disponible à l'adresse : <https://www.budapestopenaccessinitiative.org/boai-10-translations/french>

BUREAU DÉCANAL ÉLARGI DE LA FACULTÉ DE MÉDECINE DE L'UNIVERSITÉ DE GENÈVE (UNIGE), 2012. Directive de la Faculté de Médecine : Archive ouverte UNIGE. *Université de Genève* [en ligne]. 8 octobre 2012. Révisé le 27 juin 2016.

[Consulté le 3 mars 2020]. Disponible à l'adresse : https://www.unige.ch/biblio/aou/files/8114/6882/8848/Directive_Medecine.pdf

CERES RESEARCH SCHOOL FOR INTERNATIONAL DEVELOPMENT, 2020. About the valuation system [Ceres]. *CERES Research School for International Development* [en ligne]. 2020. [Consulté le 12 février 2020]. Disponible à l'adresse : <https://ceres.sites.uu.nl/about-the-valuation-system/>

CLARIVATE ANALYTICS, 2020a. The Clarivate Analytics Impact Factor. *Clarivate Web of Science* [en ligne]. 2020a. [Consulté le 15 juin 2020]. Disponible à l'adresse : <https://clarivate.com/webofsciencelgroup/essays/impact-factor/>

CLARIVATE ANALYTICS, 2020b. *Web of Science* [en ligne]. 2020b. [Consulté le 13 mars 2020]. Disponible à l'adresse : <https://apps.webofknowledge.com> [accès par abonnement]

COLLEDGE, Lisa, 2017. *Snowball Metrics Recipe Book : Standardised research metrics by the sector, for the sector* [en ligne]. 3rd edition. [S.l.] : [s. n.] [Consulté le 6 juillet 2020]. Disponible à l'adresse : <https://snowballmetrics.com/wp-content/uploads/0211-Snowball-Metrics-Recipe-Book-v7-LO-1.pdf>

COMITÉ GOUVERNEMENTAL DE LA HES-SO, 2014. *Règles communes pour le personnel d'enseignement et de recherche : Typologie des fonctions*. 20 novembre 2014. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

CONSEIL DE DIRECTION DE LA HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE - GENÈVE, 2017. *Règlement interne sur le personnel de la Haute école spécialisée de Suisse occidentale - Genève*. 6 février 2017. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

CONSEIL DE DOMAINE ECONOMIE ET SERVICES DE LA HAUTE ECOLE SPECIALISEE DE SUISSE OCCIDENTALE (HES-SO), 2015. *Délimitations pour les indicateurs I1, I2 et I4*. 28 octobre 2015. Révisés par le Conseil de domaine Économie et Services le 6 mars 2020. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

CONSEIL DE DOMAINE SANTÉ DE LA HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), 2018. *Modalités d'attribution de l'enveloppe Ra&D du domaine Santé : règles et procédures*. 20 septembre 2018. Révisées par le Conseil de domaine Santé le 14 novembre 2019. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

Constitution fédérale de la Confédération suisse du 18 avril 1999 (Cst. ; RS 101). *Les autorités fédérales de la confédération suisse* [en ligne]. 18 avril 1999. Mise à jour le 1^{er} janvier 2020. [Consulté le 5 mars 2020]. Disponible à l'adresse : <https://www.admin.ch/opc/fr/classified-compilation/19995395/index.html#a20>

Convention intercantonale sur la Haute École Spécialisée de Suisse occidentale (HES-SO) du 26 mai 2011 (RS 414.7). *Les cantons de Berne, de Fribourg, de Vaud, du Valais, de Neuchâtel, de Genève et du Jura* [en ligne]. 26 mai 2011. Mise à jour le 1^{er} janvier 2013. [Consulté le 1 juillet 2020]. Disponible à l'adresse : <https://www.hes-so.ch/data/documents/Convention-intercantonale-FR-2496.pdf>

COOMBS, Sarah, PETERS, Isabella, SCHMIDT, Birgit, AGUILLO, Isidro F., PRINCIC, Alenka, MARTINEZ, Marc, KRAKER, Peter, JAHN, Najko, HAUSTEIN, Stefanie, CORNÉE, Nathalie, ABCOUWER, Kasper, HOLMBERG, Kim, LÓPEZ, Ania, GORRAIZ, Juan, FEST, Ellen et TSAKONAS, Giannis, 2018. *Scholarly Metrics Recommendations for Research Libraries: Deciphering the trees in the forest* [en ligne]. [S.l.] : Association of European Research Libraries (LIBER). [Consulté le 6 juillet 2020]. Disponible à l'adresse : <http://doi.org/10.5281/zenodo.1289831>

COSTAS, Rodrigo, BORDONS, Maria, VAN LEEUWEN, Thed N. et VAN RAAN, Anthony F. J., 2009. Scaling rules in the science system: influence of field-specific citation characteristics on the impact of individual researchers. *Journal of the American Society for Information Science and Technology* [en ligne]. 12 janvier 2009. Vol. 60, n° 4, pp. 740-753. [Consulté le 1 mai 2020]. Disponible à l'adresse : <https://doi.org/10.1002/asi.21017> [accès par abonnement]

COUZINET, Viviane, 2016. Évaluation quantitative et évaluation qualitative de la recherche au regard de l'information-documentation: détournements et risques. *Recherches qualitatives : Hors-série « Les Actes »* [en ligne]. 2016. n° 20, pp. 252-266. [Consulté le 9 avril 2020]. Disponible à l'adresse : http://www.recherche-qualitative.qc.ca/documents/files/revue/hors_serie/HS-20/rq-hs-20-couzinet.pdf

DE BELLIS, Nicola, 2014. History and evolution of (biblio)metrics. In : CRONIN, Blaise, SUGIMOTO, Cassidy R (éd.). *Beyond bibliometrics : harnessing multidimensional indicators of scholarly impact*. Cambridge : The MIT Press, p. 23-44. ISBN 978-0-262-52551-0.

Déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales (Déclaration de Berlin), 2003. *Max-Planck-Gesellschaft : Open Access* [en ligne]. [Consulté le 11 janvier 2020]. Disponible à l'adresse : https://openaccess.mpg.de/68042/BerlinDeclaration_wsis_fr.pdf

Déclaration de San Francisco sur l'évaluation de la recherche (Déclaration de Dora) [en ligne]. 16 décembre 2012. [Consulté le 11 août 2020]. Disponible à l'adresse : <https://sfdora.org/read/fr/>

DEPARTMENT FOR THE ECONOMY (NORTHERN IRELAND), HIGHER EDUCATION FUNDING COUNCIL FOR WALES, RESEARCH ENGLAND et SCOTTISH FUNDING COUNCIL, 2019. *Panel criteria and working methods (REF 2019/02) : [assessment criteria and working methods of the main and sub-panels for the Research Excellence Framework (REF) 2021]* [en ligne]. Janvier 2019. [Consulté le 21 juillet 2020]. Disponible à l'adresse : https://www.ref.ac.uk/media/1084/ref-2019_02-panel-criteria-and-working-methods.pdf

DIGITAL SCIENCE & RESEARCH LTD., 2015. What are altmetrics ? *Altmetric* [en ligne]. 2 juin 2015. [Consulté le 20 février 2020]. Disponible à l'adresse : <https://www.altmetric.com/about-altmetrics/what-are-altmetrics/>

DIRECTION DE L'UNIVERSITÉ DE LAUSANNE (UNIL), 2020. Directive de la Direction 4.6 : Dépôt et diffusion des publications scientifiques dans le serveur institutionnel de l'Université de Lausanne, SERVAL. *Université de Lausanne* [en ligne]. 17 mars 2020. [Consulté le 27 juin 2020]. Disponible à l'adresse : <https://www.unil.ch/central/files/live/sites/central/files/textes-leg/4-rech/dir4-6-serval.pdf>

DIRECTORATE-GENERAL FOR RESEARCH AND INNOVATION (DG-RTD) (EUROPEAN COMMISSION), 2016. *Open innovation, open science, open to the world : a vision for Europe* [en ligne]. Luxembourg : Publications Office of the European Union. [Consulté le 2 juillet 2020]. ISBN 978-92-79-57346-0. Disponible à l'adresse : <http://doi.org/10.2777/552370>

DURAND-BARTHEZ, Manuel, 2009. L'évaluation des publications scientifiques. *Les Cahiers du numérique* [en ligne]. 2009. Vol. 5, n° 2, pp. 123-141. [Consulté le 10 avril 2020]. Disponible à l'adresse : <https://www.cairn.info/revue-les-cahiers-du-numerique-2009-2-page-123.htm?contenu=resume>

DUTTA, Bidyarthi, 2014. The journey from librametry to altmetrics : a look back. In : JADAVPUR UNIVERSITY. Department of Library and Information Science. *Golden Jubilee Celebration of Department of Library and Information Science, 8-10 august 2014* [en ligne]. Kolkata : Jadavpur University, Department of Library and Information Science

[p. 1-20]. [Consulté le 10 mars 2020]. Disponible à l'adresse : <http://eprints.rclis.org/23665/2/B-Dutta-JU-Golden-Jubilee-Paper.pdf>

ÉCOLE NATIONALE SUPÉRIEURE DES SCIENCES DE L'INFORMATION ET DES BIBLIOTHÈQUES (ENSSIB), 2013. Indicateur. *enssib : école nationale supérieure des sciences de l'information et des bibliothèques* [en ligne]. 2013. [Consulté le 17 avril 2020]. Disponible à l'adresse : <https://www.enssib.fr/le-dictionnaire/indicateur>

EGGHE, Leo, 2006a. An improvement of the h-index: the g-index. *ISSI newsletter* [en ligne]. Mars 2006. Vol. 2, n° 1, pp. 8-9. [Consulté le 12 juin 2020]. Disponible à l'adresse : <http://issi-society.org/media/1182/newsletter05.pdf>

EGGHE, Leo, 2006b. Theory and practice of the g-index. *Scientometrics* [en ligne]. 1^{er} octobre 2006. Vol. 69, n° 1, pp. 131-152. [Consulté le 12 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-006-0144-7>

EGGHE, Leo et ROUSSEAU, Ronald, 2006. An informetric model for the Hirsch-index. *Scientometrics* [en ligne]. 19 janvier 2006. Vol. 69, n° 1, pp. 121-129. [Consulté le 15 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-006-0143-8> [accès par abonnement]

ELSEVIER, 2004. Scopus comes of age. *ELSEVIER* [en ligne]. 3 novembre 2004. [Consulté le 21 avril 2020]. Disponible à l'adresse : <https://www.elsevier.com/about/press-releases/science-and-technology/scopus-comes-of-age>

FACULTÉ DE BIOLOGIE ET DE MÉDECINE DE L'UNIVERSITÉ DE LAUSANNE (UNIL), [2020a]. Bibliométrie, saisie des publications. *Faculté de biologie et de médecine de l'Université de Lausanne (UNIL)* [en ligne]. [2020a]. [Consulté le 28 février 2020]. Disponible à l'adresse : <https://www.unil.ch/fbm/fr/home/menuinst/recherche/bibliometrie.html>

FACULTÉ DE BIOLOGIE ET DE MÉDECINE DE L'UNIVERSITÉ DE LAUSANNE (UNIL), [2020b]. Tout savoir sur Bibliomics. *Faculté de biologie et de médecine de l'Université de Lausanne (UNIL)* [en ligne]. [2020b]. [Consulté le 28 février 2020]. Disponible à l'adresse : <https://www.unil.ch/fbm/files/live/sites/fbm/files/shared/recherche/bibliometrie/Tout%20savoir%20sur%20BIBLIOMICS-2.pdf>

FACULTÉ DE BIOLOGIE ET DE MÉDECINE DE L'UNIVERSITÉ DE LAUSANNE (UNIL), [2020c]. Données et définitions [Bibliomics]. *Faculté de biologie et de médecine de l'Université de Lausanne (UNIL)* [en ligne]. [2020c]. [Consulté le 28 février 2020]. Disponible à l'adresse : <https://www.unil.ch/fbm/files/live/sites/fbm/files/shared/recherche/bibliometrie/Données%20et%20définitions.pdf>

FACULTÉ DE BIOLOGIE ET DE MÉDECINE DE L'UNIVERSITÉ DE LAUSANNE (UNIL), [2020d]. Types de publications [Bibliomics]. *Faculté de biologie et de médecine de l'Université de Lausanne (UNIL)* [en ligne]. [2020d]. [Consulté le 28 février 2020]. Disponible à l'adresse : <https://www.unil.ch/fbm/files/live/sites/fbm/files/shared/recherche/bibliometrie/Types%20de%20publications.pdf>

FACULTÉ DE MÉDECINE DE L'UNIVERSITÉ DE GENÈVE (UNIGE), [s.d.]. *Analyse MIMOSA*. [s.d.]. Document interne à la Faculté de médecine de l'Université de Genève (UNIGE).

FONDS NATIONAL SUISSE (FNS), 2016. *Domaines de recherche et disciplines du FNS* [en ligne]. Janvier 2016. [Berne] : Fonds national suisse (FNS).

[Consulté le 7 mars 2020]. Disponible à l'adresse : http://www.snf.ch/SiteCollectionDocuments/allg_disziplinenliste.pdf

FONDS NATIONAL SUISSE (FNS), [2020a]. Recherche fondamentale. *FNS - SNF : Fonds national suisse de la recherche scientifique* [en ligne]. [2020a]. [Consulté le 17 mars 2020]. Disponible à l'adresse : <http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/recherche-fondamentale/Pages/default.aspx>

FONDS NATIONAL SUISSE (FNS), [2020b]. Recherche fondamentale orientée vers l'application. *FNS - SNF : Fonds national suisse de la recherche scientifique* [en ligne]. [2020b]. [Consulté le 17 mars 2020]. Disponible à l'adresse : <http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/recherche-fondamentale-orientee-vers-application/Pages/default.aspx>

FONDS NATIONAL SUISSE (FNS), [2020c]. Chiffres clés du FNS : projets approuvés par disciplines en 2019. *Les chiffres clés du Fonds national Suisse* [en ligne]. [2020c]. [Consulté le 4 mai 2020]. Disponible à l'adresse : <https://data.snf.ch/disciplines>

FONDS NATIONAL SUISSE (FNS), [2020d]. Encouragement de projets. *FNS - SNF : Fonds national suisse de la recherche scientifique* [en ligne]. [2020d]. [Consulté le 11 février 2020]. Disponible à l'adresse : <http://www.snf.ch/fr/encouragement/projets/projets-toutes-les-disciplines/Pages/default.aspx#FAQ>

FORTIN, Marie-Fabienne et GAGNON, Johanne, 2016. *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives*. 3^e éd. Montréal : Chenelière Education. ISBN 978-2-7650-5006-3.

FURNER, Jonathan, 2014. The Ethics of Evaluative Bibliometrics. In : CRONIN, Blaise, SUGIMOTO, Cassidy R (éd.). *Beyond bibliometrics : harnessing multidimensional indicators of scholarly impact*. Cambridge : The MIT Press. pp. 85-107. ISBN 978-0-262-52551-0.

GARFIELD, Eugene et SHER, Irving H., 1963. New factors in the evaluation of scientific literature through citation indexing. *American Documentation*. 1963. Vol. 14, n° 3, pp. 195-201. [Consulté le 27 avril 2020]. Disponible à l'adresse : <http://garfield.library.upenn.edu/essays/v6p492y1983.pdf>

GARFIELD, Eugene, 1969. Citation indexes in sociological and historical research. In : *Essays of an Information Scientist Vol. 1 1962-1973* [en ligne]. Philadelphia : Institute for Scientific Information Press, pp. 43-46. [Consulté le 25 mars 2020]. Disponible à l'adresse : <http://www.garfield.library.upenn.edu/essays/V1p043y1962-73.pdf>

GARFIELD, Eugene, 1979. Is citation analysis a legitimate evaluation tool? *Scientometrics* [en ligne]. 1979. Vol. 1, n° 4, pp. 359-375. [Consulté le 28 mars 2020]. Disponible à l'adresse : <http://www.garfield.library.upenn.edu/papers/scientometricsv1%284%29p359y1979.pdf>

GARFIELD, Eugene, MALIN, Morton V. et SMALL, Henry, 1983. Citation data as science indicators. In : *Essays of an Information Scientist Vol. 6 1983* [en ligne]. Philadelphia : Institute for Scientific Information Press, pp. 580-608. [Consulté le 4 avril 2020]. Disponible à l'adresse : <http://www.garfield.library.upenn.edu/essays/v6p580y1983.pdf>

GINGRAS, Yves, 2014a. *Les dérives de l'évaluation de la recherche*. Paris : Raisons d'agir. ISBN 978-2-912107-75-6.

GINGRAS, Yves, 2014b. Criteria for evaluating indicators. In : CRONIN, Blaise, SUGIMOTO, Cassidy R (éd.). *Beyond bibliometrics : harnessing multidimensional indicators of scholarly impact*. Cambridge : The MIT Press, pp. 109-124. ISBN 978-0-262-52551.

GROSS, P. L. K. et GROSS, E. M., 1927. College Libraries and Chemical Education. *Science* [en ligne]. 28 octobre 1927. Vol. 66, n° 1713, pp. 385-389. [Consulté le 5 mars 2020]. Disponible à l'adresse : http://garfield.library.upenn.edu/papers/grossandgross_science1927.pdf

HAMEL, Rainer Enrique, 2007. The dominance of English in the international scientific periodical literature and the future of language use in science. *AILA Review* [en ligne]. 1^{er} janvier 2007. Vol. 20, n° 1, pp. 53-71. [Consulté le 15 mai 2020]. DOI [10.1075/aila.20.06ham](https://doi.org/10.1075/aila.20.06ham). Disponible à l'adresse : <https://www.hamel.com.mx/Archivos-Publicaciones/2007%20Han%20Engl%20in%20Science.pdf>

HATCH, Anna et SCHMIDT, Ruth, 2020. *Rethinking Research Assessment: Ideas for Action - DORA* [en ligne]. 19 mai 2020. DORA. [Consulté le 19 mai 2020]. Disponible à l'adresse : https://sfdora.org/wp-content/uploads/2020/05/DORA_IdeasForAction.pdf

HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), [2020a]. Qui sommes-nous. In : *Haute école spécialisée de Suisse occidentale (HES-SO)* [en ligne]. [2020a]. [Consulté le 6 mars 2020]. Disponible à l'adresse : <https://www.hes-so.ch/fr/sommes-nous-26.html>

HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), [2020b]. Le guide RH égalité et diversité [de la HES-SO] : engagement. *Haute école spécialisée de Suisse occidentale (HES-SO)* [en ligne]. [2020b]. [Consulté le 9 juillet 2020]. Disponible à l'adresse : <https://www.hes-so.ch/fr/engagement-9215.html>

HICKS, Diana, WOUTERS, Paul, WALTMAN, Ludo, DE RIJCKE, Sarah et RAFOLS, Ismael, 2015. Bibliometrics : The Leiden Manifesto for research metrics. *Nature: International Weekly Journal of Science* [en ligne]. 22 avril 2015. Vol. 520, Issue 7548, pp. 429-431. [Consulté le 17 mars 2020]. Disponible à l'adresse : <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>

HIRSCH, Jorge E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* [en ligne]. 15 novembre 2005. Vol. 102, n° 46, pp. 16569-16572. [Consulté le 24 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1073/pnas.0507655102>

INGWERSEN, Peter, LARSEN, Birger, ROUSSEAU, Ronald et RUSSELL, Jane, 2001. The publication-citation matrix and its derived quantities. *Chinese Science Bulletin* [en ligne]. 1^{er} mars 2001. Vol. 46, n° 6, pp. 524-528. [Consulté le 17 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/BF03187274> [accès par abonnement]

INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS, 2020. Defining the Role of Authors and Contributors. *International Committee of Medical Journal Editors (ICMJE)* [en ligne]. 2020. [Consulté le 17 avril 2020]. Disponible à l'adresse : <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>

IOANNIDIS, John P. A. et KHOURY, Muin J., 2014. Assessing Value in Biomedical Research: The PQRST of Appraisal and Reward. *The Journal of the American Medical Association (JAMA)* [en ligne]. 6 août 2014. Vol. 312, n° 5, pp. 483-484. [Consulté le 21 juillet 2020]. Disponible à l'adresse : <http://doi.org/10.1001/jama.2014.6932> [accès par abonnement]

JELICIC KADIC, Antonia, KOVACEVIC, Tanja, RUNJIC, Edita, SIMICIC MAJCE, Ana, MARKIC, Josko, POLIC, Branka, MESTROVIC, Julije et PULJAK, Livia, 2020. Research methodology used in the 50 most cited articles in the field of pediatrics : types of studies that become citation classics. *BMC Medical Research Methodology* [en ligne]. 17 mars

2020. Vol. 20, n° 1, Art. n° 60, [pp. 1-9] [Consulté le 14 mai 2020]. Disponible à l'adresse : <https://doi.org/10.1186/s12874-020-00940-0>

KRISHNAN, Armin, 2009. What Are Academic Disciplines ? Some observations on the disciplinarity vs. interdisciplinarity *Debate* [en ligne]. Working Paper. Southampton. ESRC National Centre for Research Methods. [Consulté le 5 mars 2020]. Disponible à l'adresse : <http://eprints.ncrm.ac.uk/783/>

LAWRENCE, Peter A., 2003. The politics of publication. *Nature* [en ligne]. Mars 2003. Vol. 422, n° 6929, pp. 259-261. [Consulté le 10 avril 2020]. DOI [10.1038/422259a](https://doi.org/10.1038/422259a) [accès par abonnement]

LE COADIC, Yves F., 2005. Mathématique et statistique en science de l'information et en science de la communication: infométrie mathématique et infométrie statistique des revues scientifiques. *Ciência da Informação* [en ligne]. Décembre 2005. Vol. 34, n° 3, pp. 15-22. [Consulté le 5 mars 2020]. Disponible à l'adresse : <https://doi.org/10.1590/S0100-19652005000300002>

LIN, Jennifer et FENNER, Martin, 2013. Altmetrics in Evolution: Defining and Redefining the Ontology of Article-Level Metrics. *Information Standards Quarterly* [en ligne]. 2013. Vol. 25, n° 2, pp. 20-26. [Consulté le 20 juin 2020]. Disponible à l'adresse : <https://doi.org/10.3789/isqv25no2.2013.04>

Loi fédérale du 9 octobre 1992 sur le droit d'auteur et les droits voisins (Loi sur le droit d'auteur (LDA ; RS 231.1)). *Les autorités fédérales de la confédération suisse* [en ligne]. 9 octobre 1992. Mise à jour le 1^{er} janvier 2017. [Consulté le 05 mars 2020]. Disponible à l'adresse : <https://www.admin.ch/opc/fr/classified-compilation/19920251/>

Loi fédérale du 30 septembre 2011 sur l'encouragement des hautes écoles et la coordination dans le domaine suisse des hautes écoles (Loi sur l'encouragement et la coordination des hautes écoles (LEHE ; RS 414.20)). *Les autorités fédérales de la confédération suisse* [en ligne]. 30 septembre 2011. Mise à jour le 1^{er} janvier 2020. [Consulté le 05 mars 2020]. Disponible à l'adresse : <https://www.admin.ch/opc/fr/classified-compilation/20070429>

LOTKA, Alfred J., 1926. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* [en ligne]. 19 juin 1926. Vol. 16, n° 12, pp. 317-323. [Consulté le 30 mars 2020]. Disponible à l'adresse : <https://www.jstor.org/stable/24529203> [accès par abonnement]

MAKHLOUF-SHABOU, Basma, 2018. *Cours 3 – 2e phase: planifications méthodologiques - principaux types de devis, opérationnalisation des concepts et qualité de la recherche* [document Powerpoint]. 11 octobre 2018. Support de cours : Cours « Méthodologie de recherche », Haute école de gestion de Genève, filière Information documentaire, année académique 2018-2019.

