

第3回 SPARC Japan セミナー2016

「科学的知識創成の新たな標準基盤へ向けて：オープンサイエンス再考」

欧州から見たオープンサイエンス

Ron Dekker

(欧州委員会研究・イノベーション総局

(European Commission (DG Research & Innovation)))

講演要旨



オープンサイエンスは政治的な課題として重要になっている。2016年に欧州理事会でオランダが議長国となっている間に、28加盟国はオープンサイエンスに関するEU競争力担当相理事会の結論を採択し、オープンサイエンスに関するアムステルダム行動要請(仮訳、Amsterdam Call for Action on Open Science)を表明したオープンサイエンスの会議があった。欧州委員会はヨーロッパオープンサイエンスアジェンダを策定し、オープンサイエンスクラウドやオルトメトリクスといったテーマでいくつかの専門家グループを設置した。またEU加盟の数カ国では国家的にオープンサイエンスポリシーや戦略を制定している。その他、大学や出版者、資金配分機関等のステークホルダーもまた市民科学を含むオープンサイエンスの活動に関わっている。同時に、例えば学術情報流通の方法といったことへの重要な変更は先行者不利益によって阻害され、大きな財政的な再分配を必要とする。

そしていま、われわれはどのような状況にいるのか。どのようにオープンサイエンスエコシステムの変化を誘導できるか。どのようにして出版への配慮をしつつ、オープンアクセスへの移行を触媒し変化を起こすことができるか。さらに、研究データシェアリングを促進するためには何が必要とされるか。

本講演では、ヨーロッパの状況と上記の課題について述べるとともに、可能な解決策についての論議を呼び起こしたい。また出版や研究データについて主に語る一方で、オープンサイエンスの他の側面やイノベーションを含めた一般的な科学と社会の関係についても触れたい。

Ron Dekker

マーストリヒト大学で経済学を専攻し、労働市場研究のキャリアをスタートした。1995年にティルブルフ大学に入り、研究の軸足をデータマネジメントに移した。これはオランダ科学研究機構(Netherlands Organisation for Scientific Research (NOW))での業務のきっかけとなり、1997年にはデータ部門の責任者となった。その後、社会科学部と中央計画・研究所部のチームコーディネーターとして勤務した。2007年NOW研究所ディレクター。2013年にはオランダ高等教育研究イノベーション機構(仮訳、Dutch IT-innovation organisation for Higher Education & Research (SURF))でディレクターを務め、2014年にオランダのEU議長国としての準備のためにオープンサイエンスのプロジェクトリーダーとして教育・文化・科学省(Ministry of Education, Culture and Science)に出向した。2016年、欧州委員会研究・イノベーション総局オープンサイエンス担当加盟国出向専門家(仮訳、Seconded National Expert on Open Science at the European Commission, Directorate-General Research & Innovation)となった。2017年3月から、Consortium of European Social Science Data Archives (CESSDA)ディレクターとなる予定。CESSDAはヨーロッパ最大のインフラ(いわゆるESFRI(研究インフラ欧州戦略フォーラム)Landmark)のひとつで、本拠地はノルウェーのベルゲンである。



私は欧州委員会に出向している立場なので、欧州委員会を代表して発言するわけではありません。今回は欧州委員会のオープンサイエンスに関する動向を踏まえ、オープンサイエンスとは何であるか、欧州のオープンサイエンスポリシーについて、出版物とデータのオープン化に焦点を絞ってお話しし、それからオープンサイエンスの展望について述べたいと思います。

オープンサイエンスとは何か

オープンサイエンスとは、科学を変化させるものです。アイデア・概念、分析、データ収集、出版、レビューという研究サイクル全体が変わります。この状況は、デジタル化によって生じています。生産されるデータの量は指数関数的に増加しており、透明性が求められるようになってきました。例えば、心理学では公表された結果の約半分に再現性がありません。従って、説明責任を満たした研究を実現する新しい方法を見つけなければなりません。また、科学は、企業を含む社会全般とより良いつながりを持ち、特に貧困と疾病という大きな社会的課題に取り組む必要があります。

オープンサイエンスとは、データ、出版、ソフトウェア、オープンノート（研究公正）、市民が積極的に科学に参加する市民科学から構成された、包括的用語と見ることができます（図1）。オープンサイエンスを一つに定義することは難しいですが、私は「オープンサイエンスとは、あらゆる種類の科学的知識は、発見と同時ぐらいに、早いタイミングで公開されなけれ

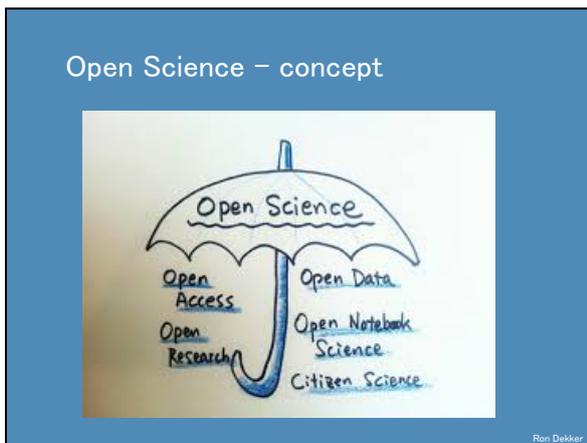
ばならないという考えである」というマイケル・ニールセン（Michael Nielsen）の定義が好きです。直ちに共有しなかったり、壁を維持したりする理由はいろいろあるでしょうが、核心はできるだけ早く公開することです。

現在、既に利用可能なオープンサイエンスのツールがたくさんあります（図2）。この内側のサイクルは、データ収集、分析、出版、レビューといった伝統的な科学です。外側のサイクルはオープンサイエンスで、内側のサイクルに比べてより外界との相互作用があります。オープンサイエンスとは、このように科学をつないでいくものなのです。

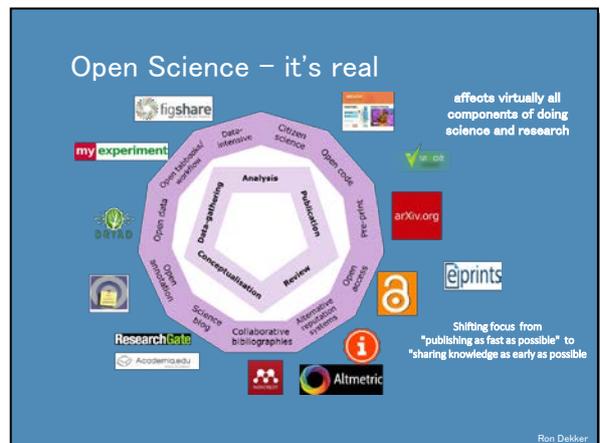
欧州のオープンサイエンス政策

2014年に欧州委員会は、研究者、出版社、資金提供者といったオープンサイエンスのステークホルダーとのコンサルテーションを開始しました。主なトピックは、出版、データおよび研究インフラでした。

