

報告

オープンデータに関する権利と義務
—本格的なデータジャーナルに向けて—



平成26年（2014年）9月30日

日本学術会議

情報学委員会

国際サイエンスデータ分科会

この提言は、日本学術会議情報学委員会国際サイエンスデータ分科会が、CODATA 小委員会、WDS 小委員会の審議結果を踏まえ取りまとめ公表するものである。

国際サイエンスデータ分科会

- 委員長 岩田 修一 (連携会員) 事業構想大学院大学 教授
- 副委員長 五條堀 孝 (連携会員) 大学共同利用機関法人 情報・システム研究機構
国立遺伝学研究所 特任教授
King Abdullah University of Science and Technology
Distinguished Professor
- 幹事 中西 友子 (連携会員) 東京大学 大学院農学生命科学研究科 教授
- 北川 源四郎 (第三部会員) 大学共同利用機関法人 情報・システム研究機構 機構長
- 安達 淳 (連携会員) 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 教授
- 植田 憲一 (連携会員) 電気通信大学レーザー新世代研究センター 特任教授
- 加藤 幾芳 (連携会員) 北海道大学 名誉教授
- 佃 栄吉 (連携会員) 独立行政法人 産業技術総合研究所 理事
- 土居 範久 (連携会員) 慶應義塾大学 名誉教授
- 濱口 宏夫 (連携会員) 台湾国立交通大學 教授、東京大学 名誉教授
- 村岡 洋一 (連携会員) 早稲田大学 理工学術院 教授
- 長島 昭 (特任連携会員) 中部大学 中部高等学術研究所 客員教授

CODATA 小委員会

- 委員長 五條堀 孝 (連携会員) 大学共同利用機関法人 情報・システム研究機構
国立遺伝学研究所 特任教授、
King Abdullah University of Science and Technology
Distinguished Professor
- 安達 淳 (連携会員) 大学共同利用機関法人 情報・システム研究機構
国立情報学研究所 教授
- 岩田 修一 (連携会員) 事業構想大学院大学 教授
- 植田 憲一 (連携会員) 電気通信大学レーザー新世代研究センター 特任教授
- 加藤 幾芳 (連携会員) 北海道大学 名誉教授
- 佃 栄吉 (連携会員) 独立行政法人 産業技術総合研究所 理事
- 土居 範久 (連携会員) 慶應義塾大学 名誉教授
- 中西 友子 (連携会員) 東京大学 大学院農学生命科学研究科 教授
- 濱口 宏夫 (連携会員) 台湾国立交通大學 教授、東京大学 名誉教授
- 福井 弘道 (連携会員) 中部大学教授・中部高等学術研究所副所長・国際GISセンター長

村岡 洋一 (連携会員)	早稲田大学 理工学術院 教授
大石 雅寿 (特任連携会員)	大学共同利用機関法人 自然科学研究機構 国立天文台 天文データセンター長・准教授
長島 昭 (特任連携会員)	中部大学 中部高等学術研究所 客員教授
渡邊 堯 (特任連携会員)	茨城大学 名誉教授 名古屋大学 太陽地球環境研究所 客員教授 独立行政法人 情報通信研究機構 招聘専門員
芦野 俊宏	東洋大学 国際地域学部 教授
荒木 徹	京都大学 名誉教授
大武 美保子	千葉大学 大学院工学研究科 准教授
国沢 隆	東京理科大学 理工学部 教授
小柳 義夫	神戸大学 計算科学教育センター 特命教授
佐藤 正樹	独立行政法人 科学技術振興機構 情報企画部 調査役
馬場 哲也	独立行政法人 産業技術総合研究所 計測標準研究部門 招聘研究員
村田 健史	独立行政法人 情報通信研究機構 統合データシステム研究開発室 統括
鎗目 雅	東京大学 公共政策大学院・特任准教授

WDS 小委員会

委員長	渡邊 堯 (特任連携会員)	茨城大学 名誉教授 名古屋大学 太陽地球環境研究所 客員教授 独立行政法人 情報通信研究機構 招聘専門員
副委員長	岩田 修一 (連携会員)	事業構想大学院大学 教授
幹事	家森 俊彦	京都大学 大学院理学研究科 教授
幹事	村山 泰啓	独立行政法人 情報通信研究機構 統合データシステム研究開発室 室長
	柴崎 亮介 (連携会員)	東京大学 空間情報科学研究センター 教授
	春山 成子 (連携会員)	三重大学 生物資源学部 教授
	大石 雅寿 (特任連携会員)	大学共同利用機関法人 自然科学研究機構 国立天文台 天文データセンター長・准教授
	荻野 瀧樹	名古屋大学 太陽地球環境研究所 名誉教授
	門倉 昭	大学共同利用機関法人 情報・システム研究機構 国立極地研究所 教授
	金尾 政紀	大学共同利用機関法人 情報・システム研究機構 国立極地研究所 准教授

北本 朝展	大学共同利用機関法人 情報・システム研究機構 国立情報学 研究所コンテンツ科学研究系准教授
楠 勝浩	海上保安庁 日本海洋データセンター長
国沢 隆	東京理科大学 理工学部 教授
桜井 隆	大学共同利用機関法人 自然科学研究機構 国立天文台 教授
篠原 育	独立行政法人 宇宙航空研究開発機構 宇宙科学研究所 准教授
柴崎 清登	大学共同利用機関法人 自然科学研究機構 国立天文台 教授
坪井 誠司	独立行政法人 海洋研究開発機構 地球情報基盤センター 地球情報技術部長
中島 英彰	独立行政法人 国立環境研究所 地球環境データベース 推進室 室長
平原 聖文	名古屋大学 太陽地球環境研究所 教授
牧野 雅彦	独立行政法人 産業技術総合研究所 地質情報研究部門 地球物理研究グループ長
村田 健史	独立行政法人 情報通信研究機構 統合データシステム研究開発室 統括
谷田貝 亜紀代	名古屋大学 太陽地球環境研究所 特任准教授
鎗目 雅	東京大学公共政策大学院・特任准教授
亘 慎一	独立行政法人 情報通信研究機構 電磁波計測研究所 宇宙環境インフォマティクス研究室 研究マネージャー

本件の作成に当たっては、以下の職員が事務を担当した。

盛田 謙二	参事官(審議第二担当)
齋田 豊	参事官(審議第二担当)付参事官補佐 (平成26年8月まで)
松宮 志麻	参事官(審議第二担当)付参事官補佐 (平成26年8月より)
冲山 清観	参事官(審議第二担当)付審議専門職 (平成26年6月まで)
加藤 美峰	参事官(審議第二担当)付審議専門職付 (平成26年5月より)