MERTON, Robert K., 1968. The Matthew Effect in Science. *Science* [en ligne]. 5 janvier 1968. Vol. 159, n° 3810, pp. 56-63. [Consulté le 6 mai 2020]. Disponible à l'adresse : <http://www.garfield.library.upenn.edu/merton/matthew1.pdf>

MINGERS, John et LEYDESDORFF, Loet, 2015. A review of theory and practice in scientometrics. *European Journal of Operational Research* [en ligne]. 1^{er} octobre 2015. Vol. 246, n° 1, pp. 1-19. [Consulté le 5 mars 2020]. DOI [10.1016/j.ejor.2015.04.002](https://doi.org/10.1016/j.ejor.2015.04.002) [accès par abonnement]

MOED, Henk F., BURGER, W. J. M., FRANKFORT, J. G. et VAN RAAN, Anthony. F. J., 1985. The application of bibliometric indicators : important field- and time-dependent factors to be considered. *Scientometrics* [en ligne]. 1^{er} septembre 1985. Vol. 8, n° 3, pp. 177-203. [Consulté le 27 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1007/BF02016935> [accès par abonnement]

MOOSA, Imad A., 2018. *Publish or Perish: perceived benefits versus unintended consequences*. Cheltenham : Edward Elgar Publishing. Social and Political Science 2018. ISBN 978-1-78897-587-2.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO), 2016. *NISO RP-25-2016 : Outputs of the NISO Alternative Assessment Metrics Project* [en ligne]. Baltimore : NISO, 14 Septembre 2016. [Consulté le 6 juillet 2020]. ISBN 978-1-937522-71-1. Disponible à l'adresse : <https://www.niso.org/publications/rp-25-2016-altmetrics>

ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES (OCDE), 1997. Document de travail de la DSTI 1997/1 : *Indicateurs bibliométriques et analyse des systèmes de recherche: méthodes et exemples* [en ligne]. Paris. Organisation de coopération et de développement économiques (OCDE). [Consulté le 2 avril 2020]. Documents de travail de la Direction de la science, de la technologie et de l'industrie (DSTI), OCDE/GD(97) 41. Disponible à l'adresse : [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(97\)41&docLanguage=Fr](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(97)41&docLanguage=Fr)

ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES (OCDE) et SCIMAGO RESEARCH GROUP (CSIC), 2016. *Compendium of Bibliometric Science Indicators* [en ligne]. Paris : Organisation de coopération et de développement économiques (OCDE). [Consulté le 2 avril 2020]. DSTI/EAS/STP/NESTI(2016)8/FINAL. Disponible à l'adresse : <http://www.oecd.org/sti/inno/Bibliometrics-Compendium.pdf>

ORGANISATION DES NATIONS UNIES POUR L'ÉDUCATION, LA SCIENCE ET LA CULTURE (UNESCO), 2016. *Rapport de l'UNESCO sur la science : vers 2030* [en ligne]. Paris : UNESCO. [Consulté le 17 mars 2020]. ISBN 978-92-3-200106-1. Disponible à l'adresse : <https://unesdoc.unesco.org/ark:/48223/pf0000246417>

PANSU, Pascal, DUBOIS, Nicole et BEAUVOIS, Jean-Léon, 2013. *Dis-moi qui te cite et je saurai ce que tu vaux: que mesure vraiment la bibliométrie ?* Grenoble : Presses universitaires de Grenoble. ISBN 978-2-7061-1780-0.

PERRET, Jean-François et KADELBACH, Thomas, 2017. *Pour une évaluation descriptive des activités de recherche en sciences humaines et sociales: Visualiser la fécondité des unités de recherche et la portée de leurs travaux* [en ligne]. Neuchâtel : Université de Neuchâtel. [Consulté le 1 juillet 2020]. Disponible à l'adresse : <https://www.unine.ch/files/live/sites/qualite/files/Recherche/Rapport%20final%20-%20CUS%20B5%20P3.pdf>

PIGNARD, Nathalie, 1999. *Les enjeux économiques et scientifiques de la publication sur Internet des revues de physique* [en ligne]. Lyon : Institut de la communication de l'université Lyon II. Mémoire pour la Maîtrise en Information Communication. [Consulté le 18 mars 2020]. Disponible à l'adresse : <http://cds.cern.ch/record/408393/files/thesis-99-069.pdf?version=1>

PRICE, Derek John de Solla, 1975. *Science since Babylon* [en ligne]. Enlarged edition. New Haven ; London : Yale University Press. [Consulté le 23 mars 2020]. Disponible à l'adresse : <http://derekdesollaprice.org/science-since-babylon/>

PRICE, Derek John de Solla, 1986. *Little science, big science...and beyond* [en ligne]. New York : Columbia University Press. [Consulté le 25 mars 2020]. ISBN 0-231-04957. Disponible à l'adresse : http://www.andreasaltelli.eu/file/repository/Little_science_big_science_and_beyond.pdf

PRIEM, Jason, TARABORELLI, Dario, GROTH, Paul et NEYLON, Cameron, 2010. *altmetrics: a manifesto*. *altmetrics* [en ligne]. 2010. [Consulté le 17 mars 2020]. Disponible à l'adresse : <http://altmetrics.org/manifesto/>

RAGOUET, Pascal, 2000. Notoriété professionnelle et organisation scientifique. *Cahiers Internationaux de Sociologie* [en ligne]. 2000. Vol. 109, pp. 317-341.

[Consulté le 6 mai 2020]. Disponible à l'adresse : <https://www.jstor.org/stable/40690859>
[accès par abonnement]

RECTORAT DE LA HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), 2016. *Règlement du fonds de recherche et d'impulsions (FRI)*. 24 octobre 2016. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

RECTORAT DE LA HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), 2017a. *Dispositions d'application du règlement du fonds de recherche et d'impulsion (FRI) applicables au domaine Santé de la HES-SO*. 19 décembre 2017. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

RECTORAT DE LA HAUTE ÉCOLE SPÉCIALISÉE DE SUISSE OCCIDENTALE (HES-SO), 2017b. *Dispositions d'application du règlement du fonds de recherche et d'impulsion (FRI) applicables au domaine Économie et Service[s] de la HES-SO*. 19 décembre 2017. Version du 19 février 2019. Document interne à la Haute école spécialisée de Suisse occidentale (HES-SO).

REEDIJK, Jan et MOED, Henk F., 2008. Is the impact of journal impact factors decreasing ? *Journal of Documentation* [en ligne]. 1^{er} janvier 2008. Vol. 64, n° 2, pp. 183-192. [Consulté le 2 juillet 2020]. Disponible à l'adresse : <https://doi.org/10.1108/00220410810858001>

ROSTAING, Hervé, 1996. La bibliométrie et ses techniques [en ligne]. Toulouse ; Marseille : Sciences de la société ; Centre de recherche rétrospective de Marseille. [Consulté le 2 juin 2020]. Outils et méthodes, Supplément au n° 38. Disponible à l'adresse : <https://hal.archives-ouvertes.fr/hal-01579948>

ROUSSEAU, Ronald, 2002. George Kingsley Zipf : life, ideas, his law and informetrics. *Glottometrics* [en ligne]. 2002. n° 3, pp. 11-18. [Consulté le 17 avril 2020]. Disponible à l'adresse : <https://www.ram-verlag.eu/wp-content/uploads/2018/08/g3zeit.pdf>

ROUSSEAU, Ronald, EGGHE, Leo et GUNS, Raf, 2018. *Becoming Metric-Wise : A Bibliometric Guide for Researchers*. Cambridge ; Kidlington : Chandos Publishing an imprint of Elsevier. ISBN 978-0-08-102475-1.

ROY, Simon N., 2009. L'étude de cas. In : GAUTHIER, Benoît (éd.), *Recherche sociale : de la problématique à la collecte des données*. 5e éd. Québec : Presses de l'Université du Québec. pp. 199-225. ISBN 978-2-7605-1600-7.

ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCE (KNAW), 2010. *Quality assessment in the design and engineering disciplines* [en ligne]. 2nd édition : April 2011. Amsterdam : Royal Netherlands Academy of Arts and Sciences (KNAW). [Consulté le 17 avril 2020]. ISBN 978-90-6984-620-0. Disponible à l'adresse : <https://www.know.nl/shared/resources/actueel/publicaties/pdf/quality-assessment-in-the-design-and-engineering-disciplines>

ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCE (KNAW), 2011. *Quality indicators for research in the humanities* [en ligne]. Amsterdam : Royal Netherlands Academy of Arts and Sciences (KNAW). [Consulté le 17 avril 2020]. ISBN 978-90-6984-629-3. Disponible à l'adresse : https://www.know.nl/en/news/publications/towards-a-framework-for-the-quality-assessment-of-social-science-research/@@download/pdf_file/20131006.pdf

ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCE (KNAW), 2013. *Towards a framework for the quality assessment of social science research* [en ligne]. Amsterdam : Royal Netherlands Academy of Arts and Sciences (KNAW). [Consulté le 17 avril 2020]. ISBN 978-90-6984-668-2. Disponible à l'adresse : https://www.know.nl/en/news/publications/towards-a-framework-for-the-quality-assessment-of-social-science-research/@@download/pdf_file/20131006.pdf

ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCE (KNAW), NETHERLANDS ORGANISATION FOR SCIENTIFIC RESEARCH (NWO) et ASSOCIATION OF UNIVERSITIES IN THE NETHERLANDS (VSNU), 2014. *Standard Evaluation Protocol 2015-2021 : Protocol for Research Assessment in the Netherlands* [en ligne]. Amsterdam ; The Hague : Royal Netherlands Academy of Arts and Sciences (KNAW); Association of Universities in the Netherlands (VSNU); Netherlands Organisation for Scientific Research (NWO). Mis à jour : septembre 2016. [Consulté le 17 avril 2020]. Disponible à l'adresse : https://www.know.nl/nl/actueel/publicaties/standard-evaluation-protocol-2015-2021/@@download/pdf_file/SEP%202015-2021%20amended%20version%20sept%202016.pdf

SAREWITZ, Daniel, 2016. The pressure to publish pushes down quality. *Nature News* [en ligne]. 12 mai 2016. Vol. 533, n° 7602, p. 147. [Consulté le 18 mai 2020]. Disponible à l'adresse : <https://doi.org/10.1038/533147a>

SCHREIBER, Michael, 2007. Self-citation corrections for the Hirsch index. *EPL : Europhysics Letters* [en ligne]. Mai 2007. Vol. 78, n° 3, p. 30002 : 1-6. [Consulté le 24 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1209/0295-5075/78/30002> [accès par abonnement]

SCHREIBER, Michael, 2008. The influence of self-citation corrections on Egghe's g index. *Scientometrics* [en ligne]. 2008. Vol. 76, n° 1, pp. 187-200. [Consulté le 12 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-007-1886-6> [accès par abonnement]

SCIMAGO, [2020]. SJR : Scimago Journal & Country Rank. *SJR : Scimago Journal & Country Rank* [en ligne]. [2020]. [Consulté le 19 juin 2020]. Disponible à l'adresse : <https://www.scimagojr.com/>

ŠEMBER, Marijan, UTROBIČIĆ, Ana et PETRAK, Jelka, 2010. Croatian Medical Journal Citation Score in Web of Science, Scopus, and Google Scholar. *sCroatian Medical Journal* [en ligne]. Avril 2010. Vol. 51, n° 2, pp. 99-103. [Consulté le 19 juin 2020]. Disponible à l'adresse : <http://dx.doi.org/10.3325/cmj.2010.51.99>

SPRINGER, [2020]. Bookmetrix. *Springer* [en ligne]. [2020]. [Consulté le 11 juin 2020]. Disponible à l'adresse : <https://www.springer.com/gp/palgrave/book-authors/author-services-resources/bookmetrix>

STAUB, Lukas, 2020. Une publication biaisée. *Journal ASMAC : le journal de l'Association suisse des médecins-assistant(e)s et chef(fe)s de clinique* [en ligne]. Avril 2020. n° 2, p. 18. [Consulté le 3 mars 2020]. Disponible à l'adresse : <https://www.yumpu.com/fr/document/read/63235458/journal-asmac-no-2-avril-2020>

STOKES, Donald E, 1997. *Pasteur's quadrant : basic science and technological innovation* [en ligne]. Washington, D.C. : Brookings Institution Pr. [Consulté le 7 mars 2020]. ISBN 978-0-8157-8178-3. Disponible à l'adresse : https://courses.cs.washington.edu/courses/cse510/16wi/readings/stokes_pasteurs_quadrant.pdf

SUGIMOTO, Cassidy R. et LARIVIÈRE, Vincent, 2018. *Measuring research : what everyone needs to know*. New York : Oxford University Press. ISBN 978-0-19-064012-5.

SWISSUNIVERSITIES, [s.d.]. Hautes écoles suisses reconnues ou accréditées. *swissuniversities* [en ligne]. [Consulté le 4 mars 2020]. Disponible à l'adresse : <https://www.swissuniversities.ch/fr/themes/etudes/hautes-ecoles-suisses-reconnues-ou-accreditees>

SWISSUNIVERSITIES, 2017. *Stratégie nationale suisse sur l'Open Access* [en ligne]. 2017. Bern : swissuniversities. [Consulté le 7 mars 2020]. Disponible à l'adresse :

https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Open_Access_strategy_final_f.pdf

SWISSUNIVERSITIES, 2018a. *Publication finale du programme «Performances de la recherche en sciences humaines et sociales» [Programme CUS P-03]* [en ligne]. Berne. swissuniversities. [Consulté le 22 février 2020]. Disponible à l'adresse : https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/CUS_P-3/Abschlusspublikation_P-3_FR.pdf

SWISSUNIVERSITIES, 2018b. *Stratégie nationale suisse sur l'Open Access: Plan d'action* [en ligne]. 8 février 2018. Bern : swissuniversities. [Consulté le 7 mars 2020]. Disponible à l'adresse : https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Plan_d_action-f.pdf

SWISSUNIVERSITIES, 2020. *Proposition de projet 315A/17 : contributions liées à des projets selon la LEHE, période 2021-2024* [en ligne]. 2020. [Consulté le 6 juillet 2020]. Disponible à l'adresse : https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/Open_Access/Antrag_P-5_Open_Science_I.pdf

UNIVERSITÉ DE LAUSANNE, [2019]. Cycle de vie et types de données. *UNIL - Université de Lausanne* [en ligne]. [2019]. [Consulté le 3 juillet 2020]. Disponible à l'adresse : <https://www.unil.ch/openscience/fr/home/menuinst/open-research-data/les-donnees-de-recherche/cycle-de-vie-et-types-de-donnees.html>

UNIVERSITÉ DE LAUSANNE, [2020]. L'Open Science à l'UNIL : FAQ Open Access. *UNIL - Université de Lausanne* [en ligne]. [2020]. [Consulté le 21 mai 2020]. Disponible à l'adresse : <https://www.unil.ch/openscience/fr/home/menuinst/open-access/faq.html>

UNIVERSITÉ LAVAL, 2020. Limites des indicateurs bibliométriques. *Université Laval* [en ligne]. 2020. [Consulté le 1 juillet 2020]. Disponible à l'adresse : <https://www5.bibl.ulaval.ca/services/soutien-a-ledition-savante-et-a-la-recherche/bibliometrie-et-impact-de-la-recherche/indicateurs-et-sources/limites-des-indicateurs-bibliometriques>

UNIVERSITY OF CANTERBURY, 2020. Open access: Latest news and trends: Publishing your research in Open Access. *University of Canterbury: Library Subject Guides* [en ligne]. 2020. [Consulté le 3 juillet 2020]. Disponible à l'adresse : <https://canterbury.libguides.com/c.php?g=894027&p=6640992>

VAN RAAN, Anthony F. J., 2006. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics* [en ligne]. 1^{er} juin 2006. Vol. 67, n° 3, pp. 491-502. [Consulté le 27 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1556/Scient.67.2006.3.10> [accès par abonnement]

VOS, Anton et MONNET, Vincent, 2007. Mimosas fait rimer fonds et publications. *Campus : le magazine scientifique de l'UNIGE* [en ligne]. Juin-septembre 2007. N° 86, p. 22. [Consulté le 12 février 2020]. Disponible à l'adresse : https://www.unige.ch/campus/files/9614/7246/8274/campus86_dossier4_dossier86.pdf

VUILLEMIN-RAVAL, Isabelle, 2018. Les principaux outils bibliométriques. *The Graduate Institute Library Blog* [en ligne]. 27 avril 2018. [Consulté le 12 juin 2020]. Disponible à l'adresse : <https://libraryblog.graduateinstitute.ch/2018/04/27/les-principaux-outils-bibliometriques/>

WALTMAN, Ludo, VAN ECK, Nees Jan, VAN LEEUWEN, Thed N., VISSER, Martijn S. et VAN RAAN, Anthony F. J., 2011. Towards a new crown indicator : an empirical analysis. *Scientometrics* [en ligne]. 1^{er} juin 2011. Vol. 87, n° 3, pp. 467-481.