図3は、コンサルテーションで出てきたオープンサイエンスの主な促進要因です。同時に促進の障壁となるものも挙げられました。その一つが品質保証です。それは、直ちに公開された場合に、どのように品質を保証するのか。データを共有したいとき、いかにクレジットを提供するのか。データと知識を共有するのに十分なインフラはあるか。研究者とそれ以外の人々は利点を認識しているか、ということでした。そこで私たちは、データと出版物について多くの優先事項を設



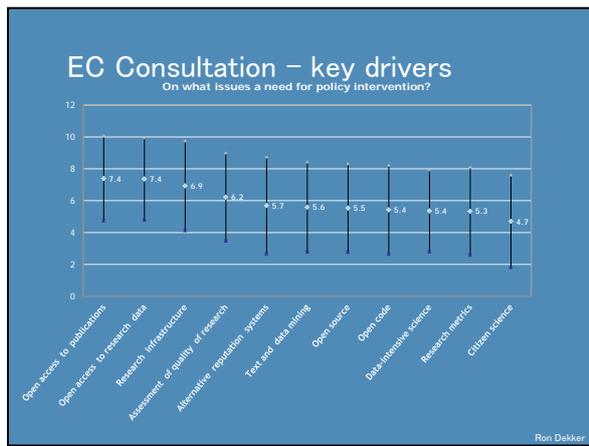
(図1)



(図2)

定しました。これは欧州委員会委員のカルロス・モエダス (Carlos Moedas) だけでなく、オランダ教育・文化・科学省副大臣のサンダー・デッカー (Sander Dekker) にも取り上げられました。

昨年、オランダは欧州理事会の議長を務め、オープンサイエンスを優先させることを決定しました。まず簡単に議長国オランダの成果を紹介し、欧州委員会の優先する政策について話します。私たちは、議長国時代にデータと政策に関する成果物をつくることに決めました。ERAC Task Force on Optimal Reuse of Research Data (研究データの最適な再利用に関するタスクフォース) は、データトレーニング、データ管理、持続可能性、IP 問題などに関する数多くの活動と勧告に取り組みました (図 4)。データの最適な再利用を追跡することはとても複雑です。私たちはここで、あえて「オープンデータ」という言葉は使いませんでした。



(図 3)

ERAC Task Force on Optimal Reuse of Research Data

TRAINING OF STAKEHOLDERS AND AWARENESS RAISING

- Promote a better understanding of open research data
- Establish training and education programs on Open Science
- Establish a reward system for data sharing activities
- Ensure sound monitoring

DATA QUALITY AND MANAGEMENT

- Make data identifiable and citable
- Promote metadata standardisation and production of metadata
- Promote innovative models for peer-review and quality assurance
- Strongly promote the use of data management plans

SUSTAINABILITY AND FUNDING

- Ensure the existence of FAIR open research data infrastructures
- Ensure funding for open research data and for data sharing activities

LEGAL ISSUES

- Make IPR issues insightful

(図 4)

「オープンデータ」という言葉には、データは直ちに公開されるべきだということを示唆する面がありますが、データの中には機密性があるものや、そのデータを最初に使用する権利があるものもあるからです。そうした場合には、自分に再利用の権利があることを明らかにしなければなりません。

これらを実現するため、アムステルダムで会議を開催し、オープンサイエンスに関するアムステルダム行動要請 (仮訳、Amsterdam Call for Action on Open Science) を表明しました (図 5)。そこで決まった目標は、一つ目は、2020 年までに科学出版物の完全なオープンアクセス化を達成することです。二つ目は、データ共有のための新しいアプローチを準備することです。三つ目は、データ共有の際に人を認識し、クレジットを与える新しい報酬と評価のシステムをつくることです。東ヨーロッパの現在のシステムでは、データの共有に対する報酬はありません。権威ある雑誌に出版すれば報酬が与えられますが、教育、データの共有、知識の共有などに対しては報酬がありません。この報酬制度を変えていかなければ、共有するインセンティブがありません。四つ目は、少し秘密めいていますが、つまり知識の共有と監視システムです。率直に言うと、国や大学には出版の費用について何らアイデアがありません。購読の契約費用は知っていますが、例えば、ゴールドオープンアクセスの論文掲載料 (APC) については知らないのです。従って、これらのデータを収集し、出版コストに関する情報を共有する必要があります。

Amsterdam Call for Action

Two important pan-European goals for 2020:

- Full open access for all scientific publications
- A fundamentally new approach towards optimal reuse of research data

Flanking policies

- New assessment, reward and evaluation systems
- Alignment of policies and exchange of best practices

(図 5)

エンクラウドは、データと全ての研究成果に対するインフラとしての役割を果たしています。研究公正に関しては、結果を再現可能なものにすることが重要です。欧州委員会は、これらの項目に関するハイレベル専門家グループを既に設置し、または今後設置する予定です。

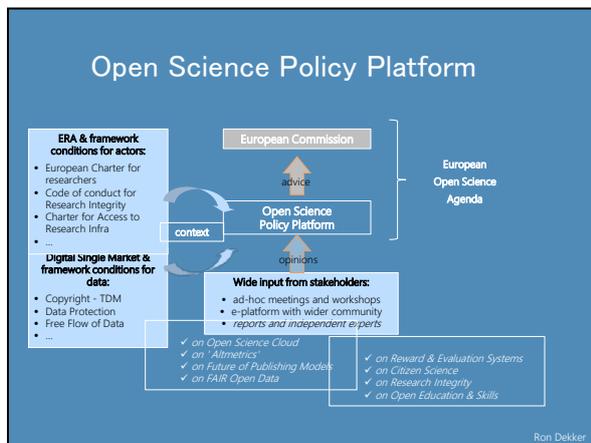
また、欧州委員会は、オープンサイエンスポリシープラットフォームを導入しました。このポリシープラットフォームは、委員会に助言する 25 人のステークホルダーから構成され、ヨーロッパオープンサイエンスアジェンダの八つの項目を、欧州委員会がオープンサイエンスへの移行を促進し、触媒することを推進できるように、包括的に把握することが期待されています。

