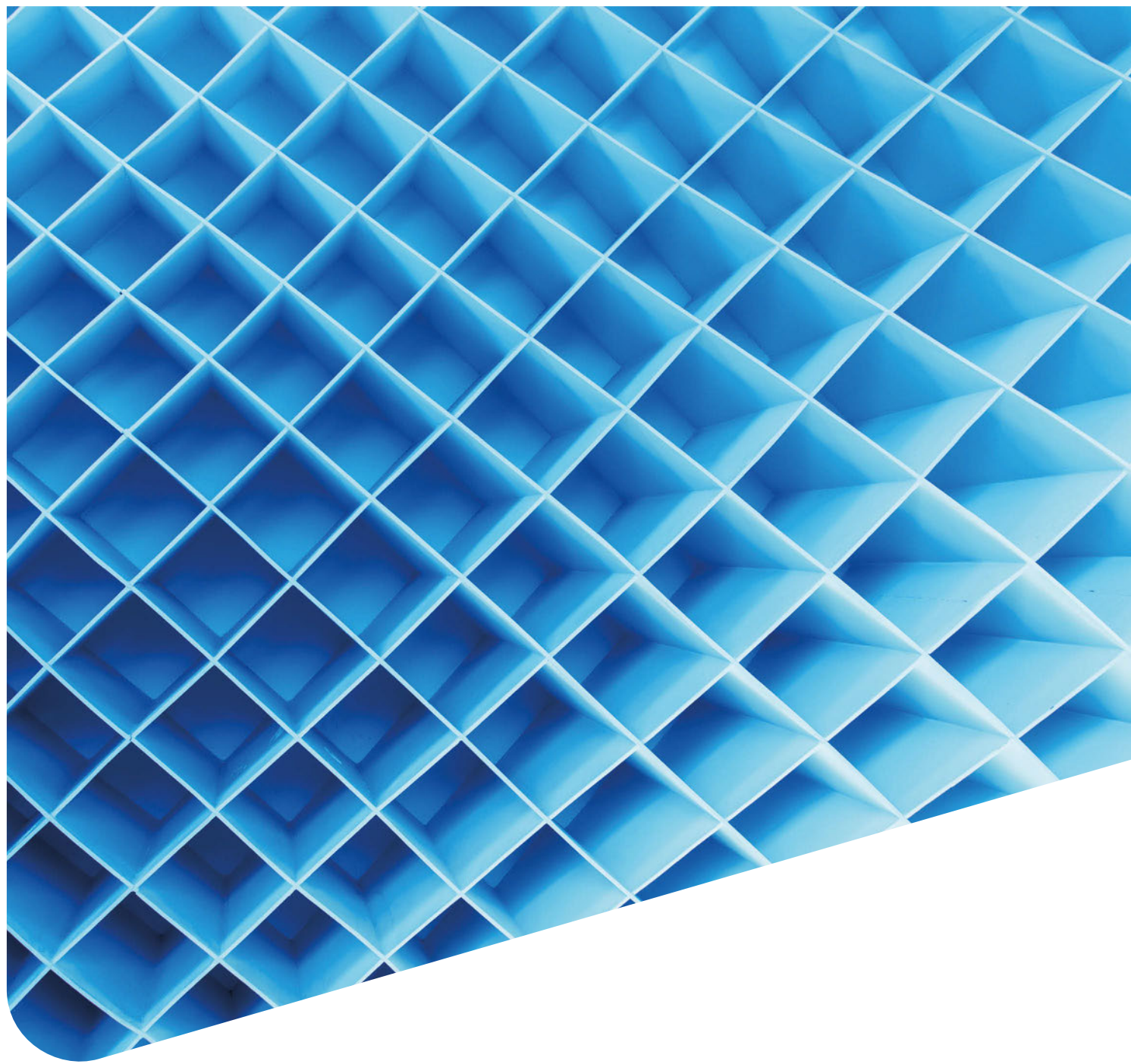




書籍の引用研究における 発展と未来

イーブック白書 Vol.10

ADVANCING
DISCOVERY

まえがき

科学計量学的分析は今まで、学術研究の主な成果としてジャーナル論文を重視してきました。一方、書籍は、人文・社会科学分野をはじめとする多くの分野で重要な役割を果たしてきたにもかかわらず、科学計量学的研究においてはジャーナル論文ほど注目されていませんでした。学術コミュニティは、これまで半世紀以上に渡り、出版・引用データを中心とする科学計量学的データを使い、ジャーナル論文がいかに利用され、後の研究に影響を与えるのかを研究してきました。対照的に、書籍に関してはデータの入手が難しく、書籍の利用状況の分析・評価を進める上で大きな障壁となっていました。最近になってようやく、科学計量学的データソースが充実し、書籍に関する大規模かつ系統的なデータが入手できるようになりました。

シュプリングer・ネイチャーは、多くの出版社から出版されたあらゆる分野の書籍についてScopus（スコーパス）のデータを用い、書籍に関する出版・引用統計を分析しました。その統計は、人文・社会科学において書籍が重要な役割を果たしていることを裏付けると同時に、また、自然科学の様々な分野においても書籍が重要であることを明らかにするものでした。自然科学分野においては、人文・社会科学に比べて書籍の点数ははるかに少ないものの、書籍の方がジャーナル論文より被引用数が高くなる傾向が見られます。

今後は、書籍の科学計量学的データも、少なくともジャーナル論文データと同じくらい幅広く利用できるようになることが望まれます。そのために、書籍やブック・チャプターに不変の識別コードを体系的に割り当てたり、書籍の様々なメタデータを可能な限り提供したりすることで、出版社は重要な役割を果たすことができるでしょう。そうすれば、学術コミュニティは適切な方法で書籍の価値を検証し、正當に評価できると考えられます。また、著者、編者、出版社も様々な形態の書籍（印刷版と電子版、オープンアクセスと購読など）を最大限活用することができるようになります。

本白書は、学術的知識の普及において書籍が果たす役割をより包括的かつ体系的に理解する一助となるものです。書籍の科学計量学的分析への関心を高める上で、本白書がお役に立てば幸いです。今回の分析は引用データに基づいていますが、今後の分析では、ダウンロード数、ソーシャルメディアでの活動、ピアレビュー（査読）に関するデータなど、より広範なデータソースを活用したいと考えています。



Ludo Waltman

ライデン大学

定量科学研究教授

科学技術研究センター副センター長

要旨

シュプリング・ネイチャーは、書籍の著者に対する調査に加え、書籍の引用分析も行い、さらに流通媒体としての学術書の利用傾向を分析しました。被引用数は、書籍の出版物としての成功、質、分野ごと、および分野を横断的に見たときの影響力（インパクト）を示す主要指標として重視されています。

分析の結果、引用のピークは研究分野によってばらつきがあり、生命科学など変化の速い分野では、引用半減期に達する期間が人文・社会科学に比べて短いことが明らかになりました。また、引用の傾向を出版社別に調べると、多くの場合、それぞれの出版社が強い分野における引用パターンに近いことが示されました。さらに、フォーマットについては、テーマ別のブックシリーズとして出版されると、多くの場合、モノグラフに比べ被引用率が高くなります。

