

第1回 データ環境推進委員会 議事要旨

日時：2024年2月26日(月) 10:30-12:30

場所：IPA 15階 委員会室1 及びオンライン

出席者：

【委員（50音順）】

瀬戸 寿一 委員
高木 祐介 委員
高田 智和 委員
武田 英明 委員
中野 一晃 委員
西村 正貴 委員
眞野 浩 委員

【傍聴】

阿部 隆弘 氏（独立行政法人国立印刷局 官報部 ベース・レジストリ準備G専門官）

【事務局】

独立行政法人情報処理推進機構

議題：

1. 出席者ご紹介、ご挨拶
2. 日本のデータ環境について
3. データ環境推進委員会について
4. 委員長選出
5. 論点
6. 今後のデータ環境推進に関するロードマップについて
7. クロージング

配布資料：

- 資料 0 議事次第
- 資料 1 委員名簿
- 資料 2 データ環境の現状について
- 資料 3 データ環境推進委員会について
- 資料 4 議論のポイント
- 資料 5 今後のデータ環境に関するアクションアイテム

議事概要：

■事務局連絡

- ・ 配布資料の確認
 - ・ 会議の趣旨及び委員会へ招集させていただいた背景説明
 - ・ 現在、日本には良質なデータ環境が求められている。
 - ・ メタバースやデジタルツインの進展により、テキストデータ中心から、地理空間データや画像データのように、求められるデータの幅が広がってきている。
 - ・ 日本にはビジョンがない。
 - ・ ビジョンを考えると合わせて、ロードマップを定めて進める。
 - ・ 定めたロードマップを日本政府のデジタル戦略等に反映することを推進したい。
 - ・ 日本のデータ環境を作っていきましょう。
 - ・ 委員のご紹介
 - ・ 高木委員 デベロッパーの視点でご意見を頂きたい。
 - ・ 高田委員 データの持つ意味、データの辞書等作成において、日本語が重要視されていることから、専門の高田先生に来ていただいた。
 - ・ 武田委員 語彙の専門家、知識処理の専門。
 - ・ 瀬戸委員 地理空間情報の活用を推進するという点でご意見を頂きたい。
 - ・ 中野委員 ベース・レジストリなどデータマネージメントの専門家としてご意見を頂きたい。
 - ・ 西村委員 オープンデータ、コード表、データ辞書を先行してやられていることからご意見を頂きたい。
 - ・ 眞野委員 データ流通という観点からご意見を頂きたい。
- ・ 資料、議事概要を事前に委員に確認の上、公開することを説明。

■1. 日本のデータ環境について

(事務局)資料 2 に基づき日本のデータ環境について説明

- ・ 本委員会では、日本に限定せず、国際情勢に関する情報を蓄積し共有し、そこを目指すためには何をしなければならないかという議論をしたい。
- ・ 今までオープンデータを進めてきたが、日本では進んでいないのが実情で、データスペースも浸透していない。
- ・ 一方、AI の普及が進み、AI へ投入するデータが重要視され始めた。
- ・ 日本の課題（スライド 6）
 - ・ 日本国内においては、まずはワンチームになることが重要。
 - ・ データサイエンティストは沢山いるが、データの根本を整備する役割が不足して