[Consulté le 27 avril 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-011-0354-5> [accès par abonnement]

WALTMAN, Ludo et VAN ECK, Nees Jan, 2013. Source normalized indicators of citation impact : an overview of different approaches and an empirical comparison. *Scientometrics* [en ligne]. 1^{er} septembre 2013. Vol. 96, n° 3, pp. 699-716. [Consulté le 23 juin 2020]. Disponible à l'adresse : <https://doi.org/10.1007/s11192-012-0913-4> [accès par abonnement]

WANG, Qi et SCHNEIDER, Jesper Wiborg, 2020. Consistency and validity of interdisciplinarity measures. *Quantitative Science Studies* [en ligne]. Février 2020. Vol. 1, n° 1, pp. 239-263. [Consulté le 7 mars 2020]. Disponible à l'adresse : https://doi.org/10.1162/qss_a_00011

WEBSTER, Charles, 1974. New Light on the Invisible College the Social Relations of English Science in the Mid-Seventeenth Century. *Transactions of the Royal Historical Society* [en ligne]. Janvier 1974. Vol. 24, pp. 19-42. [Consulté le 24 mars 2020]. DOI [10.2307/3678930](https://doi.org/10.2307/3678930) [accès par abonnement]

WHITE, Howard D., BOELL, Sebastian K., YU, Hairong, DAVIS, Mari, WILSON, Concepción S. et COLE, Fletcher T. H., 2009. Libcitations : A measure for comparative assessment of book publications in the humanities and social sciences. *Journal of the American Society for Information Science and Technology* [en ligne]. 2009. Vol. 60, n° 6, pp. 1083-1096. [Consulté le 20 juin 2020]. Disponible à l'adresse : <http://doi.org/10.1002/asi.21045> [accès par abonnement]

WILSDON, James, ALLEN, Liz, BELFIORE, Eleonora, CAMPBELL, Philip, CURRY, Stephen, HILL, Steven, JONES, Richard, KAIN, Roger, KERRIDGE, Simon, THELWALL, Mike, TINKLER, Jane, VINEY, Ian, WOUTERS, Paul, HILL, Jude et JOHNSON, Ben, 2015. *The Metric Tide : Report of the Independent Review of the Role of Metrics in Research Assessment and Management* [en ligne]. S.I. : Higher Education Funding Council for England (HEFCE). [Consulté le 8 avril 2020]. ISBN 1902369273. Disponible à l'adresse : <https://responsiblemetrics.org/the-metric-tide/>

WILSDON, James, BAR-ILAN, Judit, FRODEMAN, Robert, LEX, Elisabeth, PETERS, Isabella et WOUTERS, Paul, 2017. *Next-Generation Metrics : Responsible Metrics and Evaluation for Open Science ; Report of the European Commission Expert Group on Altmetrics* [en ligne]. Brussels : European Commission. ISBN 978-92-79-66130-3. Disponible à l'adresse : <http://doi.org/10.2777/337729>

WOUTERS, Paul, 2014. The citation : from culture to infrastructure. In : CRONIN, Blaise, SUGIMOTO, Cassidy R (éd.). *Beyond bibliometrics : harnessing multidimensional indicators of scholarly impact*. Cambridge : The MIT Press, pp. 47-66. ISBN 978-0-262-52551-0.

YANG, SiLuo et YUAN, Qingli, 2017. Are Scientometrics, Informetrics, and Bibliometrics different ? In : ISSI, The International Society for Scientometrics and Informetrics. *The 16th International Conference on Scientometrics & Informetrics (ISSI2017): ISSI 2017 Proceedings, Whuan, Wuhan University, 16-20 october 2017* [en ligne]. Whuan : ISSI, pp. 1507-1518. [Consulté le 6 mars 2020]. Disponible à l'adresse : http://www.issi-society.org/proceedings/issi_2017/2017ISSI%20Conference%20Proceedings.pdf

ZIJLSTRA, Hans et MCCULLOUGH, Rachel, 2016. CiteScore : a new metric to help you track journal performance and make decisions. *Elsevier* [en ligne]. 8 décembre 2016. [Consulté le 12 mai 2020]. Disponible à l'adresse : <https://www.elsevier.com/editors-update/story/journal-metrics/citescore-a-new-metric-to-help-you-choose-the-right-journal>

ZWEIGBIBLIOTHEK MEDIZIN, 2020. Journal Citation Score. *Nachrichten aus der Zweigbibliothek Medizin* [en ligne]. 2020. [Consulté le 1 mai 2020]. Disponible à l'adresse : <https://www.uni-muenster.de/ZBMed/aktuelles/wiki/journal-citation-score>

Ouvrages de référence

ACADÉMIE FRANÇAISE, 2019. Dictionnaire de l'Académie française. In : *Dictionnaire de l'Académie française* [en ligne]. 2019. [Consulté le 9 mars 2020]. Disponible à l'adresse : <https://www.dictionnaire-academie.fr/>

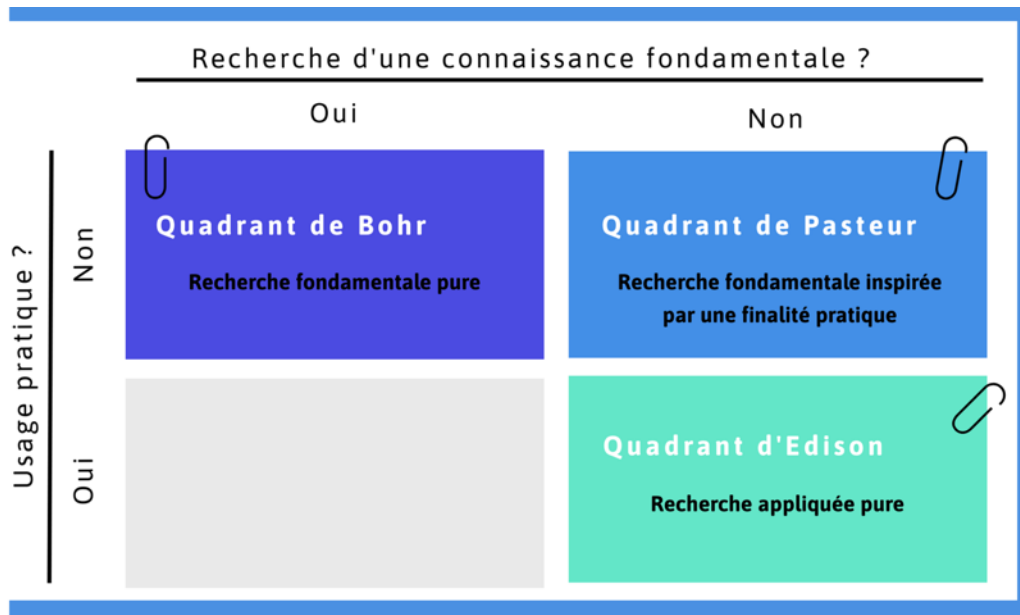
HEILBRON, John L., 2003. *The Oxford Companion to the History of Modern Science* [en ligne]. [Oxford] : Oxford University Press. [Consulté le 9 mars 2020]. ISBN 978-0-19-511229-0. Disponible à l'adresse : <https://www.oxfordreference.com/view/10.1093/acref/9780195112290.001.0001/acref-9780195112290> [accès par abonnement]

STEVENSON, Angus et LINDBERG, Christine A. (éd.), 2011. *New Oxford American Dictionary* [en ligne]. 3rd ed. New York : Oxford University Press. [Consulté le 9 mars 2020]. ISBN 978-0-19-539288-3. Disponible à l'adresse : <https://www.oxfordreference.com/view/10.1093/acref/9780195392883.001.0001/acref-9780195392883> [accès par abonnement]

VAUCHEZ, André, DOBSON, Barrie et LAPIDGE, Michael, 2000. *Encyclopedia of the Middle Ages*. Éditions du Cerf ; Città nuova. Cambridge : J. Clarke & Co. ISBN 978-0-227-67931-9.

WALLACE, Susan (éd.), 2015. *A Dictionary of Education* [en ligne]. Oxford ; New York : Oxford University Press. [Consulté le 9 mars 2020]. ISBN 978-0-19-967939-3. Disponible à l'adresse : <https://www.oxfordreference.com/view/10.1093/acref/9780199679393.001.0001/acref-9780199679393> [accès par abonnement]

Annexe 1 : Quadrants de la recherche scientifique (Stokes 1997)



Source : adapté de Stokes 1997, p. 73

Annexe 2 : Évolution du nombre de revues scientifiques depuis 1665 (Price 1975)

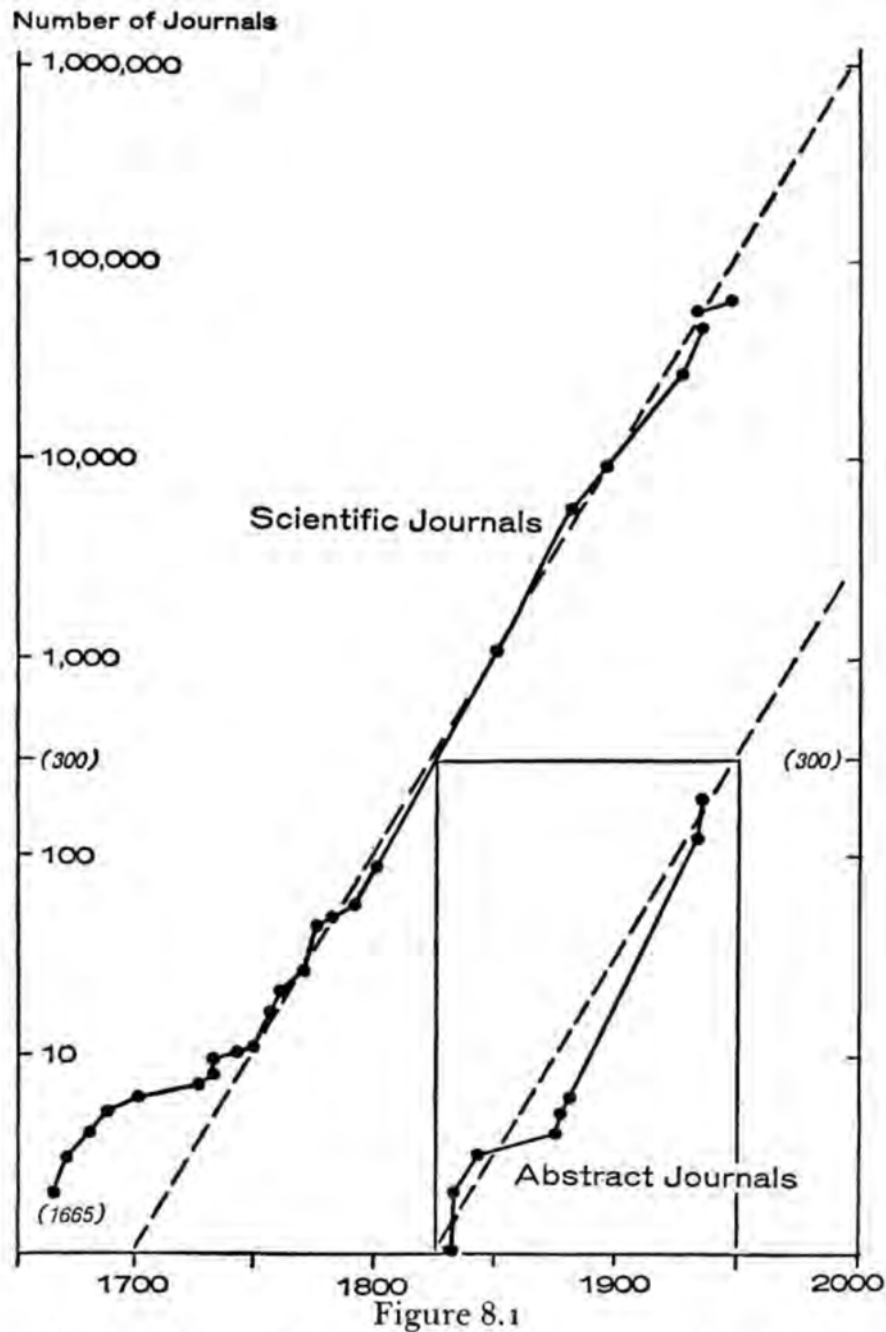
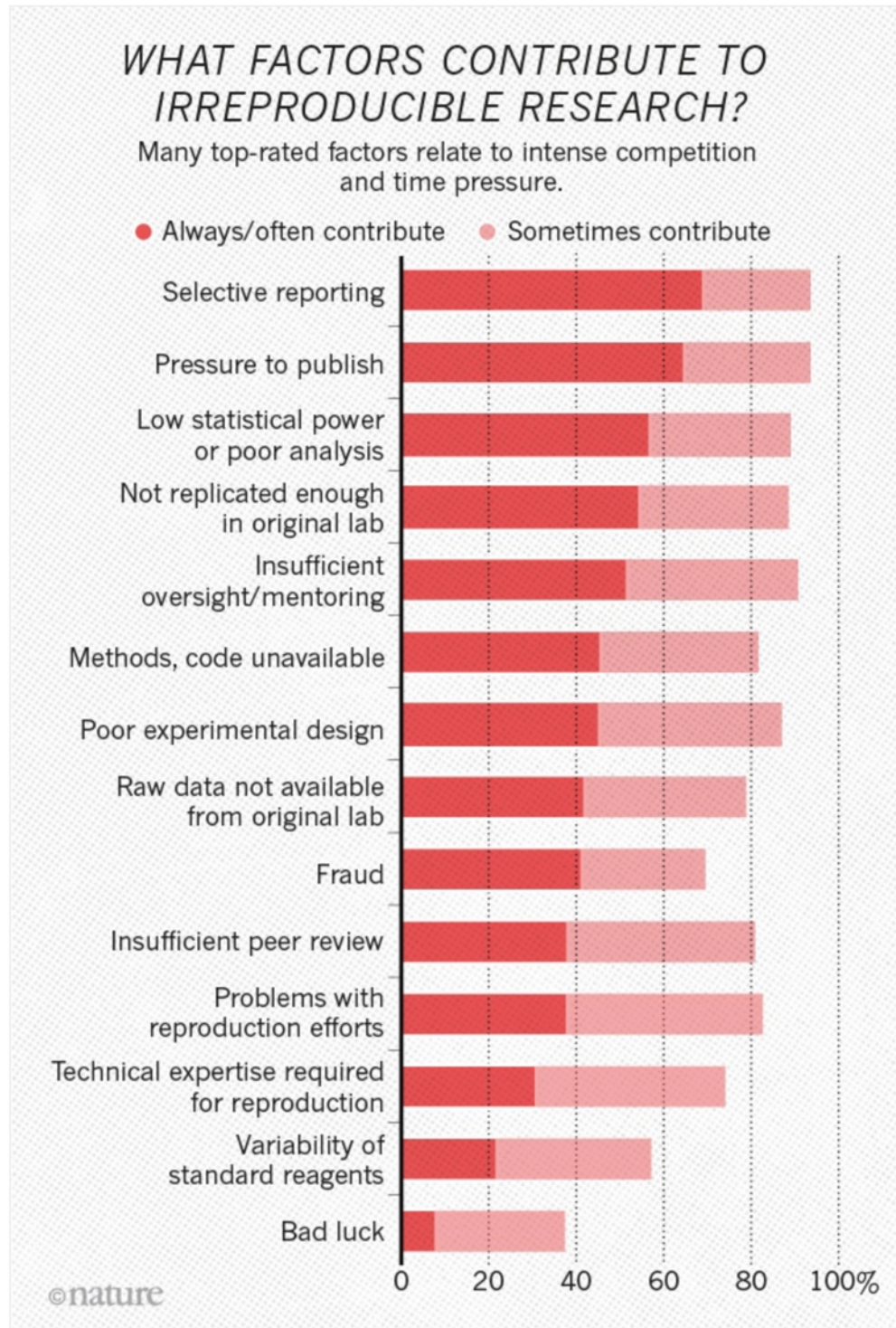


Figure 8.1
Number of journals founded (*not surviving*) as a function of date. The two uppermost points are taken from a slightly differently based list.

Sources : Price 1975, p. 166 et Price 1986, p. 8

Annexe 3 : Facteurs contribuant à l'irreproductibilité de la recherche (Baker 2016)



Source : BAKER (2016). <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970> [Consulté le 20 mai 2020]

Annexe 4 : Du Pré-print au Post-print



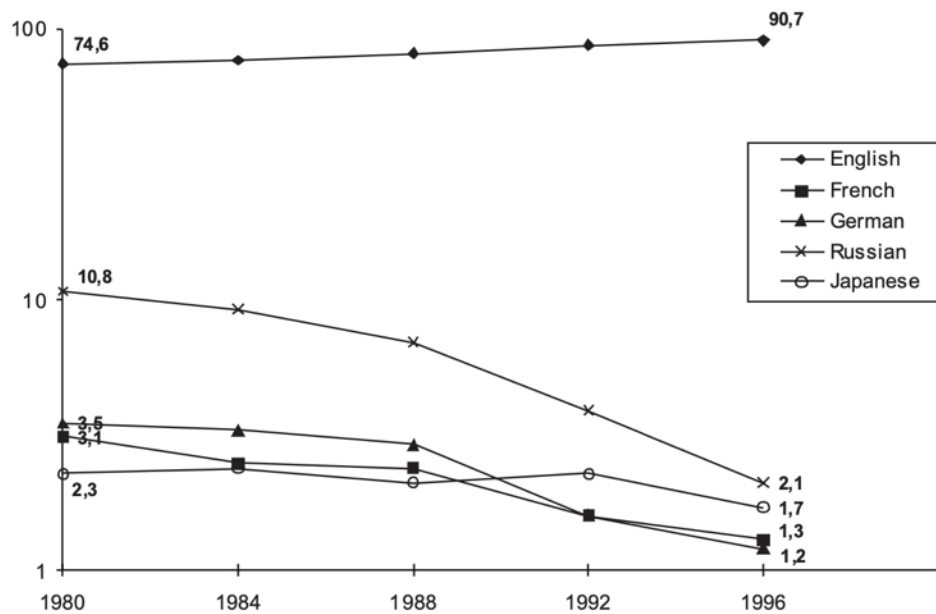
Source : UNIL, 2020. <https://www.unil.ch/openscience/fr/home/menuinst/open-access/faq.html>
[Consulté le 21 mai 2020]

Annexe 5 : Les disciplines scientifiques sous différentes perspectives (Krishnan 2009)

Perspective =>	Philosophique	Anthropologique	Sociologique	Historique	Managériale
<u>Paradigme</u>	Connaissance	Culture	Organisation sociale	Temps	Marché
<u>Facteurs favorisant la naissance d'une discipline</u>	Langage Discours	Identité culturelle et segmentation	Professionnalisation Structures de pouvoir	Leadership des pères fondateurs de la discipline	Succès passés de l'organisation disciplinaire
<u>Facteurs favorisant l'inter- ou la trans-disciplinarité</u>	Universalisation du discours	Nouvelles formes de communauté et d'identité	Changement social Déclin des professions	Maturation de la discipline Manque de leadership	Meilleure adaptation au marché
<u>Remarques de Krishnan (2009) après analyse de la vision de la "discipline" selon ces perspectives</u>	Les disciplines sont nécessaires pour valider les affirmations de vérité	Les disciplines offrent une identité stable et sont similaires à des structures tribales	Les structures des disciplines sont difficiles à franchir en raison de l'intérêt du personnel de groupes de pouvoir	Historiquement, plutôt que de se réduire, le nombre de disciplines a constamment augmenté	Les disciplines sont une forme obsolète de l'organisation des sciences et des institutions académiques

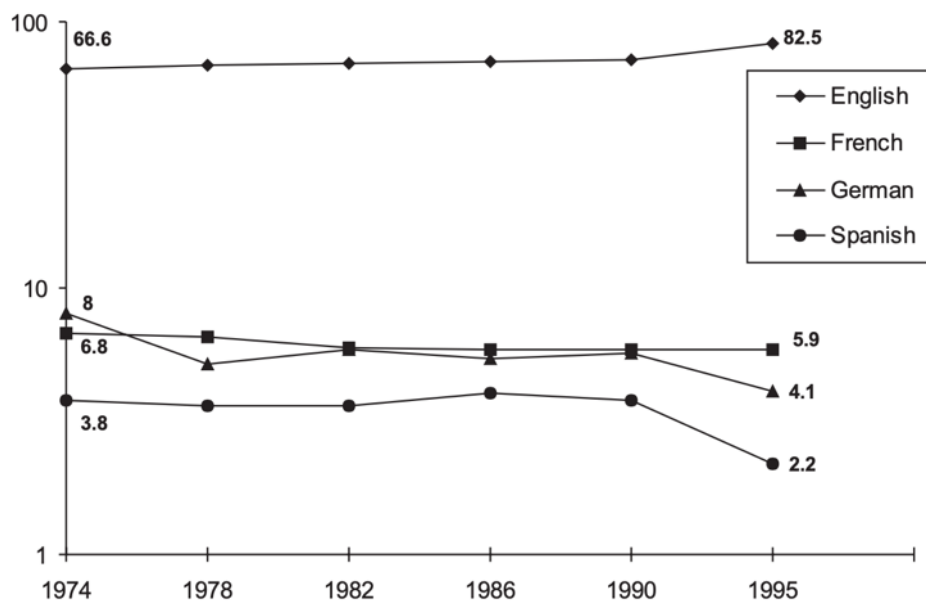
Source : adapté de Krishnan 2009, p. 4

Annexe 6 : Répartition des langues de la communication scientifique dans les sciences naturelles : 1980-1996 (Ammon 2006)



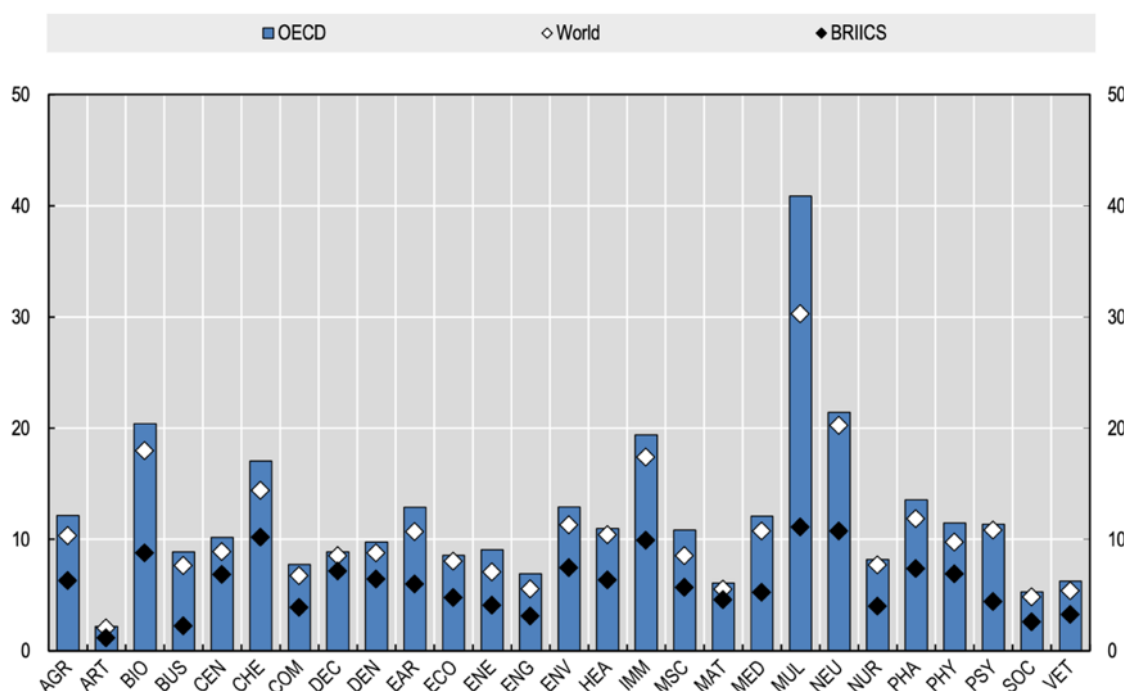
Source : Ammon 2006, p. 3

Annexe 7 : Répartition des langues de la communication scientifique dans les SHS : 1974-1995 (Ammon 2006)



Source : Ammon 2006, p. 4

Annexe 8 : Nombre moyen de citations par document et par domaine : 2003-2012 (OCDE et SCImago Research Group 2016)



Source : OCDE et SCImago Research Group 2016, p. 43.