チャートを使って説明します。まず左は入力側です(図 8)。オープンサイエンスポリシープラットフォームはワーキンググループを持ち、ステークホルダーから意見をもらうことができます。そしてプラットフォームは、最終的に欧州委員会に助言するだけでなく、結果を機関に還元します。ワーキンググループのステークホルダーがプラットフォームの方策に合意すれば、それは全てのステークホルダーによって容易に実施されることになります。

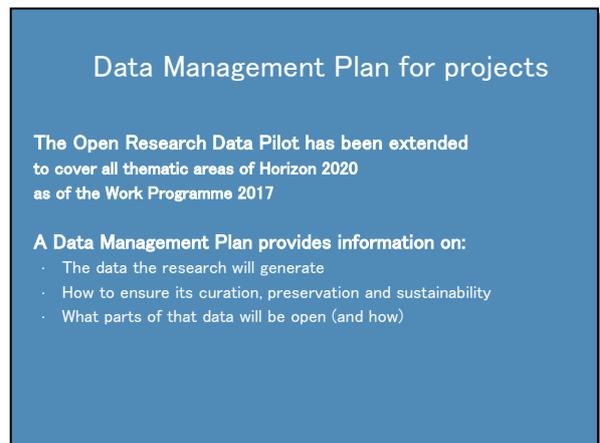
また、欧州委員会は、リサーチプログラムの中でオープンサイエンスに取り組んでおり、プログラムの中でも EU 第 8 次研究・イノベーション枠組み計画「Horizon 2020」が最も重要なものです。その中では、出版物をグリーンまたはゴールドでオープンアクセス

化することを義務付けています。できるだけ早くデポジットし、領域ごとに決められた 6~12 カ月間のうちに利用可能にする必要があります。ゴールドオープンアクセスの費用は、助成金から支払うことができます。同じことがデータにも当てはまります。助成金受領者は、データをリポジトリに保管し、アクセスを提供する必要があります。私たちは FAIR 原則にのっとり、出版物だけでなく、データも共有したいと考えています。

データ管理計画 (DMP) のパイロットも行いました(図 9)。これは若手研究者に、他人に有益なものを生み出すことへの意識づけをするためです。パイロットへの参加は任意であり、本パイロットに関わる人の約 3 分の 2 が参加しています。本パイロットの適用範囲外のプロジェクトから、さらに 10% が自発的に参加しました。このパイロットについては、Horizon 2020 の全てのプログラムをカバーするように拡張することにしました。研究によって生成されるデータ、どのようにしてデータを提供するかといった手順は、むしろ簡単でシンプルであり、研究者はデータキュレーションとセットアップに関わる情報を求められます。これによって、研究者の、他の人にとって価値のあるものをつくりだしているという意識を高め、データを研究者のラップトップから引き出し、機関リポジトリや国家・各専門分野のデータベースに取り込むのです。この DMP は非常に重要で、データを共有するための必要条件です。



(図 8)



(図 9)

データと出版物のオープンアクセスに関する欧州委員会の政策活動をご紹介します。私たちは政治声明、ポリシー、ハイレベル専門家グループ、ステークホルダーから成るプラットフォームを有しており、これらを補助金交付規程に反映しています。

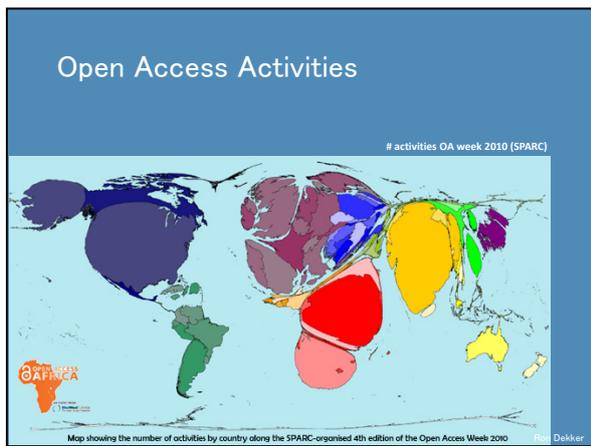
現在、ほとんどの国がオープンアクセスに関する政策を導入しています。実際に G7 では昨年、声明を発表しました。それは同時に G20 の目的でもあります。オープンサイエンスは、今度欧州で開催される G7 会議の議題になる予定で、報酬制度とクラウドインフラに焦点が当てられます。この議題は、日本と欧州委員会が共同で作成したものです。

学術出版物のオープンアクセス

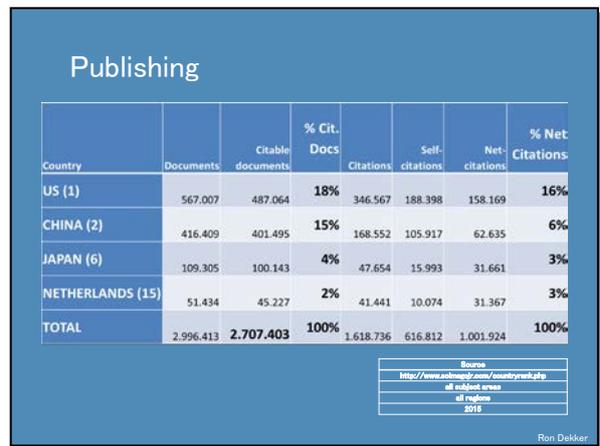
出版物とデータについてに移ります。誰がオープンアクセス政策に関与しているかを考えるため、まず私

は SPARC による OA Week 2010 の図 10 を使って、オープンアクセス政策が活発な国を紹介します。オープンアクセスは多くの国々で議題になっています。しかし、2050 年に予測される人口動向を見ると、アジア諸国への大きな移動があります (図 11)。人々が住んでいるところは、すなわち、科学、イノベーション、経済成長の可能性があるところです。アジアで、科学を社会にもっとうまくつなげることができれば、科学の枠を超えて大きな可能性があることは明らかです。

出版に話を戻すと、毎年約 300 万件の出版物があります (図 12・13)。90%が引用可能で、その約半数が引用されており、自己引用を差し引くと、270 万件の論文につき約 100 万件が引用されていることになります。つまり、3 件のうち 1 件のみ引用されているので、インパクトファクターは 0.3 です。ところが、論文によってはインパクトファクターが 30 も 40 も付与