いる。

- ・ データ基盤やデータ共有に対する理解が進んでいない。
- ・ 国内最大のデータオーナーである行政の取組が遅れている。
- ・ トラスト基盤の整備ができていないことから、本人認証の基盤が定まっていない。したがって ID の振り出しから始まることになる。
- ・ データ戦略推進における個別の課題（スライド 7）
 - ・ データモデルが注目される傾向が強い。
 - ・ フレームワークを作ったが、IMI との整合性がわかりにくいという意見もある。
- ・ 世界を取り巻く大きな流れ（スライド 8）
 - ・ 日本のデジタル競争力の低下。
 - ・ ランキング結果は、WEF（世界経済フォーラム）が日本の経営者に聞いたものを集計しているため、日本人経営者の主観的な回答が反映されている。
 - ・ もっと普及、啓発活動をしないといけない。
- ・ コンセプトの具体化に課題（スライド 10）
 - ・ 日本は Society5.0 や DFFT などのコンセプトのみで具体化ができていなかった。
 - ・ 一方で、欧州ではビジョンが明確で、具体的な交換方法や言語変換まで整備が進められている。
- ・ 今後、目指すべき姿：データスペース（スライド 13）
 - ・ 今後は、データスペースが重要と考える。
 - ・ 既に民間独自ルールに従ったものや、独自の接続方式等により取引が実現している。
 - ・ 国際的な枠組みをフォローするだけでなく、日本独自の枠組みを作っていく必要がある。
- ・ 今後、目指すべき姿：データマチュリティとデータマネジメント（スライド 15）
 - ・ 日本ではデータマチュリティについて議論がされていない。
 - ・ 各社保持しているデータをどのように経営に生かすか、活用するかということを考え、これまでのカルチャーを変えることによって、経営者がデータマネジメントに注力することに繋がる。
 - ・ 各社がデータを公開することにより、連携や品質向上につながり、組織毎のデータマネジメントを強化しなければならない。
 - ・ 英国政府のデータマチュリティアセスメントに習い、日本も取り組まなければならない。
- ・ 検討されている体制（スライド 18）
 - ・ データ推進に関する今後の体制。
 - ・ IPA は、これまで経済産業省の所管であったが、今後はデジタル庁との共管になるため、デジタル庁とも戦略などすり合わせを行い、デジタル庁の下、国立印刷局、

J-LIS、NICT とも組み、ワンチーム体制としたい。

- ・ データ戦略の具体化（スライド 19）
 - ・ 2023 年 12 月に「AI 時代の官民データの整備・連携に向けたアクションプラン」を公表した。
 - ・ 例えば、政府調達に GIF を含めることにより、前段のオープンデータの整備までつながる。
 - ・ 各省がベース・レジストリの計画を作ることが求められている。それを支援するような体制が必要。
 - ・ AI 学習ニーズがある行政保有データについても、なんらかの支援が必要。
- ・ ワンチームで目指すゴール（スライド 26）
 - ・ 数年以内にプラス 10 位以上の上昇を目指す。
 - ◇ 意識改革
 - ◇ 事例整備

■2. データ環境推進委員会について

(事務局)資料 3 に基づきデータ環境推進委員会について説明

- ・ 委員会の趣旨（スライド 2）
 - ・ 有識者の方々とビジョンを議論し、ロードマップを作成したい。
 - ・ 2023 年 6 月に閣議決定されたデジタル社会の実現に向けた重点計画において、IPA はデジタル基盤の基準・標準策定の機関として機能することを求められている。
 - ・ 標準化団体（SDO、Standards developing organizations）までは難しいが、オーサライズしてガイドを出すような体制にもしたい。
 - ・ ワーキンググループの設置や事務局で作成する計画や成果物について、審議・承認を行う機関としても機能することとしたい。
- ・ 委員会の活動概要（スライド 4）
 - ・ 国際情勢及び国内情勢を共有し、ビジョンの正しさ、優先的に取り組むことなどを計画し、議論したい。
 - ・ 公開前のガイドライン群の内容確認。
 - ・ 委員会配下に設置予定のデータモデルワーキンググループの技術的課題について検討。
- ・ 委員会の主な成果物（スライド 8）
 - ・ 次の重点計画が 2024 年 6～7 月頃公開される可能性があるため、それまでにそれらの支援となるようなものを打ち出したい。
 - ・ 委員会議事及び資料は特段の指定のない限り公開する。