要 旨

1 作成の背景

近年、情報通信技術（ICT）の発達と普及に支えられて、科学技術論文に関してはオリジナル論文の電子ジャーナルへの掲載と電子化された情報へのオープンアクセスが新たな潮流を形成しつつある。一方、科学技術データに関しては、流通するデータの量は飛躍的に増加し、多様化するとともに、迅速な処理が要求されている。その結果、データに関連する研究活動は大きく拡大し、先端的な学術分野からだけでなく、経済、社会、政治等の伝統的な諸学術分野からも大きな期待を集めるようになった。“ビッグデータ”という言葉と期待が先行し、データの時代への予兆を告げる論説が氾濫している。しかしながら、普遍性と正確性を持った科学技術データの多くは、活用されないままに棄却されてきたという歴史と現実がある。

本報告では、科学技術データの生産、蓄積、共有と活用に関する活動を、原則論と実践論の両面から見直した。原則論に関してはオープンアクセス、オープンデータに関する権利と義務について検討し、実践論に関しては、従来からのデータベース専門家を中心にした大規模なデータベースやアーカイブの構築の活動ではなく、データの生産者、データ専門家、そしてデータの利用者が連携して科学技術データを構築するための新たな方法としてのデータジャーナルを検討した。

2 現状及び問題点

これまでの科学技術データの共有や活用においては、標準化されたデータモデルが存在し、データベースを組織的に構築し易い分野に焦点が当てられてきた。具体的には、基礎科学技術分野がそれであり、基礎物理定数、核データ、原子分子データ、スペクトル化学、物質・材料、バイオデータ等々のデータベース構築されてきた。

一方、防災、安全、エネルギー・経済、環境・気候変動、健康リスク等といったデータモデルが容易に標準化できない分野についても、それぞれの分野への社会的な関心の高さから、基礎科学技術分野以上にデータの共有や活用についての議論が盛んになっている。こうした分野の科学技術データは、データの種類、規模、精度、網羅性等が多様で、データの発生も散発的で、内容的にも断片的であることも多いため、基礎科学分野のデータベース構築のように組織的かつ集中的に行うという方法がとり難い。そこでは、複数の学術分野の連携も必要であり、分野や対象によってデータの持つ正確性や不確実性が異なり、表現方法も異なる。かつ、種々の学説によってデータの解釈が異なるといった、多様性を的確に包み込むことも必要となる。このためデータの整備は容易でなく、客観性の高い科学技術データによって事実を把握するという理想と、実際に提供できるデータとの間に依然として大きな隔りがある。

こうした隔りを狭める一つの方策は、質の高い科学技術データを棄却しないで継続的に収集・蓄積するための制度設計とその体制の構築である。世界各国で行われている多分野でのデータ活動の成果を、データ生産者が連携することによって出版するデータジャーナルが発刊されるようになった。これは、データを継続的に収集・蓄積するための有効な手段である。また分野

別に開始されたデータジャーナルの次の展開としては、分野別のデータジャーナルを統合することによるデジタルデータを通じた学術分野の連携も期待される。その意味で、個別の学術分野におけるデータジャーナルの拡充とともに、分野を超えたデータの活用への新たな道を拓くデータジャーナルの検討が必要となる。

3 報告の内容

(1) オープンデータの目標

電子化された学術情報を、オープンアクセスによって無料で入手する方法と、単に情報にアクセスするだけではなくデータの意味を機械判読することまで可能にするオープンデータによる方法とを検討し、学術分野内での利用だけではなく、学術による情報蓄積を広く社会に利用してもらうための開かれた手段へとオープンデータを発展させることによって、学術が社会に貢献し得る新たな道が拓けることを示した。

(2) データジャーナル

オリジナル論文の発表を中心とした分野別の学術研究のこれまでの成果発表の方法に加えて、データ生産者が分野を超えて連携して、オリジナル論文に埋め込んだデータや論文投稿時に棄却した高品質のデータを学術の成果として集積するための新たな場としてデータジャーナルを検討した。

(3) データに関する権利

データを公共財としてとらえる考え方に基づいた制度と、データの生産者や提供者が果たした貢献を適正に評価する制度を構築することが必要であることを示した。

(4) データに関する義務

データの共有と活用を促進するために専門家が行うべきことを明らかにした。

(5) データによる分野連携

客観性、普遍性に基づく説得力のある科学技術データの活用をベースにした学術分野の融合について検討した。

以上の検討をもとに、第 22 期の国際サイエンスデータ分科会の活動のまとめとして、以下のような今後の議論への問題提起と方向付けを行い、報告とした。

- ① 学術研究の成果を事実と解釈に分けて、前者をデータジャーナル、後者を論文誌として分けて、公開することの妥当性。
- ② 誰もがデータジャーナルに自由にアクセスし、解釈し、活用する権利を持つという考えの妥当性。
- ③ データの生産者はデータを提供するだけでなく、データの多面的活用を促進するための方法を提示する義務があるという考えの妥当性。
- ④ データジャーナルを発刊し、既往の情報資源との相互運用を可能としながら、これまでのデータ活動を補完し、データの生産からそれを活用して得られる価値の創出にいたる循環を促すことができる事例を蓄積することの妥当性。
- ⑤ データジャーナルの発刊を通じて、価値中立と価値創出とを両立させる新たな学術研究成果の公開制度を確立するべきという考えの妥当性。

目 次

1.	はじめに	1
2.	現状及び問題点	3
3.	報告の内容	5
	(1) オープンデータの目標	6
	(2) データジャーナル	7
	(3) データに関する権利	9
	(4) データに関する義務	10
	(5) データによる分野連携	11
4.	まとめ	11
	<用語の説明>	14
	<参考文献>	16
	<審議経過>	16

1. はじめに

近年の科学技術データに関わる活動は大きな変革の渦中にある。書誌学を起点とした文献データベースは、文献カード、文献データベース、電子ジャーナルへと発展し、さらには強力なブラウザを用いることによって、クラウド上に仮想的な情報の共有地、“情報コモンズ”を作るという形をとるようになってきている。一方、数値データに関しては、定形のデータカードやデータファイルによる記録に始まり、電子化されたハンドブック、数値データベース、さらにはネットワーク上で統合可能な分散型データベースの構築へと発展してきた。複数のデータ源を統合する仕組みとしては専門用語や背景となる概念の電子化や構造化が行われてきた。こうした仕組みに関する研究開発には、大別して、索引、シソーラス、デスクリプター、メタデータ、セマンティックウェブ、オントロジー等といった専門知識に関するデータの定義・分類に関わるものと、統計解析をはじめとしてニューラルネット、ファジー論理、自己組織化マップ、ラフセット等々といったデータの持つ意味の発見を通じた知識獲得に関わるものがある。こうした研究開発の成果はネットワーク上で利用可能なツールや知的インターフェイスとして提供され、科学技術データを共有し、活用する範囲を大きく拡大してきた。後述するオープンアクセス、オープンデータ、そしてデータジャーナルをめぐる議論はこうした大きな変化の文脈の中で理解する必要がある。

日本学術会議では、情報学研究連絡会を中心に 1980 年代から情報学シンポジウムを主催するとともに、国内外の科学技術関連のデータ活動について継続的に検討を進めてきた。1990 年代に入ってから、特に、バイオサイエンス分野で行われたデータの知財権をめぐる国際的な激論を踏まえて、オープンデータへの道を拓く報告「データベースに関して新たに提案されている知的所有権について」と提言「データベースに関して提案されている独自の権利 (*sui generis right*) について」を取りまとめ、それぞれ平成 10 年 3 月 31 日、平成 13 年 10 月 17 日に公表した (1)、(2)。