書籍の引用分析は、比較的新しい研究分野です。書籍に関する最初のサイテーションインデックスがリリースされてからまだ10年しか経っていません。今後は各種のサイテーションインデックスを活用し、研究のライフサイクルに書籍がどのように利用されているのか、徐々にわかってくることで新たな可能性が拓かれていくでしょう。

■ オリジナル研究

はじめに



書籍は、長きに亘って学術コミュニケーションのチャンネルとして重要な役割を果たしてきました。しかしながら、2020年の現在でも書籍のサイテーションインデックスはその登場からわずか10年しか経っていません (Torres-Salinas, et al, 2012)。また、様々な指標がある中で、特に引用は、過去数百年にわたり研究者や研究機関の間の科学コミュニケーションの重要な指標であり続けてきました。価値の度合いを示すとともに、ホットな分野に焦点を当てたり、時間や空間を越えて研究成果を結び付けたりする役割を果たしてきたのです。実際に、正当に評価されるべき研究に対し、正式にその功績を認めるという意味において、引用は、現代の研究・出版文化における要とみなすことができます。

モノグラフや他の種類の書籍は、大半の研究分野において科学的知識を深める上で、間違いなく重要な役割を担ってきましたが、書籍やその被引用に関する科学計量学的分析は、あまり研究が進んでいない分野です (Zhu, et al., 2020)。とりわけ、ジャーナルの引用に関する研究の多さと比べると、その差は歴然としています。

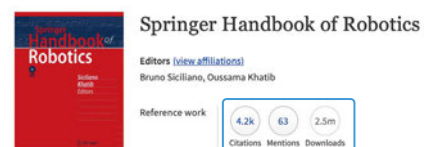
10年前のClarivate (現社名) によるBook Citation Index (BKCI) のリリース以降、書籍の引用実績が分析できるツールは増え続けています。例えば、ClarivateやElsevierが提供するツールやDigital ScienceのDimensions、最新のものとしてはInitiative for Open Citations (I4OC) が挙げられます。これらのデータベースは、書籍引用データのインデックス化および分析の手法はもとより、書籍の収録範囲や分野の分類方法に大きなばらつきがあるとはいえ、これまではほとんど結び付けられなかった書籍と引用とを組み合わせることで、興味深く新しい分析を行うことが可能になりました。

書籍の引用状況を評価すれば、研究のインパクトや質に加え、各種出版プログラムの実績に対し知見が得られます。そこで、2019年にシュプリング・ネイチャーは、書籍の著者に対し大規模な調査を行いました。その結果、調査に応じた著者の3分の2が、自著の成功度を測定するために被引用数を売り上げ実績など他の指標と併せて利用していました。また、利用可能な指標の中でも、引用は書籍の実績を測るための最重要かつ最も頻繁に参考にする指標であるとされていました。

以上を踏まえ、シュプリング・ネイチャーはScopusの書籍引用情報について調査を行いました。本白書は、その研究結果をまとめたものです。主な成果は以下の通りです。

- 学術書は、研究の伝達や進展に大いに役立つツールです。その中で、被引用数は書籍の成功、質、さらには、分野ごと、および分野を横断的に見たときのインパクトを示す主要指標として研究者、出版社、図書館が利用しています。
- テーマ別のブックシリーズとして出版された書籍は、多くの場合モノグラフと比較して被引用率が高くなります。
- 引用された書籍の割合が全体的に高ければ、分野を問わず、そのフォーマットが科学コミュニケーションにおいて重要であることが明らかになりました。
- 引用のピークに至る期間は、分野によってばらつきがあります。生命科学など変化が激しい分野は、人文・社会科学分野に比べると引用半減期に達する期間が短いことが明確に示されました。これは、人文・社会科学以外の分野における書籍の意義をはっきりと表わしています。
- 書籍の引用分析は、比較的新しい研究分野です。書籍に関する各種のサイテーションインデックスを活用し、研究のライフサイクルを延ばすためどのように書籍が利用されているかがわかり、そこには新たな機会が得られることでしょう。

SpringerLinkでは各書籍のページに被引用数、ダウンロード数、SNS等での言及数が表示され常時更新されている



link.springer.com

メソッドおよびデータ



今回の引用研究に際してはScopusの書籍データを使用しました。Scopusは、他の計量書誌学的データベースよりも、少し広い範囲の書籍を対象としています。また、言語フィルターや「source type」と呼ばれるフィルターがあり、データベース内の全ての書籍（ブックシリーズ、それ以外の書籍も含む）について、インデックス化された全ての項目をフィルタリングできます。こうした点が、類似のデータベースがいくつかある中から、Scopusを選択した主な理由です。さらに、本研究には幅広い研究分野の書籍の著者に対する調査結果も取り入れています。本セクションでは今回の研究で用いたデータおよびメソッドの概要を記しています。より詳しい内容は付属資料Iを参照してください。

データ

データは、2019年10月にScopusからダウンロードしました。出版年は2015年から現在までを網羅し、あらゆる種類・分野の書籍の引用情報が得られるよう、十分な期間を確保しました。引用期間は、2015年の書籍の出版日から2019年10月のデータ検索日までの期間としています。

対象は英語の書籍のみとしました。ローカル言語の出版物のダイナミクスの違いによって、ひずみが生じるのを避けるためです。また、「source type」の属性を「books」「book series」でフィルタリングしました。その結果、データ検索時において、合計221,688件のレコードが得られました。

これらの書籍のうち「volume（巻）」フィールドにデータがある書籍は、ブックシリーズとみなしました。また、Scopusのソースファイルを用いて、出版社のインプリントおよび所有者のデータを追加しました。分野に関する情報は、Scopusの「All Science Journal Classification Codes」(ASJC) に従い、Scopusのソースファイルを用いて、エクスポートしたデータセットに追加しました。

分析方法

書籍全体の引用実績を完全に網羅し分析するため、書籍の引用およびブック・チャプターの引用両方を用いました。この方法は、科学計量学の代表的な文献（Torres-Salinas, et al., 2014など）に倣っています。書籍およびチャプターレベルの引用データは、抽出したフィールド「source title」を用いてひとつのブック・エントリーにまとめました。このフィールドは、書籍レベルのエントリーかチャプターレベルのエントリーか見分けがつくようになっています。結果として、2015年に英語で出版され、Scopusでインデックス化された書籍、18,791点のデータセットが得られました。引用データは合計で531,557件です。このデータセットに基づいて、以下の分析を行いました。

調査データおよび調査方法

シュプリングer・ネイチャーは、2019年に、世界中の書籍の著者に調査を行いました。調査期間は2019年7月1日から2019年7月21日までです。調査のためのリンクは、Eメール、TOCアラート、ソーシャルメディアを通じて配信しました。初期調査の回答者11,688人のうち、過去5年間にシュプリングer・ネイチャーに限らず、書籍を出版した著者8,557人に対し、続く本調査への回答を依頼しました。調査に最後まで回答した著者は合計で4,375人でした。



レコード数

221,688



書籍数

18,791



引用データ数

531,557

研究結果



前述のとおり、学術書の引用分析を実施した結果、あらゆる学術研究分野や書籍フォーマット、さらに出版社においても研究や出版内容に傾向があることがわかりました。また、モノグラフ（書き下ろし）シリーズ、レビューシリーズ、会議録シリーズなど、テーマ別のブックシリーズとして出版される書籍の重要性が明らかになりました。今回の定量的研究は、人文・社会科学の多くの分野においてモノグラフに優位性があることが確認されました。このような違いがあるとはいえ、幅広いトピックに及ぶ書籍出版プログラムを有する出版社が、発行する書籍の点数が最も多いだけでなく引用数も最も大きくなっています。

