

河川環境データベースのオープンデータ化

Release as Open Data on River Environmental Database

東京都立大学都市環境科学研究科 准教授 大澤 剛 士

2013年にオープンデータ憲章が合意されてから10年が経過した。この10年で政府、自治体を主とするオープンデータの取り組みが進んだ結果、インターネット上で入手できるオープンデータは量・質ともに飛躍的に向上した。しかし、オープンデータは一度データを公開したら終わりというものではなく、さらに量・質ともに向上させていく必要がある。2021年8月、「河川環境データベース」は、そのリニューアルに合わせて利用規約を改定し、システムから公開されているデータは原則として全てがオープンデータとなった。本稿は、「河川環境データベース」を例に、webにおいて公開されているデータをオープンデータとする手続きを紹介する。さらに、オープンデータの推進を妨げ得る理解違いを提示し、なぜオープンデータ化に関する理解違いが発生するのか、その要因を論じる。これらを通して、今後さらにオープンデータ化を推進する上での課題について議論する。

キーワード：オープンデータ、DX、政府標準利用規約第二版、クリエイティブ・コモンズ

Ten years have passed since the Open Data Charter was agreed upon in 2013. Over these ten years, initiatives centered around Open Data mainly by governments and local municipalities have advanced significantly, resulting in a remarkable improvement in both the quantity and quality of Open Data available through the internet. However, Open Data initiatives are not merely about the initial release of data; there is an ongoing need to enhance its volume and quality. In August 2021, the "River Environment Database" underwent a renewal. Concurrently, its terms of use were revised, stipulating that data released from the system should, in principle, be considered Open Data. This report uses the "River Environment Database" as a case study to introduce the procedures for making web-published Open Data. Additionally, it delves into the reasons behind misunderstandings related to Open Data transformation, discussing the factors involved and deliberating on the challenges of promoting Open Data a conversion for web databases in the future.

Key words : Creative Commons, DX, Government of Japan Standard Terms and Conditions (Version 2.0), Open Data

1. はじめに

2013年のG8(当時)サミットにおいてオープンデータ憲章(https://www.mofa.go.jp/mofaj/gaiko/page23_000044.html, 2023年9月20日確認)が合意されてから10年が経過した。オープンデータとは、Open Knowledge Foundationの定義では自由な利用、再利用、再配布が保証されたデータとされているが(https://okfn.de/en/themen/open_data/, 2023年9月20日確認)、日本における基本指針では、利用性を担保するため、機械判読が可能であることも要件に含められている(デジタル庁 2021)。日本ではオープンデータ憲章の合意を皮切りに、2016年に制定された「官民データ活用推進基本法」、翌年に閣議決定された「世界最先端IT国家創造宣言・官民データ活用推進基本計画」(高度情報通信ネットワーク社会推進戦略本部官民データ活用推進戦略会議 2017)、そして2021年

に発足したデジタル庁より公表された「デジタル社会の実現に向けた重点計画」(デジタル庁 2023)等、国の政策文書類に「オープン」という言葉がしばしば登場するようになった。各種政策において、オープンデータを含む「オープン」という概念は、社会のデジタル化、いわゆるDXの実現に向けた鍵の一つと捉えられている。これらを受け、国はもちろん、民間や個人も含めて様々なデータのオープン化が進められ、インターネット上で入手できるオープンデータは量・質ともに、この10年で飛躍的に向上した。デジタル庁が実施している調査によると、内容は問わず、2023年6月時点でオープンデータに取り組んでいる自治体は都道府県では47全て、市区町村を含めると1449自治体と全体の8割超に上り(https://www.digital.go.jp/resources/data_local_governments, 2023年9月20日確認)、オープンデータという考え方は、我が国の社会に浸透しつつある。