- ・ 委員会議事について公開はするが、発言者については匿名とする。

■3. 委員長選出

- ・ 武田委員が委員長に選出された。
- ・ 次回以降は、委員長である武田氏に主導して頂く。

■4. 論点

(事務局)資料 4 に基づき議論のポイントについて説明

- ・ 主な成果物について、不足している実施項目や成果物はないか、また中長期的な課題等についてコメントを頂きたい。

■5. 今後のデータ環境推進に関するロードマップについて

(事務局)資料 5 に基づき今後のデータ環境推進に関するロードマップについて説明

- ・ 仮ロードマップ (スライド 2)
 - ・ FY2024-1Q6 月にデジタル戦略が改正。
 - ・ FY2023-4Q から FY2024-2Q に向けては、データカルチャ変革とデータマチュリティ強化。
 - ・ 2023 年度の積み残しがあるため、FY2024-1Q がタイトなスケジュールになっている。
 - ・ データ辞書の整備については、2つのフェーズに分け長期的に推進する。
- ・ GIF モデルの改定 (2.位置情報表記) (スライド 7)
 - ・ 緯度経度の表記について、統一した表記法がない
 - ・ 例えば、公園を表現する場合は、ポリゴンで表すのか、中心点でよいのか、公園の入り口とするかなど。
 - ・ 空間 ID については、3次元化の検討。
 - ・ POI (Point of Interest) は作ったが、普及していない。
- ・ GIF データモデルの改定 (3.IMI との整合) (スライド 9)
 - ・ 精緻に作られた IMI と実装のためのモデルである GIF との擦り合わせが必要。
 - ・ 実装モデルを GIF と置き、普及を進めたい。
- ・ データ品質の向上 (スライド 11)
 - ・ GIF と内閣府からデータ品質ガイドが公開されているがあまり知られていない。
 - ・ だれでも使える簡易評価ツールを作る。
- ・ ベース・レジストリでの GIF 活用 (スライド 13)
 - ・ ベース・レジストリは、再利用が重要。
 - ・ ベース・レジストリは、GIF に則って作ることを訴えなければならない。
- ・ ワンスオンリーの推進 (スライド 15)

- ・ ショーケースとしてワンスオンリーを推進したい。
- ・ 現在の証明書等の多くは PDF 形式のため、自動審査ができない状況。
- ・ ヨーロッパでは、2023 年末時点で、全域でワンスオンリーシステムが稼働している。
- ・ gBizINFO 改訂のタイミングで、法人の証明情報を集め、自動審査をやすくしショーケースの一つとしたい。
- ・ ショーケースの推進（スライド 17）
 - ・ 各省庁からのベース・レジストリ整備やデータ整備の依頼についての支援も担う。
- ・ ツールの整備（スライド 19）
 - ・ 海外の有用なツールを紹介することなど、面でフォローできるようなツールカタログなどの整備を行う。
- ・ データ辞書の整備（スライド 21）
 - ・ GIF コアデータモデルの中でも、言葉の揺れがある。
 - ・ 各省庁で使う単語の揺れを整備する。
 - ・ コード、統制語彙などの整備は統計局と一緒に統一していきたい。
- ・ データ人材の育成（スライド 23）
 - ・ 専門家集団のレベルアップと合わせて、中核的なデータ人材の強化も必要。
 - ・ 大学との連携、例えば、大学で作った資料を公開するなどのもその一つと考える。
- ・ オープンデータの推進（スライド 27）
 - ・ オープンデータの質を向上するために何が必要なのか検討する。
 - ・ 啓発活動の一つとして、IPA で「オープンデータリポート会議@巣鴨」を 3/7(木) 16:30 から 18:00 に開催予定

■6. ここまでの説明に対する質問

(委員) 資料を作ったが内容が十分でなく、公開できないケースはどうなるのか。

(事務局) 有用なものは公開する。タイトなスケジュールになっているが、作るのが大変なのは、例えば入門編として 5 ページくらいのもを作り公開していくことになる。

(委員) 標準化の話が出なかったがどのように考えているのか。

(事務局) やらなければならないことは認識している。データスペースの標準化についても ISO の標準を目指したいところ。本来ならば、標準化チームも作りたかったが人材が不足している。今後、一年ぐらいかけて体制を整える。

(委員) デリバラブルが重要。公開物については、アウトラインや用語を統一するところから始めるべき。例えば、誰を対象にし、何を訴求して、どのような書きぶりにするのかなどを最初に作るべき。

(事務局) 認識している。十分かは別として、IPA サイトで公開している「データに関する用語集」などは、これまで IPA が不足していたユーザ視点によるもの。それらに DSA やデジ