今回の研究結果は、科学計量学および関連分野で出版された文献に見られる傾向に概ね似通っています。研究から、書籍の引用はジャーナル論文に比べて、ピーク（書籍の被引用数が最大となる年）に達するまでにより長い時間を要することが明らかになっています。ジャーナルの場合は、わずか数か月のうちに引用が急増するケースもあります。ある研究によると、書籍の大半は引用のピークに達するまでに8年を要していました（Zhu, et al., 2020）。これらの指標を筆者らが作成したデータセットに適用したところ、様々な研究分野、出版フォーマット、出版社において、書籍の引用傾向に関し注目すべき結果が得られました。

研究分野

Scopusにインデックス化されているタイトルの半数以上が、人文・社会科学（以下HSS）分野の書籍に分類されています。これは、HSS分野では研究上のコミュニケーションを書籍というフォーマットに大きく依存しているという歴史的証拠にも合致しています。

今回の研究から、HSS分野の書籍は、1書籍当たりの平均被引用数が科学・技術・医学（以下STM）分野に比べて少ないことが明らかになりました。データの構造的な理由である可能性が非常に高いと考えられます。今回の研究では、2015年に出版された書籍について、2015年の出版日から2019年10月までの全ての引用を収集しました。この比較的短い期間のデータに基づいた場合、変化が激しく引用半減期が短いSTM分野の方が、引用半減期が長いHSS分野よりも見た目には有利で、引用回数が多くなります。したがって、HSSとSTMとを比較するとひずみが生じてしまう可能性があり、HSSとSTMを比較するよりも、HSS、STMそれぞれで分野間の違いを分析の方がより有用な洞察が得られる可能性が高いと考えられます。

健康科学、物理科学、生命科学の幅広い専門分野を網羅するSTMのタイトルは、それぞれ大きく異なる傾向を示しています。コンピュータサイエンス分野においては、書籍というフォーマットの人気が高いことがみてとれます。同分野は、学術書の発行点数で上位5以内にランクされ、その多くがトピック別シリーズです。広く引用されている会議録シリーズの典型例として、シュプリングer・ネイチャーが1973年から出版している「Lecture Notes in Computer Science」¹が挙げられます。年間600タイトル以上が出版される同シリーズは、コンピュータサイエンスのあらゆる分野を対象に、研究開発の最新情報を読者に提供し続けており、同分野における学術コミュニケーションの重要な手段として高い評価を受けていることが分かっています。

■ 全ての研究分野および書籍における
出版・研究手法の傾向

Scopus

¹ Lecture Notes in Computer Scienceについては、下記のURLを参照。

<https://link.springer.com/bookseries/558>

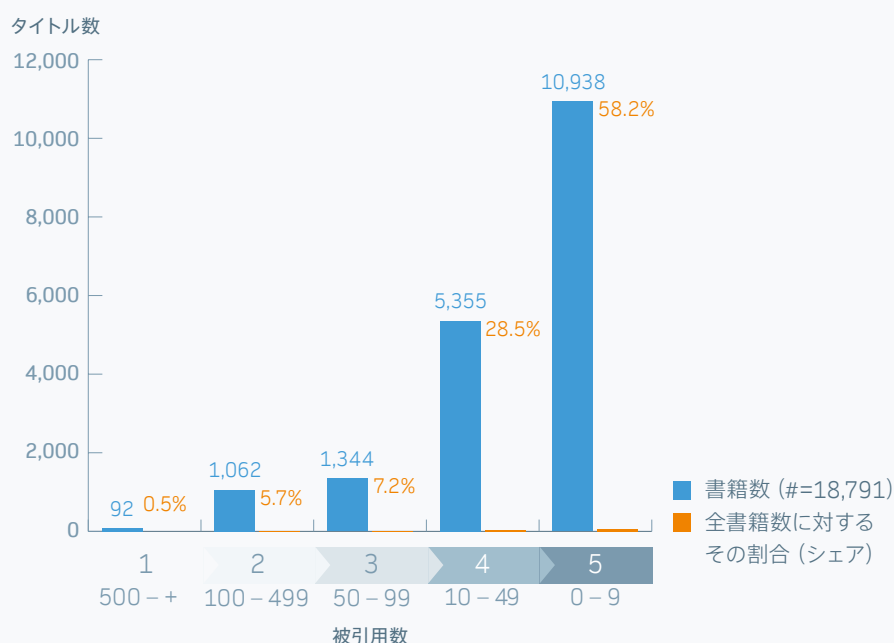
<https://www.springer.com/jp/computer-science/lncs>

注目すべき傾向として、コンピュータサイエンス分野の書籍は、出版社全体で被引用数の合計が最大であり、1書籍当たりの平均被引用数でも地球科学と並んで最も高くなっています。一方、コンピュータサイエンスの対極に位置しているのが化学工学と環境科学です。これらの分野は、他の分野に比べ書籍の出版点数の合計が最も少なく、被引用数の合計も最低となっています。しかしながら、両分野ともに、1書籍当たりの被引用数は必ずしも最低ランクではありません。

#	分野	Scopusに 収録されて いる書籍数 (2015年)	書籍数 合計に 占める 割合 (%)	チャプター 数の合計	チャプター 数合計に 占める 割合 (%)	1書籍 当たりの 平均 チャプター数	被引用数 の合計	被引用数 合計に 占める 割合 (%)	1書籍 当たりの 平均 被引用数
1	Arts and Humanities	5,068	27.9%	25,842	13.5%	5	50,015	9.9%	10
2	Social Sciences	3,493	19.3%	26,874	14.1%	8	55,040	10.9%	16
3	Business, Manage- ment and Accounting	1,634	9.0%	12,816	6.7%	8	26,554	5.3%	16
4	Computer Science	1,599	8.8%	36,695	19.2%	23	117,606	23.4%	74
5	Medicine	1,354	7.5%	25,142	13.1%	19	30,575	6.1%	23
6	Engineering	1,243	6.9%	17,569	9.2%	14	39,518	7.9%	32
7	Biochemistry, Genetics and Molecular Biology	867	4.8%	12,610	6.6%	15	58,032	11.5%	67
8	Agricultural and Biological Sciences	842	4.6%	11,877	6.2%	14	42,338	8.4%	50
9	Psychology	416	2.3%	4,265	2.2%	10	20,754	4.1%	50
10	Mathematics	357	2.0%	2,017	1.1%	6	6,877	1.4%	19
11	Earth and Planetary Sciences	317	1.7%	4,421	2.3%	14	22,984	4.6%	73
12	Chemistry	305	1.7%	2,996	1.6%	10	12,989	2.6%	43
13	Physics and Astron.	230	1.3%	2,647	1.4%	12	8,057	1.6%	35
14	Environmental Sc.	207	1.1%	2,253	1.2%	11	4,181	0.8%	20
15	Chemical Engin.	204	1.1%	3,217	1.7%	16	7,744	1.5%	38
TOTAL		18,136	100%	191,241	100%	11	503,264	100%	28