とはいえ、公的機関等のwebサイトを見てみると、まだオープンデータへの取り組みがなされていない組織は多数あるし、取り組みを始めたものの、初期にカタログサイト等のポータルを設置し、そのまま更新されずに数年が経過しているという状況も散見される。例えばデジタル庁から提供されているオープンデータ取組自治体一覧（令和5年3月1日時点）をもとに、各都道府県オープンデータサイトのURLを確認してみると、2023年9月時点で確認できた範囲で、3年以上更新されていないと思われるサイトが3件確認された。オープンデータは一度既存データを公開したら終わりというのではなく、その量・質を向上させていく必要があるため（https://www.mofa.go.jp/mofaj/gaiko/page23_000044.html, 2023年9月20日確認）公開の仕組みを確立した後も、継続的にデータを更新していかなければならない。

2021年8月、河川水辺の国勢調査のデータを格納したwebシステムである「河川環境データベース」（<https://www.nilim.go.jp/lab/fbg/ksnkankyo/>, 2023年9月20日確認）は、そのリニューアルに合わせて利用規約を改定し、システムから公開されているデータは原則として全てがオープンデータとなった。「河川環境データベース」は、公的機関によるデータ配布サイトであるにも関わらず、これ以前は公開データをオープンデータとしていなかった。ただし、これは意図をもってオープンデータとしていなかったのではなく、オープンデータ化の手続きに対する認識不足に起因するものであったと筆者は考えている。そこで本稿は、「河川環境データベース」を事例に、webシステム等を通じて公開しているデータをオープンデータ化する最も簡単な手続きを紹介する。さらに、なぜオープンデータ化に関する認識不足が発生するのか、その要因を論じ、今後webデータベースのオープンデータ化を推進する上での課題について議論する。なお、本稿の内容は、筆者の個人的な見解に基づくものであり、主張等に対して国土交通省、水源地環境センター等関連機関およびその上位機関は関知していないことを申し添える。

2. 河川環境データベースのオープンデータ化と、その方法

「河川環境データベース」は、国土交通省が実施している「河川水辺の国勢調査」成果のうち、公開可能なデータを配布しているwebシステムで、2023年9月現在、ここに格納されているデータは、出典を明らか

にする限り商用を含め、自由に利用、再利用、再配布が可能なオープンデータとなっている（<https://www.nilim.go.jp/lab/fbg/ksnkankyo/index.html>, 2023年9月20日確認）。では、2021年8月以降、何をもって公開データがオープンデータになったのであろうか？そもそも、オープンデータとしてデータを公開するには、いったい何が必要なのだろうか？あくまで筆者の経験に基づく実感だが、オープンデータの考え方はかなり社会に浸透したものの、これに比して実際に何をどうすればオープンデータ化したとみなせるのかという手続きについて整理が不十分で、行政担当者等、実務者の認識も不十分になっている。

何をどうすればオープンデータ化したとみなせるのかに対する答えは、利用規約（ライセンス）、すなわち、データ提供者の権利義務や利用者が遵守すべき内容を設定する文書において、オープンデータの定義に当たる自由な利用、再利用、再配布を認めることとなる。データが著作物として認められる場合（注釈：この判断はときに困難であるが、一般的に創作物としての創造性を含むものは著作物と認められるため、この条件を満たす場合、複数のデータをまとめ上げたデータセット、データベースは著作物と認められる、<https://ec-houmu.com/right/bigdata#i-2>, 2023年9月20日確認）、その取り扱いに関する権利義務は原則として著作権法によって規定される。利用規約とは、法律上で保証された権利義務の主張範囲を定めるものと理解すればよい。ただし、適切な利用規約の設定には著作権法を中心とした法的な知識が必要になるため、当然これにオープンデータという要件を正確に設定するためにも、一定の専門知識が必要になる。このため、知的財産を管理する部門等を持たない組織や、個人が自身で適切な利用規約を設定することは容易ではない。これが、権利者はオープンデータとして公開したつもりであっても、実際はオープンデータになっていないという状況を作り出してしまいう一因になっていると考えられる。

この問題を解決するために、オープンデータにおいては自由に利用できる利用規約および、その「ひな形」が用意されている。自由に利用できる利用規約とは、国際的に広く利用されているクリエイティブ・コモンズ 表示 ライセンス (CC BY) であり、「ひな形」とは、クリエイティブ・コモンズ 表示 ライセンス (CC BY) に互換し、各府省ウェブサイトの利用ルール統一のために定められた政府標準利用規約（第2.0版）である。河川環境データベースは、2021年8月に政府標準利用規約（第2.0版）という「ひな形」を利用して利

