

オープンサイエンス政策の背景と情報科学への期待

林 和弘¹

概要: オープンサイエンスは ICT によるデジタル化とネットワーク化された情報基盤およびその基盤が開放する多量で多様な情報を様々な活用して科学研究を変容させる活動であり、産業を含む社会を変え、科学と社会の関係も変える活動である。オープンサイエンスはイノベーションを駆動するものとして重要な位置づけとなっている。本発表では、既報の会誌記事を踏まえてオープンサイエンス政策の背景と日本が関わる政策の最新状況を概説し、歴史から見た情報科学が果たす役割と期待について述べる。

キーワード: オープンサイエンス, G7 科学大臣会合, 統合イノベーション戦略, 研究データ基盤, デジタルトランスフォーメーション

Background of Open Science Policy and Potential of Information and Computer Science

KAZUHIRO HAYASHI^{†1}

Abstract: Open Science is a movement to transform Science, Society, and “Science and Society” driven by making various use of ICT’s digitalized and networked information infrastructure and the vast and diverse information that the infrastructure opens up. Open science has become an important position as a driving force for innovation in science and technology policies. In this presentation with referring the published article on the membership magazine, this report outlines the background and current status of open science policy, and describes the potential of Information and Computer Science from a historical perspective.

Keywords: Open Science, G7 Science Minister Meeting, Integrated Innovation Strategy, Research Data Infrastructure, Digital Transformation

1. はじめに

オープンサイエンスにはまだ広く一般的に通用する定義はないが、オープンサイエンス政策は、現在、ICT の発展による社会変革と研究データの持つ可能性に着目し、世界の主要国において重要なテーマとなっている。日本においても、研究データはイノベーションを生み出す源泉の一つとして重要視している。

本稿は、2019 年に情報処理学会誌（情報処理）に寄稿したオープンサイエンス政策とその背景に関する総論[1]（以下会誌記事と呼ぶ）を、研究データマネジメントを中心に、アップデート、強化、補足するものであり、情報処理学会オープンサイエンスと研究データマネジメント（RDM）研究グループの発足に伴う基調講演報告とするものである。オープンサイエンスについては、本稿においては、“ICT によるデジタル化とネットワーク化された情報基盤およびその基盤が開放する多量で多様な情報を様々な活用して科学研究を変容させる活動であり、産業を含む社会を変え、科学と社会の関係も変える活動である” [2]とする。

なお、本項の主旨から、オープンサイエンスの可能性の一つである、科学と社会の関係性の変化や共創型シチズン

サイエンスの可能性については触れないので既報[2]を参照されたい。

2. オープンサイエンスが持つビジョンと変化の兆し

オープンサイエンスは科学や社会にどのような変化をもたらすのか、そのビジョンと変化の兆しについて最新の状況を報告する。

(1) “オープン” の誤解と意味するもの

オープンサイエンスの持つ“オープン”の意味については、ある情報を“すべからず誰でも自由に再利用可能”という狭義の意味合いで議論されることもあるが、社会全体の変革の観点からは、このオープン化は情報の“開放”を意味するものであって、この開放の仕方が、ICT の発展によって、新しく、また、多様になっていると解釈すべきである。すなわち、オープンサイエンスは“知の開放”による科学および、社会の変革とすることが望ましいと考える。歴史的にみても、活版印刷によって宗教革命や科学革命がもたらされたように、情報基盤のオープン化が社会変革をもたらしたとして解釈が可能である。実際、政策的にも特に産学連携や産業育成において、このコンセプトを提示することにより現実的な政策的な議論が始まり、後に説明する統合イノベーション戦略の策定に繋がっている。 [3]

¹ 文部科学省科学技術・学術政策研究所
National Institute of Science and Technology Policy (NISTEP)

(2) 科学のデジタルトランスフォーメーションと位相変化

ICT によるオープン化が進むと、情報流通の基盤が根底から変化するため、科学自身もデジタルトランスフォーメーションする必要がある。その位相変化を一例として示したのが図 1 となる。