本研究で対象とした書籍全体における被引用数の分布を見ると、全ての分野において引用される頻度の高い書籍の割合が相当数あることが分ります。これは全ての分野の科学的議論において、書籍の重要性とその価値が高いことを示しています。とりわけ、本データセットは他の研究 (Zhu, et al., 2020 など) に比べ、引用される頻度の低い書籍 (被引用数 0~9) の割合が小さくなっています。この点もまた専門分野に特化した読者を有する高度な研究テーマを扱う書籍の重要性を示しています。ニッチな分野における被引用数の少なさは、研究のインパクトが低いとか意義がないことを示している訳ではなく、むしろ、専門分野の規模が小さいと被引用数も必然的に小さくなるということであり、読者の規模が大きい分野の被引用数とは単純には比較できないことが今回の研究から示唆されます。

書籍の被引用数を5つの範囲に分類 書籍数とそのシェア

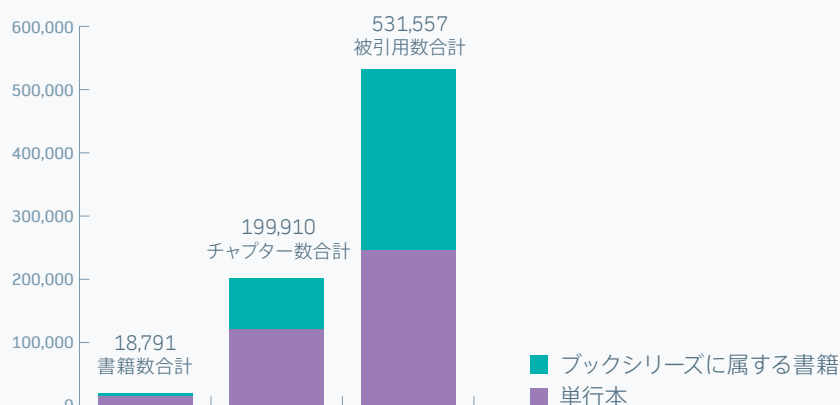


出版フォーマット

今回の研究で用いたデータセットは、77%が単行本に、23%がブックシリーズに分類されており、対象となった書籍のフォーマットやタイプは多岐に亘ります。ブックシリーズの最も一般的なタイプは、会議録、レビュー、そして昔ながらのモノグラフなどです。データに収録された書籍のうち、占める割合が高いのはモノグラフですが、ブックシリーズに属する書籍は引用実績が高いことが注目すべき点です。この結果は、シリーズで編集された書籍は、シリーズに属さない書籍（単行本）に比べ、チャプター数がより多く被引用数も全体的に多いことを明らかにした他の計量書誌学的研究（Torres-Salinas, et al., 2014）とも一致しています。



Scopusに収録されている書籍数（2015年）



単行本として出版された書籍の上位25位までは、生物学からビジネスまで、幅広い分野に分類されます。それとは対照的に、ブックシリーズで出版された書籍のうち、被引用頻度が最も高いのは、主にコンピュータサイエンスであることが注目に値します。これは他の研究でも示された傾向と一致します。コンピュータサイエンスが比較的新しい専門分野であり、研究発表が会議を通じて行われることが多く（Fortnow, 2009）、その会議録がブックシリーズとして出版されることが多いことが理由の1つです。会議録を書籍のカテゴリーに分類するかどうかは出版社によってばらつきがあり、会議録をジャーナルに分類している出版社もあります。

出版社の傾向

学術書を出版している出版社を分析すると、最も高い実績をもつ書籍は、多くの場合複数の分野のプログラムを有する出版社から発行されていました。書籍の被引用数の合計が最も高い上位10社のうち7社が複数の分野を扱っています。3社は、人文科学、社会科学、法律など研究成果や研究手法の共有を書籍に依存するところが大きい分野を中心に出版しています。

しかし、複数の分野を出版する出版社の中にも際立った書籍プログラムがあり、特定の分野に強みがある出版社も存在します。例えば、大学出版局などは様々なテーマの書籍で成功していますが、最も高い実績があるのは特定の分野であることが分かりました。一例として、Cambridge University Pressは芸術や数学の分野でとりわけ優れています。

全体として、大手出版社の引用実績には典型的なポートフォリオの特徴が反映されています。例えば、Taylor & Francisは、1書籍当たりの平均被引用数が低いように思われますが、これは同社の出版分野が人文・社会科学に大きく偏っているためです（人文・社会科学分野は、STM分野の書籍よりも引用半減期が長いことが明らかになっています）。一方、Elsevierは1書籍当たりの平均被引用数が高くなっています。これは、同社のポートフォリオのほとんどが医学や生命科学といったとりわけ変化の激しい分野を中心としているためです。

STM分野のトップ10出版社

#		タイトル数	チャプター数	合計被引用数	1タイトルあたりの平均被引用数
1	Springer Nature	3,786	72,558	198,434	52
2	Elsevier	931	12,184	49,175	53
3	Wiley-Blackwell	634	7,214	19,892	31
4	Nova Science Publishers	583	4,646	1,928	3
5	Taylor & Francis	561	5,039	10,938	19
6	Cambridge University Press	179	3,007	10,180	57
7	IGI Global	116	2,286	2,339	20
8	Wolters Kluwer Health	81	3,039	560	7
9	Scientific.net	81	5,266	6,456	80
10	US Nat. Research Council	68	801	3,864	57

HSS分野のトップ10出版社

#		タイトル数	チャプター数	合計被引用数	1タイトルあたりの平均被引用数
1	Taylor & Francis	2,838	16,986	39,291	14
2	Springer Nature	1,928	17,137	31,642	16
3	Cambridge University Press	752	3,750	15,086	20
4	Peter Lang Publishing	418	2,376	1,223	3
5	Brill	399	2,769	1,720	4
6	Nova Science Publishers	398	2,039	519	1
7	Edward Elgar Publishing	213	2,505	4,046	19
8	Wiley-Blackwell	193	1,683	6,746	35
9	Elsevier	183	6,021	8,277	45
10	Princeton University Press	183	371	3,414	19

上記の表では、書籍のタイトル数の順位と被引用数の合計の順位とが往々にして一致していることが分かります。しかし、タイトル数は比較的少ないものの平均被引用数が高い出版社もあります。この場合、その分野において高い評価を受けている高品質のコンテンツを提供していることを示していると考えられます。

今回のデータセットの中で被引用数が最も多い書籍を個別に検討してみると、複数の分野にまたがる書籍であることが分かりました。そうした強みを持つ書籍の多くは、産業や専門職へ応用できる内容です。最も高い被引用実績を持つ書籍のいくつかは、理論を実践へと応用したハンドブックです。こうしたいわゆる参考書の類は、近年経済的・技術的な面で苦戦しているものの、学生や研究者にとっては今も重要であることが分かります。

今回の、また他の研究からも、書籍はあらゆる領域において研究の伝達や科学の進展に欠かせないツールであることが証明されており、書籍の引用は研究のインパクトを測る重要な指標とみなすことができます。今や、書籍のサイテーションインデックスが利用できるようになって10年近く経ち、今後、科学計量学的評価は研究コミュニティに価値をもたらす研究として成長していくでしょう。