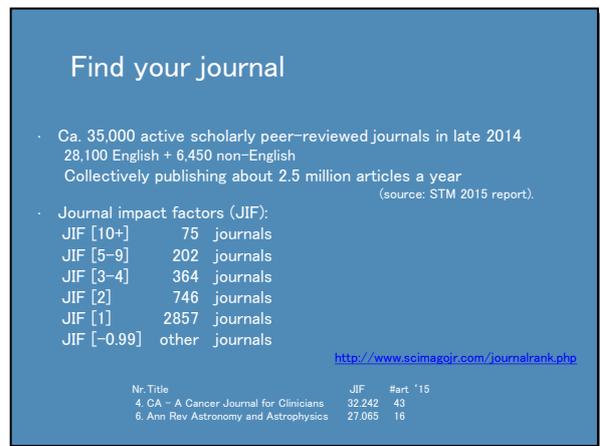
(図 10)



(図 12)



(図 11)



(図 13)

されます。これはつまり、多くの人々が引用を全くされていないということを意味します。これらの論文は誰のために出版されているのでしょうか。誰とこの知識を共有しているのでしょうか。データを見ると、引用率がオランダと日本で同じぐらいであることがわかります。

3万5千種のジャーナルがあるので、ジャーナルを見つけることは難しいでしょう。さらに、これら3万5千種のうち1%だけが5以上のインパクトファクターを持ち、3~4のインパクトファクターを持つものも1%です。オランダでは、研究者は2以上のインパクトファクターを持つジャーナルに掲載することが推奨されていますが、これは全研究ジャーナルのわずか4%にすぎません。

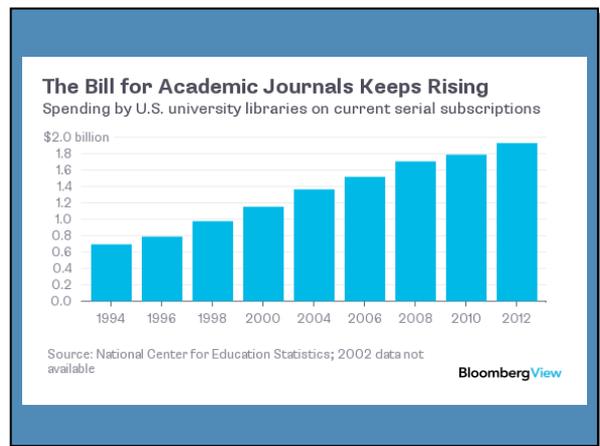
その関連性を述べると、インパクトの高いジャーナルに掲載したい場合はレビューを書かなければなりません。これが最も多く引用されるからです。しかし、天文学と物理学を例に取ってみると、2015年にアクセプトされた論文は16件しかないで、その16件の中の一つでなければならないということです。この中に入れば論文は引用されるのですが、これが科学でしょうか。私たちは論文をこのようなジャーナルの一つに掲載するために、公的資金を費やしているのでしょうか。私には、これらのジャーナルに掲載されること自体が目標になってしまっているように思えてなりません。

出版社は、IPアドレスの助けを借りてマーケットを細分化し、インターネットを最大限に活用しています。多くのジャーナルを購読できることが魅力のパッケージ購読がありますが、これにより、一つの購読をやめることが、全てのジャーナルの購読をやめることになるため、難しくなっています。ジャーナルのインパクトファクターはまた、外に出ることが非常に難しい経済的ロックインをもたらします。ですから、他のステークホルダーらはジャーナルのインパクトファクターを品質の代わりとして使用するのです。これは、リサーチカウンシルで助成金を決定する際に関係しま

す。大学のランキングを決めるときにも関係します。ここには先行者不利益があります。そのシステムから出ることを決めたら、もはや大学のランキングに入ることはできないでしょう。それは政治的または金銭的不利益をもたらすかもしれません。

従って、有意で肯定的な結果を得なければならないために、リスクの高い研究を行わず、権威あるジャーナルに掲載することにインセンティブが働きます。有意ではなく否定的な結果しか出なければ、トップレベルのジャーナルには掲載されません。これは出版バイアスにつながります。一例として、臨床試験では、公表されやすいのはポジティブな結果が出た試験ですが、インターネットで検索すれば、ネガティブな結果が出た試験や、有意でなかった結果が多く見られます。これでは私たちは情報のほんの一部を取得するだけになってしまうので、この報酬制度を変える必要があります。一方、現在のシステムは研究者と出版社には利益をもたらしています。研究者は出版社に著作権を与え、出版社はそれによってビジネスを行う機会を得て、優れた論文を提供してジャーナルの価値を高めています。出版社は研究者の評判を高め、その評判は研究者が大学内で助成金や威信を得るのに役立ちます。

これは win-win であるとも言えるかもしれませんが、誰かがお金を支払う必要があります (図 14)。各誌の価格上昇やジャーナルの数の増加によって、上昇し続けるジャーナルのコストを払わなければならないのは、図書館や大学です。このことは他のモデルを考えてみ



(図 14)

る理由の一つです。

私たちはなぜオープンアクセスに移行したいのでしょうか。第一に、インターネット技術を利用すれば、コピーや配布は簡単で安価です。また、資金提供者はより多くの利益を望んでいます。それは、出版においてのみ言っているのではなく、知識や成果を共有するという意味でもあります。そのインパクトは、科学的なインパクトだけでなく、社会やイノベーションに対するインパクトまで拡大しなければならないのです。

重要な原則の一つが、公的資金を受けた研究の結果は公的に利用可能であるべきだということです。資金提供に値する研究は分かち合う価値があるはずで、科学を社会とつなげなければ、出版物や成果物へのアクセスを改善する必要があります。人々は今、専門分野内の成果ばかりに焦点を当てているので、アクセスの改善は科学に役立つかもしれません。エルゼビアのSTM Digestは、各専門分野で何が起きているかを説明するためのものです。これは一般の人が使用するものではなく、研究者が他の分野について学ぶために使用されています。従って、これは科学が前進する助けとなっています。私たちは、科学の公正性と信頼の向上に取り組むことができるのです。

ゴールド、グリーン、ハイブリッドなどさまざまなバリエーションのオープンアクセスがあります。しかし、オープンサイエンスには多くの誤解があります。公表する義務はありません。それは特許と相反するものではありません。特許を取りたいのであれば、伝統的ジャーナルであろうとオープンジャーナルであろうと、まず特許を取ってから出版する必要があります。また、ピアレビューのプロセスにも相違ありません。どちらにもピアレビューがあります。オープンアクセスには悪質なジャーナルもありますが、それは伝統的なシステムでも同じです。一部のジャーナルでは既存の論文をコピーして購読モデルにすることさえあります。