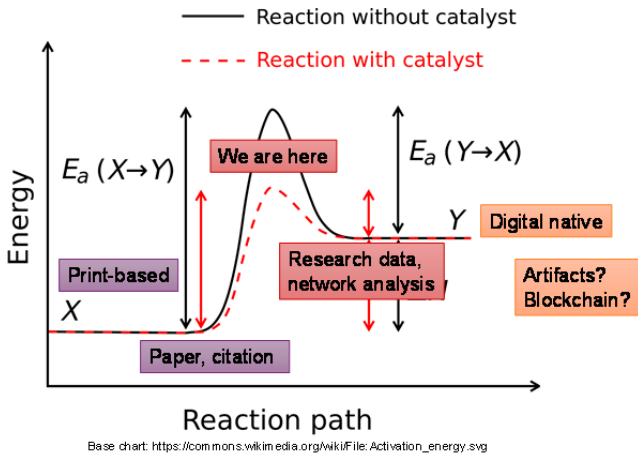


図 1 オープンサイエンスパラダイムへの位相変化

印刷と物流による情報流通の位相 (パラダイム) から、デジタルネイティブな位相に変化するためには、多くのエネルギー (ヒト・モノ・カネを中心としたコスト) が必要で、また、触媒機能として政策や知恵を活用することで、位相の変化を加速することになる。今現在は、その端緒として、印刷基盤から生まれた論文や被引用数を用いた研究インパクト計量の位相から、研究データを用い他の研究に関するアイテムを絡めたネットワーク分析による研究インパクト計量の開発を目指している状況にあり、その先は、研究成果に限らず、研究プロセスを含めて必要に応じて保存・共有・公開できるパラダイムをオープンサイエンスのビジョンは持っている。すなわちこの位相変化は多段階であると言える。

この、デジタルネイティブな位相における学術情報流通として考えられるのが、研究の着想から成果公開までの活動を、その個々の貢献者の把握と共に記録できる研究プラットフォームである。その一例として図 2 に示す。このようなプラットフォームが実現すれば、論文と被引用に過度に依拠した研究評価をより健全なものにできるだけでなく、著者に限らない、研究データ作成や研究費獲得に関わる貢献者を定量的に評価できる可能性を持つ。その一方、商業出版社が、各研究プロセスにおいて便利なツールを開発あるいは開発した企業を買収してこの研究プラットフォームを構築しようとしており、オープンサイエンス政策の一つの目的は、学術ジャーナルの論文で起きた寡占化や価格高騰を避け、研究者の主権の元、より健全な学術情報流通基盤を構築しようとするにある。

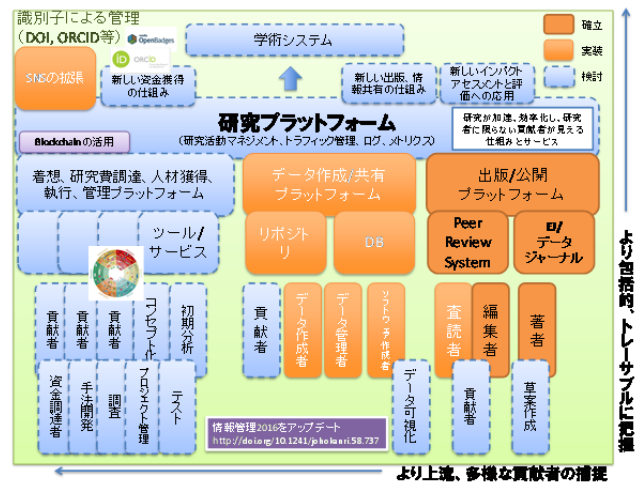


図 2 出版/公開プラットフォームから研究プラットフォームへ[4]

(3) 破壊的変化のきざし

オープンサイエンスパラダイムへの変化の兆しについては、会誌記事でも述べた通り、研究者にとってより身近な論文を中心とした世界の中でも起きています。例えば、プレプリントサーバーを多くの分野が受け入れていることや、研究助成団体が研究成果公開プラットフォームを持つことで、出版者や図書館を介すること無く研究者に研究資金を助成してその成果を公開できるようになるなどが起きています。本報告では、新たに事前登録報告 (Registered Report) とクラウド共同執筆について詳しく述べる。

まず、事前登録報告 (Registered Report) は、着想、研究計画の段階で一旦査読を受け、通った研究計画に基づいて得られた研究成果は、例えそれが作業仮説に反していても、原則出版・公開されることが保証されているものである。現在は臨床医学系を中心に取り組まれている。