最も引用の多い書籍トップ20

	タイトル	ブックシリーズ か否か	出版社	分野	被引用数 合計
1	Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science, Volume 9351	Yes	Springer Nature	Computer Science	4,443
2	Climate Change 2014 Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects	No	Cambridge University Press	Earth & Plan. Sciences	2,652
3	International Encyclopedia of the Social & Behavioral Sciences: Second Edition	No	Elsevier	Social Sciences	2,586
4	National Vital Statistics Reports, Volume 64	Yes	US Dep. of HHS	Social Sciences	1,798
5	Marine Anthropogenic Litter	No	Springer Nature	Agric. & Bio Sc.	1,559
6	Recommender Systems Handbook, Second Edition	No	Springer Nature	Computer Sc.	1,470
7	Climate Change 2014: Impacts, Adaptation and Vulnerability: Part B: Regional Aspects	No	Cambridge University Press	Earth & Plan. Sciences	1,247
8	Developmental Psychopathology: Volume One, Second Edition	Yes	Wiley-Blackwell	Psychology	1,077
9	Advances in Cryptology -- EUROCRYPT 2015, Lecture Notes in Computer Science, Volume 9056	Yes	Springer Nature	Computer Science	1,052
10	Chaos Modeling and Control Systems Design, Studies in Computational Intelligence, Volume 581	Yes	Springer Nature	Computer Science	1,026
11	MEDINFO 2015: eHealth-enabled Health, Studies in Health Technology and Informatics, Volume 216	Yes	IOS Press	Engineering	1,012
12	Tools and Algorithms for the Construction and Analysis of Systems, Lecture Notes in Computer Science, Volume 9035	Yes	Springer Nature	Computer Science	997
13	National Health Statistics Reports, Volume 2015	Yes	U.S. NCHS	Medicine	943
14	Computer Vision - ECCV 2014 Workshops, LNCS, Volume 8926	Yes	Springer Nature	Computer Sc.	928
15	Advances in Cryptology - EUROCRYPT 2015, Lecture Notes in Computer Science, Volume 9057	Yes	Springer Nature	Computer Science	922
16	12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering, Computer Aided Chemical Engineering, Volume 37	Yes	Elsevier	Chemical Engineering	910
17	Handbook of Unmanned Aerial Vehicles	No	Springer Nature	Computer Sc.	880
18	Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015, Lecture Notes in Computer Science, Vol.9349	Yes	Springer Nature	Computer Science	873
19	Handbook on Business Process Management 1: Introduction, Methods, and Information Systems	Yes	Springer Nature	Bus., Man. & Accounting	791
20	Asteroids IV, Space Science Series	Yes	Uni. of Arizona P.	Physics & Astr.	790

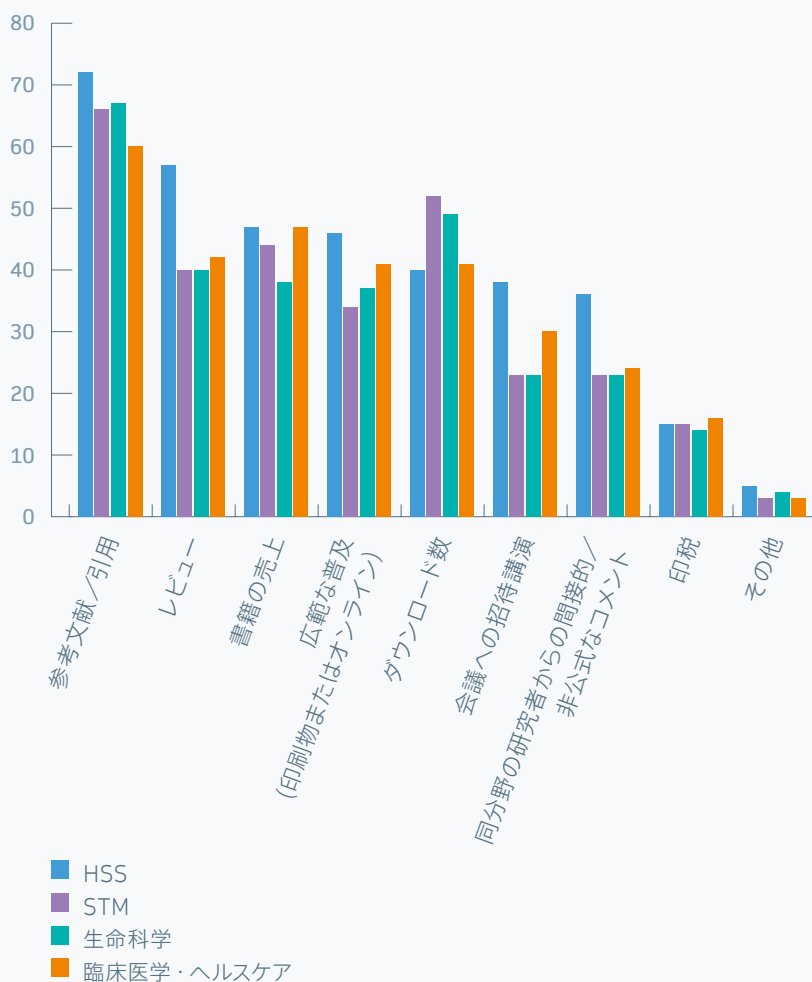
書籍著者の視点

以上の結果について考察する際には、書籍著者の視点を考慮すること、そして研究者が成功を測る指標として書籍の被引用数を活用していることを認識することが重要です。過去5年間に書籍を出版した著者4,375人に対して、調査を実施したところ（出版社は全てを対象とし、シュプリング・ネイチャーに限定しない）、調査回答者の大半は、出版した書籍の成功を測る最も重要な指標は引用であると考えていました。分野による違いを見ると、HSS分野の著者は、引用をとりわけ重要な指標と評価していました。それに対して、STM分野の著者は、指標として引用を重視する割合がHSS分野の著者よりもやや低く、ダウンロード数を重視する割合がわずかに高くなっていました。

どのような指標を元に出版した書籍の成功を測定していますか？

当てはまるものを全て選んでください（ $n=4,375$ ）

HSS分野の著者は、参考文献／引用とレビューを高く評価しています。それに対し、生命科学やSTM分野の著者は、ダウンロード数をやや高く評価しています。



分析結果



大半の研究分野において、引用は主要な実績評価指標としての役割を果たしていますが、それぞれの分野で書籍は異なる形で利用されています。今回の研究結果とこれまでの研究とを勘案すると、学術書の被引用頻度が高くなる要因に同様の傾向が見られ、「時間」が引用指標に影響を与えられそうです。

コンピュータサイエンスなどの研究分野では、学会で発表した研究を会議録に掲載し、研究成果を短期間のうちに発表します。これらの分野は、引用されるまでの時間も同様に短くなります。これに対し、歴史学などの人文科学分野では、書籍が引用されるようになるまで、より長い時間がかかっています (Zuccala, et al., 2015)。引用半減期もまた、出版サイクルに左右されるのではないかと推察できます。書籍の出版は、最初の企画書から完成原稿を経て出版に至るまでに3年を要することがあります。出版サイクルの長さが引用期間に強い影響を及ぼすのは、おそらく、歴史学など出版フォーマットとして書籍が好まれるHSS分野だと考えられます。