つまり、われわれがオープンアクセスを望むなら、

政策が必要です。お伝えしたとおり、ほとんどの国は既にオープンアクセスに関する政策を持っており、多くはグリーンに関してですが、一部はゴールドに関するものです。恐らく最も見事な政策はゲイツ財団のものだと思います。それは1ページで書かれ、五つの項目からなります。オープンアクセスで直ちに公開し、データを共有し、ゲイツ財団がコストをカバーして、公正な価格を支払うというものです。この方針には例外がなく、2年間浸透に費やした後、発効しています。従って、出版物のオープンアクセスについては、私たちはどの方向性に進むべきかもう分かっているのです。

研究データの最適な再利用

研究データの最適な再利用については、さらに少し複雑になります。少なくとも欧州では、出版は一つの出版モデルから新しい出版モデルへの移行期間にあります。

研究データは項目の組み合わせです。欧州委員会には三つの柱があります。一つ目はインフラを管理し、コンテンツと連結する欧州オープンサイエンスクラウド (EOSC) です。研究インフラ欧州戦略フォーラム (ESFRI) は大量のデータをつくる大規模な研究インフラであり、コンテンツの一部です。二つ目の柱は、高性能コンピューティング (HPC) ネットワーキング ソフトウェアを提供する欧州データインフラストラクチャー (EDI) です。これは、大規模な欧州 HPC、ネットワーキング、ソフトウェアを持ち、これらのサービスと活動を組み合わせて一つのサービスに集約するものです。つまり、サービスとしてのインフラです。研究者として、私はストレージと計算機能が必要です。同僚とつながる必要もあるので、これらはサービスとして提供されて然るべきです。三つ目の柱はアクセスを拡大することで、それには中小企業、産業界、政府が含まれます。

EOSC に焦点を当てると、委員会の前議長はかつて、これは欧州のものではなく国際的なものであると言いました。一部のデータはクローズドなので完全なオー

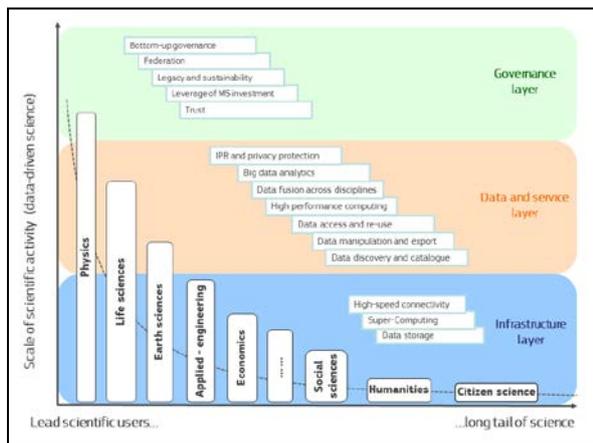
ブンではありません。また、科学のみならず、イノベーション、公共的なものでもあります。これはクラウドでありながら、現実のものであり、設備を提供する地上のデータサービスです。研究者がデータを保存、管理、再利用するための環境をつくり出しています。私たちは、大学や研究所に、そして国家的または欧州レベルでスーパーコンピュータとともに多くのインフラが既に存在していることを認識しています。課題は、既存および新しいインフラといかに連携するかということです。ガバナンスの問題とこれらの施設との連携については、昨年 10 月に発表された EOSC レポートに記載されています。

項目の一つを見ると、図 15 はガバナンスの設定、データまたはコンテンツと新しいサービスの提供、またはインフラの稼働についてのものです。これはこのクラウドの挑戦です。オランダの DANS 研究所の試みは、データに対する一種のミシュランガイドを提供することです。私たちは、FAIR 原則の最初の三つの部分（探索可能・アクセス可能・相互運用可能）のそれぞれに対して五つの質問をし、成績をつけることができます。平均を取ることで、一般的な成績がわかります。これはデータを知らせる最初の方法です。私は、課題の一つはデータを発見可能にすることだと思います。たくさんのデータの中から、どうすれば見つけることができるのでしょうか。

私はクラウドは既に存在していると強調したいと思います。例えば、アメリカ国立衛生研究所 (NIH) の

NIH Commons においてです。アメリカ国立科学財団にもクラウドがあり、マイクロソフトとアマゾンには商業クラウドがあります。私が印象的に思うのは、ほとんどがアメリカの例であるということです。彼らはまず、NIH が運営している 600 万ドルのクラウドのようなパイロットから始めています。欧州ではもっとトップダウンで行っています。私たちはガバナンスについて考え、始める前に全てを整えたいと思っています。しかし、このアプローチを取れば、日本には既にバックボーンと設備があるため、有利なスタートをすることができます。アメリカがクラウドに投資し、アメリカの研究がクラウドに投資すれば、国家レベルの投資になるでしょう。欧州がクラウドに投資したいと思ってプロバイダに進めば、国際的な投資になるでしょう。また、これらのサービスを提供するために国内または欧州のプロバイダを調整し準備する必要があります。このために、このパイロットが必要なのです。

先ほど述べたように、データは見つけることができなければなりません、再利用するためにはデータは信頼できるものでなければなりません。DANS 研究所と日本の World Data System (WDS) による「Data Seal of Approval」は、データが良好なフォーマットで見つけ出され、信頼できるものであるという保証を与えるツールを提供します (図 16)。また、データセットの識別子も提供します。データを作成した研究者にクレジットを与えたいならデータを参照できるようになっていることが大切です。



(図 15)



(図 16)

しかし、データを共有し最適に再利用するには、共有を推進するための信頼を勝ち取る文化的な変化が必要です。このためにはデータを参照できる必要があるため、識別子を利用すべきです。しかし、共有するデータを生産する者に対して報酬やインセンティブも与えなければなりません。また、データ生産者とユーザーの認証も必要です。理想としては、データを提供する、使用する際に自分が何者かを知らせるために、データにアクセスするためのある種のシングルサインオンがあったらいいでしょう。特に生命科学や社会科学では、機密データをどう扱うかに注意を払わなければなりません。一部のデータはインターネット上で公開することはできませんが、それでも研究には関係してきます。われわれはそれにどのように取り組んでいけばいいのでしょうか。データインフラについて結論付けると、その方向性はあまり明確ではありません。私たちは、共有のためのインセンティブ、報酬の仕組み、識別子が必要であることは分かっていますが、これらをどう実現するかはまだ明確になっていません。

オープンサイエンスの展望

オープンサイエンスに関するビジョンについては、まず出版物から話を始めて、次にデータに移りたいと思います。国際 STM 出版社協会が出した「STM Tech Trends 2015」では、「The Article in a Hub and Spoke Model (ハブ&スポークモデルにおける論文)」と「Data as first class Research Object (第一級研究対象と