さらに第 20 期に発足した国際サイエンスデータ分科会では、情報学研究連絡会における科学技術データ関連の活動を引き継ぎ、科学技術データは人類全体で共有すべき公共財であるという理念と科学技術データの活用を通じた社会貢献のあり方を検討した。また科学技術データに関わるオープンアクセスやオープンデータの議論や、さらに国連情報社会サミット WSIS (World Summit on Information Society) の機会を通してグローバルな視点での情報コモンズ GICSI (Global Information Commons for Science Initiatives) 実現への国際的な議論に参加してきた (3)、(4)。

第 22 期からは、ICSU (International Council for Science) の常置委員会である CODATA (Committee on Data for Science and Technology) と WDS (World Data System) との密接な連携をはかるために、国際サイエンスデータ分科会に小委員会を設置した。この小委員会は、基礎物理定数に代表される普遍性のある基盤データの整備と分野連携を推進してきた CODATA と、世界各国に分散するデータセンターの連携を目的にした WDS の活動を国内においても連携させる役割を果たしたのである (5)、(6)。また、国際的には、データの公開、共有、活用、知財権の保護等の問題に関して、議論が継続さ

れている（例えば（7）～（10））ことを受けて、小委員会でも、単にオープンアクセスによってデータ、情報へのアクセスを無料化するという理想論だけではなく、理想を実現するための財源問題やビジネスモデル、法制度に関する検討も行き、国際的な連携を視野に入れた具体的なアクションも提案してきた。

こうした作業が開始されて間もない時期に、東日本大地震が発生した。地震・津波、そして福島第一原子力発電所事故は、膨大で多様な一次データを生み出したが、同時に、データの流通に関する多くの課題が明らかとなった。事実に対応した一次データとバイアスのかかった加工データとが、さまざまな形態で、各種メディアやインターネットを通じて流された結果、信頼度の異なるデータと情報が飛び交うことになり、社会に混乱ともいえる影響を与えることになった。大規模災害が起こってみると、データが適正に共有され、十分に活用される体制が整っているとはいえず、データに基づいた被害の予測も対策も十分には行うことができる体制にはなかったことが浮き彫りになった。

分野横断的な活動としては、社会・産業的活用に目標を絞って過去 10 余年にわたって行われてきた知的基盤整備計画の実績があるものの、多くの場合、専門分野の枠組みの中で閉じたデータベースの構築とその活用に留まるプロトタイプレベルの事例に留まってきた。データベースの活用も単なる検索や可視化までで、科学技術データが含んでいる学術的コンテンツの本格的な活用事例は極めて限定的であったといわざるを得ない。

本分科会では、国内外の科学技術データ活動の事例調査をもとに、一次データの有効な活用、付加価値創出、適切な公開ポリシーのあり方、データベースの評価体制、データ構造の標準化、メタデータの提供等の問題を系統的に検討してきた。また科学技術データを介した異分野の融合という極めて困難な課題についても、データ科学が果たし得る役割を明らかにしてきた。こうした検討作業の過程でデータが社会的に注目されるようになり、多くの学術分野から新たなデータ活動が提案され、国際的な活動への積極的な参加を企図するものも現れるようになった。このような“データの時代”を表す種々の動きを受けて、新たな視点でオープンアクセス、オープンデータについて検討することが必要になった。

データジャーナルは、既往の数値データベースやオリジナル論文に含まれない科学技術データを対象とし、テキスト主体、オリジナル論文主体の情報基盤と緊密に連携したジャーナルとして新たに登場した。そこでは、データ活動の範囲を特定の学術分野ではなく、全ての学術分野とすることが前提とされている。また、データ活動をデータベースの構築に限定するのではなく日常的な学術活動の一環として位置付け、複数の学術分野が連携してデータ活動を行うことが想定されている。その意味でデータジャーナルは学術諸分野の連携に資する本格的な電子的学術基盤である。この基盤構築のためには、情報科学の専門家に加えて、生命科学や材料科学などといったそれぞれの専門分野に精通したデータ活用の専門家が協働することが必要である。またオリジナル論文の掲載を中心とする学術誌が個人の成果を発表する場であるのに対して、公共的な性格を持つデータを発表する場であるデータジャーナルは社会性に富んだ学術基盤という性格を持つ。

観測方法や測定方法の標準化が進んだ分野では大量の均質なデータが生産されるが、一方で、オリジナル論文の作成に直接使われなければ、多くの一次情報データが棄てられる

ことになる。データに関する活動領域は大きく拡大し、先端的な学術分野だけでなく、例えば、経済、社会、政治を対象とした伝統的な学術分野等にも及ぶようになった。“ビッグデータ”という言葉と期待が先行し、データの時代への予兆を告げる論説が氾濫しているが、普遍性と正確性を備えて作られてきた科学技術データの多くは活用されないまま棄てられているという現状にある。

以上の考察を踏まえ、本報告では、広汎な科学技術データの生産、蓄積、共有と活用を原則論と実践論の両面から見直すことにした。オープンアクセス、オープンデータに関する権利と義務を検討し、従来のデータベース構築の専門家だけではなく、データの生産者、研究者、専門家、そしてデータの利用者が連携して公共財としての科学技術データの本格的な整備を図るための実践という観点からデータジャーナルの可能性を検討した。

2. 現状及び問題点

これまでの科学技術データの共有と活用に関する議論は、標準化されたデータモデルが存在するためにデータベースを組織的に構築することが容易な分野を対象としてきた。基礎科学技術分野がもっぱらこれに該当し、基礎物理定数、核データ、原子分子データ、スペクトル化学、物質・材料、バイオデータ等が具体例である。

科学技術データの本質的な価値である普遍性を確保するためには、一次データの発生から目的に応じたデータ・情報としての活用に至るまで、データの編集過程を論理的に追跡可能な形式で記述することが必要である。この編集過程を付加価値生成のデータライフサイクル、略してデータのライフサイクルと言うことにし、ライフサイクルが追跡可能な状態をトレーサビリティが確立していると定義する。このトレーサビリティのあるデータを一定の目的で集積させることは、特定の学術分野をデータを通して表現することを意味し、その学術分野の対象領域を論理に基づいてデジタル処理する可能性を与える。学術分野相互間がデータを通して連携するには、データのライフサイクルのトレーサビリティを確立するという準備が必要であり、そのことによってデータの信頼性とデータへのアクセシビリティが高まり、デジタル処理による解釈や活用の多様性が生まれる。