その分野の変化の速度もまた書籍の引用期間の長さに影響を与えられそうです。書籍の引用期間の長さや、引用のピークに達する書籍数を評価する場合には、引用半減期という指標が役立ちます。

書籍の被引用数は、良くも悪くも特定の研究分野におけるその書籍の相対的インパクトを明確に表わしています。しかしながら、被引用数は書籍の功績を示すひとつの指標に過ぎません。著者の評判や販売方法などによる影響度は、科学計量学だけで評価はできません。学術書の引用分析によって明らかになる傾向をより深掘りするには、他のデータポイントや視点を加えることが役立つかもしれません。ソーシャルメディアや書評などのオルトメトリクスを使った研究では、書籍の成功度合を測る基準を定め、インパクトをはかる指標と品質をはかる指標の違いを議論する上で、他のデータを追加利用できる可能性があることが実証されています (Leydesdorff, et al., 2016)。

シュプリング・ネイチャーの調査では、様々な指標の中でも、特に被引用数によって自分の書籍の成功を測定していると3分の2の著者が回答²しています。引用動向は研究分野によってばらつきがあり、研究の時間軸、優先事項、コミュニケーションの方法が引用行動に影響を及ぼします。引用以外の指標が、書籍のもたらす影響や書籍に対する読者の関わり合い方について新たな視点をもたらす一方で、引用分析は今後も学術書の価値やインパクトについて重要な知見を提供してくれるでしょう。

書籍の成功を表す指標とは

² Springer Nature book author survey, 2019.

まとめ



書籍の科学計量学はまだ比較的新しい研究分野であり、書籍出版や研究コミュニケーションの多くの側面に関してさらなる研究が必要です。また引用自体についても、例えば、証拠や裏付けのために引用されたのか、あるいは研究に対する反証として引用されたかなど、引用の目的や機能についてもより分析が必要です（例：scite.ai）。オルトメトリクスやマルチプラットフォームを用いた研究が盛んな今日、書籍の科学計量学でもより幅広いデータポイントを活用することが有益であると考えられます。他のインパクトを測る指標を取り入れれば、相対的なランキングや、書籍を引用している出版物に関する詳しい知見を評価できるようになるでしょう。

各社の書籍のサイテーションインデクスが、引用分析に使う書籍をどのように分類しているかも検討する必要があります。というのも、多くの（書籍）収載サービスでは、研究分野の分類は独自の方法を用いているからです。例えば、人文・社会科学の分野がまとめられているため、人文科学の分野間での傾向分析があいまいになり、分野間の違いが見分けられなくなることもあります。また、特定の研究分野が人文科学に分類されているか、はたまた社会科学に分類されているかも、データベースによって異なります。とりわけ、ビッグデータ研究や意思決定科学などの新しい研究分野でこの傾向が顕著です。こうした分野の分類は、研究機関や研究者自身が用いているスキームと一致していないことが多く、このような重要なズレがあると、既存のデータを歪める可能性があり、更なる調査が必要です。

■ 今後求められること

今後、書籍の引用を長期的に分析することは、将来の電子出版イノベーションに役立つと考えられます。例えば、書籍を電子化した場合に、もともと印刷版で出版された同じ書籍よりも、全体的な被引用頻度の割合が高くなるか評価できるかもしれません。また、その傾向から書籍コンテンツの電子化に伴う読者層や販路の拡大に影響を与える可能性もあります。しかし、こうした長期的研究は、数十年にわたって行う必要があるため、今はまだ、将来の研究に期待するという段階です。

書籍は数百年の歴史を有する遺産ですが、科学計量学を用いて、学術書の出版・研究パターンを測るようになったのはわずか10年ほど前に過ぎません。書籍の電子化が進展し、引用分析の技術・科学が進化を続ける中で、シュプリングer・ネイチャーは、今後も新たな知見の提供に取り組んでまいります。

付属資料I

調査上の制約

今回の研究で用いたデータセットには、2015年に出版されたすべての書籍が網羅されていないことに留意する必要があります。Scopusに収録されていない書籍は含まれていません。また、Scopusがブックシリーズを索引付けしており、今回の分析は完全にScopusに依存しています。この点については、Annual Reviewsの例に示される通りあいまいである可能性があります (Torres-Salinas, 2013)。さらに、分野のデータはScopusの分類に基づいており、1タイトルに対して複数の分類が割り当てられている場合、最初にあがっている分野を選択しています。

以上に加え、データ作成中に判明したこととして、かなりの数の書籍が2015年より前に出版された文献に引用されていました。その理由は、過去のタイトルを2015年に電子化した出版社があり、Scopusではこれらの書籍を2015年の出版物として収録し、引用データは、オリジナルの情報を使っているからです³。そのため、2015年より前の出版物の割合が2015年にインデックス化した書籍の5%を上回っている出版社については、過去のタイトルをデータセットから除外することとしました。影響を受けた出版社、ならびに、除外した書籍数は、以下の通りです。Princeton University Press (80%、書籍数：829点)、Oxford University Press (23%、書籍数：30点)、Cambridge University Press (12%、書籍数：122点)、Wiley-Blackwell (8%、書籍数：68点)。

³ 例：Morris Rosenbergの“Society and the Adolescent Self-Image”はもともとPrinceton University Pressより1965年に出版され、電子版は2015年de Gruyterによって出版されました。こうしたタイトルをScopusでは2015年の出版物として収載しています。本書は、Scopusのデータでは2020年1月の時点で14,584回の被引用数がありましたが、そのうちの8,224回(56%)が2015年より前に引用されたものでした。

参考文献

- Fortnow, L. (2009), “Viewpoint: Time for computer science to grow up,” *Communications of the ACM*, Vol. 52 No. 8, pp. 33–35.
- Leydesdorff, L., Bornmann, L., Comins, J. A., & Milojević, S. (2016), “Citations: Indicators of Quality? The Impact Fallacy,” *Frontiers in Research Metrics and Analytics*, 10.3389/frma.2016.00001.
- Torres-Salinas, Daniel; Robinson-García, Nicolás & Delgado López-Cózar, Emilio. “Towards a ‘Book Publishers Citation Reports’. First approach using the ‘Book Citation Index’.” Ec3 Working Papers 7, Sun, 29 Jul 2012.
- Torres-Salinas, D., Robinson-García, N., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2014), “Analyzing the citation characteristics of books: Edited books, book series and publisher types in the book citation index,” *Scientometrics*, Vol. 98 No. 3, pp. 2113–2127.
- Zhu, Y., Yan, E., Peroni, S., & Che, C. (2020), “Nine million book items and eleven million citations: A study of book-based scholarly communication using OpenCitations,” *Scientometrics*, Vol. 122 No. 2, pp. 1097–1112.
- Zuccala, A., Guns, R., Cornacchia, R., & Bod, R. (2015), “Can we rank scholarly book publishers? A bibliometric experiment with the field of history,” *Journal of the Association for Information Science and Technology*, Vol. 66 No. 7, pp. 1333–1347.