(図 3)

事前登録報告 (Registered Report) の特徴は、これまで、投稿者が投稿する時点で自分が良いと思った成果、あるいは学術ジャーナルに通りそうな結果だけが選別される出版バイアスが軽減されることにある。学術論文をビッグデータ化して AI が解析することに注目が集まっているが、現状偏った論文データに基づいた解析となっているのが現状であり、このデータがより中立になることは非常に望ましい。いわゆる失敗データも研究成果として認められることになれば、研究者の行動も変容することになる。また、着想後の研究プロセスが透明化することで、研究の再現性の確保に寄与するだけでなく、想定外の結果が得られた過程を追跡し、セレンディピティを補足することも容易になる。一方、着想研究計画の段階で査読を受けることは、アイデアの流出と見ることもでき、現状大きな抵抗が想定される分野もある。この点

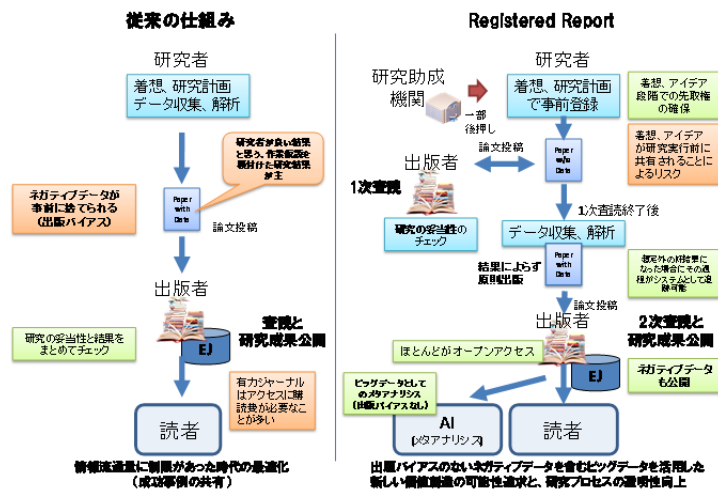


図 3 Registered Report の概要, 出版バイアスと AI との親和性

は、判断が分かれるところであるが、将来実験はロボットクラウドに任せて、研究者はそのロボットに研究を執行するプログラム（実験設計書）を書けば良いという考えも出始めている中[5]、誰よりも早く着想したことを第3者に認めてもらうという機能は今後大きな可能性を持つことも考えられる。

続いてクラウド共同執筆の可能性について述べる。クラウド共同執筆では、文字通り、クラウド上で、論文の執筆と編集を行うものであり、ツールによっては、版を訴求し誰がどの記述を行ったかを追跡することも可能となっている。

クラウド共同執筆が持つ破壊敵なポテンシャルは、まず、共同執筆したものは、そのままプレプリントとしていつでもどのバージョンでも公開が可能となる。そして、共同執筆の過程で、誰が執筆に貢献したかをログとして把握できる。さらに、執筆のどの段階でも共同研究者以

外の研究者の意見を取り入れ、あるいは学術ジャーナルの査読者に相当する人を入れて論文を直接改善することも技術的には可能であり、ピアレビューの新たな形態を生み出す可能性がある。すなわち、論文を編集あるいは質を上げるという役割において、共著者、査読者の境界があいまいになっていく可能性がある。学術論文の目的は研究で得られた成果を一定の質保証をした上で早く広く世に伝えることであり、その目的に最適化して、ツールやシステムと共に業務フローが再構成されることが既に起きようとしているといえる。

会誌記事で紹介したプレプリントサーバー、研究助成団体による研究成果公開プラットフォームと同様に、これらの変化は、

学術ジャーナル、学術論文、査読のあり方を根本から問い直しており、学術ジャーナル自身もデジタルトランスフォーメーションする必要がある。例えば、学術論文についてはではないが、Google が提供する Assignments と呼ばれるプラットフォームでは、先生が生徒を指導する目的のもと、Google Docs や Google Drive を活用して宿題などの作成と管理ができるようになってきている。その機能の中には、宿題を集めて点をつけ、生徒に対しコメントを付けてフィードバックするというものがあり、盗作チェックもできるようになった。[6]この盗作チェックは事前に生徒が行うことができ、そのレポート (Originality Reports) を先生に提出することで、双方の手間も省ける。この Assignments で起きている、教育業務の再構成は一つの参考イメージとなる。