一方、防災、安全、エネルギー・経済、環境・気候変動、健康リスク等々のデータモデルの標準化が容易でない分野についても、それぞれの分野への社会的な関心が高いことを背景に、基礎科学技術分野以上にデータの共有と活用についての議論が盛んになっている。こうした分野の科学技術データは、データの種類、規模、精度、網羅性等々が多様で、データの発生も散発的で、内容的にも断片的であることが多いため、組織的かつ集中的にデータベースを構築することは容易でない。学術分野の連携にあっても、分野や対象によるデータの正確性や表現方法の違いや、学術分野によるデータの解釈の違いを調整することも必要となる。このためデータベースの整備は容易でなく、客観性の高い科学技術データによって事実を把握しようとするという理想論と実現可能なデータサービスの間に依然として大きな隔りがある。

こうした隔りを狭める一つの方策は、小規模ではあっても質の高い科学技術データを棄却しないで段階的かつ継続的に収集・蓄積するための制度の設計と体制の構築である。

世界各国で行われている分野毎のデータ活動の成果を、データの生産者の連携に基づいて出版するデータジャーナルが発刊され始めた。これは、データを段階的かつ継続的に収集・蓄積するための有効な手段となる。また分野別データジャーナルの次の展開としては、これらを統合し、電子化されたデータ（以下、デジタルデータ）を公表することによって、学術分野の連携を図ることも期待できる。その意味で、個別の学術分野の拡充とともに、分野を超えたデータの活用への新たな道を拓くデータジャーナルについての検討が必要となる。

データのライフサイクルの起点となる一次データは国内各研究機関・研究者の活動によって日々生み出されている。現在、一次データのデータベースとしての整備・公開には個々の研究機関や研究者が携わっているが、そうした努力を適切に評価する仕組みは確立されていない。それでも一次データの蓄積作業においては、我が国においても研究者の共通基盤が整備されつつある。具体的には、データの共同利用体制の確立のために必要となる、データ公開ポリシーの策定、各分野で開発されたデータのリンク、データの評価、評価・改訂履歴、インターネット上でのアクセスのためのデータへの URI、DOI といった識別子の付与などである。

しかし、オリジナル論文や技術レポート類はキーワードでしか検索できず、多くのデータは文献データベースとは別の数値データベースとして整備されている。分散したデータベースにアクセスするためのメタデータの整備も十分ではない。利用可能なデータベース群を組織的に活用する手段も限られている。識別子を介して漸くたどりついたデータベースもトレーサビリティが不十分でデータの意味が明快でないことが多い。データの意味をデジタル的に処理することも容易ではない。したがって複数の学術分野にまたがる科学技術データを整備するためには、これまでのデータベース構築の手法では限界があり、新しい方法が必要であると考えられてきた。

新しい方法はオープンアクセスの論文誌の編集作業の過程で生まれた。科学技術データを必要なだけ論文中に掲載することができないこと、データだけの投稿を受け付ける場が無いことが問題となるようになった。質の高い科学技術データを棄却しないで収集・蓄積するための制度の設計と体制の構築が必要であることも議論されてきた。こうした議論は、これまでの特定の分野を対象とした大規模データベース構築の際には、あまり問題にならなかった小規模で多数のデータベースの構築という問題に取り組むことに繋がった。

データが必要となる場面は多様である。災害時や事故時においては、想定外の事象が連続することが多く、専門分化した学術分野だけでは対処できない。新たな領域に挑戦する場合も同様で、入手可能なデータには必然的に欠損や時間的な制限が付きまとう。また、どのようなデータが欠損するか、どのようなデータが得られるかを事前に想定することは困難である。時々刻々得られるデータを基に、不足している情報を適切な仮定や手法で補い、適切なアクションにつなげることが大切である。インターネット、モバイルデバイスなどの普及によってデータを測定し、共有するためのインフラは格段に整っている。断片的であれ、必要なデータが入手できれば、減災に役立つ適切なアクションの決定や、新たな価値の創出の可能性が高まる。

データ活動の理想的な状態は、事象の発生と同時に、事象にかかわる科学技術データを、メタ

データを付加してクラウドにアップロードすることである。アップロードに際しては、オープンにアクセス可能な状態に整備されていることが望ましい。これが、誰にでも利用可能なようにデータが集積された共有地:データコモンズの形成に当たる。さらに、データコモンズでは、利用者からの要請に応じて事象に関連するデータが集約・加工され、デジタルデータとして、適切な精度で、適切な時間内に提供されるようになっていくことが期待される。そうした社会的要請に応えるためには関連学術分野の総力を結集したデジタル化作業が必要である。多くの学術分野には、それぞれの分野に特有の深化と進化の歴史があり、それぞれ固有のデータベースを形成してきているので、各学問分野のデータがもつ意味の論理的な記述と、利用者がそれらを正しく活用するための支援を行うことが重要である。

原子炉事故、環境問題、健康リスク等々の社会的影響の大きい事象では、見解の相違を互いに理解しながら対処していくことが求められている。そのためには事実関係を客観性の高いデータで確認しながら、異分野の専門家が協働して問題解決のための方策を作り上げていくことが求められる。

3. 報告の内容

電子化された学術情報の入手可能性とは、アクセスを無料にするオープンアクセスばかりではなく、アクセスだけでなくデータの意味の機械判読まで含んだオープンデータをも意味する。災害や事故のリスク管理の場面を想定しながら科学技術データの公開について考えてみよう。大規模災害時の対策を検討するためには数多くの種類のデータを総合的に扱う必要があることはいうまでもない。このためには、インターネットを始めとする情報技術を駆使することが不可欠である。東日本大震災の原子炉事故時に手書きの数値をスキャンした画像データが公開されて、その扱い難さが問題となったが、データを電子的に再利用可能な形で公開するといっても多様性がある。画像データで公表する場合でも、圧縮の形式、画像の意味についてのデスクリプターの付け方、数理的な等価性にも違いがある。破損した機器のスケッチを公表する場合、設計データから製造時データ、メンテナンスデータまで電子的に関係付けて入手可能とすればデータの意味はかなり明らかになるが、知財権、情報セキュリティ、データリテラシーの壁が存在する。何処までの公開が求められるかは目的や状況に依存し、発信者と受信者のデータリテラシーの違いによって実現方法も異なる。受信者が受信したデータを加工して発信することが混乱を招くこともあるので、信頼性の高いデータとその解釈を広く周知することが重要となる。風評被害が問題になっているが、もし、信頼できるデータが提供され、それらが多くの専門家によって多角的に分析されて、真実に近い情報が卓越することになれば、風評は最終的に収まっていくが、それには時間がかかる。その時間を短縮するためにも、流通するデータのトレーサビリティを維持する不断の作業が必要となる。

科学技術データを電子的に公開する方法に関するルールは未確定であるといっている。現状では、Web上でデータを共有することについて、行政データなどを中心に Semantic Web 技術をベースとした LOD (Linked Open Data) によるデータ表現が開発されているが、科学技術データを標準化するには至っていない。