参考

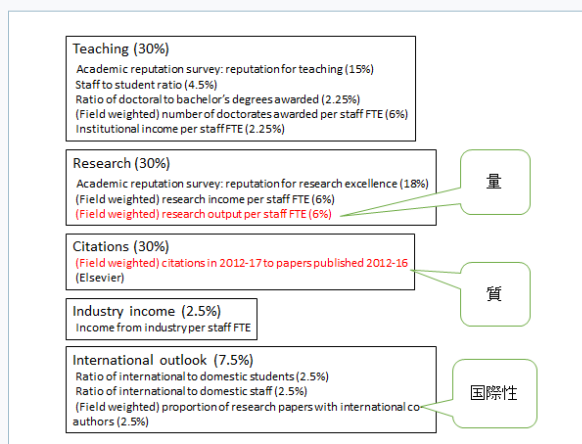
研究力と書籍の引用

ジャーナル論文と同様に、ブックの引用情報は様々なところで活用されています。代表的なものがランキングの指標です。例えば、Times Higher Education (THE) の大学ランキングでは、その大学から出版された文献や出版物が世界の研究者にどのくらい引用されたか、その平均回数を把握することで、研究の影響力を調査しています。研究の影響力を示すこの引用情報は、ランキング指標のうち3割を占めています (図1)⁴。また、この引用情報は、2016年から書籍も含めて数えられるようになりました。分野によって異なる引用の文化や行動があるために生じるばらつきについては補正が施されます (Field Weighted Citation Impact)。これは、伝統的に被引用数の多い分野で研究活動を行っている機関が不当に有利になることはないことを意味しています。

文献の種類による違いについてもこの点は補正がかけられます。例えば、レビュー論文が被引用数を稼ぐには有利、と言われていますが、本ランキングはその違いが補正されます。ブックについても、ブック同士で比較され、本と論文を比較したりはしません。よって、論文と比べて本は引用が少なく、と断じてしまうのは、少なくとも大学ランキングを正しく理解するという文脈においては間違った見方となります。

日本語版
オリジナル

図1 THE 世界大学ランキングの指標



出典：自然科学研究機構小泉周氏

「ポストコロナの図書館運営を考える—大学ランキングと助成金・運営戦略・書籍への応用」発表資料より

大学ランキングに関する解説は、2020年7月に開催されたオンラインイベント「ポストコロナの図書館運営を考える—大学ランキングと助成金・運営戦略・書籍への応用」でも自然科学研究機構小泉周氏によって詳述されている。

https://go.sn.pub/20200701_webinar

とはいえ、電子化が進み、文献データベースに収録されるブックが増えるにつれ、ブックが様々な出版物に引用されることが増えてきたのも事実です。シュブリンガー・ネイチャーのイーブック・コレクションは、その多くのタイトルが論文など、何らかの学術文献に引用されています (図2)。ブックが最先端の研究を下支えしていることを示す興味深いデータです。

前述のとおり、引用のピークに至る期間は分野によってばらつきがあり、生命科学など変化が激しい分野は、人文・社会科学分野に比べると引用半減期に達する期間が短いことが示されています。とはいえ、論文⁵に比べるとSTMの分野であっても引用期間は比較的に長いこと

⁴ <https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2020-methodology>

⁵ <https://www.nber.org/papers/w23447.pdf>

がわかりました。図3は、2012年から2018年にかけて出版されたイーブック・コレクション（STM）に関する、2018年7月時点での引用状況です。新刊であればあるほど、引用できる期間が短いため、一見引用数が減少しているように見えますが、これは逆の見方が必要です。期間が経てばたつほど、引用されており、出版後何年にもわたって引用され続けるということがわかりました。書籍の利用はロングテールと言われていますが、引用面においても、長きに亘って必要とされるコンテンツであることがここでも証明されました。

図2 研究の「質・厚み」を支える書籍

最も引用されたトップ25タイトルがどの種類のコンテンツに引用されているか？ STM／HSS別

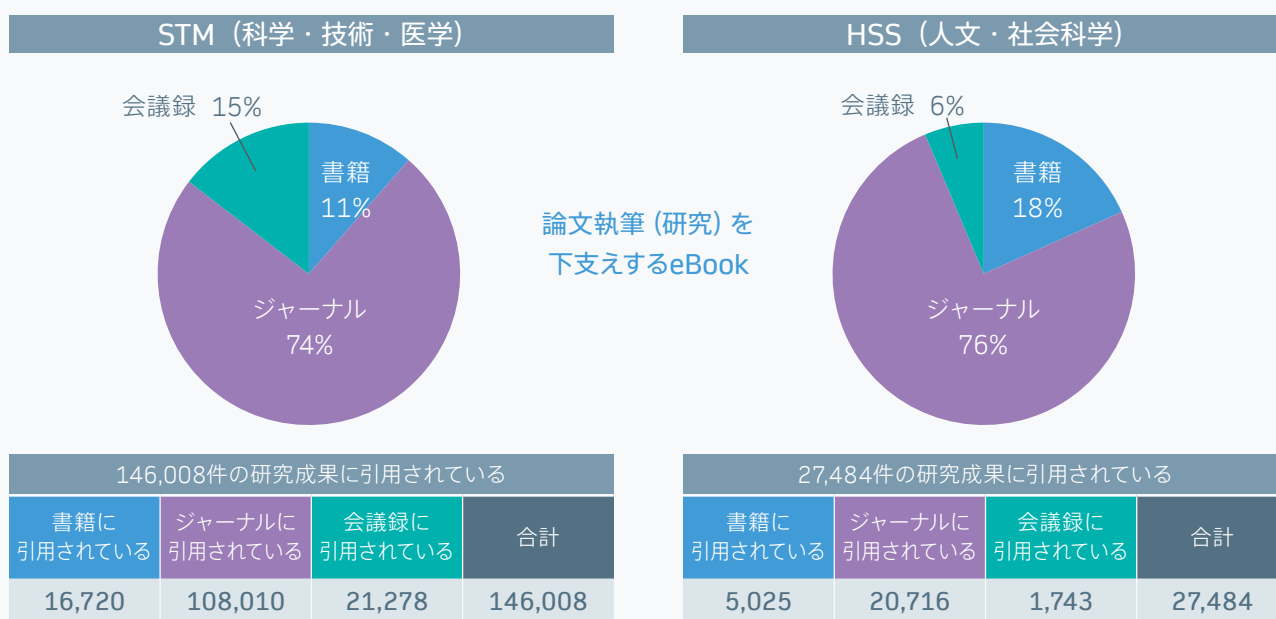
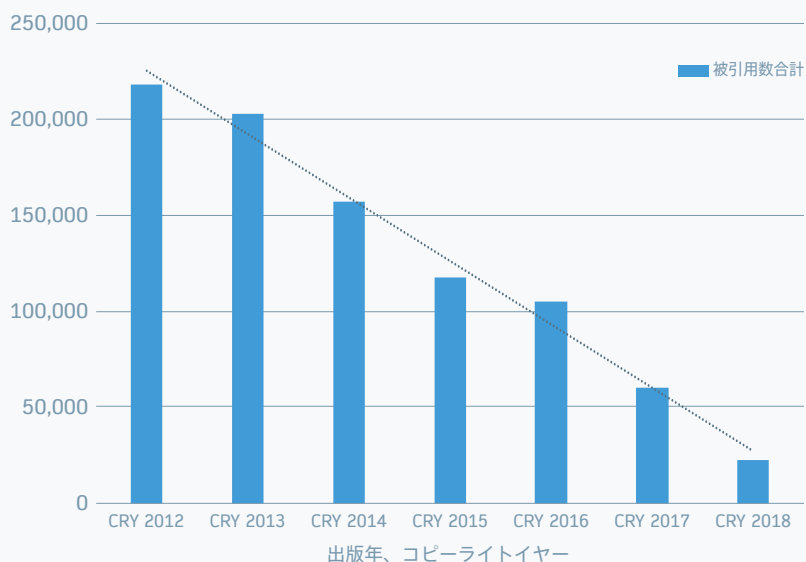