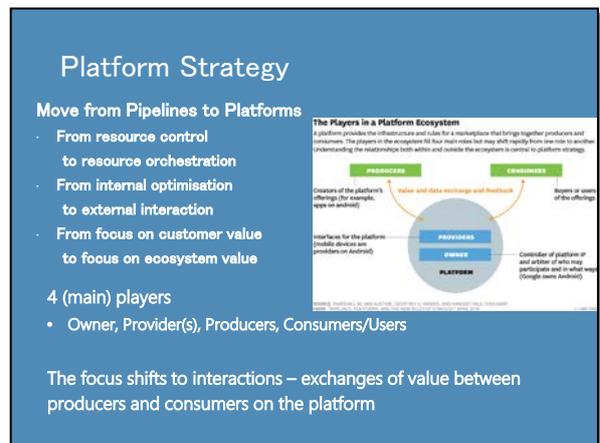
(図 17)

してのデータ)」とあります。データは研究成果であって、共有する必要があるということになったのです。

「STM Tech Trends: Outlook2020」は、ビッグデータと、データと人工知能の組み合わせに関するものです(図 17)。機械は、論文やデータを読む新しい読者になるかもしれないし、研究者を助けることになるかもしれない。そのためには、正確な情報を提供することが非常に重要です。私は 1 年間に発行される 300 万の論文の中から 300~400 件の論文しか読むことができないので、選択の際に何らかの助けが必要です。以前は 1~2 誌だけしか読めなかったものが、たくさんの雑誌が出てきました。関連論文が他誌に出たりすることもあり、フィルタリングが必要です。

出版物に追加された価値はフィルタリング機能です。例えば、私の 15 歳の息子はコンテンツに対してお金を支払おうとはしませんが、インターネットで音楽や映画を見つけるときには助けが欲しいかもしれません。そのために彼はソーシャルネットワークを利用します。一方、研究では、関連情報を見つけるのに助けが必要です。出版物において、読者は関連する論文を探したいと思っていますが、著者はできるだけ多くの読者にアプローチしたいと考えるのは明らかです。出版物の供給と需要という市場があるのです。

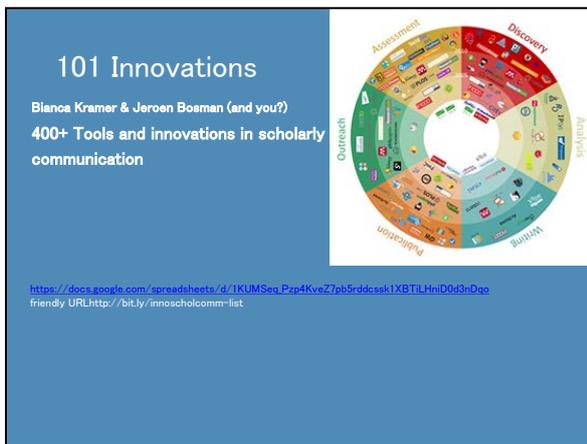
新しい方向は、プラットフォーム戦略を持つことです。図 18 は、戦略を持ちたければ異なった考えをしなければならないと言う「ハーバード・ビジネスレビュー」の論文に基づいています。私たちは焦点をシフ



(図 18)

トしています。それは限界コストに関するものではなく、プラットフォーム上の生産者とユーザー間の相互作用と価値交換に関するものです。これらのプラットフォームには、所有者、プロバイダ、生産者、ユーザーがいます。一つの大きな例は Apple です。彼らは iPhone と iPad というプラットフォームを所有していて、これは電話のためでなく、そこでしか使えないアプリケーションを用意しています。ユーザーを見つける App Store と iTunes のことです。Apple の成功は、このつながりをつくったことにあります。

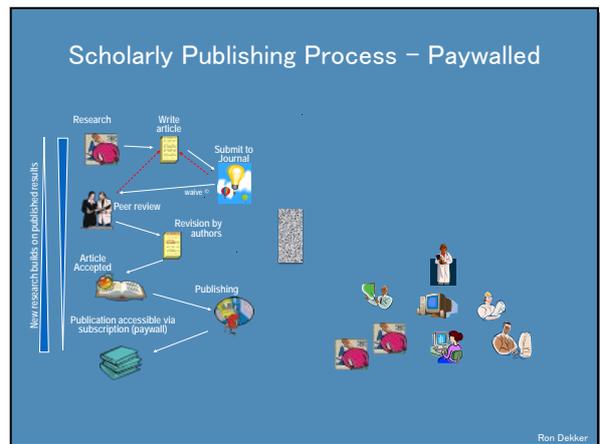
もう一つの例はソニーの PlayStation です。PlayStation 4 では、月会費を払えば、オンラインで世界中の他のプレイヤーとプレイできます。このように、プラットフォーム上の付加価値は重要です。私は、ユーザーと生産者が良好なつながりを持つことが、出版物の方向であるべきだと思います。



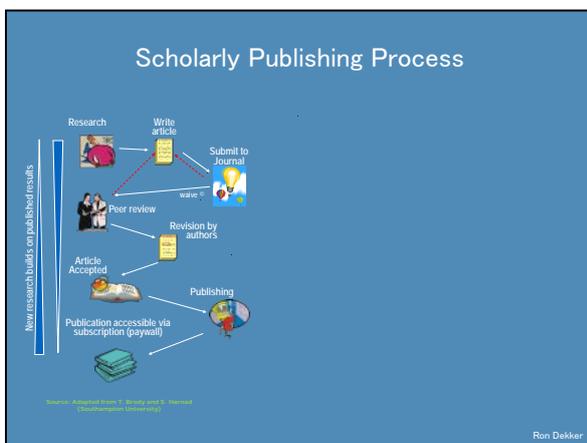
(図 19)

第 3 回 SPARC Japan セミナー 2015 で、図 19 を見たことがあると思います。「101 Innovations」は、利用可能な全てのツールを記録したのですが、現在はオープンサイエンスを行うときに利用可能なツールの数であるため、「400+ Innovations」と呼ばれるべきです。