デジタルデータを公開するにあたっては、それらはすべて異なる背景を持ち、異なる品質・粒度を持っている。一次データが時宜を得て提供されれば、専門家が、たとえ被災地から遠く離れた場所においても、データの信頼性についての評価を行うことができる。しかし、福島第一原子力発電所の事故においては、電源の喪失により炉内の状況を知ることができなくなった。このように、大災害においてはどのようなデータが欠損するかを予め想定することは困難である。放射線量のモニタリングのデータの公開においては、計測方法や測定機器の較正方法の不明なデータが混在してしまった。測定値と計算値も混在することになった。緊急時には、限られたデータから推算することが求められる。専門家が参加して、公開すべきデータの範囲や、デジタル化のレベルなどを決定し、その際の判断の根拠を明示しなければならない。知財権、情報セキュリティ、データリテラシー等の問題を迅速に調整して、適正なデータサービスを提供することも大切である。いずれにしてもデータのトレーサビリティを確保する必要がある。

データの品質を高めるためには、データの生産者側が品質管理に努めるとともに、利用者からのフィードバックが重要である。前者に関しては、定義に基づくデータの意味、データ間の関係に基づく電子的な異常値検出ツール、シミュレーションツールが有効である。一次データに、計算モデル、評価システムを適用する際には、結果のみを公開するのではなく、その計算の前提となったデータ、計算モデル、プログラムソースコード、コードの内容・意味、精度、性能、限界なども一つのセットとして開示されなくてはならない。後者の利用者からのフィードバックに関しては、データの引用情報、因果関係の抽出と検証（Data Mining & Knowledge Discovery）ツール、錯綜した議論の可視化と議論のフレーム評価ツールが活用できる。共通のプラットフォーム上でデータが提供されれば、多数の利用者の眼を通して品質が向上するような“情報エコロジー”とでも呼ぶべき生きた情報環境の創出が可能になる。以下、項目毎に検討事項を要約する。

（１）オープンデータの目標

科学者コミュニティにおいては、学術研究を推進するための一次データの重要性、それを使った分析結果等を社会へ還元することの重要性は広く認識されている。日本学術会議のマスタープラン 2011 においても各分野におけるデータベースの開発、分散した個別のデータベースの統合化などデータに関わるインフラストラクチャの構築に関連するものが多く提案されている。また、学術データを公共財として共有し、活用する上での格差の解消に関する提言が国連情報社会サミットの原則宣言にも反映されている。オープンデータの目標は、個別の学術分野の成果を提示することにあるだけではなく、社会との連携を通して学術の新生を促すことにあると考えられている（3）。データのデジタル化による変化を再確認するために、スピードを要求される事故時の教訓を示そう。

- ① データの信頼性の確保と品質保証へ向けた集合知の活用：デジタルデータはコピーと転送が容易であるというデジタル社会の特徴を踏まえて、一次データに対する専門家のアノテーションの付与、信頼性の確認などを通じて正確なデータを流布に努める必要がある。また、信頼に足る情報を公開し、保全するデジタルデータ拠点を確保する

必要がある。

- ② データの公開と緊急時においてデータをリアルタイムで活用できるようにすること：データは電子的に処理可能な形式で公開されるべきである。大災害発生時にはリアルタイムで大量のデータが発生するため、事前に分野横断的に標準化されたメタデータを設けておく必要がある。ただし、当然ながら標準化の及ばないデータも発生する可能性があるため、これらを随時取り込むことが可能なデータ生産者参加型の仕組みを用意しなくてはならない。
- ③ データの活用に関するシナリオ、詳細なシミュレーションとその計算過程の公開：データの取得・編集・解釈や活用方法は多様である。同じ事象に関するデジタルデータでも観察方法、実験方法、計算方法によって意味が異なる。特に、緊急時における計算データについては、インプットされた観測データと実験データに加えて計算過程を公開し、データに基づいて行われるリアルタイムのシミュレーションの全体像を明らかにする必要がある。災害発生時には多くの欠損データが生じることは稀ではなく、これらを補うためのデータの補完や暫定的なモデルの活用が必要となることもある。メタデータから計算モデルの仕様までデジタル化を要する項目は多い。
- ④ 第三者視点による総合的データ分析の活用：大規模災害が発生した際には、現場から離れた専門家の方がリソースを自由に活用することができ、包括的な見方をすることができる場合がある。異なる専門家が行う分析が矛盾する結果をもたらす場合もあるが、これらを公開し、調整するための場が必要である。

(2) データジャーナル

科学データをデータベース化するには、データを熟知した研究者が、時間をかけて品質チェックを行うことが必要である一方、大量で複雑なデータのデータベース化には時間や労力を要し、研究者が研究の合間に出来るものではない。結論的に、情報専門家の協力が必要であるが、日本にはそういう専門家のためのポストが極めて少ない。そのために、研究者に過重な負担がかかっている。ところが、研究者がデータベースを作成しても、業績として評価されにくいという問題がある。こうした事情が重なって我が国のデータベースの構築は停滞した状態が続いていた。

そうした状況の中で、Nature は、事前の膨大な議論（例えば（7）、（8）、（9））を踏まえて、データセットに光を当て、その再利用の促進を目指すオープンアクセスジャーナル、Scientific Data を 2014 年 5 月に創刊した。グローバルなデータベース活動に大きな影響を与えることが想定される。このことに関連して、我が国のデータ活動の今後の展開を考えてみたい。

データセットが広く再利用されるためには、上述したように、データのトレーサビリティを維持したデータを保管しておく必要がある。研究データがどのように生成され、また、どのようにデータの品質管理が行われたかが分かることが大切である。しかしながら品質管理のために詳しい情報：メタデータを記述する作業は大変な労力を必要とする。また、メタデータの記述作業に対する評価は必ずしも十分とは言えないため、メタデータの記述

は不完全なものとなることが多い。

メタデータの重要性はデータベース管理システムが開発された初期から議論され、計算機ネットワークによって実現した分散型のデータベースの開発／普及によって、さらに広く関係者の間で認識されるようになってきている。しかし、情報システムの発展／普及、学術コンテンツそのものの先端化と進歩、メタデータ作成作業の生産性等々が爬行的に進んできた上、知財権をめぐるコンフリクトも加わって、貴重なデータセットとなり得る情報（コンテンツ、メタデータ、データ活用コンテキスト）が公表されない、あるいは、部分的にしか公開されない、また、公開されていても詳細な記述がないために再利用できないという状態が続いていた。

2013年のG8会合でデータ公開に関する勧告も出されたこともあり、オリジナル論文との連携が容易という長所を持つネイチャー・パブリッシング・グループによるオンライン限定のオープンアクセスジャーナル、Scientific Dataの創刊は学术界に大きなインパクトを与えると考えられる。加えて同誌は、査読付き科学出版物であること、主要なインデックスサービスに索引登録されること、データが再利用されることで著者の功績が相応に認められこと、といった長所もある。同誌は、“Data Descriptor”と呼ぶメタデータを付加することを論文受理の必要条件としている。これは、科学的に貴重なデータセットが共有され、再利用されることを促すよう、データを理解するための情報の提供を求めるという役割を持つものである。データジャーナル誌上でのデータ公表においては、電子化された著作物の分野で国際的に普及しているクリエイティブ・コモンズ・ライセンスが適用されている。また投稿者に“Data Descriptor”の再利用、再配布と改変を認めることによって、投稿者の連携を通じたメタデータの段階的かつ自律的な進化を促進し、メタデータに関する上述したような困難を解決しようとしている。また、最終的なメタデータの編集・改訂作業については、全てを研究者にアウトソーシングするのではなく専任のデータ専門家を雇用し、アノテーションの一貫性と有用性を図るための体制も整えられている。