図3 シュプリンガー・ネイチャーのSTM分野の書籍における出版年ごとの被引用数（世界全体、2018年）



* Data: Worldwide citations as of July 2018 for Copyright Year 2012-2018

引用される書籍とは

ブックを長きに亘って利用する読者がいるということは、出版社にとっては質の高い内容を担保していくことが非常に重要です。シュプリンガー・ネイチャーは1840年代から現在までに26万タイトルほどの学術書を出版してきました。イーブックを本格的に出版し始めたのは2005年です。現在まで蓄積された大量のデータを元に、多く利用される本に共通する要素を明らかにする調査プロジェクトを2019年に開始しています。今まで出版された英語のモノグラフを対象とし、20項目に亘って調べたところ、特に以下の5項目に高い相関が見られるという結果になりました。

5大要素

1 書籍の分野	
2 類似する書籍の利用状況	
3 書籍に含まれるキーワード	より一般的なトピックであるほど、幅広い読者に受け入れられやすい
4 チャプター数	該当する分野のうち、類似する書籍のチャプター数の幅におさまっているほど、読まれやすい
5 シリーズ	シリーズに収録されているほど、読まれやすい

1つ目は、書籍の分野です。その書籍の主な分野が、伝統的に書籍が使われてきた分野であるのか、近年勃興してきた新しい分野であるのかに関わらず、書籍として使われやすい環境が整った分野であれば、よく利用される傾向にあります。2つ目に、類似する書籍がどのように利用されているかも利用に影響する結果となりました。3つ目は、キーワードです。より一般的なキーワードであるほど、幅広いトピックとして多くの層に受け入れられやすいことがうかがえます。4つ目はチャプターの数です。その分野において、ちょうどいい分量であることが利用の増加につながります。利用統計を単なるチャプター数の累計で見ていくのではなく、データを正規化した上で分析すると、その書籍に該当する分野の中で、類似する他の書籍のチャプター数の範囲に収まっているほど読まれやすいということがわかりました。5つ目の要素は、ブックシリーズに含まれるかどうかです。シュプリンガー・ネイチャーでも多くの書籍をシリーズとして出版していますが、シリーズに属するものほど読まれやすい傾向にあるという結果もわかりました。

調査はまだ継続中ですが、前述の5大要素を含むブックが（モノグラフの中では）より利用されやすいということがわかりました。こういった分析から、出版社は予め出版前に需要を予想することが可能になるでしょう。現在、シュプリンガー・ネイチャーでは、この多く利用される書籍に共通する5つの要素を軸とし、機械学習を用いながらベンチマークを設定することで、出版を計画している書籍に対し定量的な評価を与える仕組みを実験的に開始しています。このように、今までの査読を中心とした定性的な「お墨付き」に加え、ビッグデータに基づいた機械学習型の予測ツールも使いながら、質を担保していけるのかどうか、書籍に新たな付加価値を与えることができるのか、出版社自身もより定量的な指標で評価される時代になっていると言えます。

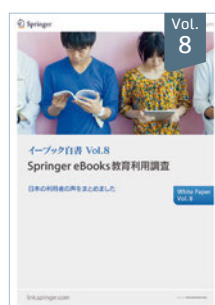
6 エントリー数の多いレファレンス・ブック（辞書等）は含まれておらず、ここでは対象はモノグラフ（単行本）に限定

シュプリングer・ネイチャーのイーブック白書

Vol.1	電子ブックが大学図書館にもたらすもの その価値とコスト
Vol.2	ユーザーの声をまとめました
Vol.3	ビッグディールによるイーブック導入：利用統計調査 リバプール大学のイーブック研究 Part 1
Vol.4	リバプール大学のイーブック利用者アンケート リバプール大学のイーブック研究 Part 2
Vol.5	学術書としてのイーブック 図書館における投資効果 (RoI) を理解する
Vol.6	日本の研究者の声をまとめました
Vol.7	シュプリングer・ブック・アーカイブ 研究における情報格差解消のために
Vol.8	シュプリングerイーブック教育利用調査 日本の利用者の声をまとめました
Vol.9	ブック・アーカイブ導入によるモンタナ州立大学図書館の変化と ブック・アーカイブがもたらす5つの価値

ダウンロードはこちら

https://go.sn.pub/ebook_white_paper_jp



Best in Books

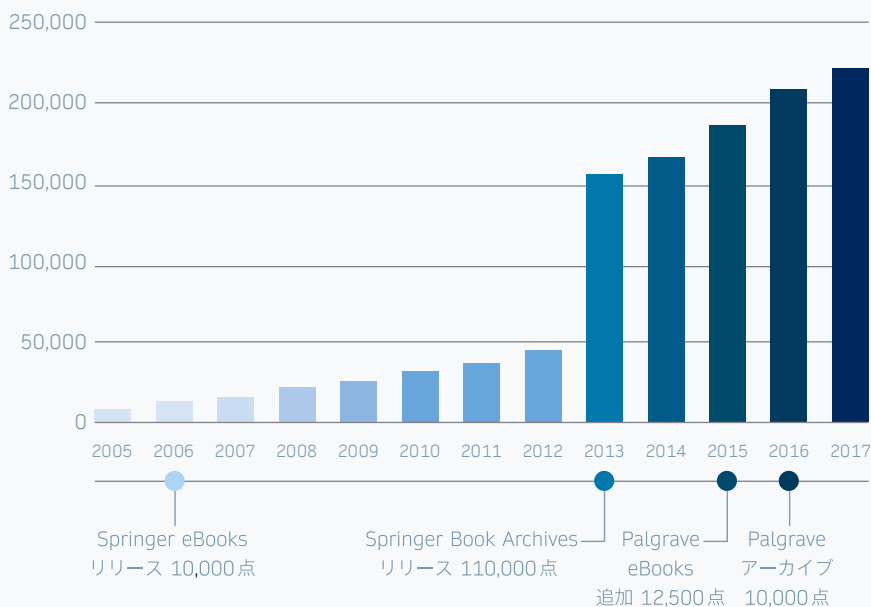
Springer Nature eBooks (シュプリンガー・ネイチャー・イーブックス) は、STM分野 (科学・技術・医学) とHSS分野 (人文科学・社会科学) をカバーし、ピアレビューを経た研究・教育情報リソースへのオンラインアクセスを提供します。

Springer Nature eBooksは、アーカイブ・コレクションを含む25万点以上の多彩なタイトルを収載しています。年間1万2,000点を超える新刊が加わり、コレクションは年々拡充しています。



palgrave
macmillan

Springer Nature eBooksのタイトル数



260,000
タイトル
2020年10月現在