用規約を定めることで、公開データをオープンデータに変更することができた。

3. クリエイティブ・コモンズ 表示 ライセンス (CC BY) と政府標準利用規約 (第 2.0 版)

クリエイティブ・コモンズ・ライセンスとは、インターネット時代における著作物の適切な流通を支援するために国際的非営利組織クリエイティブ・コモンズによって定められた利用規約のテンプレートで、2023年9月現在で最新のバージョンは4.0となっている (<https://creativecommons.jp/>, 2023年9月20日 確認)。クリエイティブ・コモンズ・ライセンスは、いくつかの記号の組み合わせによって、どういった権利義務を主張するかを設定できる構造になっており、著作者は、いずれかの組み合わせを選定し、図-1に示すマークを表示することで、著作物に対しこのライセンスを適用するという意思表示になる (<https://creativecommons.jp/licenses/>, 2023年9月20日 確認)。記号の組み合わせには様々なものがあるが (図-1) 一般にオープンデータとみなされるのは CC BY、CC BY-SAの2つである。これに加えて、著作権の保護期間が終了している、あるいは権利を放棄していることを意味するパブリックドメイン (PD) や (<https://creativecommons.jp/licenses/>, 2023年9月20日確認)、権利を保有しないことを宣言するCC 0というライセンスもあるが (<https://creativecommons.jp/sciencecommons/aboutcc0/>, 2023年9月20日 確認)、

データ公開機関が権利を放棄あるいは保有しないというケースは多くないため、本稿では取り扱わない。データにCC BY、CC BY-SA のいずれかが表示されている場合、該当データはオープンデータとみなされる。各ライセンスにおける詳細な許諾内容はインターネット上で公開されているため (例えばCC BY, <https://creativecommons.org/licenses/by/4.0/legalcode.ja>, 2023年9月20日確認)、利用者は必要に応じて許諾内容の詳細を確認することができる。オープンデータとみなされるライセンスであるCC BYとCC BY-SAの違いを簡単に説明すると、CC BYが付与された場合、そのデータを加工する等した派生品等を発表する際に何の制約も発生しないが、CC BY-SAは、ライセンスを付与されたデータを利用した派生品等についても、同じライセンスを付与することが求められるという点である。

政府標準利用規約 (第2.0版) とは、日本政府によって定められた各府省ウェブサイトの利用ルールのひな形であり、これを適用するためのテンプレートがデジタル庁から公開されている (https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/70143e67/20220523_resources_data_betten_03.pdf, 2023年9月20日確認)。このテンプレートに必要な部分を追記し、ウェブサイト上で公開するだけで、ウェブサイトに掲載されたコンテンツをオープンデータにするという利用規約になる。政府標準利用規約 (第2.0版) は、基本的にクリエイティブ・コモンズ 表示 (CC BY) 4.0と互換するものとされているが、省庁等をはじめとするデータ権利者の個別の事情に応じて、追加で適切な利用条件を定められる仕組みとしている (https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/a0f187e6/20220706_resources_data_betten_01.pdf, 2023年9月20日確認)。ただし、政府情報は原則としてオープンデータにするという方針 (「二次利用の促進のための府省のデータ公開に関する基本的考え方 (ガイドライン)」 (平成25年6月25日 各府省情報化統括責任者 (CIO) 連絡会議決定), https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/7c57e1a9/20220523_resources_data_guideline_01.pdf, 2023年9月20日確認) が存在しているため、これを政府機関が利用する際には、「個別法令に根拠のない利用制約を課すような別の利用ルールを設ける場