また、大型科研費など、ある程度以上の予算を使うプロジェクトについては得られたデータを公開する（公開体制も含む）ことをプロジェクト採択の条件とすることを推奨している。占有期間が過ぎたデータは、データ取得者が責任を持って公開する流れができるきっかけを作ることも考慮されており、“Data Descriptor”付きデータセットの投稿には、関連するジャーナル掲載論文とデータリポジトリに保管されたデータファイルとのリンクが設定され、データセットの読者が研究論文やデータに関する記述と実際のデータとを容易に対照できるように工夫されている。すなわち、それぞれの“Data Descriptor”付きデータセットには、機械処理が可能な（コンピューターで読み取れる）メタデータが付いており、オリジナル論文のコンテンツのマイニング（網羅的分析による情報抽出）や検索を行う際に役立つようになっている。

Natureによるデータジャーナルの刊行は、データの共有と活用に対する資金面と制度面の支援を骨子としてマックスプランク研究所が中心になってまとめたオープンデータのガイドラインBerlin Declaration（9）を基礎として、同誌のブランド価値を活用した巧みなビジネスモデルの展開といえる。個々の科学技術分野で生産される数値データは内容

的にも多様で規模も異なるため、従来のデータベース構築体制には組み込むことは容易ではない。一方、データジャーナルは、科学技術データ生産者とデータの専門家との密接な連携を可能とし、中小規模のデータ共有が可能である。データの利用者との連携も可能であることからデータジャーナルはデータの新たな共有地：コモンズとしての機能する可能性を有している。

繰り返しになるが、データベースは、長期間継続して作り続け、そして維持し続ける必要がある。我が国の現状は、科学研究が、短期の競争的資金に頼らざるを得ず、研究で得られたデータが、データベース化されないために第三者には使えない未整備ものがかなりあり、データベースの構築中止や破棄が生じている。また、研究基盤としては重要でも、短期的業績を示しにくいデータセンターやデータ部門が廃止される可能性が大きい。このため日本では、は国際的に評価の高い大規模データセンターが存在していない。一方で、研究、教育、産業、行政、医療など各分野で多くのデータが生み出されており、これらはデータジャーナルを通じて公開するのに適している。

しかし、政府としてのデータポリシーが確立していないため、それらは、有機的に繋がることなくバラバラに作られていて、日本全体としての状況が把握されていない。文科科学省傘下に限っても、全体のデータ活動は十分に把握されていない。つまり海外に比べて、日本のデータ利用体制は貧弱であり、その原因は、国力の差というよりは、国としての明確なデータポリシーが確立されていないことにある。国家としての適切な戦略を策定するためのデータの管理を海外に依存するのか、国際的な競争力のない中途半端なプロジェクトを継続して適切な時機が到来するのを待つのか、あるいは世界を先導しうる新たなデータプロジェクトを設計し、実行に移すのか、厳しい選択を迫られているといえる。データジャーナルの整備は、科学技術データ分野での我が国の可能性や応分の世界貢献が問われる試金石として考えることもできる。

(3) データに関する権利

日本から発信される高品質のデータや知識は、欧米の大手出版社、学術組織や ICT ビジネスが形成しているレジームに吸収され知財化され、必ずしも日本国内へ適切に還元されているとはいえない。長年指摘されてきたこの課題の改善のためには、データに関する権利に関する考察に基づいた制度設計とそのための社会システムの構築が必要である。国内の産官学の主要分野において、科学技術一次データの発生から、データの整備・評価・再編成、情報資源化のためのリンクから活用に至るライフサイクルを事例別に分析することが必要である。ほとんど未編集の一次データをクラウドにアップロードする場合、機関レポジトリに登録する場合、データジャーナルに投稿する場合、さらには特定の開発方針やビジネスモデルで構築するデータベースに登録する場合等、データ公開にはさまざまなスタイルがあり、それぞれデータに関する権利は異なる。科学技術データは誰のものかという根源的な問いかけに答えることは容易ではないが、少なくとも、データのデジタル化を通してデータを迅速かつ戦略的に社会に役立てることが必要であり、社会におけるデータによる価値の創出事例を通してデータへの権利を見直す必要がある (1)、(2)。デー

データの生産者、データサービス担当者、出版社、データの利用者、全体の調整を期待される「データ専門家」等々、データに関わるステークホルダーは多様化し、権利関係も複雑になってきている。議論の方向としては、学術全体の生産性の向上に資することが重要であり（例えば（7））、また、ダイナミックに変化するデータの流通現場の状況を反映させることが重要である。

（4）データに関する義務

観察、実験、計算によって新たに取得したデータに編集を加えずに公開し、データの利用に関しては利用者に任せるスタイルを特徴とするデータの共有と活用が“ビッグデータ”ブームの中で増えてきている。その一方で、特定の学術分野の成果を踏まえたデータの評価と形式に基づいてデータベースを作り、公開するという伝統的なスタイルのデータ活動も継続している。データに関する義務は公開のスタイルあるいはポリシーによって異なるので、データ公開にあたってはデータ提供側とデータ利用側双方のデータ公開のスタイルあるいはポリシーについての理解が一致することが大切である。データに関する義務を定義づけるにはそうした作業が必要である。

国内外の個々の学会では学術データの構築や研究推進が行われているが、統合連携する機能は欠けており、それに向けた研究も行われていない。また、多くの分野で、取得したデータに編集を加えない一次データを提供することへの評価や、データの信頼性を高める作業への評価や、インセンティブ、データ公開ポリシーの問題などを深めることが共通の課題として挙げられる。

今回の東日本大地震・津波・原子炉事故におけるデータの利活用に関する教訓としては、多くの貴重な科学技術データが特定の学術分野内でのみ流通していて、他分野からのアクセスが極めて困難であったことである。この困難を克服することは容易ではないが、科学技術データは公共財である正論を実現するという観点から改善を図ることが必要であることは言うまでもない。重要なことは、研究開発に従事する専門家が、それぞれの社会における役割を再認識し、自らの専門家としての義務を研究開発の成果の公開を通してより積極的に果たすことである。また、データを検証し、信頼でき利用しやすいデータに整備する「データ専門家」の養成も緊急課題である。このことは日本学術会議 18 期学術データ情報専門委員会などからもアピールされているが、正論を繰り返すだけで目標を達成することは困難であるため、実施可能な分野から、データの共有と活用に関する実証実験を開始し、運用経験の蓄積を通して、データ専門家を育成する上でのインセンティブと専門家の義務と使命を明らかにしていくことが必要である。