Code	Field	Code	Field
ART	Arts & Humanities	HEA	Health Professions
AGR	Agricultural & Biological Sciences	IMM	Immunology & Microbiology
BIO	Biochemistry, Genetics & Molecular Biology	MSC	Materials Science
BUS	Business, Management & Accounting	MAT	Mathematics
CHE	Chemistry	MED	Medicine
CEN	Chemical Engineering	MUL	Multidisciplinary
COM	Computer Science	NEU	Neuroscience
DEC	Decision Sciences	NUR	Nursing
DEN	Dentistry	PHA	Pharmacology, Toxicology & Pharmaceutics
EAR	Earth & Planetary Sciences	PHY	Physics & Astronomy
ECO	Economics, Econometrics & Finance	PSY	Psychology
ENE	Energy	SOC	Social Sciences
ENG	Engineering	VET	Veterinary
ENV	Environmental Science		

Source : OCDE et SCImago Research Group 2016, p. 7.

Annexe 9 : Nombre moyen de références par article et par domaine : 1970-1980 (Moed et al. 1985)

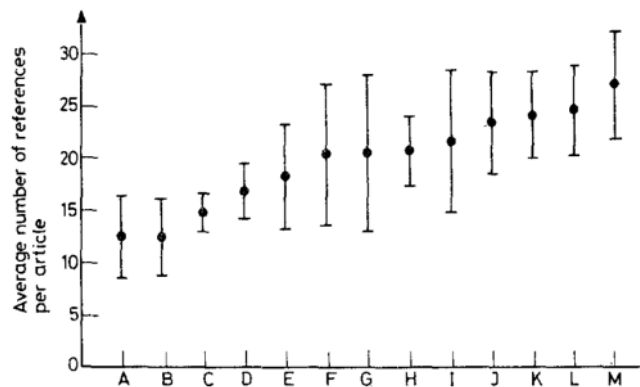


Fig. 1. A – Pure and Applied Mathematics (n=15), B – Statistics and Informatics (n=9), C – Pharmacognosy and Pharmaceutical Technology (n=7), D – Inorganic Solid State Chemistry (n=9), E – Experimental and Molecular Physics (n=13), F – Taxonomy and Bio Systematics (n=11), G – Pharmacology and Pharmaco-therapeutics (n=11), H – Plant Physiology and Botanic Morphogenesis (n=12), I – Organic Chemistry (n=13), J – Physical Chemistry (n=10), K – Biophysics (Photo Synthesis) (n=5), L – Cell Biology and Morphogenesis (n=8), M – Biochemistry (Protein Bio Synthesis) (n=8)

Source : Moed et al. 1985, p. 180

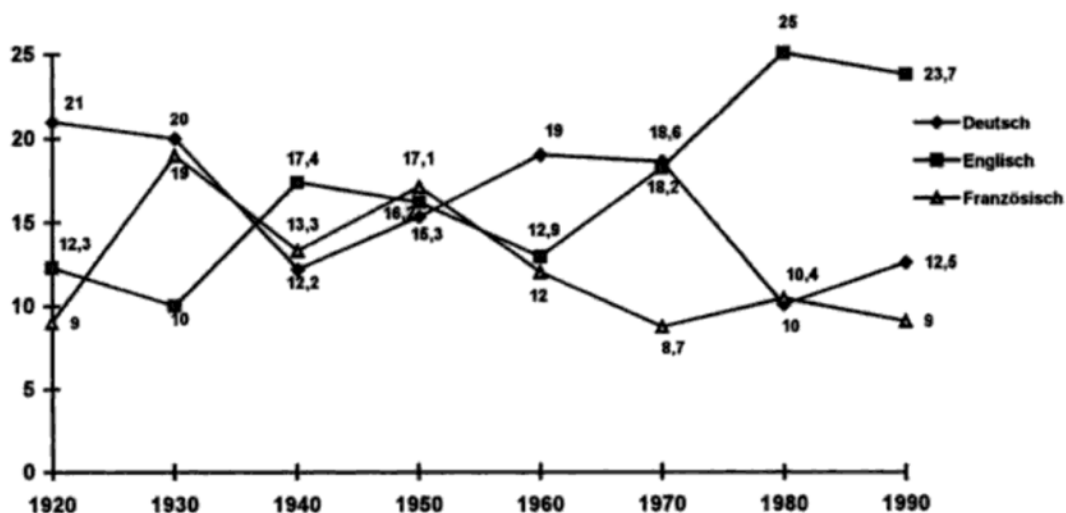
Annexe 10 : Taux de citation à court terme dans certains domaines (sciences dures) (Moed et al. 1985)

Disciplines	Taux de citation à court terme
Biochimie Biologie cellulaire et morphogénèse Biophysique	ÉLEVÉ
Physique expérimentale et moléculaire Chimie physique Chimie organique	MOYEN
Physiologie des plantes et morphogénèse botanique Pharmacologie et la pharmacothérapie	
Mathématiques, Statistiques et informatique, Pharmaceutique et technologie pharmaceutique taxonomie et bio-systématique, Chimie des solides inorganiques	BAS

Source : adapté de Moed et al. 1985, p. 182

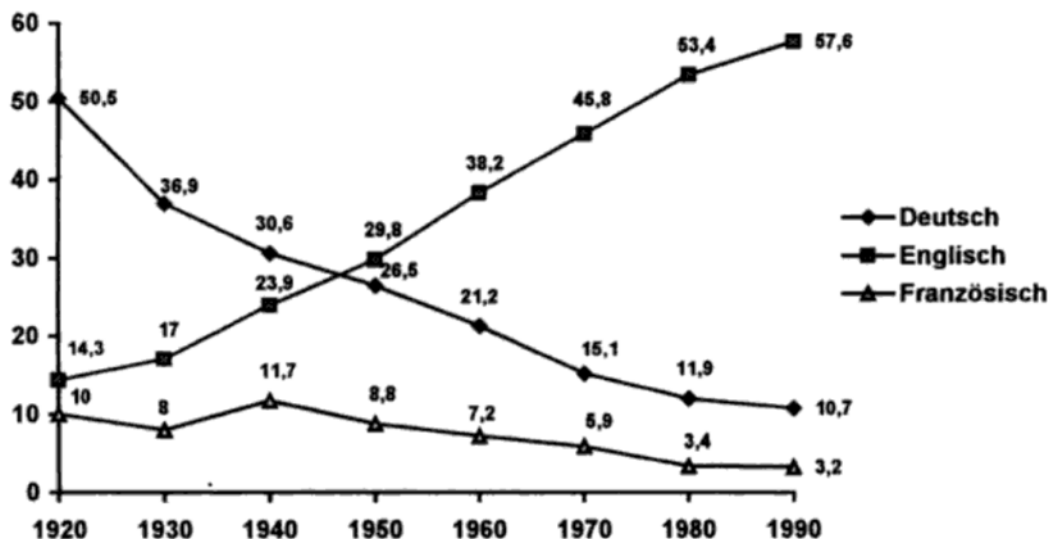
Annexe 11 : Évolution de la moyenne des parts de citations dans les langues allemande, française et anglaise dans les revues d'histoire et de chimie : 1920-1990 (Ammon 1998)

Revue d'histoire



Source : Ammon 1998, p. 75

Revue de chimie



Source : Ammon 1998, p. 53

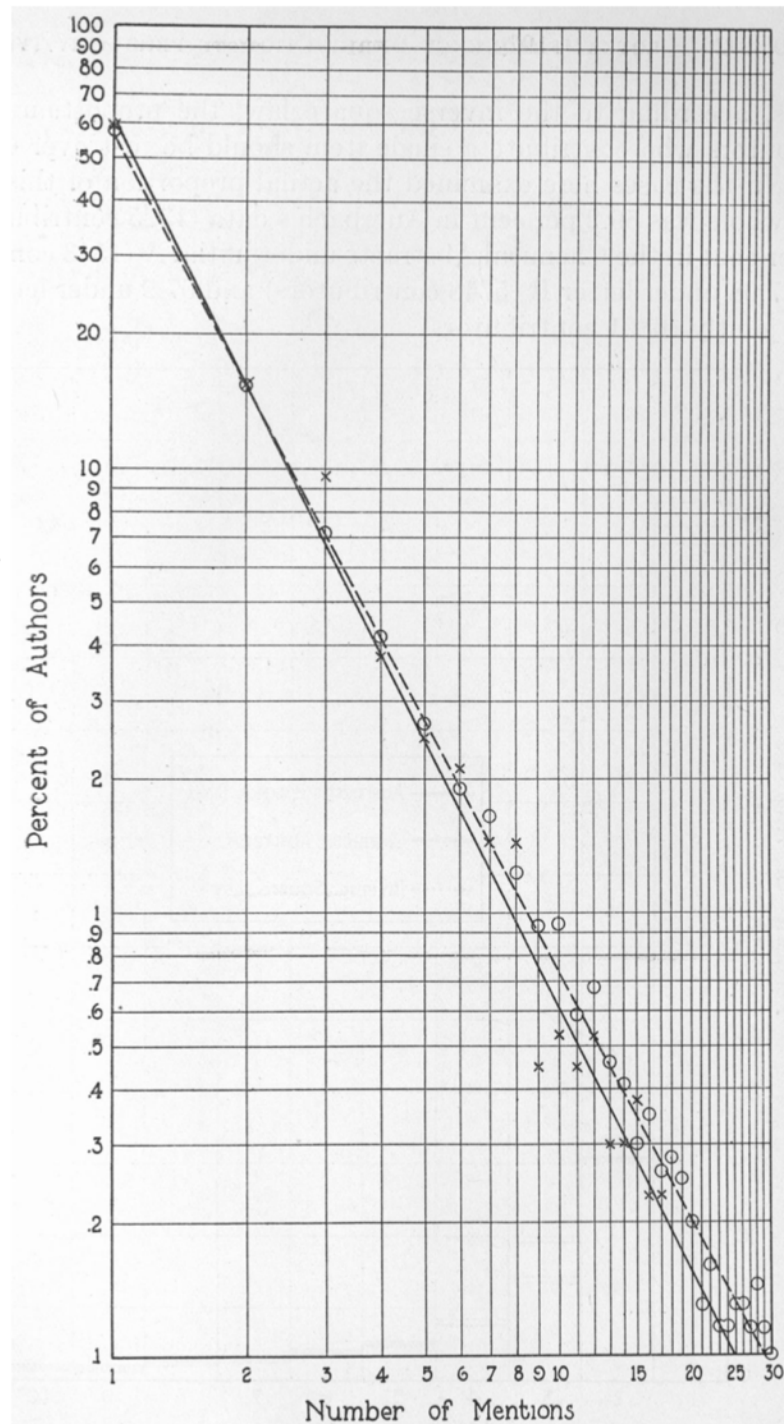
Annexe 12 : Langues des publications indexées dans le WoS

Recherche effectuée : toutes les publications disponibles dans le WoS publiées entre 2010 et 2018.



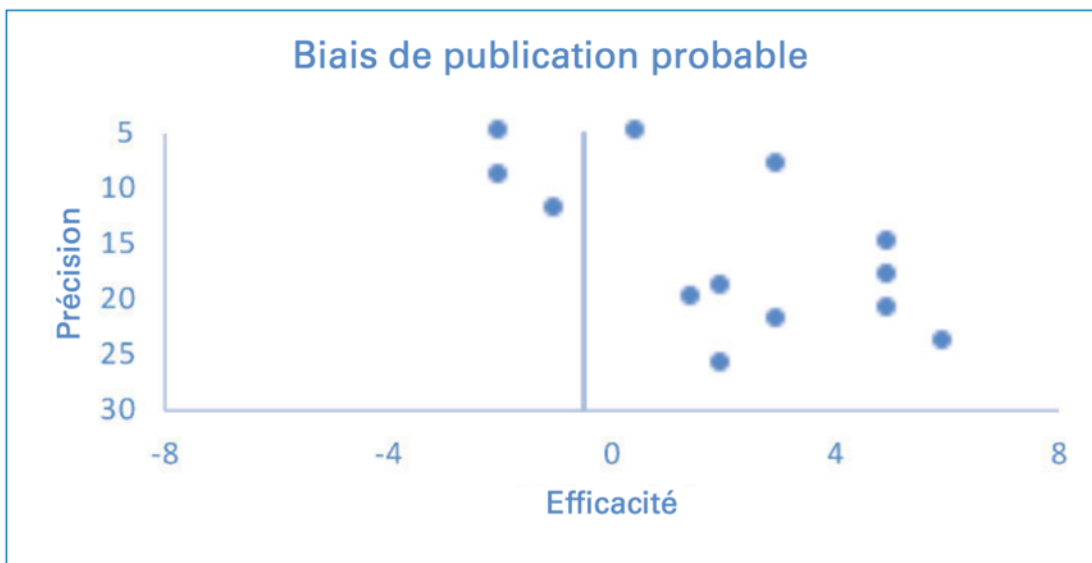
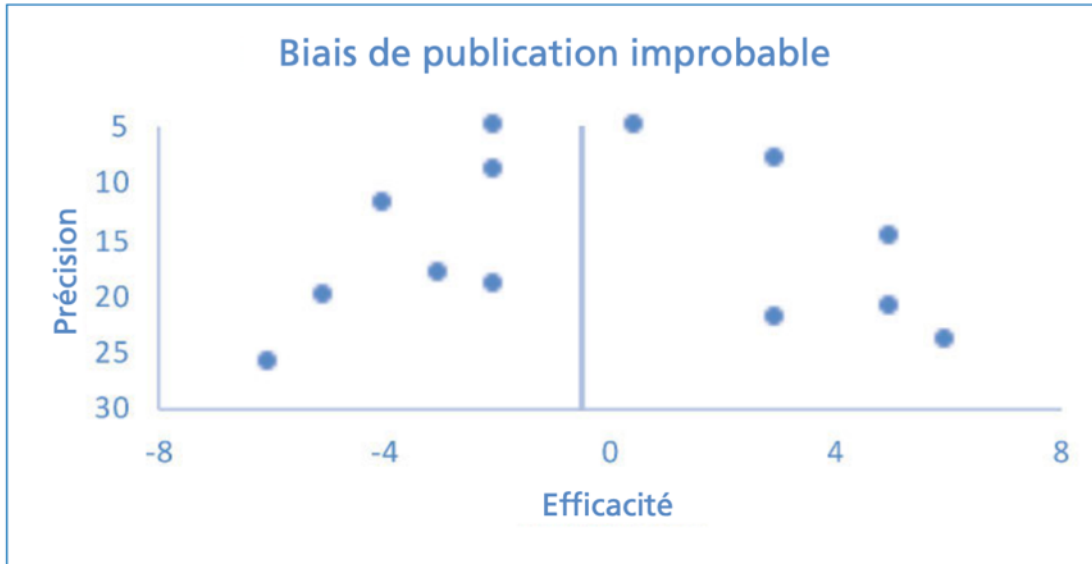
Source : CLARIVATES ANALYTICS 2020b. <https://apps.webofknowledge.com> [accès par abonnement].
[Consulté le 13 mars 2020]

Annexe 13 : Graphique illustrant la loi de l'inverse du carré appliquée aux citations par auteur (Lotka 1926)



Source : Lotka 1926, p. 322

Annexe 14 : Biais de publication (Staub 2020)



Source : Staub 2020, p. 18

Annexe 15 : Nombre moyen de citations par publication, deux ans après leur parution, livres inclus ou exclus : 1970-1978 (Moed et al. 1985)

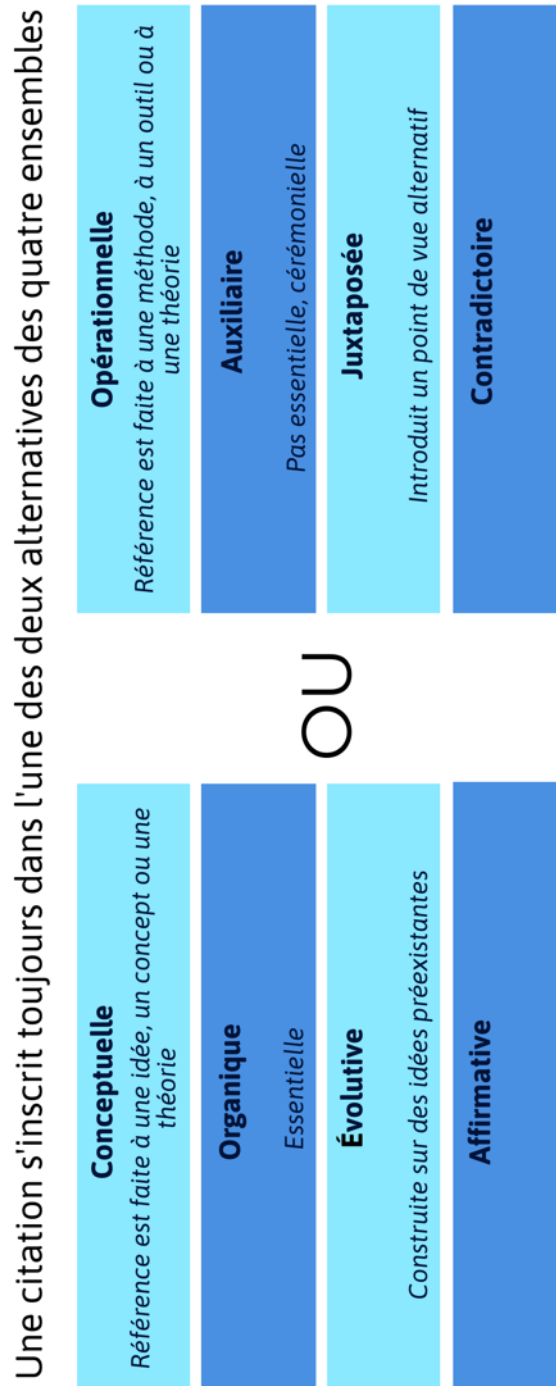
Table 2
Average number of citations (two years after publication-date) per publication for all publications from the Basic Leiden File, published between 1970 and 1978 (aggregated into four-year blocks). From 1977 ISI added source books to the *SCI* data-base. This table illustrates the effect of publications and citations from source books on the citation per publication ratio

Average number of citations per publication (two years after publication-date)	Publication-year blocks					
	70-73	71-74	72-75	73-76	74-77	75-78
including publications from books including citations from books	2.0	2.0	2.3	2.5	2.6	2.7
excluding publications from books including citations from books	2.0	2.0	2.3	2.5	2.7	2.8
excluding publications from books excluding citations from books	2.0	2.0	2.1	2.1	2.2	2.3

Source : Moed et al. 1985, p. 193

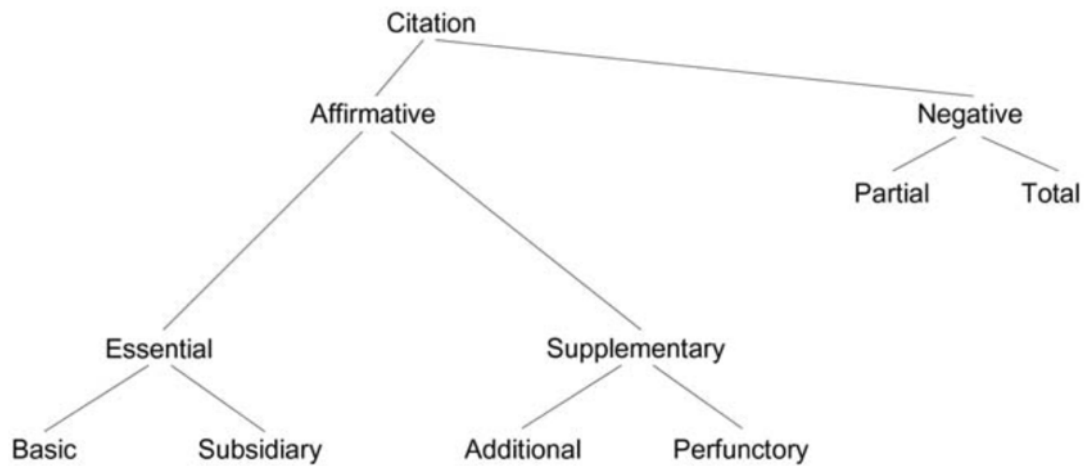
Moed et al. (1985) réalisent ce tableau avec les données du *Basic Leiden File* qui regroupe les publications de la faculté de médecine puis celles de la faculté de mathématiques et sciences naturelles. "This table illustrates the effect of publications and citations from source books on the average number of citations per publication." (Moed et al. 1985, p. 193)

Annexe 16 : Système de classification des citations (Moravcsik et Murugesan)



Source : adapté de Moravcsik-Murugesan (1975, cités dans Rousseau, Egghe et Guns 2018, p. 108)

Annexe 17 : Système de classification des citations (Chubin et Moitra)



Source : Chubin et Moitra (1975, cités dans Rousseau, Egghe et Guns 2018, p. 109)

Annexe 18 : Les quatre contextes de citation de White

Perfunctory-positive	"[...] pointing to related studies, providing background, invoking a prestigious name."
Perfunctory-negative	"[...] justifying a study by pointing to omissions in prior studies."
Organic-positive	"[...] discussing previous research in detail ; acknowledging concepts, methods, hypotheses introduced in earlier research ; pointing to research results which are essential for the author's research."
Organic-negative	"[...] showing errors in previous publications ; refuting the conclusions of earlier research."