3. オープンサイエンス政策の近況

オープンサイエンス政策については会誌記事で多くの紙面を割いたので、ここでは、日本が関わる政策の最新状況をと研究データを中心にその概要、会誌記事との差分や補足を中心に述べる。

3.1 オープンアクセス政策、学術ジャーナルと研究データ

オープンサイエンス政策には、学術ジャーナルと論文のオープン化を中心としたオープンアクセス政策も含まれるが、本研究会は研究データマネジメントを主題にするために、本稿では取り上げない。

ただし、多くの学術ジャーナルで、原稿提出時に論考を支えるデータ (エビデンスデータ) の提供を求めるところが増えていることには注目すべきである。こ

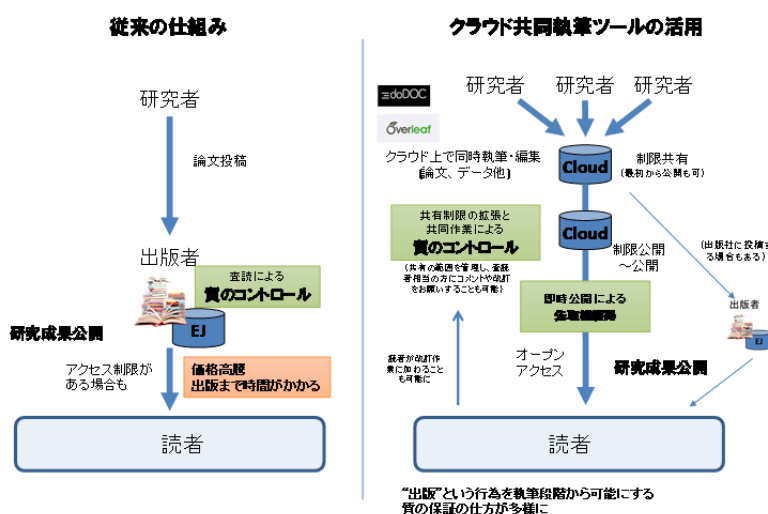


図 4 クラウド共同執筆による即時出版と査読の変革

れは、単に、学術ジャーナルと研究データマネジメントに接点があるという以上の意味を持つ。すなわち、研究者にとって、慣習的かつ研究者として現在必要不可欠な論文投稿という行為に付随しているが故、研究データ管理に関わる者にとっては、研究者に欠かせないサービス開発における重要な機会（チャンス）であると考えべきである。実際、結晶学のように論文投稿時に国際標準のデータファイル（結晶学の場合は CIF ファイル）を必ず登録するような分野では、時に論文と研究データのメディアとしての逆転現象が起こり、研究データを publCIF というツール内で整えつつ、論文の内容も記述してコンパイルすると国際結晶学連合の学術ジャーナルのフォーマットに沿った論文ができあがるようなサービスも存在する。[7]

3.2 研究データ基盤整備としてのオープンサイエンス政策

オープンサイエンス政策の現在は、研究データ基盤整備、研究データの研究成果としてのメディア化、そして研究データ流通によるエコシステムの構築を目指していると言っている。

(1) G7 科学技術大臣会合

2013 年ロンドン G8 における、研究データオープン化合意に始まり、2016 年つくば G7 では、オープンサイエンス推進合意と共に日・EU 共同議長国のワーキンググループが発足、WG 発足して、研究データのインフラ整備とインセンティブの付与を中心テーマとして、2017 年トリノ G7 におけるトリノコミュニケにてこの主題が明記された。2019 年パリ G7 においては、6 月に本ワーキンググループのワークショップを開催して、各国の取り組みを共有し、また、CODATA など関連するイニシアチブの専門家を招いて議論を行った。引き続きリーディング国として、研究データのインフラ整備とインセンティブづくりに取り組むこととなり、来年米国で開催予定の G7 における検討に注目が集まる。