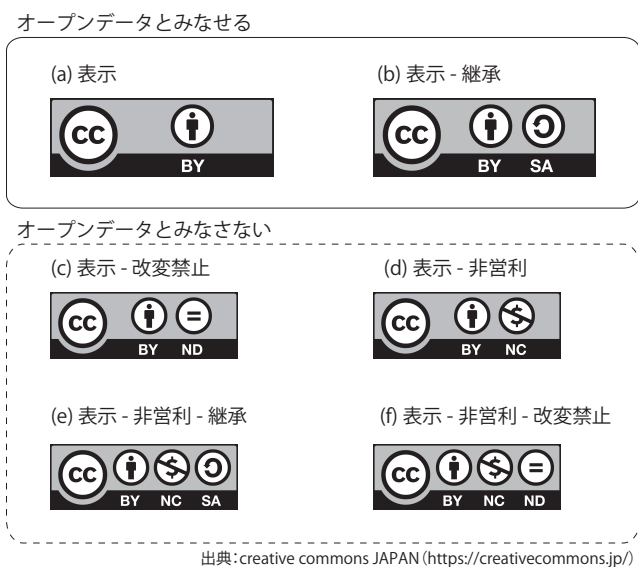


図-1 クリエイティブ・コモンズ・ライセンスでは、4つの記号を組み合わせることで6種類のライセンスを設定できる。このうち (a) 表示と (b) 表示-継承 が付与されているデータはオープンデータとみなせる。

合は、そのコンテンツの範囲を具体的に示した上で、別の利用ルールを設ける具体的かつ合理的な根拠を示す」と明示されている点に注意する必要がある (https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/70143e67/20220523_resources_data_betten_03.pdf, 2023年9月20日確認)。すなわち、政府機関は、過去の経緯や前例のような合理的な根拠がない理由によってデータをクローズにすることはあってはならないと明示されているのである。この点も勘案し、河川環境データベースでは利用に関する免責事項等のみ追記し、特例の利用制約はかけないという方針で、政府標準利用規約(第2.0版)を用いた利用規約を設定した (https://www.nilim.go.jp/lab/fbg/ksnkankyo/Mokj_chui.htm, 2023年9月20日確認) (図-2)。

4. 政府標準利用規約(第2.0版)の適切な利用と認識の違いに起因する課題

政府情報には個人情報につながりうるもの等、扱い

に注意が必要なものが含まれることがあるため、個別事情に応じ、根拠を示した上で利用制約等を設定できるようにした政府標準利用規約(第2.0版)テンプレートの配布および運用は、オープンデータに取り組む政府機関や行政機関にとって非常に利用性が高く、実際に様々な省庁や機関で利用されている(例えば総務省電波利用ホームページ, <https://www.tele.soumu.go.jp/j/musen/readme/>, 国際農林水産業研究センターウェブサイト利用, <https://www.jircas.go.jp/ja/about/guide>, ともに2023年9月20日確認)。しかし、残念なことに、「個別事情に応じて制約等が設定できる」という部分が独り歩きし、追加で設定した規約がオープンデータの趣旨から逸脱してしまっているように見えるケースも散見される。あえて具体例を挙げると、法務省が公開を開始した登記所備付地図データに関する項目がこれに該当する。登記所備付地図データはオープンデータであると明示されているが (https://www.moj.go.jp/MINJI/minji05_00494.html, <https://www.digital.go.jp/news/4b7250a3-3fcf-4b83-8d52-4bb131e1ba9d>, ともに2023年9月20日確認)、政府標準利用規約(第2.0版)を利用して策定された利



出典: 国土交通省河川水辺の国勢調査 河川環境データベースシステム (<https://www.nilim.go.jp/lab/fbg/ksnkankyo/index.html>)