Source : adapté de White (2001, cité dans Rousseau, Egghe et Guns 2018, p. 109)

Annexe 19 : Indicateur de Hirsch-Banks pour les composés et sujets de recherche

Tableau des indicateurs Hirsch-Banks (*h-b* index) pour les composés

Compound name	<i>h-b</i> index	<i>m</i>
PrPb ₃	6	0.26
TmGa ₃	6	0.30
Si ₂₈	17	0.31
CeB ₆	32	0.76
V ₃ Si	39	0.77
Ni ₂ MnGa	37	0.82
Nb ₃ Sn	48	0.94
MgB ₂	67	1.31
CeCu ₂ Si ₂	39	1.44
SrTiO ₃	94	1.96
GaN	144	2.12
C ₆₀	182	5.20

Source : Banks 2006, p. 164

Tableau des indicateurs Hirsch-Banks (*h-b* index) pour les sujets de recherche

Topic	<i>h-b</i> index	<i>m</i>
Borides	46	0.44
Pyrochlore	61	0.62
Spin flop	34	0.83
Optical lattice	43	0.90
Antiferroquadrupolar	18	1.00
Amorphous silicon	116	1.10
Spin frustration	30	1.36
Ferroelectricity	78	1.39
Spin liquid	45	1.55
Kondo and lattice	63	1.97
Perovskites	103	2.10
Spin ice	17	2.13
Magnetoresistance	172	2.39
Quantum information	65	2.41
Geometrical frustration	21	2.63
Quantum critical point	42	2.63
Porous silicon	104	3.25
Spin glass	108	3.38
Spin valve	48	3.43
Heavy fermion	97	3.73
Superstrings	99	3.96
Teleportation	61	5.08
Quantum computation	73	5.21
M-theory	79	6.58
Giant magnetoresistance	116	6.82
Fullerenes	140	7.78
Quantum dots	149	7.84
Nanowires	105	8.75
Carbon nanotubes	167	12.85

Source : Banks 2006, p. 166

Annexe 20 : Le *h*-index affiné par Schreiber (2007), exemples concrets

Data set	Total no.	Max. count	Kind of SCCs				Hirsch index	
			none	1st	2nd	3rd	h	h_s
								h_o
A	268	178	27	24	23	22	18%	
B	280	420	38	34	33	33	13%	
C	86	24	13	7	7	7	46%	
D	183	468	50	47	46	45	10%	
E	322	73	20	17	16	16	20%	
F	63	279	19	16	15	15	21%	
G	66	149	15	14	13	12	20%	
H	51	112	15	14	13	12	20%	
I	72	55	14	11	11	10	29%	
J	77	47	13	10	10	9	31%	
K	47	108	13	13	11	11	15%	
L	61	40	12	10	9	9	25%	
M	46	53	12	11	10	10	17%	
N	60	79	10	7	7	7	30%	
O	44	41	10	9	8	8	20%	
P	15	25	5	4	4	3	40%	

Où “*Total no.*” est le nombre total de publications et “*Max. count*” le nombre de citations le plus élevé.

Source : Schreiber 2007, p. 3002-5

Annexe 21 : Les *h*-index versus *g*-index de Leo Egghe et Henry Small

Les *h*-index (=13) et *g*-index (=19) de L. Egghe

TC	<i>r</i>	Σ TC	<i>r</i> ²
47	1	47	1
42	2	89	4
37	3	126	9
36	4	162	16
21	5	183	25
18	6	201	36
17	7	218	49
16	8	234	64
16	9	250	81
16	10	266	100
15	11	281	121
13	12	294	144
13	13	307	169
13	14	320	196
13	15	333	225
12	16	345	256
12	17	357	289
12	18	369	324
12	19	381	361
11	20	392	400

Source : Egghe 2006a, p. 8

Les *h*-index (=18) et *g*-index (=39) de H. Small

TC	<i>r</i>	Σ TC	<i>r</i> ²
305	1	305	1
239	2	544	4
127	3	671	9
109	4	780	16
86	5	866	25
80	6	946	36
77	7	1023	49
75	8	1098	64
67	9	1165	81
49	10	1214	100
44	11	1258	121
36	12	1294	144
26	13	1320	169
26	14	1346	196
25	15	1371	225
22	16	1393	256
22	17	1415	289
18	18	1433	324
18	19	1451	361
15	20	1466	400
12	21	1478	441
10	22	1488	484
9	23	1497	529
8	24	1505	576
8	25	1513	625
7	26	1520	676
6	27	1526	729
5	28	1531	784
5	29	1536	841
5	30	1541	900
3	31	1544	961
3	32	1547	1024
2	33	1549	1089
2	34	1551	1156
2	35	1553	1225
1	36	1554	1296
1	37	1555	1369
1	38	1556	1444
1	39	1557	1521
1	40	1558	1600

Source : Egghe 2006a, p. 9

Annexe 22 : Les *h*-index versus *g*-index et les valeurs de *g/h* des chercheurs décorés de la médaille de Price

Name	<i>h</i> -index	Name	<i>g</i> -index
Garfield	27	Garfield	59
Narin	27	Narin	40
Braun	25	Small	39
Van Raan	19	Braun	38
Glänzel	18	Schubert	30
Moed	18	Glänzel	27
Schubert	18	Martin	27
Small	18	Moed	27
Martin	16	Van Raan	27
Egghe	13	Ingwersen	26
Ingwersen	13	White	25
Leydesdorff	13	Egghe	19
Rousseau	13	Leydesdorff	19
White	12	Rousseau	15

Source : Egghe 2006b, p. 143

Name	<i>g/h</i>
Garfield	2.19
Small	2.17
White	2.08
Ingwersen	2.00
Martin	1.69
Schubert	1.67
Braun	1.52
Glänzel	1.50
Moed	1.50
Narin	1.48
Egghe	1.46
Leydesdorff	1.46
Van Raan	1.42
Rousseau	1.15

Source : Egghe 2006b, p. 143

Annexe 23 : le *g*-index affiné par Schreiber (2008) et comparé au *h*-index affiné

Data set	<i>h</i>	<i>h_s</i>	<i>h_s/h</i>	<i>g</i>	<i>g_o</i>	<i>g_c</i>	<i>g_s</i>	<i>g_s/h_s</i>	<i>g_s/g</i>	<i>g_s/h</i>
A	27	22	0.81	45	41	40	38	1.73	0.84	1.41
B	39	34	0.87	67	60	58	55	1.62	0.82	1.41
C	13	7	0.54	15	8	8	7	1.00	0.47	0.54
D	8	7	0.88	18	17	17	17	2.43	0.94	2.13
E	5	3	0.60	10	9	8	8	2.67	0.80	1.60
F	7	6	0.86	9	8	8	7	1.17	0.78	1.00
G	14	10	0.71	22	19	17	16	1.60	0.73	1.14
H	14	10	0.71	22	18	17	17	1.70	0.77	1.21
I	14	13	0.93	21	20	19	19	1.46	0.90	1.36

Source : Schreiber 2008, p. 197

Annexe 24 : Définition du h -index et de certaines de ses variantes par Bornmann, Mutz et Daniel (2008)

Index	Definition
h index	“A scientist has index h if h of his or her N_p papers have at least h citations each and the other ($N_p - h$) papers have fewer than $\leq h$ citations each” (Hirsch, 2005, p. 16569)
m quotient	$\frac{h}{y}$ where $h = h$ index, $y =$ number of years since publishing the first paper
g index	“The highest number g of papers that together received g^2 or more citations” (Egghe, 2006b)
$h(2)$ index	“A scientist's $h(2)$ index is defined as the highest natural number such that his $h(2)$ most-cited papers received each at least $[h(2)]^2$ citations” (Kosmulski, 2006, p. 4)
a index	$\frac{1}{h} \sum_{j=1}^h cit_j$ where $h = h$ index, $cit =$ citation counts
m index	The median number of citations received by papers in the Hirsch core (this is the papers ranking smaller than or equal to h)
r index	$\sqrt{\sum_{j=1}^h cit_j}$ where $h = h$ index, $cit =$ citation counts
ar index	$\sqrt{\sum_{j=1}^h \frac{cit_j}{a_j}}$ where $h = h$ index, $cit =$ citation counts, $a =$ number of years since publishing
h_w index	$\sqrt{\sum_{j=1}^{r_0} cit_j}$ where $cit =$ citation counts, $r_0 =$ the largest row index j such that $r_w(j) \leq cit_j$

Source : Bornmann, Mutz et Daniel 2008, p. 832

Annexe 25 : Avantages de certains indices par rapport au *h*-index (Bornmann et al. 2011)

Full name of <i>h</i> index variant	Short name of <i>h</i> index variant	Takes into account field dependence	Takes into account self-citations	Takes into account multi-authorship	Takes into account career length or age of the publications	Measures citation intensity in <i>h</i> core or gives more weight to highly cited papers	Author(s)
A-index	A index					X	Jin (2006)
AR-index	AR index				X	X	Jin (2007)
ch(2)-index	ch(2) index		X			X	Kosmulski (2006)
Contemporary <i>h</i> -index	ch index		X		X	X	Kosmulski (2006)
Degree <i>h</i> -index h_A	cont <i>h</i> index						Sidiropoulos et al. (2007)
Degree <i>h</i> -index h_p	degree h_A						Schubert et al. (2009)
<i>f</i> -index	degree h_P						Schubert et al. (2009)
<i>g</i> -index	<i>f</i> index					X	Tol (2009)
$h(2)$ -index	<i>g</i> index					X	Egghe (2006a)
<i>hg</i> -index	$h(2)$ index					X	Kosmulski (2006)
h_1 -index	<i>hg</i> index					X	Alonso, Cabrerizo, Herrera-Viedma, and Herrera (2010)
h index of first authored papers	h_1 index	X					Battista et al. (2006)
h_m -index	<i>h</i> first author			X			Opthof and Wilde (2009)
h_{maj} -index	hm index			X			Schreiber (2008b)
h_{ms} -index	hml index			X			Molinari and Molinari (2008)
h_s -index	hms index		X				Schreiber (2009b)
Index of quality and productivity	hs index		X				Schreiber (2007b)
major contribution <i>h</i> -index	hw index				X	X	Egghe and Rousseau (2008)
Maxprod	IQP	X			X	X	Antonakis and Lalive (2008)
<i>m</i> -index	<i>h</i> maj			X			Hu et al. (2010)
Modified impact index	Maxprod					X	Kosmulski (2007)
Multidimensional <i>h</i> index (second component: h_2)	<i>m</i> index					X	Bornmann et al. (2008)
Multidimensional <i>h</i> index (third component: h_3)	MII					X	Sypsa and Hatzakis (2009)
Prathaps h_2 index	<i>m</i> quotient			X			Garcia-Perez (2009)
Raw <i>h</i> rate	multidim h2				X		Garcia-Perez (2009)
R-index	multidim h3						Vinkler (2009)
Specific impact index	Pi index					X	Prathap (2006)
Tapered <i>h</i> -index	Prathap h2					X	Cabrerizo et al. (2010)
Weighted <i>h</i> -index	$q(2)$ index				X		Burrell (2007)
Wohlin index	Raw <i>h</i> rate					X	Jin et al. (2007)
Wu index	R index				X	X	de Visscher (2010)
	s index					X	Anderson, Hankin, and Killworth (2008)
	ht index					X	Tol (2009)
	t index					X	Lee et al. (2009)
	<i>h</i> weight			X			Wohlin (2009)
	wohlin index					X	Wu (2010)
	wu index					X	
	Total	2	4	6	6	17	

Source : Bornmann et al. 2011, p. 349

Annexe 26 : Tableau des avantages et inconvénients des principaux indicateurs bibliométriques relevés dans la présente recherche

Indicateur	P	Avantages	Inconvénients	Remarques
<p>Nombre total de publications</p>	<p>P</p>	<p>i. Mesure la productivité (Hirsch 2005 ; Van Raan 2006) ii. Comparer ce nombre à celui d'autres sujets d'études permet d'obtenir « des mesures plus significatives de la puissance relative des sujets étudiés. » (OCDE 1997, p. 26)</p>	<p>i. Considéré individuellement, ce nombre est un « indicateur bibliométrique grossier » (OCDE 1997, p. 26). ii. Ne mesure pas l'impact des publications (Hirsch 2005 ; Van Raan 2006) iii. Les « biais de publication » sont occultés (Staub 2020) iv. Cette seule mesure ne rend pas compte de la qualité des travaux (OCDE 1997) ; v. Cet indicateur ne représente pas la quantité de travail fournie par le chercheur, d'autant plus s'il a des coauteurs (OCDE 1997)</p>	<p>i. Si [...] les données de base sont divisées par le nombre des chercheurs ou par les sommes d'argent investies, on obtient des indicateurs "dérivés" pouvant permettre, en quelque sorte, une analyse de la "productivité" des travaux en question. (OCDE 1997, p. 26). ii. Certains font référence au nombre de publications dans des revues intégrées dans le « citation index » (WoS) du domaine scientifique du chercheur (Van Raan 2006 ; Costas et al. 2009). Price (1986) dénonce le seul recours au nombre de publications dans des revues considérées prestigieuses, notamment en raison de ses effets sur la culture POP.</p>
<p>Nombre de publications considérées comme importantes, soit celles ayant reçu plus de c citations</p>		<p>i. Élimine les inconvénients des mesures que sont le nombre de publications, le nombre total de citations, le nombre de citations par publications (Hirsch 2005) ii. Renseigne quant à l'impact global dans le temps (Hirsch 2005)</p>	<p>i. c est arbitraire (Hirsch 2005) ii. c doit être ajusté selon le niveau « d'ancienneté » (Hirsch 2005) iii. Il favorise ou pénalise le chercheur de façon aléatoire (Egghe 2006a)</p>	

Indicateur	C	Avantages	Inconvénients	Remarques
<p>Nombre total de citations</p>	<p>C</p>	<p>i. Mesure l'impact global (Hirsch 2005) ii. L'indicateur est considéré comme objectif du fait qu'il serait dénué de toute interprétation personnelle (Brown et Gardner 1985 puis Beattie et Goodacre 2006, cités dans Moosa 2018) iii. Les mesures de citations seraient une forme valide du jugement par les pairs et apportent un élément objectif à l'évaluation (Garfield 1979) iii. Il est une mesure globale du niveau de contribution d'un chercheur à la science (Garfield 1979)</p>	<p>i. Il n'est pas évident à trouver et peut être biaisé si certaines publications largement relayées ont été co-rédigées (Hirsch 2005) ii. Les articles de synthèse, plus fréquemment cités ne sont pas distingués des autres publications (Hirsch 2005) iii. Il « signifie généralement peu de chose en lui-même ou en comparaison avec d'autres disciplines [...] » (Gingras 2014a, p. 33)</p>	<p>i. Pour avoir recours à cet indicateur, Garfield (1979) précise qu'il faut [1.] connaître le domaine scientifique en question [2.] être conscient des subtilités et des limites de ces données [3.] observer la méthodologie pertinente et [5.] être capables d'interpréter les résultats ii. Il faut être conscient des facteurs pouvant influencer l'indicateurs tels que [1.] l'évolution du nombre de chercheurs [2.] des modifications effectuées dans les bases de données [3.] les habitudes de publication et de citation selon les domaines scientifiques (Moed et al. 1985) ou encore [4.] la tendance de certains chercheurs de se référer à la réputation de leurs confrères en vue de la sélection des publications qui leurs servent de référence (Merton 1968) iii. Les intentions derrière les citations peuvent être de différentes natures (Le Coadic 2005). iv. L'autocitation permet, théoriquement, de manipuler le taux de citation, taux qui n'est pas facilement déterminable (cf. les taux avancés par Garfield et Sher 1963 ; Garfield 1979 ; Schreiber 2008), d'autant plus que le terme n'est pas toujours défini de la même manière (Moed et al. 1985) v. Le taux de citation ne devrait pas être comparé entre les différents types de publication (Université Laval 2020) vi. « [...] il doit toujours être rapporté aux pratiques de la discipline du chercheur à une période donnée. » (Gingras 2014a, p. 33) vii. Certains calculent le nombre de citations reçues par les publications dans des revues intégrées dans le « citation index » (WoS) du domaine scientifique du chercheur (Van Raan 2006 ; Costas et al. 2009)</p>

Indicateur	Avantages	Inconvénients	Remarques
Nombre moyen de citations par auteur	Est la mesure la plus précise du potentiel de citation – lequel peut varier de manière significative entre les domaines scientifiques (Garfield 1979[p], cité dans Moed et al. 1985)		« Dépend fortement de la taille des disciplines et des spécialités » (Gingras 2014a, p. 33)
Nombre de citations par publication	Permet de comparer les scientifiques toutes générations confondues (Hirsch 2005)	i. N'est pas évident à trouver (Hirsch 2005) ii. Récompense une faible productivité (Hirsch 2005) iii. Pénalise une productivité élevée (Hirsch 2005)	i. « [...] dépend du nombre de références que les articles contiennent. Plus ce nombre est élevé, plus les chances d'être cité augmentent. » (Gingras 2014a, p. 33) ii. Il peut être calculé sans les autocitations (Van Raan 2009 ; Costas et al. 2009). Ce qui est entendu par « autocitation » doit être clairement défini (Moed et al. 1985). iii. Il est également possible de calculer le nombre moyen des citations par publication (CPP), y compris sans les autocitations (Van Raan 2009 ; Costas et al. 2009)
Nombre de citations de chacune des q publications les plus citées	Permet de maîtriser plusieurs des inconvénients des autres mesures (Hirsch 2005)	i. q est arbitraire et favorise ou pénalise le chercheur de façon aléatoire et ii. q n'est pas une valeur unique, l'indicateur ne peut aisément être comparé (Hirsch 2005 ; Egghe 2006)	

Indicateur		Avantages	Inconvénients	Remarques
h -index	h	<p>i. Il permet d'estimer l'impact global (Schreiber 2007)</p> <p>ii. "Le h-index est robuste car il est insensible [1.] à un excès accidentel d'articles non cités [...] [2.] à un ou plusieurs articles extrêmement cités." » (Bommann et al. 2005, Braun et al. 2005, Glänzel 2006a cités dans Egghe et Rousseau 2006, p. 122 ; Egghe 2006a)</p> <p>iii. « Il combine les effets de la quantité (nombre de publications) et de la qualité (nombre de citations). » (Bommann et al. 2005, Braun et al. 2005, Glänzel 2006a cités dans Egghe et Rousseau 2006, p. 122 ; Egghe 2006a)</p> <p>iv. Il pourrait rendre les procédures de nominations plus objectives et transparentes (Van Raan 2006)</p>	<p>i. C'est un indicateur complémentaire, non un substitut (Egghe et Rousseau 2006)</p> <p>ii. Est un indicateur invalide du fait qu'il ne remplit pas deux des trois critères de validités de Gingras (2014b) à savoir [1.] L'adéquation entre l'indicateur et ce qu'il doit mesurer et [2.] L'homogénéité des dimensions de l'indicateur (Gingras 2014b, p. 118-119)</p> <p>iii. Il est <i>de facto</i> plus élevé pour les chercheurs plus avancés dans leur carrière, quand bien même ils ne publient pas de façon continue (Hicks et al. 2015)</p>	<p>i. Sa valeur varie d'une base de données à une autre (Hicks et al. 2015)</p> <p>ii. Sa valeur de référence varie selon les disciplines (Hirsch 2005)</p> <p>iii. Son rapprochement ne peut se faire que pour les chercheurs d'un même domaine scientifique (Schreiber 2007)</p> <p>iv. Il ne faut pas comparer les h-index de chercheurs qui n'en sont pas au même stade de développement de leur carrière (Vuillemin-Raval 2018)</p> <p>v. Il doit être combiné avec d'autres facteurs (Hirsch 2005)</p> <p>vi. Il ne devrait pas être utilisé en matière d'évaluation (Gingras 2014b)</p>
h -index sans les autocitations	h_o			Il ne fait pas considérablement varier le h -index (Schreiber 2007)
h -index (auto)citations des auteurs	h_c			Il réduit plus encore le h -index (Schreiber 2007)
h -index (auto)citations entre plusieurs co-auteurs en étudiant chacune des publications citant	h_s			Fastidieux à calculer et des erreurs peuvent plus facilement être faites. Toutefois, selon Schreiber (2007), l'effet sur le h -index entre h_o et h_s est infime, si ça n'est nul.