データの公開においてはトレーサビリティを確保した公開が原則である。理想としてはデータの恣意的な選択や棄却は避け、事実に対応した必要十分なデータ群を提供することが必要である。データについての解釈とデータの記述と独立させることで、データの普遍性を確保するとともに解釈の自由度を担保できる。データの公開を先取権につなげることができれば、オリジナル論文で拙速な解釈を施すことを防ぐことができる。共通のデータを基にさまざまな解釈を試みる自由が担保できれば、議論は活性化し、学術分野の進歩に

貢献する。データに関する義務に関しては多くの議論があるが、実例と実働を通じた原則の改善が必要である。

(5) データによる分野連携

今世紀は、学術と社会との新たな関係を確立し、環境、エネルギー、格差是正等々にかかわる地球規模の課題の解決に挑戦する時代である。軸となるべき新たなスタイルの学術活動が必要で、学術には、個々の分野の展開に関係するデータを創出するだけでなく、多分野の成果をまとめる総合力とグローバルなリーダーシップを発揮することが求められる。IPCC や生物多様性（2）での激しい国際競争で明らかのように、合意形成のための議論の展開の基点となる科学技術データの役割は大きく、客観性、普遍性に基づく説得力のある科学技術データの活用が極めて重要である。

以上の議論をまとめて、以下、データに関する権利と義務についての議論を総括する。

4. まとめ

一次データの生産、そしてその共有と活用は独自の学術や文化を形成するための要件である。我が国は、明治以降、多くの学術分野で先行する西欧の科学技術にキャッチアップするため、データから技術標準まで丸ごと導入し、さまざまな工夫を加えて価値を付加し、最終的には安価で優れた性能をもつ工業製品を製作し、それらを輸出して富を蓄積してきた。

データに関しては最初から大幅な輸入超過が続き、一次データではなく社会的な価値につながりやすい加工データが中心であった。その結果、一次データから加工データまでの付加価値生成の経験が不足し、独自に生産した一次データの提供や一次データに価値を付加した情報や知識を提案することにも積極的にはなれなかった。また我が国は応用特許の数では世界をリードする位置にいるが、競争相手の追随を許さないような質的に高い基本特許の数は極めて限定的であった。工業標準についても基盤となる一次データの整備がなければ提案することもできないため、他国で準備された標準の利用者にならざるを得なかった。一次データの生産やその知的基盤としての整備には人的資源も含めて多大な投資と時間を必要とするが、そうした基盤分野の拡充での努力を惜しみ、結果的にはデータの利用者に徹するという安直な選択をすることが多かった。つまり技術立国のための大事な要件に欠落があったといわざるを得なかった。

科学技術には、利用目的、人、場所等々を選ばない普遍性があり、その普遍性の故に科学技術データの拡散速度は速く、科学技術の進歩を加速させる。一次データに関する非開示や知財権をめぐる過剰な競争は、科学技術の発展を阻害するだけでなく、データを保有する国と保有しない国との間の格差を広げ、不必要な軋轢の原因となる。近年は気候変動やエネルギー資源といった地球規模の課題が顕在化し、これらの関する様々な意見が発信される。信頼性の高い科学技術データを基にした意見の調整が要請されるが、独自の一次データを持たないと提示された加工データの信頼度や不確実性を評価できない。そうした弱点を抱えながら十分な検討もせずオープンアクセスやオープンデータへのグローバル

な動きに追随することは危険であるが、だからといって知財権の保護を教条的に主張することも賢明ではない。

以上の科学技術データに関する課題は、従前の科学技術の成果の公開方法への抜本的な見直しを要請し、我が国の科学技術の進展におけるデータの役割についての歴史的な検討を踏まえた深い議論を要請する。一次データを輸入し、欧米の主導で整備された知的基盤をユーザとして活用し、価値を付加して製品を輸出する国の富を蓄積するという我が国の伝統的な発展モデルから脱皮することも考えなければならない。グローバルな視点で一次データの生産から価値創出にいたるライフサイクルを提案し、主導することも必要である。また環境、放射線、食品リスク、医療サービス、工業製品の安全性等々、科学技術データの信頼性や不確実性に起因する国内の社会的な混乱は、自分達で一次データを評価し、評価結果を積み上げ、自分達で安全基準を定め、自分達で合理的な行動を決定する経験の不足に起因する。こうした争点を超えて前に進むためには、輸入型の科学技術では問題は解決しない。グローバルに展開するオープンアクセス、オープンデータに参加することの意義は、科学技術データから価値を創出するための知力と獲得した価値を世界に広める知性を強化する絶好の機会を得ることにあると捉えるべきである。データや情報は瞬時にコピーされ世界に拡散するが、価値創出のための知力や知性は簡単には複製できない。したがって、クラウド上での共有に適した一次データとも、専門性の高いデータ編集作業が必要な従来型の分野別データベースとも異なり、両者をつなぐ役割が期待されるデータジャーナルの活用方法を発見することに大きな意味がある。

以上の検討をもとに、第22期の国際サイエンスデータ分科会の活動のまとめとして、以下のような今後の議論への問題設定と方向付けを行い、今期の報告とする。

- (1) 学術研究の成果を事実とその解釈に分けて、前者をデータジャーナル、後者を論文誌として公開する。データジャーナルではデータを完璧にデジタル化して掲載し、データをオリジナル論文で引用する時にはデータの出典を明示する。研究評価においては論文誌の内容や引用状況だけでなく、データの引用状況も対象とする。研究費の配分においては、当該プロジェクトで作成したデータのデータジャーナルへの提供を義務付ける。
- (2) 誰もがデータジャーナルに自由にアクセスし、解釈し、活用する権利を持つ。
- (3) データの生産者は、データを提供するだけでなく、投稿データの多面的な活用を促進する義務を負う。
- (4) データジャーナルを発刊し、既往の情報資源との相互運用を可能としながら、データの生産から価値の創出にいたる循環を促す。
- (5) データジャーナルを、価値中立と価値創出とが両立する学術研究成果の新たな発表

制度として位置づける。

<用語の説明>

CODATA: The Committee on Data for Science and Technology

科学と技術に関するデータの管理、信頼性、使用法などを改善することによって社会のために役立つことを目的にして設置された ICSU の常設委員会。活動は多岐に亘るが、特に物理学定数について推薦する値を公表していることで知られている。日本からは日本学術会議が加盟し、1966 年の ICSU の決議からの創設時からのメンバーである。

DOI : Digital Object Identifier の略称。ある対象に付与されるデジタル識別子でありのことである。多くの場合、インターネット上のデジタルコンテンツを示す識別子として用いられており、その本文到達性・持続性・一貫性は国際 DOI 財団によって保証されている。2000 年ごろから学術出版物に付与されはじめ、現在では 4600 社以上の出版社が参加している。ここ数年は、科学データに対して DOI を付与し、データ引用(Data Citation)を可能にしようという議論が活発になり、特にヨーロッパを中心にすでに実動し始めている。