(2) 内閣府 統合イノベーション戦略

第 5 期科学技術基本計画に明記されたオープンサイエンスの推進を踏まえて、目標を掲げて実践する統合イノベーション 2018 において、オープンサイエンスを“知の源泉”を担う柱の 1 つとして位置づけ、研究データ基盤（インフラ）の整備、研究開発法人のデータポリシー策定、研究助成機関におけるデータマネジメントプラン提出の要請等を促し、研究開発法人向けのデータポリシーガイドラインや研究データリポジトリの整備・運用ガ

イドラインを作成している。さらに 2019 年版の統合イノベーション戦略においては、2019 年 6 月に閣議決定された統合イノベーション 2019 では、研究データ基盤の整備を進めつつ国際的な連携を進め、グローバルに、また、相互通用性を担保しながら研究データの利活用を押し進めるとしている。

(3) 文部科学省 情報委員会

文部科学省では、科学技術・学術審議会学術分科会の下に学術情報委員会を置き、主に図書館や学術ジャーナルに関連する学術情報流通に関する施策を議論してきたが、平成 31 年度に第 9 期の学術情報委員会を終えると、科学技術・学術審議会の下に情報委員会をおき、学術ジャーナルの問題を含めてより包括的に取り組むことになった。まだ研究データ基盤やその利活用に関する具体的な課題は設定されていないが、内閣府の動きと歩調を合わせた対応が見込まれる。

(4) 内閣府 ムーンショット型研究開発プログラムにおける先進的な研究データマネジメント

我が国発の破壊的イノベーションの創出を目指し、従来の延長にない、より大胆な発想に基づく挑戦的な研究開発（ムーンショット）を、司令塔たる総合科学技術・イノベーション会議（CSTI）の下、関係省庁が一体となって推進する新たな制度として、ムーンショット型研究開発制度が立ち上がっている。統合イノベーション戦略 2019 にも明記されているこのムーンショット型研究開発プログラムにおいては、先進的な研究データマネジメントも行うとしている点が、大きな特徴である。

研究データ管理を行う際の大きな障壁の一つとして、研究者はすでに何らかの形で研究データを管理し続けており、極端に言えば、個々人によって研究データの管理が違う。このため、研究データ基盤を整備したとして、そちらに合わせることは躊躇しやすい。また、規模の小さな研究において、研究データ管理のコストが研究費を超えるものには、その意味はない。この二つの観点を踏まえ、これから行う大型研究において、先進的研究データマネジメントを取り入れることには一定の妥当性があると考えられる。