Indicateur		Avantages	Inconvénients	Remarques
Indice Hirsch-Banks	<i>hb</i>	<ul style="list-style-type: none"> i. Il permet d'identifier les sujets et composés « en vogue » (Banks 2006) ii. Il permet d'évaluer l'état d'avancée de la recherche dans un domaine donné sur un sujet ou un objet donné (Banks 2006) iii. Il peut servir lors de l'attribution de fonds pour la recherche (Banks 2006) 		
<i>g</i> -index	<i>g</i>	<ul style="list-style-type: none"> i. Il conserve les avantages du <i>h</i>-index et corrige son insensibilité aux articles les plus cités (Egghe 2006a) ii. Il facilite la comparaison entre les chercheurs évoluant dans un même domaine (Egghe 2006a) iii. Il rend mieux compte de la visibilité de l'ensemble des publications du chercheur (Egghe 2006b) 		<p>À l'instar de son étude sur le <i>h</i>-index (Schreiber 2007), Schreiber (2008) propose d'étudier les indicateurs que sont [1.] le <i>g</i>-index sans les auto-citations (g_o) [2.] le <i>g</i>-index sans les (auto)citations des co-auteurs (g_c) et [3.] le <i>g</i>-index sans les (auto)citations de tous les co-auteurs (g_s)</p>

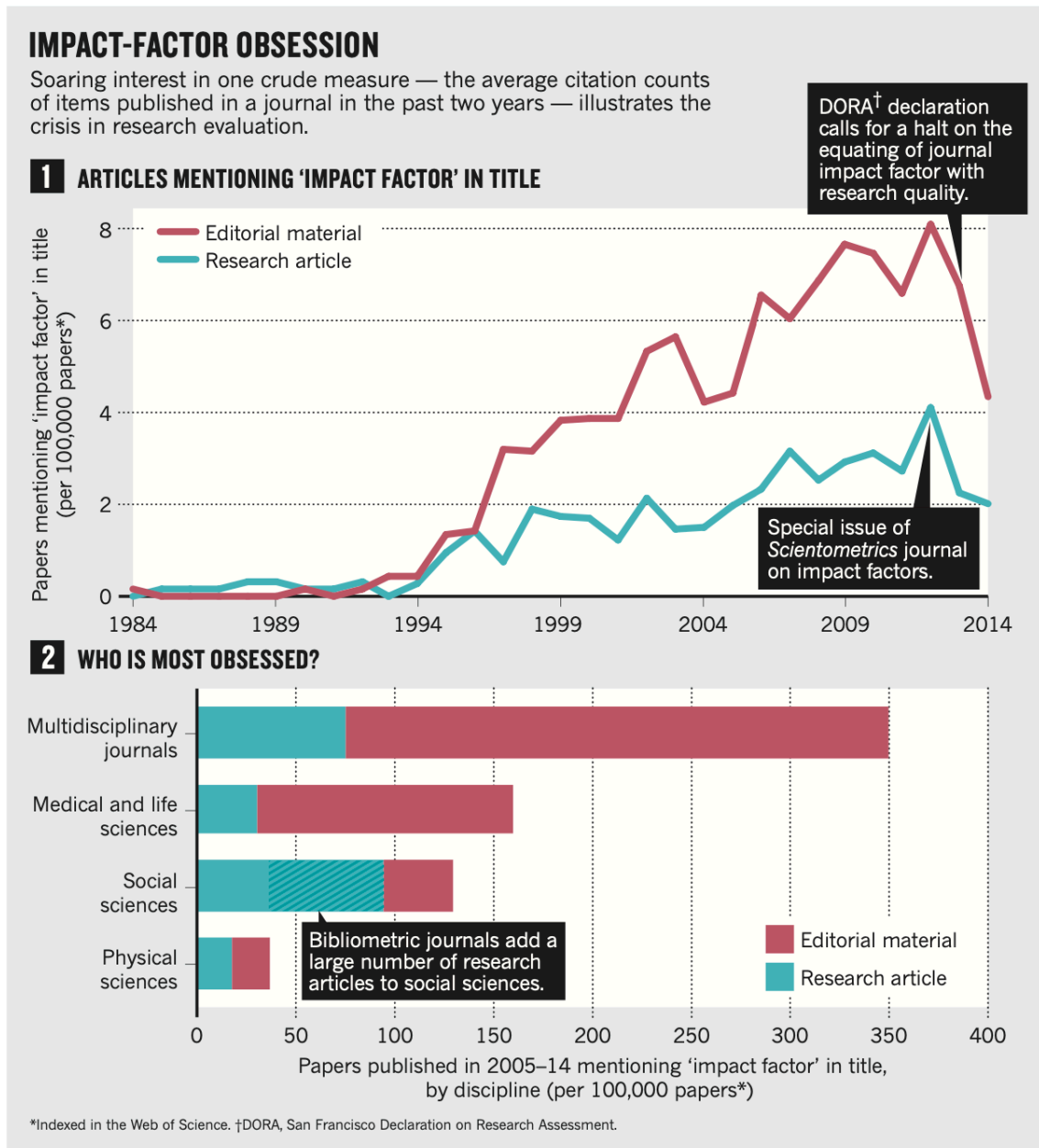
Indicateur	Avantages	Inconvénients	Remarques
Facteur d'impact *	FI Il est plus robuste que le FID lorsqu'il s'agit d'évaluer la revue et son impact global (Ingwersen et al. 2001)	i. « les distributions des citations dans les revues sont très asymétriques » (Déclaration de Dora 2012, [p. 1]) ii. « les propriétés du facteur d'impact sont propres à chaque domaine : il s'agit d'un agrégat de types d'articles multiples et très divers, avec des articles primaires comme de synthèse » (Déclaration de Dora 2012, [p. 1]) iii. les facteurs d'impact peuvent être manipulés (ou « instrumentalisés ») par une politique éditoriale (Déclaration de Dora 2012, [p. 1]) iv. « les données utilisées pour calculer les facteurs d'impact ne sont ni transparentes ni ouvertement accessibles au public » (Déclaration de Dora 2012, [p. 1])	
Facteur diachronique d'impact	FID i. Par rapport au FI, le FID est plus à même d'exprimer l'impact actuel d'une revue (Ingwersen et al. 2001) ii Il recoupe des articles publiés la même année que la publication évaluée (Ingwersen et al. 2001)		

* N.B. nous ne signalons pas ici les nombreux arguments formulés à l'encontre du recours au FI en matière d'évaluation. Ils se trouvent dans le chapitre 4.5.2.1.

Indicateur	Avantages	Inconvénients	Remarques
Journal Citation Score sans les autocitations	JCS		<p>i. Il prend une période de citation plus longue que l'IF – généralement 5 ans (Zweibibliothek Medizin 2020)</p> <p>ii. Il peut être calculé avec ou sans les autocitations (Van Raan 2006 ; Costas et al. 2009 ; Šember, Utročić et Petrak 2010 ; Zweibibliothek Medizin 2020)</p>
Mean Journal Citation Score	JCS _m	<p>i. Il permet de comparer des domaines avec des habitudes de publication et de citation différentes (Zweibibliothek Medizin 2020)</p>	<p>i. Le JCS est déterminé pour l'ensemble des revues dans lesquelles une institution académique ou un groupe de recherche ont publié. Il représente la valeur de référence d'un groupe [...] basé sur les revues dans lesquelles il publie (Zweibibliothek Medizin 2020).</p> <p>ii. Il peut aussi être appliqué aux chercheurs (Costas et al. 2009)</p>
Comparaison du nombre de citation par publication (CPP) par rapport au taux de citation moyen de la revue (JCS _m) sans les autocitations	CPP/JCS _m	<p>Il permet d'observer si l'impact d'un chercheur est au-dessus ou au-dessous la moyenne internationale dans les revues dans lesquelles il publie (Costas et al. 2009)</p>	<p>i. Constitue le crown indicator (Van Raan 2006 ; Costas 2009 ; Waltman et al. 2011)</p> <p>ii. Si CPP/JCS_m < 1, les chercheurs obtiennent moins de citations que la revue dans laquelle ils publient (Costas et al. 2009)</p>
Field/subfield specific citation score soit le score de citation spécifique au domaine/sous-domaine	FCS		<p>Il peut être généré avec ou sans les autocitations (Van Raan 2006 ; Costas et al. 2009)</p>
Score moyen (mean) de citation spécifique au domaine (soit au sous-champ pour Costas et al. 2009)	FCS _m	<p>i. Il rend compte de la densité de citation (Van Raan 2006 ; Costas et al. 2009)</p> <p>ii. Il est pertinent lorsque différentes disciplines sont regroupées au sein d'un domaine scientifique (Van Raan 2006 ; Costas et al. 2009)</p>	<p>Costas et al. (2009, p. 741) se réfèrent à cet indicateur comme au « field-specific citation density* » (p. 741)</p> <p>* La densité de citation = le nombre de citations divisé par le nombre d'années depuis la publication (Jelicic Kadic et al. 2020)</p>

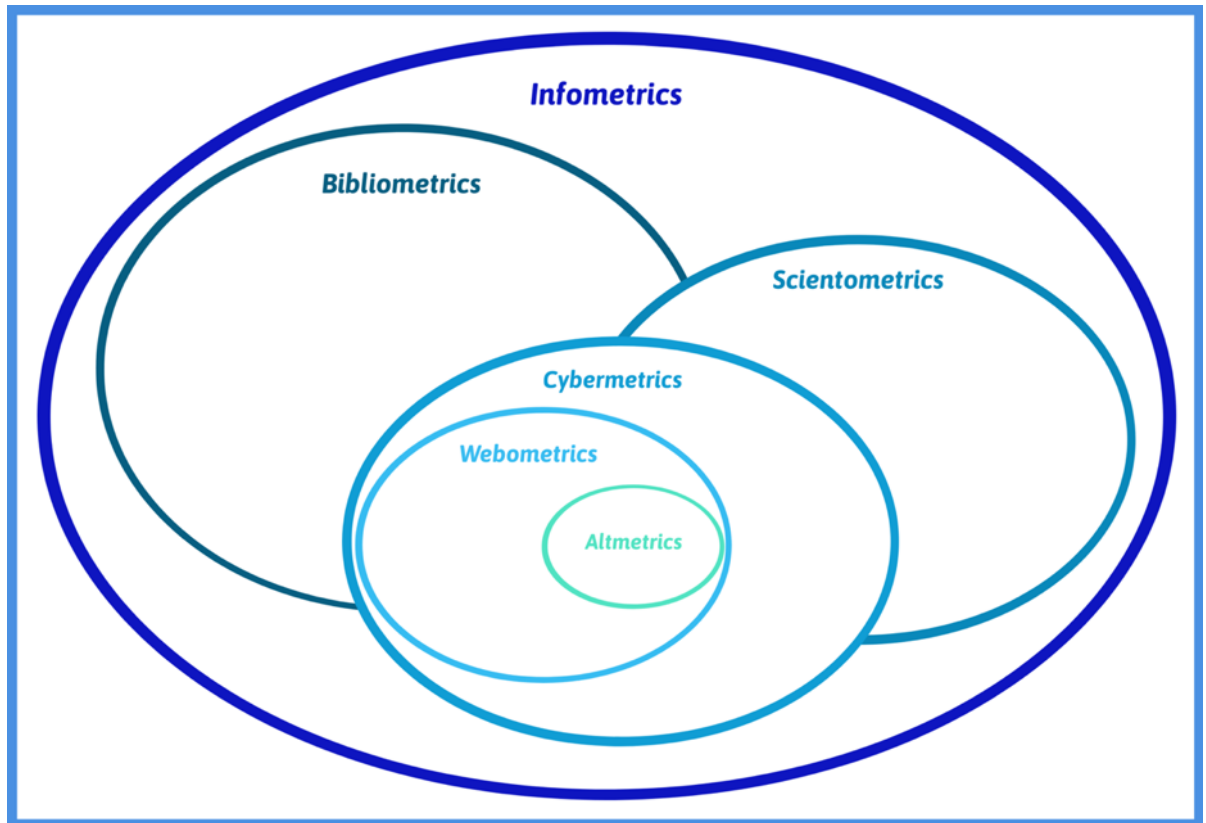
Indicateur	Avantages	Inconvénients	Remarques
<p>Comparaison du nombre de citation par publication (CPP) par rapport au score moyen (mean) de citation spécifique au domaine (FCSm) sans les autocitations</p>	<p>Il permet d'observer si la performance de l'unité d'analyse est bien au deçà (<0.5), au deçà (0.5-0.8), proche (0.8-1.2), au-dessus (1.2-1.5) ou bien au-dessus (>1.5) de l'impact standard international dans le domaine concerné (Costas et al. 2009, p. 742).</p>		<p>i. Certains considèrent qu'il existe une bonne corrélation entre le CPP/FCSm et une révision de qualité par les paires (Rinia, van Leeuwen, van Vuren et Van Raan 2001, cités dans Costas et al. 2009) ii. L'indicateur est forcément bas pour les publications récentes (Waltman et al. 2011). Les auteurs testent donc le MNCS (mean normalized citation score) qui accorde le même poids à toutes les publications, quelle que soit l'année de leur publication : les plus récentes introduisent du bruit => Waltman et al. 2011 proposent de ne pas en tenir compte ni dans le CPP/FCS_m, ni dans le MNCS. iii. Si CPP/FCSm < 1 : le nombre de citations reçues est inférieur à la moyenne du domaine (Costas et al. 2009). Si JCS_m/FCSm > 1, le chercheur a tendance à publier dans des revues dont l'impact est supérieur à la moyenne dans son domaine (Costas et al. 2009)</p>
<p>Indicateur d'impact relatif de la revue normalisé par rapport au domaine (relative, field-normalized journal IF)</p>	<p>Il indique si le chercheur publie dans des revues ayant un fort ou faible impact dans le domaine (Costas et al. 2009)</p>		

Annexe 27 : L'obsession du FI (Hicks et al. 2015)



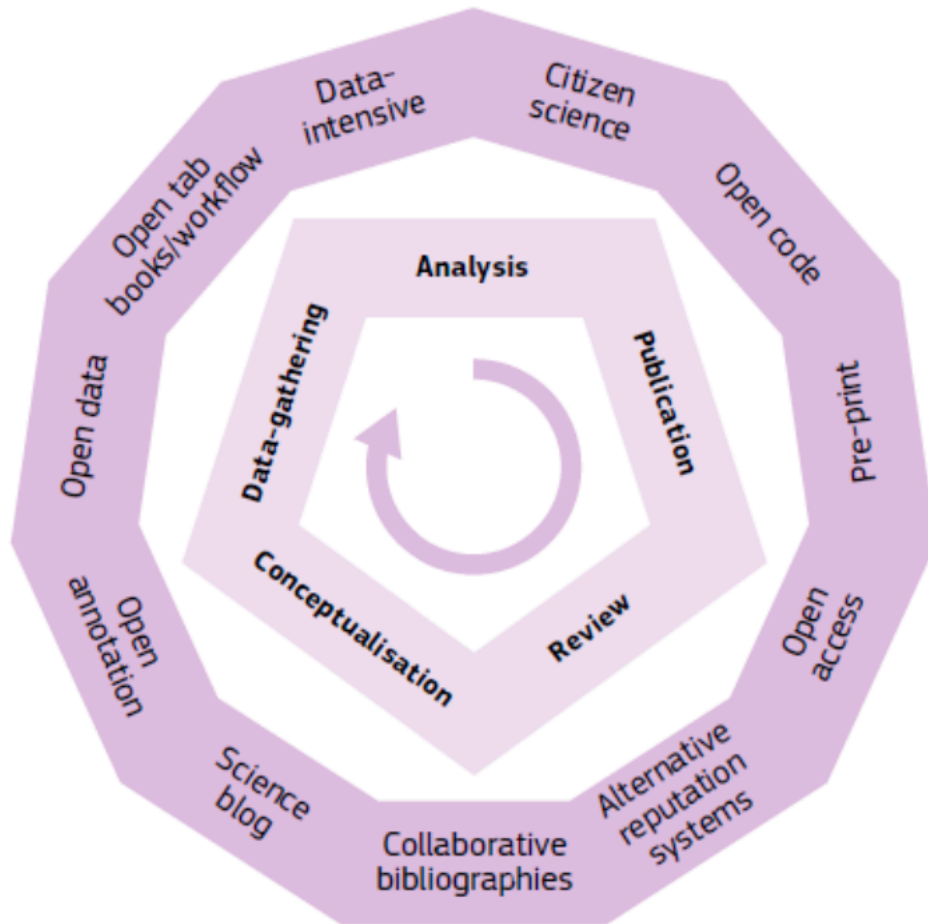
Source : Hicks et al. 2015, p. 43

Annexe 28 : Situer les *webometrics* et *altmetrics*



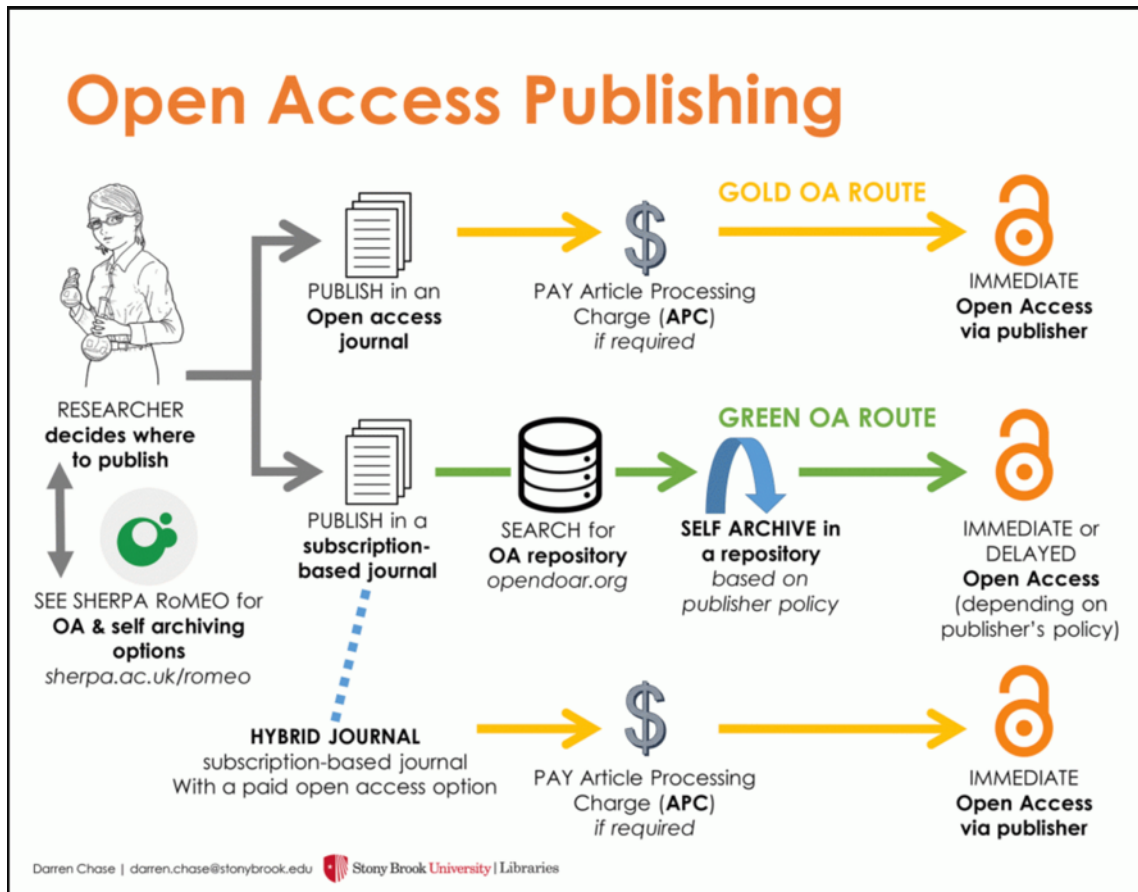
Sources : Traduit et adapté de Wilsdon et al. 2017, p. 8 – lesquels avaient adaptés un schéma produit par Björneborn et Ingwersen 2004