図-2 河川環境データベースのトップページと、利用規約の表示箇所

用規約を見てみると、個別事情に応じた利用制約を説明する項目の中に「法務省の判断によって、コンテンツの一部及び全部の利用制限、又は今後一切のコンテンツ利用を認めないことがあります」という項目と、その具体事例の一つに「ウ その他法務省が不適切と判断した場合」という内容が示されている。管理主体が不適切と判断した場合に～という制約は、利用規約を設定する際にしばしば使われる表現である。しかし、この表現は、オープンデータの本来の趣旨に反すると筆者は考える。オープンデータの本質とは、データが著作権や用途等、社会に存在する様々な制御メカニズムから開放され、利用者はデータの品質以外の何にも制限されないことが保証されることである(大澤 2017)。この点において、事前に利用について法務省に判断を求める必要があるように読める制約を設定することは、オープンデータの定義から考えると、首をかしげざるを得ない。これに加えて、先述したとおり、デジタル庁が公開している政府標準利用規約(第2.0版)の解説には、「個別法令に根拠のない利用制約を課すような別の利用ルールを設ける場合は、そのコンテンツの範囲を具体的に示した上で、別の利用ルールを設ける具体的かつ合理的な根拠を示すべき」と明記されている(https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/a0f187e6/20220706_resources_data_betten_01.pdf, 2023年9月20日確認)。「法務省の判断によって」という曖昧な表現は、制約の対象となるコンテンツ範囲が不明瞭であるし、個別法令に根拠があるとも思えない。この制約は関係者のオープンデータに対する認識不足に起因して設定されたものと推察されるが、権利者がオープンデータにする意思があるにも関わらず、ライセンス上オープンデータになっていないというミスマッチは、データ提供者、利用者の双方に不利益をもたらす。ライセンス設定の困難さを解決することに加え、個別事情に応じて規約内容に柔軟性を持たせることでオープンデータの推進を目指す政府標準利用規約(第2.0版)は非常に優れた仕組みであるが、柔軟性を持たせたがために、本来の趣旨であるオープンデータ化が達成できなくなるケースが存在するというのは大きな問題である。これを解決するために、形式的なものではなく、オープンデータそのものの内容と本来の意義をより広く周知する努力を講じることと同時に、その実現に向けた方法である政府標準利用規約(第2.0版)の適切な利用方法も、より広く周知していく必要があるだろう。

5. オープンデータ化に向けた各種資料と課題

先述の課題解決に向けた普及活動の第一歩として、適切な資料を作成し、広く配布することが挙げられる。この取り組みはかなり進んでおり、2023年9月現在、web上を検索すると基礎自治体やデジタル庁発行のオープンデータ作成ガイドライン、手引き書を容易に見つけることができる。例えばデジタル庁は自治体がオープンデータに取り組む際に参考となるガイドライン・手引書を公開している(https://www.digital.go.jp/resources/open_data, 2023年9月20日確認)。発行年が2018年と多少古いものの、農研機構は研究者や研究機関を対象に、研究成果、データをオープン化するための手引きを公開している(https://www.naro.go.jp/publicity_report/publication/laboratory/niaes/manual/083173.html)。九州大学は、オープン化を含めた研究データの管理方法に関する教材を公開している(https://catalog.lib.kyushu-u.ac.jp/opac_detail_md/?lang=0&amode=MD100000&bibid=6796147, 2023年9月20日確認)。これらはあくまで一例で、地方自治体やNPO等による手引き等も多数公開されており、各主体が実践する際の参考資料として重要な役割を果たしていると考えられる。

これら各手引きは基本的にオープンデータの意義、それを実現するための具体的な方法を適切に記述しており、オープンデータに取り組む自治体等の職員が参考にすることはもちろん、学校機関における教材としても利用できる質であるものが多い。特にデジタル庁発行の自治体向けガイドライン・手引きはオープンデータ化の実施に留まらず、それを継続するための組織的な対応等にも言及しており、オープンデータに取り組む多くの自治体が参考にしており、考えられる。ただし、現在流通しているガイドライン、手引きの多くは一般的なもの、すなわち政府、行政等が公開するデータ全般を対象としたものが多く、適用するデータの内容や性質を絞り、個別的な課題に対応する方法等が記述されたものはほとんど見られない。先述の登記所備付地図データについても、例えば不動産に関わるデータをオープンデータ化する際に留意すべき事項を、関連する個別法令とともに提示するガイドラインがあれば、政府標準利用規約(第2.0版)の解説にあるように、個別的な利用制約について、具体的かつ合理的な根拠に基づいたルールを設けることができた可能性がある。今後ますますオープンデータを推進するためには、ある程度対象を絞って具体的な対応方法を示した個別マニュアルのようなものが必要なのかもしれない。