GICSI: Global Information Commons for Science Initiative

2014 年チュニスで開催された第 2 回の国連情報社会サミットで、CODATA 、 International Council for Scientific and Technical Information (ICSTI) 、 International Network for the Availability of Scientific Publications (INASP) 、 World Data Centers (WDC) が中心になり、ICSU、Inter Academy Panel on International Issues (IAP) 、 Academies of Science in Developing Countries (TWAS) 、 Organization for Economic Co-operation and Development (OECD) 、 United Nations Economic, Scientific, and Cultural Organization (UNESCO) と Science Commons との支援のもと提案した行動計画である。公的資金の支援を受けて獲得した科学技術データのオープンアクセスによるデータ共有を提案した。学術そのものの知的生産性向上のための仕組みとしてだけでなく、グローバルな課題の解決や、先進国と開発途上国との情報格差の解消等々を目標にした、国際的なデータ活動の連携への提案である。

ICSU: The International Council for Science

科学が社会のために役立つことを目的にして、国際的共同研究の推進、科学ポリシーについての提言、科学についての格差解消などについて活動している非政府の国際的学術団体。現在、141 ケ国の科学技術団体と 31 の国際学術連合が加盟している。日本からは日本学術会議が日本の科学者を代表して加盟している。

LOD: Linked Open Data

文書が互いにリンクされた現在の WWW に対して、標準的なフォーマットを用いて記述されたデータ同士を互いにリンクさせて計算機で処理することの出来る「データの Web」を作ろうとするのが Linked Data の考え方である。中でもオープンな、誰でも利用できるデータをリンクさせたものを Linked Open Data と呼ぶ。行政データなどを LOD として構造化・

公開することで活用が広がり、新たな付加価値が生まれるものと期待されている。

Open Access :

データ（研究情報を含む）の国際的共有を目指して、国際地球観測年（1957-58）を期に ICSU が設定した理念。研究目的であれば、データの利用に特別な制限は設けず、無償または必要最小限の経費で提供することが原則。多くの国際的な研究事業では、この理念に沿ったデータ利用規約が制定されており、' full and open access ' と呼ばれることが多い。例えば遺伝子配列の特許などの商業化に対して、科学界から自由な研究が阻害されるのではないかという懸念から、ICSU は科学技術データへの ' full and open access ' が必要であるという決議を表明している。これとは別の流れとして最近では、科学雑誌の高騰化に対抗する手段として学術論文をネット経由で誰もが無償で閲覧できる状況を指す用語としても使われている。博物館展示品などを誰もが自由に閲覧できるような状況もオープンアクセスという。

RDA: Research Data Alliance (<http://www.rd-alliance.org/>):

オープンなデータの流通のための社会的・技術的な基盤を構築するためのコンソーシアムとして、EU、米国、オーストラリアの主導により 2013 年 3 月に設置された。インターネット技術の標準化手順を参考に、自由参加・コミュニティベースの活動として組織され、CODATA や WDS とも協調している。2014 年 4 月時点で Data Citation、Metadata Standard Directory など 10 のワーキンググループが活動している。

Sui Generis Right

計算機・通信技術の進歩により、データ集やデータベースが簡単にコピーできるようになり、作成者の権利保護が情報社会の一つの課題になっている。データベースはそのデータの配列や集め方などに創意工夫があれば著作権により法的に保護されている。EU では、著作権とは別のデータベース独自の権利(sui generis right)を定めることにより、データベース作成者・投資者を保護する法律を制定している。

URI: Uniform Resource Identifier

一定の書式によってリソース（資源）を指し示す識別子で、リソースの「場所」を識別する URL (Uniform Resource Locator) とネットワーク内の位置を示してリソースを同定する URN (Uniform Resource Name) から構成される。

WDS : World Data System

ICSU の傘下にある、科学研究に向けたデータセンター等の国際連携組織。品質管理されたデータの長期保全と、特別な制約の無い提供を理念とする。1957 年に設置された「世界資料センター (WDC)」と「天文地球物理恒久事業連盟 (FAGS)」を母体として、2008 年に発足した。(独) 情報通信機構に国際プログラムオフィス (WDS-IPO) が設置されており、約

90ヶ所のデータセンター等が加入している。

WSIS: World Summit on the Information Society

国連世界情報社会サミット。ITU(International Telecommunication Union)が幹事団体になり情報と通信における問題について2003年(ジュネーブ)と2005年(チュニス)で開催された会議で、先進国と開発途上国との情報格差の解消を基軸にした決議案が採択された。ICSUやCODATAはCivil Societyのグループで参加した。

<参考文献>

- (1) 日本学術会議 報告「データベースに関して新たに提案されている知的所有権について」平成10年3月31日
http://www.scj.go.jp/ja/info/kohyo/17htm/17_1.html
- (2) 日本学術会議 提言「データベースに関して提案されている独自の権利 (*sui generis right*) について」,平成13年10月17日。
<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-18-k136.pdf>
- (3) 岩田修一: 学術の動向「科学技術データとオープンアクセス」2005年12月号 pp. 40-43。
- (4) 岩田修一: 「CODATA の活動と展望について Data-centric Science and Society へ」2009年5月号 pp. 79-83。
- (5) WDS に関しては、学術の動向 2012年6月号 特集「科学データの長期保全とグローバルな共有—ICSU世界データシステムの構築—
- (6) 学術の動向 2013年9月号 特集1「データと発見」
- (7) BITS OF POWER, National Academy Press, NATIONAL ACADEMY PRESS
Washington, D. C. 1997
- (8) <http://www.budapestopenaccessinitiative.org/read>
- (9) <http://openaccess.mpg.de/Berlin-Declaration>
- (10) <http://www.icsu.org/general-assembly/news/ICSU%20Report%20on%20Open%20Access.pdf>

<審議経過>

平成24年(2012年)

- 1月18日 国際サイエンスデータ分科会(第1回)
第22期の活動計画について
- 3月9日 国際サイエンスデータ分科会(第2回)
CODATA小委員会、WDS小委員会の発足について
- 3月23日 WDS小委員会(第1回)
今期における活動方針、CODATA小委員会との連携
- 9月1日 国際サイエンスデータ分科会(第3回)

学術フォーラムについて、関連分野の活動について
CODATA 小委員会 (第1回)
第22期の活動について

平成25年(2013年)

- 3月11日 国際サイエンスデータ分科会(第4回)
CODATA 小委員会 (第2回)
マスタープラン2014「学術大型研究計画」について
- 3月25日 WDS 小委員会 (第2回)
WDS-SC 報告、WDS 国内連携組織の構築

平成26年(2014年)

- 2月14日 WDS 小委員会 (第3回)
極域データフォーラム報告
CODATA 小委員会 (第3回)
国際サイエンスデータ分科会からの提言案について
- 4月25日 国際サイエンスデータ分科会 (第5回)
対外報告について、CODATA 役員の推薦について
- 9月11日 日本学術会議 (第200回)
情報学委員会国際サイエンスデータ分科会 報告
「オープンデータに関する権利と義務 - 本格的なデータジャーナルに
向けて -」について承認