4. 情報科学への期待・課題とデジタルネイティブ

本項では、会誌記事と同様、オープンサイエンスパラダイムに向けた情報科学への期待を別の角度も交えて改めて述べる。

4.1 印刷革命と数理論理学の関係にみる可能性

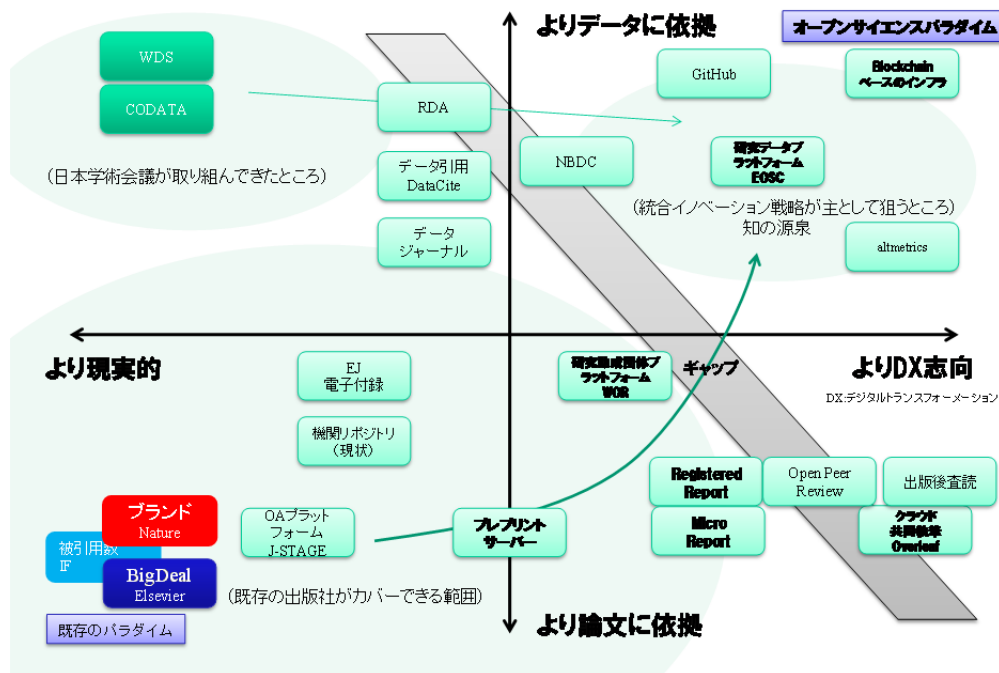


図 5 論文からデータへのデジタルトランスフォーメーションとそのギャップ

冒頭や会誌記事でも述べたとおり印刷革命による知の開放によって科学革命も起こり、数理物理学が生まれ、近代科学につながる発展を遂げている。この歴史を繰り返すとすれば、ICT による革命をもたらす新しい科学と社会の発展をオープンサイエンスは数百年ぶりに生みだそうとしているという見方が可能であり、情報科学が果たす役割大きい。そして、次に述べる通り、今の情報科学そのものがその新しい学問とは言えない状況にある。

4.2 情報科学自身のデジタルトランスフォーメーションの必要性

現在の情報科学自身の運営基盤としては、17 世紀に生まれた学会の仕組みで研究者コミュニティを形成し、同じく 17 世紀に生まれた学術ジャーナルよりは国際会議のプロシーディングス等別メディアを指向しつつも、その執筆、発行プロセスが学術ジャーナルとその電子化されたものと大きな差がない仕組みの中で動いているのが現状である。これは、現状のコミュニティに必要なサービスをコストを含めて最適化した結果であって、何も否定されるものではない。これまでの学会や学術ジャーナルが果たしてきた役割や、今後の本質的な貢献は揺るぎないものとして、その上で、情報科学とそのコミュニティが自らデジタルトランスフォーメーションする必要がある。そのためには、今ある学会が担ってきた機能をメタ化して、将来の研究者コミュニティを想定した再構成とサービスデザインを行う必要があると考えられる。研究者の SNS である Researchgate や SSRN(Social Science Research Network)は Web ネイティブな研究者コミュニティとして一定の地位を確立し、また、

ブロックチェーン技術を活用した、学術情報流通のシステム作りも試行が始まっている中、次世代の研究者サービスデザイン自体が情報科学の重要なテーマであると拝察する。

4.3 過渡期におけるギャップ、ハイブリッド戦略とアカデミアの意味

デジタルネイティブな社会に向けて我々は 100 年単位でみれば過渡期に居ながら、3-5 年単位では局所最適化を繰り返して安定点を確保する必要にも迫られることは容易に想像がつく。それは、研究データに特化した成果公開メディアを開発しながら、学術ジャーナルと論文を中心とした現在のエコシステムを改革する必要があるという例としても表現できる。しかし、本質的には、学術ジャーナルというレガシーフレームの漸次的な変化の先に、デジタルネイティブなオープンサイエンスパラダイムが構築されるとは考えにくく、そこにはギャップが存在する。(図 4) 研究データの利活用の観点からみても、日本学術会議が支援する形などで世界科学データシステム(World Data System, WDS)や科学技術データ委員会 (Committee on Data for Science and Technology, CO-DATA) の活動に積極的に参画して世界の研究データコミュニティの中で日本の一定のプレゼンスを保ってきたが、こちらの活動も、研究データのデータベース化、アーカイブ化等、ストックとしての研究データサービスに特化した歴史を持ち、オープンサイエンスパラダイムが指向する、研究の着想段階からを含めた研究プロセス全体のプラットフォーム化に一足飛びに進化するのは難しいと考えられる。

従って、大局的には、研究、事業、教育において与えられている前提が変わりうることを念頭に、局所的には、何

を守り、何を変えていくかを、個人、組織のそれぞれの粒度で主体的に考えるハイブリッド戦略が必要であり、その活動の集合が、いつかギャップを乗り越え、近未来の社会を構成していくことになる。