筆者が知る限り、適用するデータの性質を絞っているオープンデータ化にむけたガイドラインの例はまだ多くないが、徐々に策定は進みつつあるようだ。例えば2023年6月には、都市計画情報のデジタル化・オープン化ガイダンスが公表された (https://www.mlit.go.jp/toshi/tosiko/toshi_tosiko_tk_000187.html, 2023年9月20日確認)。2023年5月には、自然環境調査に係る生物多様性情報の整備と発信のガイドライン (<https://www.biodic.go.jp/kiso/masterplan.html>) が公表された。後者は筆者も作成に関わったことから、その概要を紹介したい。このガイドラインは、法定業務である自然環境保全基礎調査の基本方針・調査計画として策定された「自然環境保全基礎調査マスタープラン 令和5～14(2023～2032)年度」の成果データの取り扱いに関する実施方針として活用することと同時に、本調査に関わらない生物多様性に関わる各種データの整備およびオープンデータ化に広く貢献することを狙って作成したものである。ガイドラインには、オープンデータの意義や先述のライセンスに関する概説に加え、生物多様性情報に関する個別的な留意事項として、生物多様性データにおける標準データ記述フォーマット、位置情報の扱い、希少種情報の扱いについての概説も含めており、生物多様性情報に絞ってオープンデータ化の具体的な対応方法を示している。ガイドラインが乱立することは避けるべきであろうが、潜在的なデータリソースが膨大である対象や社会的ニーズが高い対象については、ある程度焦点を絞った個別マニュアルの整備を進めることで、適切で利用性が高いオープンデータがさらに増加していくことが期待できる。

6. おわりに

本稿は、河川環境データベースを事例に、既にwebシステム等を通じて公開しているデータをオープンデータ化する簡単な手続きの紹介を通し、オープンデータ化の実現に向けた対応が極めて簡便化していること、その反面、オープンデータの本来の趣旨を逸脱してしまうという課題が発生していることを論じてきた。さらに、この問題を解決するための手段として、ある程度対象を絞ったガイドラインの必要性を提示した。オープンデータの推進において最も重要なことは、オープンデータ化はあくまで行政情報の透明化や民間の業務等を活性化させる手段であり、それを目的化してはいけないという点である。何のためにデータをオープン化するかという本質を押さえておけば、自

ずと実施すべきこと、避けるべきことは見えてくると筆者は考える。河川環境データベースに続き、今後も様々な行政調査のデータ類のオープン化が進むことを期待したい。

謝辞

本稿の執筆にあたり、(研)農研機構の岩崎亘典氏、(株)MIERUNEの古川泰人氏との議論は極めて有意義であった。両氏および駒沢大学の瀬戸寿一氏には原稿に対して有益なコメントをいただいた。国土交通省河川環境課の天羽淳氏には、河川環境データベースのライセンス変更に多大なる尽力をいただいた。ここに名を示して謝意を表したい。

参考文献

- 1) デジタル庁(2021) オープンデータ基本指針. https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/f7fde41d-ffca-4b2a-9b25-94b8a701a037/20210615_resources_data_guideline_01.pdf, 2023年9月20日確認
- 2) デジタル庁(2023) デジタル社会の実現に向けた重点計画. https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/5ecac8cc-50f1-4168-b989-2bcaabffe870/b24ac613/20230609_policies_priority_outline_05.pdf, 2023年9月20日確認
- 3) 高度情報通信ネットワーク社会推進戦略本部 官民データ活用推進戦略会議(2017) 世界最先端IT国家創造宣言・官民データ活用推進基本計画 <https://warp.ndl.go.jp/info:ndljp/pid/12187388/www.kantei.go.jp/jp/singi/it2/kettei/pdf/20170530/honbun.pdf>, 2023年9月20日確認
- 4) 大澤 剛士(2017) オープンデータがもつデータ開放の意味を再考する：自由な利用と再利用の担保に向けて. 情報管理 60(1): 11-19.