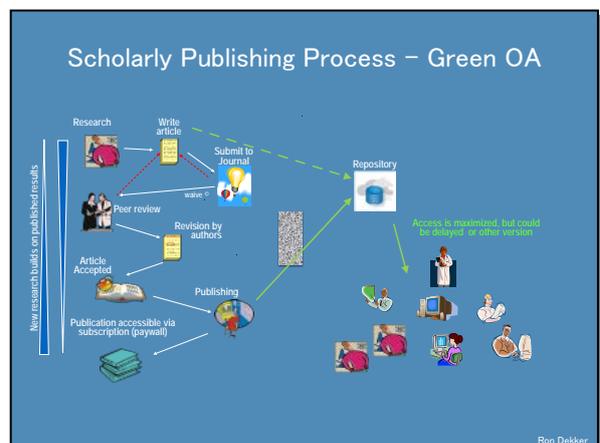
図 20 は伝統的な出版プロセスを示しています。論文を書く、提出する、却下される、別のジャーナルに再度提出する、査読される、拒否されるか改善する、論文が受け入れられる、発表する、これがジャーナルに掲載されるということです。ペイウォールモデルでは、科学界の外の人々は見ることができず、購読していなければ論文にアクセスするために 30 ドル支払わなければなりません (図 21)。グリーンオープンアクセスでは、エンバゴ方式とプレプリント方式によって利用可能です (図 22)。従って、私たちはプロセスの最後を変えるゴールドオープンアクセスを考え出し



(図 21)



(図 20)



(図 22)

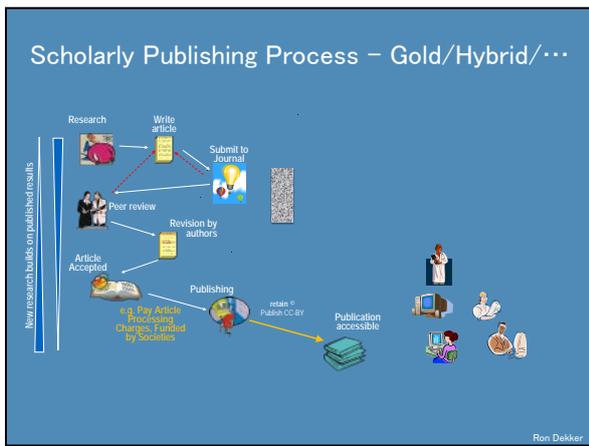
ました (図 23)。私たちが支払いをし、誰でも利用できるようにします。このゴールドモデルは標準になるでしょうか、それとも他のモデルが出るでしょうか。

出版ではウェルカムトラストによる新しいアイデアがあります (図 24)。これは「Open Research」と呼ばれ、論文執筆後 1 週間以内に論文が発表されるものです。その後、論文が公開の場で査読され、結果を利用できるようになります。また、これは出版物に加えて検査結果やデータにも適用されます。どのようにしてデータが調査されるかは分かりませんが、これによって、データセットの品質を保証する新しい方法が見つけれられると思います。最初に出版し、次に品質評価を行うというように、物事の順番が変わってきています。さらに進むと、研究を再び途中の段階に戻すことができます (図 25)。研究して、それを公開するかどうかを決定します。公開する場合、伝統的なジャーナル、

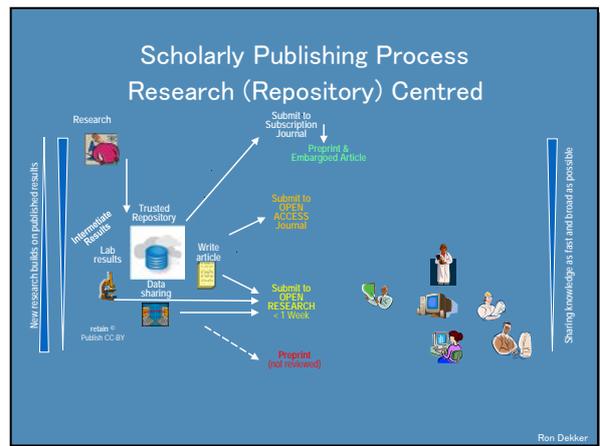
ゴールドオープンアクセス、またはオープンリサーチに送ることができます。私の場合は査読を選ばないで、プレプリントで投稿します。そうして、アクセスを得ることができます。これは研究者にとって非常に簡単な方法です。

この方法は、オープンアクセスをさらに容易にするため、資金提供者にとってもとても簡単です (図 26)。このモデルは、約 750 ドルという FAIR な価格で査読を受けられます。これは新しい方法かもしれませんが、いまだ既存のジャーナルシステムの中にあります。

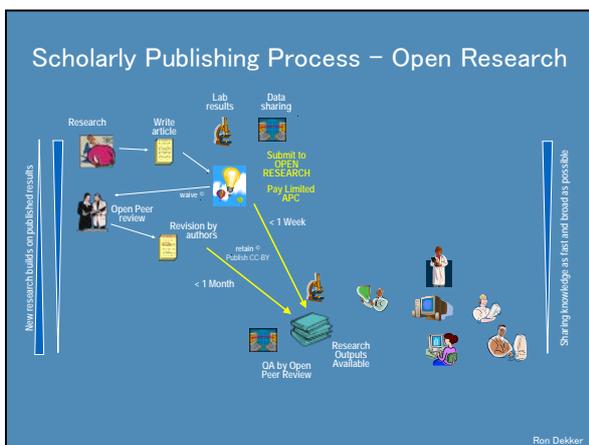
別の動きはプレプリントの中に見いだすことができます。資金提供者が識別子を取得するプレプリントを喜んで認め、プレプリントに関する助成金を容認すると言えば、若手研究者にとって助けとなります。2 年間の博士研究員の助成金を受けている場合、発表されるまでに 1 年かかる論文を提出する時間がありません。



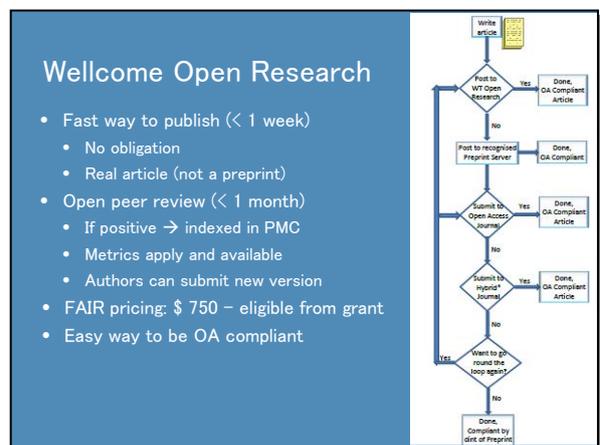
(図 23)



(図 25)



(図 24)



(図 26)

このプレプリントでタイムスタンプを取得した場合、助成金に対する新しい申請でそれを使用することができます。これによって出版や知識の共有もスピードアップします。私たちはプレプリント用の信頼できるリポジトリのシステムを持つことになるでしょう。このように、出版における知識共有の動きを追うことで、将来の方向性を見いだすことができます。