Annexe 29 : L'OS et l'ouverture du processus de recherche



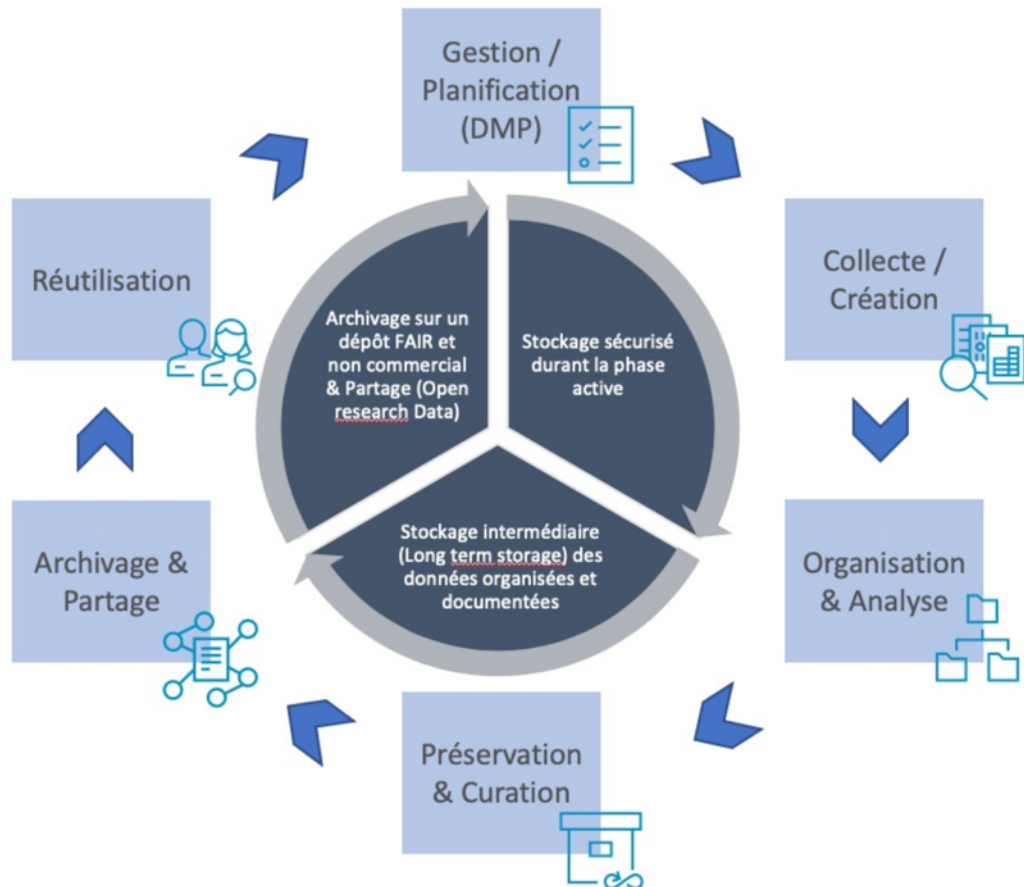
Source : Wilsdon et al. 2017, p. 6

Annexe 30 : Publier en OA



Source : UNIVERSITY OF CANTERBURY, 2020.
<https://canterbury.libguides.com/c.php?g=894027&p=6640992> [Consulté le 3 juillet 2020].

Annexe 31 : Cycle de vie des données de la recherche (UNIL – UNIRIS)



UNIL - UNIRIS 2019



Source : UNIVERSITÉ DE LAUSANNE (UNIL) ET SERVICE DES RESSOURCES INFORMATIONNELLES ET ARCHIVES DE L'UNIL (UNIRIS). Cycle de vie des données de la recherche. *UNIL* [en ligne]. [Consulté le 3 juillet 2020]. Disponible à l'adresse : <https://www.unil.ch/openscience/fr/home/menust/open-research-data/les-donnees-de-recherche/cycle-de-vie-et-types-de-donnees.html>

Annexe 32 : Guide d'entretien et questionnaire de base

1. Au sein de la Faculté *FFFF*, procédez-vous à l'évaluation des publications
 - Au niveau de la Faculté
 - Et/ou au niveau des départements ?
 - Le cas échéant, il y a-t-il des critères d'évaluation communs ?

2. Dans quel(s) contexte(s) et à quelles fins procédez-vous à l'évaluation des publications de vos chercheurs

- Avancée de la carrière académique (promotions...) OUI NON
Précisions :

- Octroi de fonds pour la recherche OUI NON
Précisions :

- À d'autres fins ?

- Est-ce le Décanat de la Faculté *FFFF* qui est en charge de cette/ces évaluation(s) ?

- Merci de préciser s'il y a des évaluations qui se font sur la « seule » base des publications :

- Les critères d'évaluation des publications sont-ils les mêmes si la finalité de l'évaluation n'est pas la même ? (mêmes types de publications, mêmes indicateurs...)

3. L'évaluation des publications est-elle individuelle ou peut-elle aussi, dans certains cas, concerner les groupes de recherches ?

- L'évaluation des publications se fait-elle au niveau individuel ? OUI NON
Précisions :

- L'évaluation se fait-elle aussi au niveau de groupes de recherches ? OUI NON
Précisions :

- Les finalités de l'évaluation sont-elles les mêmes pour les chercheurs et pour les groupes de recherche (avancée académique, octroi de fonds...) ? OUI NON
Précisions :

4. Le dépôt de la publication dans le dépôt institutionnel est-il une condition pour sa prise en considération dans l'évaluation ?

• Seules les publications déposées dans l'archive ouverte sont prises en compte dans l'évaluation OUI NON

• Les publications prises en compte dans l'évaluation ne se limitent à celles déposées dans l'archive ouverte. OUI NON

Précisions :

• La liste des publications est transmise par le chercheur OUI NON

• La Faculté *FFFF* dispose de sa propre base de données des publications des chercheurs OUI NON

• Les publications archivées dans d'autres bases de données sont également prises en considération OUI NON

Précisions :

• Commentaires :

5. Les types de publications concernés aux fins de l'évaluation individuelle des chercheurs

- Quelles sont les publications prises en considération aux fins de l'évaluation des chercheurs ?
- Les publications prises en considération varient-elles en fonction du type d'évaluation (cf. question 2) ? Dans quelle mesure ? Merci de préciser :

6. Les types de publications concernés aux fins de l'évaluation des groupes de chercheurs [si applicable]

- Le cas échéant, les publications prises en considération sont-elles les mêmes que pour l'évaluation individuelle ?
- Les publications prises en considération varient-elles en fonction du type d'évaluation (cf. question 2) ? Dans quelle mesure ? Merci de préciser :

7. Quels sont les critères d'évaluation des publications ?

• Avez-vous recours à des indicateurs bibliométriques ? OUI NON

• Quels sont ces indicateurs ?

• - Le nombre de publications OUI NON

Merci de bien vouloir préciser si toutes les publications sont considérées de la même manière ou si la « valeur » attribuée dans

le cadre de l'évaluation varie en fonction du type de publication :

- - Le nombre de citations

Merci de bien vouloir indiquer si, et alors de quelle manière, vous tenez compte de la place du chercheur dans la liste des auteurs dans les cas où ils seraient plusieurs : OUI NON

- - Le *h*-index

Si la réponse est oui, merci de bien vouloir préciser :

- la base de données dont est issu le *h*-index : OUI NON
- s'il est calculé pour l'ensemble de la carrière et/ou pour une période donnée :

- - Autres indicateurs

Merci de bien vouloir préciser : OUI NON

- Les mêmes indicateurs bibliométriques sont-ils utilisés dans le cadre de chacune des évaluations dans lesquelles les publications scientifiques sont prises en considération (avancée de la carrière académique et octroi de fonds pour la recherche p. ex.) OUI NON

- Si tel n'est le cas, merci de bien vouloir préciser :

- Avez-vous recours aux mêmes indicateurs bibliométriques si l'évaluation des publications était également réalisée à différents niveaux (« individuel » et groupes de recherche p. ex.) OUI NON

- Si tel n'est le cas, merci de bien vouloir préciser :

- Il y a-t-il d'autres critères pris en considération pour l'évaluation des publications de vos chercheurs ? OUI NON

- Si tel est le cas, merci de bien vouloir préciser :

8. Comment procédez-vous à l'évaluation en cas de publications co-écrites par plusieurs chercheurs ?

9. Candidatures de chercheurs externes

Les publications des chercheurs candidats à un poste de recherche ou d'enseignement sont-elles prises en considération au moment de l'étude de son dossier ?

- Sur quelle base évaluez-vous sa production scientifique ?
 - Curriculum vitae ?
 - Liste des publications ?
 - Avec recherche dans les bases de données (WoS, Scopus...) ?
- Avez-vous recours à des indicateurs bibliométriques ?

Annexe 33 : Synthèse comparative des données concernant l'évaluation des publications au sein des institutions académiques échantillonnées

Qui est en charge de l'évaluation ?					
Domaine Santé HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)	Faculté de biologie et de médecine UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté de médecine UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté des lettres UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté des lettres UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Domaine Economie et Services HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)
Attribution des points par le Conseil de domaine sur proposition du Comité Ra&D <i>Source : Conseil de domaine Santé de la HES-SO (2018)</i>	Décanat de la Faculté <i>Source : Faculté de biologie et de médecine de l'UNIL [2020a]</i>	Commission de la relève et de la recherche de la Faculté <i>Source : Entretien avec la Faculté de médecine de l'UNIGE</i>	Décanat de la Faculté <i>Source : Questionnaire envoyé à la Faculté de lettres de l'UNIL</i>	Nomination : commission de nomination Avancée académique : commission de renouvellement <i>Source : Entretien avec la Faculté des lettres de l'UNIGE</i>	Comité Ra&D du domaine (délégation du Conseil de domaine) <i>Source : Rectorat de la HES-SO (2017b)</i>
Contexte :					
Octroi de fonds pour la recherche					
Niveau groupes de recherche					
Décanat de la Faculté via MIMOSA					
<i>Source : Entretien avec la Faculté de médecine de l'UNIGE</i>					

Le dépôt dans l'archive institutionnelle est-il requis ?					
Domaine Santé HES-SO	Faculté de biologie et de médecine UNIL	Faculté de médecine UNIGE	Faculté des lettres UNIL	Faculté des lettres UNIGE	Domaine Économie et Services HES-SO
Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)
Non précisé	Oui	Non	Aujourd'hui non, mais dès la fin de l'année [2020], seules les publications déposées dans SERVAL (OA ou pas) intégreront l'évaluation.	Non	Les publications 11 et 12 doivent être archivées dans AODES. Les publications et apparitions médias 14 : doivent être transmises à l'institut.
Source : <i>Conseil de domaine Santé de la HES-SO (2018)</i>	Source : <i>Faculté de biologie et de médecine de l'UNIL [2020a] et [2020b]</i>	Source : <i>Entretien avec la Faculté de médecine de l'UNIGE</i>	Source : <i>Questionnaire envoyé à la Faculté de lettres de l'UNIL</i>	Source : <i>Entretien avec la Faculté des lettres de l'UNIGE</i>	Source : <i>Conseil de domaine Économie et Services de la HES-SO (2015)</i>
		Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche			
		Oui			
		Source : <i>Entretien avec la Faculté de médecine de l'UNIGE</i>			

Publications entrant dans le champ de l'évaluation (vue épurée pour le Domaine Économie et Services de la HES-SO)				
Domaine Santé HES-SO	Faculté de biologie et de médecine UNIL	Faculté de médecine UNIGE	Faculté des lettres UNIL	Faculté des lettres UNIGE
<p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p> <p>Articles peer-reviewed (0.2 points par publication) ;</p> <p>Livres (0.2 points par publication) ;</p> <p>Articles non peer-reviewed (0.1 points par publication) ;</p> <p>Chapitres de livres (0.1 points par publication).</p> <p>Source : <i>Conseil de domaine Santé de la HES-SO (2018)</i></p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p> <p>Articles ("journal articles, brief/short communications, reports, letter as journal article, systematic review, meta-analysis", Complément à 100%. N.B.: "historical articles" : pondération = 50%) ;</p> <p>Synthèses (reviews) ("Review - mini review, guidelines - recommandations - protocols (clinical trials)", Complément à 50%) ;</p> <p>Études de cas ("Case reports - journal article, case reports - review, case reports - letter", complément à 50%) ;</p> <p>Editorial (Editorials, perspectives, commentary as an editorial, news and views (<i>Mature</i>)), Complément à 50%) ;</p> <p>Letres ("Letter, comment, comment-news, news, news and features, research highlights, correspondance". Complément à 20%) ;</p> <p>Comptes-rendus ("lecture-book review", Complément à 20%)</p> <p>Source : <i>Faculté de biologie et de médecine de l'UNIL [2020d]</i></p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p> <p>La liste des publications entrant dans le champ de l'évaluation n'est pas rigide. Le CV du chercheur est lu dans son intégralité.</p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p> <p>Le dossier du chercheur est lu dans son intégralité. Les monographies, chapitres de livres, actes de conférences, articles, articles peer-reviewed, ouvrages collectifs édités ont été nommés. A l'exception, la liste des publications est transmise par le chercheur. Dès la fin de l'année [2020], seules les publications déposées dans <i>SERVAL</i> intégreront l'évaluation.</p> <p>Procédure de nomination : "les commissions lisent les publications de certains candidats".</p> <p>Source : <i>Questionnaire envoyé à la Faculté des lettres de l'UNIL</i></p>	<p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p> <p>11 : "Tul papers dans les conférences internationales validés par le Comité Ra&D" lesquels remplissent les critères cumulatifs (affiliation, répond à un appel à publication scientifique international, peer-reviewed, dispose de proceedings ou d'un full paper, fait au moins 4p. (sans la p. de garde, l'abstract et la bibliographie) et est en lien avec le domaine).</p> <p>12 : Articles dans les revues scientifiques validés par le Comité Ra&D du domaine. Incluent les livres scientifiques et chapitres de livres résultant des travaux de Ra&D et les thèses de doctorat (12). Doivent contenir l'affiliation, être peer-reviewed. Les livres et chapitres doivent être considérés "scientifiques" et résulter de travaux de Ra&D.</p> <p>N.B. : pour les livres composés de chapitres répondant individuellement aux critères 12 : chaque chapitre compte comme 12. L'introduction et la conclusion sont comptabilisées en 14.</p>
<p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p> <p>Articles peer-reviewed (0.2 points par publication) ;</p> <p>Livres (0.2 points par publication) ;</p> <p>Articles non peer-reviewed (0.1 points par publication) ;</p> <p>Chapitres de livres (0.1 points par publication).</p> <p>Source : <i>Conseil de domaine Santé de la HES-SO (2018)</i></p>	<p>Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche</p> <p>"Toutes les publications des 3 années précédentes du groupe sont considérées du moment où : - 1 membre du groupe figure dans la publications ; ET - la publication mentionne l'UNIGE ou les HUG</p> <p>Nouveaux professeurs : toutes les publications des 3 années précédentes</p> <p>Source : <i>Faculté de médecine de l'UNIGE [s.d.]</i></p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p> <p>Les chercheurs fournissent la liste de leurs publications aux commissions (nomination ou renouvellement).</p> <p>Tous les types de publication, papier comme numérique, sont pris en considération.</p> <p>Les oeuvres de "vulgarisation" sont aussi prises en considération.</p> <p>Les traductions entrent aussi dans le champ de l'évaluation (traductions des publications du chercheur et traductions effectuées par le chercheur)</p> <p>Procédure de nomination : les chercheurs relenus pour le 2e tour présentent une monographie puis trois articles de leur choix. Ils sont lus par les membres de la commission.</p> <p>Source : <i>Entretien avec la Faculté des lettres de l'UNIGE</i></p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p> <p>14 : Apparitions média et présentations publique aux niveaux régional et national portant sur les résultats de recherches effectuées remplissant les conditions cumulatives (affiliation à la HES-SO ; - être répertoriés dans l'Argus ou diffusés par l'ATS ou publiés dans un média de la presse écrite ou diffusés sur un site Internet de la presse [...] ou rédigés et publiés par un journaliste ou diffusés sur un site Internet professionnel avec rédacteur ou comité de lecture ; traitent d'un projet concret ou de résultats directs de Ra&D en lien avec le domaine.)</p> <p>N.B. : pour les communications traduites dans plusieurs langues et les différentes versions linguistiques d'un même média : seule la communication dans la langue originale compte.</p> <p>Source : <i>Conseil de domaine Économie et Services de la HES-SO (2015)</i></p>	

Pour de nouveaux instruments d'évaluation des publications scientifiques : état des lieux théorique et scénarios applicatifs

Publications entrant dans le champ de l'évaluation (vue détaillée pour le Domaine Économie et Services de la HES-SO)

<p>Domaine Santé HES-SO</p> <p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>	<p>Faculté de biologie et de médecine UNIL</p> <p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté de médecine UNIGE</p> <p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté des lettres UNIL</p> <p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté des lettres UNIGE</p> <p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Domaine Économie et Services HES-SO</p> <p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>
<p>Articles peer-reviewed (0.2 points par publication) ;</p> <p>Livres (0.2 points par publication) ;</p> <p>Articles non peer-reviewed (0.1 points par publication) ;</p> <p>Chapitres de livres (0.1 points par publication).</p> <p><i>Source : Conseil de domaine Santé de la HES-SO (2018)</i></p>	<p>Articles ("Journal articles, brief/short communications, reports, letter as journal article, systematic review, meta-analysis". Comptent à 100%. N.B. : "historical articles" : pondération = 50%) ;</p> <p>Synthèses (reviews) ("Review - mini review, guidelines - recommandations - protocols (clinical trials)". Comptent à 50%) ;</p> <p>Études de cas ("Case reports - journal article, case reports - review, case reports - letter" comptent à 50%) ;</p> <p>Editorial (Editorials, perspectives, commentary as an editorial, news and views (<i>Nature</i>). Comptent à 50%) ;</p> <p>Lettres ("Letter, comment, comment-news, news, news and features, research highlights, correspondence". Comptent à 20%) ;</p> <p>Comptes-rendus ("lecture-book review". Comptent à 20%)</p> <p><i>Source : Faculté de biologie et de médecine de l'UNIL [2020d]</i></p>	<p>La liste des publications entrant dans le champ de l'évaluation n'est pas rigide. Le CV du chercheur est lu dans son intégralité.</p> <p><i>Source : Entretien avec la Faculté de médecine de l'UNIGE</i></p>	<p>Le dossier du chercheur est lu dans son intégralité. Les monographies, chapitres de livres, actes de conférences, articles, articles peer-reviewed, ouvrages collectifs édités ont été nommés. Actuellement, la liste des publications est transmise par le chercheur. Dès la fin de l'année [2020], seules les publications déposées dans <i>SERVAL</i> intégreront l'évaluation.</p> <p>Procédure de nomination : "les commissions lisent les publications de certains candidats".</p> <p><i>Source : Questionnaire envoyé à la Faculté des lettres de l'UNIL</i></p>	<p>Les chercheurs fournissent la liste de leurs publications aux commissions (nomination ou renouvellement).</p> <p>Tous les types de publication, papier comme numérique, sont pris en considération.</p> <p>Les oeuvres de "vulgarisation" sont aussi prises en considération.</p> <p>Les traductions entrent aussi dans le champ de l'évaluation (traductions des publications du chercheur et traductions effectuées par le chercheur)</p> <p>Procédure de nomination : les chercheurs retenus pour le 2e tour présentent une monographie puis trois articles de leur choix. Ils sont lus par les membres de la commission.</p> <p><i>Source : Entretien avec la Faculté des lettres de l'UNIGE</i></p>	<p>I1 : "full papers dans les conférences internationales validées par le Comité Ra&D" lesquels :</p> <ul style="list-style-type: none"> - Contiennent l'affiliation HES-SO (à une exception près) ; - "Répondent à un call scientifique lancé au niveau international" ; - En plus du référencement dans ArODES, I1 doit disposer : de proceedings avec ISSN/ISBN - ou de proceedings disponibles en ligne - ou d'un full paper disponible en ligne sur le site de la conférence - ou d'un full paper à disposition du Comité Ra&D ; - 4 pages A4 au minimum, hors page de garde, abstract et bibliographie ; - En lien avec les champs de recherche du domaine. <p>N.B. :</p> <ul style="list-style-type: none"> - 1 même publication présentée dans plusieurs conférences n'est comptée qu'une fois, à moins de comporter de nouveaux éléments. - la seule sélection des publications par l'organisateur n'est pas considéré comme peer-review. <p>I2 : Articles dans les revues scientifiques validés par le Comité Ra&D du domaine. Incluent les livres scientifiques et chapitres de livres résultant des travaux de Ra&D et les thèses de doctorat (I2).</p> <p>Les articles : doivent, outre le référencement dans ArODES :</p> <ul style="list-style-type: none"> - contenir l'affiliation HES-SO ; - être retenus après peer-review. <p>Les livres, chapitres de livres et thèses doivent, outre le référencement dans ArODES :</p> <ul style="list-style-type: none"> - être une thèse de doctorat ou - être un livre ou chapitre de livre scientifique résultant de travaux de Ra&D. <p>N.B. : pour les livres composés de chapitres répondant individuellement aux critères I2 : chaque chapitre compte comme I2. L'introduction et la conclusion sont comptabilisées en I4.</p> <p>I4 : Apparitions média et présentations publiques aux niveaux régional et national portant sur les résultats de recherches effectuées lesquels, outre avoir été transmises à l'Infothèque :</p> <ul style="list-style-type: none"> - Contiennent l'affiliation à la HES-SO ; - Sont répertoriés dans l'Argus - ou diffusés par l'ATS - ou publiés dans un média de la presse écrite - ou diffusés sur un site Internet de la presse sous forme d'article, de séquence vidéo/radiophonique - ou rédigé et publié par un journaliste - ou diffusé sur un site Internet professionnel avec rédacteur ou comité de lecture ; - traite d'un projet concret ou de résultats directs de Ra&D en lien avec le domaine. <p>Sont aussi considérés comme I4 :</p> <ul style="list-style-type: none"> - Articles d'Hémisphères, les shorts papers (- de 4 p.), conférences nationales répondant aux mêmes critères que les conférences internationales ayant un comité de lecture national, workshops, keynotes avec un résumé faisant partie des proceedings, numéros de revues dont le chercheur est le rédacteur en chef, comptes-rendus medias de colloques organisés par la haute école, rapports avec numéro ISBN, résumés peer-reviewed, articles publiés dans des revues professionnelles, introduction d'actes de conférences, présentations lors de conférences disponibles sous forme vidéo et diffusées sur le site Internet/une chaîne de vidéo de la conférence d'une université du moment que le full paper n'est pas disponible. <p>N.B. : pour les communications traduites dans plusieurs langues et les différentes versions linguistiques d'un même média : seule la communication dans la langue originale compte.</p> <p><i>Source : Conseil de domaine Économie et Services de la HES-SO (2015)</i></p>
<p>Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche</p> <p>*Toutes les publications des 3 années précédentes du groupe sont considérées du moment ou :</p> <ul style="list-style-type: none"> - 1 membre du groupe figure dans la publications ; ET - la publication mentionne l'UNIGE ou les HUG <p>Nouveaux professeurs : toutes les publications des 3 années précédentes</p> <p><i>Source : Faculté de médecine de l'UNIGE [s.d.]</i></p>					