ここで、オープンサイエンスパラダイムの大局的な安定点（ポテンシャル）に落ち着くまで社会は変革し続けることになり、また、デジタルネイティブな特性である、人や組織のネットワーク化が双方向性をもって進むことを考慮すると、その過渡期の研究者は、独立したアカデミアという閉じた世界にだけ止まる必然性がほとんど無くなる。このことを前提に、常に社会との接点と変革を意識して研究を行うことにより大きな意味があると考えられる。

4.4 デジタルナイチンゲールの可能性

ナイチンゲールは近代看護の創始者として有名であり、看護や医療改革に尽くした人物として一般的には有名であるが、黎明期にあった統計学の研究者としての実績が大きいことでも知られる。[8]この逸話の重要なポイントは、社会変革をもたらす医療改革の実質的キードライバーである統計学とそのスキルが、一般社会的には前面に出ていないことである。すなわち、オープンサイエンスパラダイムに向けた変革においては、アカデミアの変革も踏まえ、情報科学を駆使しながら、それを前面に出さずに社会変革を行うデジタルナイチンゲールとも呼べる研究者人材が増える可能性がある。例えば、文理融合を前提とした脳科学研究、ビッグデータを活用した計算機社会科学、あるいは、多様なステークホルダーを巻き込む超学際的（Transdisciplinary）環境学など、オープンサイエンスの特性を生かした個別の取り組みの中から、デジタルナイチンゲールが現れる可能性を期待しているが、研究データのマネジメントに関するスキルが、その属性の一つになることを期待している。そして、そのスキルセットの開発や実際に研究データをマネジメントするインフラづくりにおいて、今回立ち上がった研究データマネジメント（RDM）研究グループが主要な役割を果たし、科学および社会の変革を加速することを期待してやまない。

5. おわりに

研究データをマネジメントすることは、研究を効率化し、科学や社会を発展させ、貢献者をより健全に認めやすくするために行うものであって、合目的にあるいは近視眼的にデータ管理だけをするものであってはならない。今回の報告は研究データマネジメントに直接関係しないトピックや論点も多く含まれているが、研究データに関わる周辺トピックから関係者の思考を刺激することや、あるいは、論点を包括的に整理する一助になれば幸いである。

参考文献

- [1] 林 和弘: オープンサイエンス政策の狙いと情報科学への期待. 情報処理. Vol.60, No.5, pp.399-406 (2019). https://doi.org/10.5363/tits.23.11_12, (参照 2019-08-19).
- [2] 林 和弘: オープンサイエンスの進展とシチズンサイエンスから共創型研究への発展. 学術の動向. Vol.23, No.11, pp.12-29 (2018). https://doi.org/10.5363/tits.23.11_12, (参照 2019-08-19).
- [3] 林 和弘: 統合イノベーション戦略におけるオープンサイエンス—研究データの戦略的開放による「知の源泉」を担う基盤づくりに向けて—. STI Horizon. Vol.4, No.3, pp. 42-47 (2018). <http://doi.org/10.15108/stih.00145>, (参照 2019-08-19).
- [4] 林 和弘: オープンサイエンスが目指すもの: 出版・共有プラットフォームから研究プラットフォームへ. 情報管理. vol.58, No.10, pp.737-744 (2016). <http://doi.org/10.1241/johokanri.58.737>, (参照 2019-08-19).
- [5] Nozomu Yachie, Robotic Biology Consortium & Tohru Natsume, “Robotic crowd biology with Maholo LabDroids,” Nature Biotechnology, Vol. 35, pp.310-312 (2017). <https://www.nature.com/articles/nbt.3758>, (参照 2019-08-19).
- [6] Sarah Perez, “Google's new 'Assignments' software for teachers helps catch plagiarism,” Tech Crunch, <https://techcrunch.com/2019/08/14/googles-new-assignments-software-for-teachers-helps-catch-plagiarism/>, (参照 2019-08-19).
- [7] pubICIF, <https://www.iucr.org/resources/cif/software/publCIF>, (参照 2019-08-19).
- [8] 丸山健夫 (2008). ナイチンゲールは統計学者だった！—統計の人物と歴史の物語— 日科技連出版社