データに関しては、皆さんの力が必要です（図 27）。データを再構成するために必要な労力を考慮しても、データには莫大な価値があります。研究者がデータを準備し共有することに労力を割けば、他の誰かが、論文や新しい知識を生産するのに多くの時間を節約して収入を得ることになります。従って、データ共有に関する市場をどのように設定したらいいかというのが大きな疑問の一つになります。私たちにはインフラがあるので、パイロットプログラムも行うことができます。私たちは、全ての学問分野の国家政策を立てるべきか、国際的にそれぞれの学問分野ごとに政策を立てるべきか、あるいはその両方を行うべきかということについて議論することもできます。

市場を創出するに当たって重要な win-win の一つの例は、Structural Genomics Consortium (SGC) です（図 28）。SGC は、創薬研究に関する大規模な官民連携による研究団体であり、創薬ターゲットとなるヒトのタンパク質の 3 次元構造を同定・解析する研究を行っています。製薬企業は、競合同士であっても、対立するそのときまではこのコンソーシアムで協働しています。



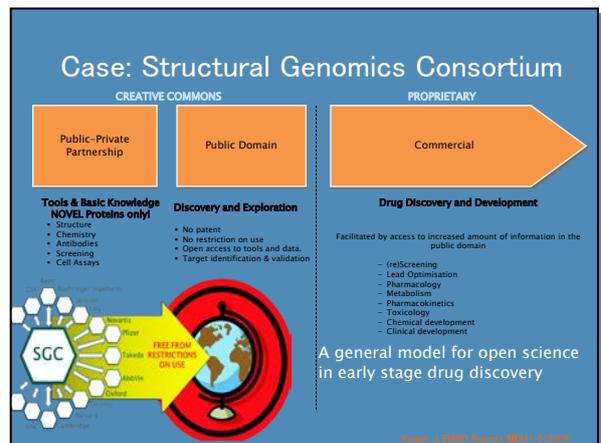
(図 27)

このコンソーシアムでは、信頼できる組織を設立し、10 の研究グループを持っています。組織はそこに調査したいと考えている上位 3 種のタンパク質を伝えます。その中から、興味を持っている参加団体が多かったトップ 30 のタンパク質について、どの団体がそれに興味を持っているか、それについてどのような取り組みをしているかは明らかにせずに研究しています。参加団体はこれに同意しています。これは win-win です。多くのデータと研究能力を持っているのですから、需要と供給、生産者とユーザーをマッチさせることが、データに対する課題です。

膨大なデータの中から、価値あるデータを探そうとすることができるようにしなければなりません。これが基本です。図書館は、(データと研究者の) 仲介役を果たすことができるかもしれません。図書館同士はつながっているため、他の図書館で何が起きているかを知っています。データに関する質問がある場合、図書館員は手助けすることができます。

結論として、私たちはオープンサイエンスの新しい時代の一部とならなければなりません。私たちはより多くの人々にアプローチし、科学だけでなく社会にも大きな影響を与えたいと考えています。それは二度手間や出版バイアスを防ぐことができます。科学、社会、イノベーションに影響を与えるためには、将来使うためにデータを保存する必要があります。

課題は、いかに協力し、イノベーションを生み出し、共有するかということです。私は日本に来られてとて



(図 28)

もうれしいです。私が日本を知るようになったきっかけは、伝統的な映画「楢山節考」、伝統的な企業でどのようにイノベーションを起こすかについて説明した『知識創造企業』という本です。私たちをオープンサイエンス、データの再利用、出版物のオープンアクセスに導いてくれるのは、伝統の融合、過去の理解、イノベーションと変革への勇気だと考えています。最後に、南アフリカ共和国のことわざをご紹介します。終わります。「早く行きたいなら、一人で行きなさい。遠くに行きたいなら、一緒に行きなさい」。

● **深貝** 横浜国立大学の深貝と申します。SPARC Japan の運営委員をしています。オープンサイエンスについて、このところ日本でも非常に活発に議論しています。歴史を振り返れば、近代の入り口の科学革命のころ、グーテンベルクの活版印刷を通じて、こもっていた局所的な知識がオープンになって、そして科学者が振る舞い方を変えていきました。それと似たようなことが今起きはじめています。ネットワークを使って、今までの印刷物ではない情報が飛び交うようになった中で、人々の振る舞い方が変わってきているのです。

電子ジャーナルが出たというのは、単に印刷物が電子版で出ただけですから、ある意味で伝統的です。しかし、最近では、人々がネットワークを使って、知識を交換し合っています。知識を交換することによって、多くの人々が知識の生産の場に関わり、アイデアがひらめくということになりつつあるのです。それを一層加速しようというのがオープンサイエンスだと思います。

ただ、その場合のバリアは、学術の成果が誰かに帰属し、その人がインパクトファクターによって評価を受けるという方向に加速していることです。そうすると、科学者・研究者は、高い評価を受けるために、知識を狭めた上で成果を挙げることに走りがちです。しかし、広い目で見て人間の知識を多く豊かなものにしていくためには、オープンにしていっていい方いい。そ

のような揺らぎの中にあるわけです。

その中で、公的資金を使って得られた成果をオープンにするということを今やっていますが、それだけで十分なのかどうか、お考えをお示しいただけるとありがたいです。

● **Dekker** 現在は、印刷物をインターネットで利用できるようにするのではなく、そもそも成果を最初からデジタルで出すようにしなければならない、そういうパラダイム転換の最中にあると言えます。私はウェルカムトラストの「Wellcome Open Research」が好きなのですが、その理由は伝統的な論文と中間結果の両方に焦点を当てているからです。そこでは重要なデータを公開し共有したいと思えば、することができます。ジャーナル「Science Matters」では、研究コミュニティにとって重要であると思われる場合は、中間結果を公表することができます。私たちは、伝統的な出版を越えて、成果の新たな発信方法を見つける必要があるのです。

これらの成果を測定する際は、ジャーナルのインパクトファクターを超えた新たな指標が必要になります。資金提供者は、代替物や新しい成果の発信方法に報酬を与えるべきです。イノベーションまたは社会的議論に重要な貢献をした場合も、資金提供者や大学によって報酬が支払われるべきです。研究者に対する独自の報酬システムを持つ大学や医療機関もあります。資金提供者は、出版物の数で評価することをやめるよう、所属する委員会のメンバーに求めるべきです。代替案としては、意味のないインパクトファクターを積んでいくのではなく、査読者がトップ3またはトップ5の出版物を精査することが挙げられます。異なるものを比較しているのですから、分野を超えてインパクトファクターを比較することはできません。DORA イニシアチブは、研究を評価するために定性的な情報を使用すると述べています。それが進むべき道であり、また、私たちは、プレプリントや他の公開方法の価値を認めるべきです。