Publications exclues du champ de l'évaluation					
Domaine Santé HES-SO	Faculté de biologie et de médecine UNIL	Faculté de médecine UNIGE	Faculté des lettres UNIL	Faculté des lettres UNIGE	Domaine Économie et Services HES-SO
Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)
"Présentations dans des congrès publiées sous forme de résumés"; "Travaux de bachelior/master/doctorat" <i>Source : Conseil de domaine Santé de la HES-SO (2018), art. 2.2.6 in fine</i>	"Les publications de l'année courante, "Epub ahead of print" ou de type "author reply", ainsi que les abstracts de meetings, les livres et/ou chapitre de livres, ne sont pas pris en considération". <i>Source : Faculté de biologie et de médecine de l'UNIL (2020a), 3e paragraphe de la lettre C.</i>	Aucune information allant dans ce sens n'est en notre possession	Aucune information allant dans ce sens n'est en notre possession	Aucune information allant dans ce sens n'est en notre possession	Des listes non exhaustives renseignent quant aux publications et communications exclues de 12 (p. ex. les nouvelles éditions, remaniées, d'un livre; les supports et manuels de cours, quand bien même ils sont en lien avec la Ra&D; ; l'édition d'un livre (N.B. l'introduction et la conclusion pourraient être complabilisés sous 14...) et de 14 (p. ex. dispositives et posters, publications sur les réseaux sociaux et sites Internet de la HES-SO, critiques et comptes-rendus de livres...) <i>Source : Conseil de domaine Économie et Services de la HES-SO (2015)</i>
		Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche Aucune information allant dans ce sens n'est en notre possession			

Période de référence

Domaine Santé HES-SO	Faculté de biologie et de médecine UNIL	Faculté de médecine UNIGE	Faculté des lettres UNIL	Faculté des lettres UNIGE	Domaine Économie et Services HES-SO
<p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>
<p>"[Les] publications acceptées qui sont parues au cours de l'année précédente".</p> <p>Source : <i>Conseil de domaine Santé de la HES-SO 2018, art. 2.1.1</i></p>	<p>L'année civile en cours n'est pas considérée, que les mesures soient effectuées pour la carrière du chercheur ou pour les 6 dernières années</p> <p>Source : <i>Faculté de biologie et de médecine de l'UNIL (2020b)</i></p>	<p>La carrière du chercheur</p> <p>Source : <i>Entretien avec la Faculté de médecine de l'UNIGE</i></p>	<p>Non précisé dans les informations en notre possession</p>	<p>Non précisé dans les informations en notre possession</p>	<p>Pour l'année de financement n : "Du 01/10/n-2 au 30/09/n-1"</p> <p>Source : <i>Rectorat de la HES-SO (2017b), art. 15 al. 3 let b</i></p> <p>Prise en compte : - Date de la conférence pour les "full papers dans des conférences internationales" - Date de publication pour les articles, chapitres de livres, livres et thèses</p> <p>Source : <i>Conseil de domaine Économie et Services de la HES-SO (2015), [pp. 2-3]</i></p>
		<p>Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche Les 3 années précédentes</p> <p>Source : <i>Faculté de médecine de l'UNIGE [s.d.]</i></p>			

Indicateurs utilisés (vue détaillée)

<p>Domaine Santé HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>	<p>Faculté de biologie et de médecine UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté de médecine UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté des lettres UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Faculté des lettres UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)</p>	<p>Domaine Économie et Services HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)</p>
<p>Nombre de publications</p> <p><i>Source : Conseil de domaine Santé de la HES-SO (2018)</i></p>	<p>Nombre de citations : "pour les articles originaux, les "reviews, les autres types de publications (lettres, commentaires etc.) et pour la totalité des publications". Calculés : Pour la carrière du chercheur et pour les 6 dernières années. Puis pour la totalité des publications et celles où le chercheur est 1er ou dernier auteur.</p> <p>h-index : i. de l'ensemble de la carrière du chercheur ; ii. pour les 6 dernières années.</p> <p><i>Source : Faculté de biologie et de médecine de l'UNIL [2020b] et [2020c]</i></p>	<p>Nombre de publications</p> <p>Nombre de citations</p> <p>h-index (WoS) : transmis par le chercheur</p> <p>Ces 3 indicateurs sont ensuite comparés avec les données de la base de données des Prof. de la Faculté de médecine.</p> <p><i>Source : Entretien avec la Faculté de médecine de l'UNIGE</i></p>	<p>Nombre de publications. À la question "la valeur attribuée varie-t-elle en fonction du type de publication" la réponse est qu'il n'y a pas de critères formels, la totalité du dossier du chercheur est prise en compte.</p> <p><i>Source : Questionnaire envoyé à la Faculté de lettres de l'UNIL</i></p>	<p>Le nombre et la variété des publications</p> <p>Recours non systématique au nombre de citations</p> <p>FI des revues et prestige des éditeurs</p> <p>La qualité des publications prime. Elle est évaluée au regard des habitudes de la discipline concernée.</p> <p><i>Source : Entretien avec la Faculté des lettres de l'UNIGE</i></p>	<p>Le nombre de publications répondant aux critères</p> <p>Le nombre "apparitions média et présentations publiques" répondant aux critères</p> <p><i>Source : Conseil de domaine Économie et Services de la HES-SO (2015)</i></p>
<p>Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche</p> <p>N.B. le système d'évaluation des groupes de recherche clinique et de recherche fondamentale <i>MMOSA</i> est en cours de réévaluation. Les indicateurs actuellement utilisés ne le seront forcément par la suite !</p> <p>FI total du groupe de recherche soit la somme de ses FI corrigés (en fonction du type de publication et de la place des chercheurs dans la liste des auteurs).</p> <p>FI moyen du groupe de recherche : FI total/nombre total des publications.</p> <p>RPU total du groupe (recherche clinique) corrigé</p> <p>RPU moyen du groupe de recherche</p> <p>La totalité des "salaires" perçus par le groupe concernant certaines sources de financement (DIP)</p> <p>Productivité des groupes de recherche : calculée sur la base des FI ou RPU au regard de la totalité des "salaires" du groupe</p> <p>Corrections du FI total du groupe en fonction du type de publication ("originaux : 100 ; revues : 50 [à l'exception de certaines revues] ; lettres : 20 (si plus de 2 pages : 100) ; édits : 30 ; articles courts, clinical trials simples : 20)</p> <p>Corrections en fonction de la place du chercheur dans la liste des auteurs : 1er, 2e, avant-dernière et dernière place = 100%. Sinon au prorata du nombre total des auteurs.</p> <p><i>Source : Faculté de médecine de l'UNIGE [s.d.]</i></p>					

Autres - remarques					
Domaine Santé HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)	Faculté de biologie et de médecine UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté de médecine UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté des lettres UNIL Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Faculté des lettres UNIGE Contexte : Nomination et promotion académique Niveau individuel (chercheur)	Domaine Économie et Services HES-SO Contexte : Octroi de fonds pour la recherche Niveau individuel (chercheur)
	<p>Les données bibliométriques des chercheurs dont les publications sont évaluées sont comparées avec celles des autres chercheurs "fondamentalistes ou cliniciens" de la FBM. La comparaison est illustrée sous forme de Box Plots.</p> <p><i>Source : Faculté de biologie et de médecine de l'UNIL [2020q]</i></p>				<p>Une même publication ne peut être comptabilisée sous 2 indicateurs différents. Le responsable Ra&D du domaine propose le 1 retenu.</p> <p>11 : la date de la conférence fait office de date de référence pour la publication.</p> <p>12 : en cas de publication online ou de pré-print et version papier, seule la version papier compte. La date de la version papier fait office de date de référence.</p> <p><i>Source : Conseil de domaine Économie et Services de la HES-SO (2015)</i></p>
		<p>Contexte : Octroi de fonds pour la recherche Niveau groupes de recherche</p> <p>L'évaluation <i>MIMOSA</i> des groupes de recherche fondamentale et clinique s'opère une fois par an.</p> <p><i>Source : Entretien avec la Faculté de médecine de l'UNIGE</i></p>			

Annexe 34 : SEP 2015-2021 – Catégories et 4 degrés d'évaluation (KNAW, NWO et VSNU 2014)

Category	Meaning	Research quality	Relevance to society	Viability
1	World leading/ excellent	The research unit has been shown to be one of the few most influential research groups in the world in its particular field.	The research unit makes an outstanding contribution to society.	The research unit is excellently equipped for the future.
2	Very good	The research unit conducts very good, internationally recognised research.	The research unit makes a very good contribution to society.	The research unit is very well equipped for the future.
3	Good	The research unit conducts good research.	The research unit makes a good contribution to society.	The research unit makes responsible strategic decisions and is therefore well equipped for the future.
4	Unsatisfactory	The research unit does not achieve satisfactory results in its field.	The research unit does not make a satisfactory contribution to society.	The research unit is not adequately equipped for the future.

Source : KNAW, NWO et VSNU 2014, p. 8

Annexe 35 : SEP 2015-2021 – *Output* indicators (KNAW, NWO et VSNU 2014)

		Quality Domains	
		Research quality	Relevance to society
Assessment Dimensions	Demonstrable products	<p>1. Research products for peers</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Research articles (refereed vs. non-refereed) - Scientific/scholarly books - Other research outputs (instruments, infrastructure, datasets, software tools or designs that the unit has developed) - Dissertations - ... - ... 	<p>4. Research products for societal target groups</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Reports (for example for policymaking) - Articles in professional journals for non-academic readers - Other outputs (instruments, infrastructure, datasets, software tools or designs that the unit has developed) for societal target groups - Outreach activities, for example lectures for general audiences and exhibitions - ... - ...
	Demonstrable use of products	<p>2. Use of research products by peers</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Citations - Use of datasets, software tools, etc. by peers - Use of research facilities by peers - Reviews in scientific/scholarly journals - ... - ... 	<p>5. Use of research products by societal groups</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Patents/licences - Use of research facilities by societal parties - Projects in cooperation with societal parties - Contract research - ... - ...
	Demonstrable marks of recognition	<p>3. Marks of recognition from peers</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Science awards/scholarly prizes - Research grants awarded to individuals - Invited lectures - Membership of scientific committees, editorial boards, etc. - ... - ... 	<p>6. Marks of recognition by societal groups</p> <p>Examples of indicators:</p> <ul style="list-style-type: none"> - Public prizes - Valorisation funding - Number of appointments/positions paid for by societal parties - Membership of civil society advisory bodies - ... - ...

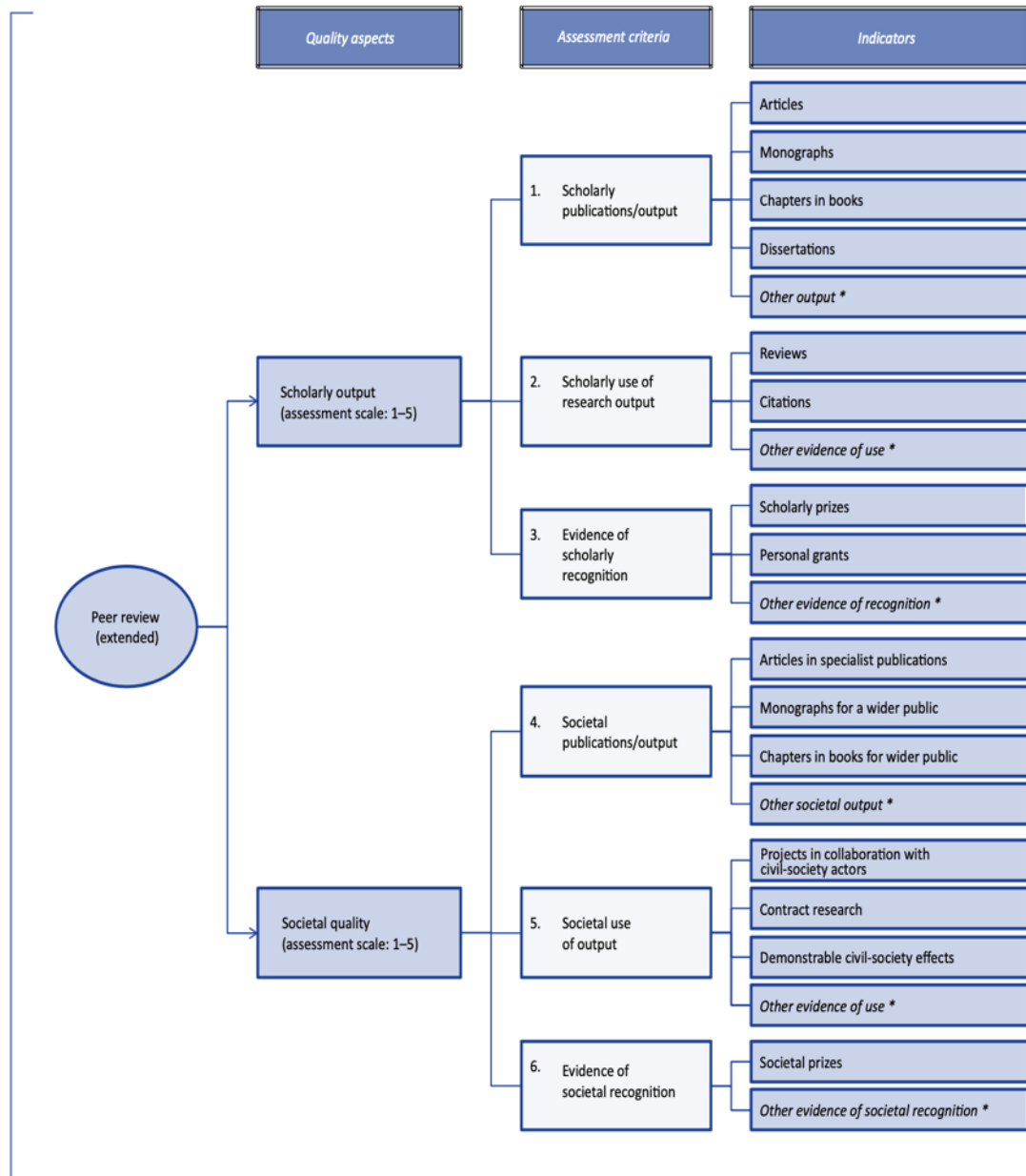
Source : KNAW, NWO et VSNU 2014, p. 25

Annexe 36 : 6 critères d'évaluation de la qualité (KNAW 2013)

		QUALITY DOMAINS	
		SCIENTIFIC QUALITY: Conducting scientific research: carrying out research within the context of the research organisation and ensuring that peers consider this research of outstanding quality	SOCIAL RELEVANCE: Communicating and/or collaborating with external target groups about one's own peer-reviewed research
ASSESSMENT DIMENSIONS	DEMONSTRABLE OUTPUT	(1) Demonstrable output that peers regard as of outstanding quality	(4) Demonstrable output for external target groups
	DEMONSTRABLE UTILISATION	(2) Demonstrable utilisation by peers of researcher's output	(5) Demonstrable utilisation by external target groups
	DEMONSTRABLE RECOGNITION	(3) Demonstrable recognition by peers for researcher's output	(6) Demonstrable recognition by external target groups

Source : KNAW 2013, p. 21

Annexe 37 : Critères et indicateurs de la qualité dans le domaine des humanités (KNAW 2011)



N.B. L'astérisque signale les indicateurs spécifiques aux disciplines ou contextes.

Source : KNAW 2011, p. 47

Annexe 38 : Grille d'évaluation CERES (2020)

Category	A		B		C		D		E	
single or multiple authors	S	M	S	M	S	M	S	M	S	M
Journal articles	5	4	4	3	3	2	2	1	1	0.7
Edited book volumes	7	5	6	4	4	2	2	1	1.5	1
Books >300 pp.	18	13	15	11	12	9	8	6	4	3
Books 100-300 pp.	15	10	12	8	9	6	6	4	3	2
Books <100 pp.	12	7	9	5	6	3	4	2	2	1
Book chapters	4	3	3	2	2	1	1	0.7	0.7	0.5
Successfully supervised PhD theses, number of supervisors	1	2	3 lead	3 other and 4						
P	4	2	2	1						
Reports	>100 pp.		25-100 pp.		<25 pp.					
(Single or Multiple authors)	S	M	S		M	S		M		
R	3	2	2		1	0.5		0.3		
O (other activities)										

Other categories of activities for which 'CERES credits' can be 'earned': for book reviews in A, B or C journals: 0.3 credits; for other book reviews (in D journals): 0.1 credits; for scientific film and video productions (with professional distributors): maximum 5 credits; for working papers (often published by own organisations): 0.3 credits; acquired NWO, KNAW or EU funds for a research programme: 1.5 credits; acquired NWO, KNAW or EU funds for a research project: 0.5 credits; Membership of KNAW, NWO organisation (e.g. WOTRO or MAGW Board or programme committee): 1 credit; CERES or graduate school PhD course co-ordination: 0.3 credits; contribution to CERES or graduate school PhD training: 0.1 credits per 10 hours; CERES summer school organisation: 0.5 credits; member of an editorial board of an ISI-rated journal or of a book series published by an A- or B-rated publisher: 0.5 credits; same other journal or book series (C, D, E): 0.3 credits; CERES organisation: Director and Adjunct-Director (0.4 fte) 6 credits; Board Chair 0.5 credits; member of CERES Board 0.2 credits

Source : CERES 2020. <https://ceres.sites.uu.nl/about-the-valuation-system/> [consulté le 12 février 2020]

Annexe 39 : Exemples d'indicateurs pour le modèle PQRST de Ioannidis et Khoury (2014)

Item in PQRST Index	Operationalization	
	Example	Data Source
P (productivity)	Number of publications in the top tier % of citations for the scientific field and year	ISI Essential Science Indicators (automated)
	Proportion of funded proposals that have resulted in ≥ 1 published reports of the main results	Funding agency records and automated recording of acknowledged grants (eg, PubMed)
	Proportion of registered protocols that have been published 2 y after the completion of the studies	Study registries such as ClinicalTrials.gov for trials
Q (quality of scientific work)	Proportion of publications that fulfill ≥ 1 quality standards	Need to select standards (different per field/design) and may then automate to some extent; may limit to top-cited articles, if cumbersome
R (reproducibility of scientific work)	Proportion of publications that are reproducible	No wide-coverage automated database currently, but may be easy to build, especially if limited to the top-cited pivotal papers in each field
S (sharing of data and other resources)	Proportion of publications that share their data, materials, and/or protocols (whichever items are relevant)	No wide-coverage automated database currently, but may be easy to build, eg, embed in PubMed at the time of creation of PubMed record and update if more is shared later
T (translational influence of research)	Proportion of publications that have resulted in successful accomplishment of a distal translational milestone, eg, getting promising results in human trials for intervention tested in animals or cell cultures, or licensing of intervention for clinical trials	No wide-coverage automated database currently, would need to be curated by appraiser (eg, funding agency) and may need to be limited to top-cited papers, if cumbersome

Source : Ioannidis et Khoury 2014. <https://jamanetwork.com/journals/jama/article-abstract/1881107> [accès par abonnement] [consulté le 21 juillet 2020]