タル庁が作ったガイドを統合するとアウトラインも描けるのではないか。

■7. 論点についてのご意見

(委員) 配布資料3「データ環境推進委員会について」の委員会の推進方法について、開催頻度が四半期に一度とすると出来ることは限られる。IPA 内のアクションは理解したが、具体的に委員会が2時間 x4 回でできることを明確にしたい。

(事務局)本来は、6回程度の開催にしたかった。事務局の成果や配下に設置するデータモデルWGの実施状況を確認してもらうことや、社会全体の状況を見て、アドバイスを頂きたい。

(委員)配布資料 3「データ環境推進委員会について」のスコープの中で、日本の体制がない分野については IPA が作ったものを委員会内で議論するのか、ゼロから委員会で議論することになるのか。

(事務局)コミュニティやワークショップが活発になることが望ましいが、実際には人材が不足している。IPA 職員、専門委員のアウトリーチが必要。ワークショップも含めて検討したい。

(委員)標準化の推進において、デファクトでもよいので、国内で提供するような意識はあるのか。

(事務局)おそらく、政府調達に組み込むことになるのではないかと。政府標準を基に、民間に波及することを狙いたい。閣議決定されたデジタル戦略において IPA の位置付けとして「IPA は、政府の標準・基準を作る機関である」と言われている。この委員会で議論したものを政府へ打ち込むものとしたい。過去の成功事例として、文字情報基盤がある。例えば、自治体標準化の中では住所ベース・レジストリに従って、というような指示が可能ではないか。IMI で作った施設語彙についても、GIF の基となっており、それが文科省の教育データ標準に活用されている。

(委員)組織の命令によるとその指示は絶対となり、更に今年度中という指示があると、100 年後というよりも目の前の仕事に注力することになる。現実の問題も考慮しなければならない。

(委員) 最終ゴールとして政府標準に適用されること、あるいは政府が自治体に対して、法的な根拠に基づいて指示できるようにするために、我々が何をインプットできるかということと理解した。国や自治体が主導してデータを整備・流通させるようなインフラ的な側面のみならず、民間企業が持っているデータを公益目的で提供し充実化する流れも考慮できると良い。本委員会では、細かいレベルまでのチェックは難しいと考えるが、将来的にあるべき姿の議論は必要。

(事務局)本来は法律から変えるべきなのだが、ガイドラインで留まっているのが現実。そこに対し、中長期的な視点でアドバイスをしていきたい。

(委員) 地理空間情報の分野では、内閣官房内に各省庁を横繋ぎする「地理空間情報活用推進室」が立ち上がっている。デジタル庁とも連携されているようなので、当該分野の検討に際しては、本委員会でもそれらの組織との連携があってもよいのではないかと。また、3次元都市モデルに関しては、

Project PLATEAU で、国際標準の CityGML でデータ整備を行っており、その点も公開情報が非常に充実しているのでインプット可能と考える。

(事務局)日本政府の中で先進的に進んでいる国土交通省では、PLATEAUや国土数値情報などオープンデータとしての側面にも注力している印象がある。他にも防災マネジメント分野での実績や国土交通プラットフォームなどがあり、是非、一緒に取り組んでいきたい。

(委員)提示されたロードマップはかなりタイトという印象を受けた。デジタル庁と情報共有している既存の関係もあるので、ワーキンググループ設置以外でも協力していきたい。

(委員)自治体は、年度区切りの制限があるため、成果物の精度に影響が出ている。自治体標準データセットは委員会の対象となるのか。

(事務局)「GIFに準拠してください」と言っているが、具体的な準拠の基準が明確でなかった。また、デジタル田園都市国家構想にも記載されたが、準拠の基準が明確でなかった。文字の場合はわかりやすい。

(委員)郵便番号の表記について、GIFと自治体標準データセットに違いがあり、準拠の基準が曖昧になっているため、標準データセット毎に準拠先が変わっている。標準化の基準の徹底が必要。

(委員)データ整備を進める上で、実際にデータ整備をした現場レベルの意見についてフィードバックを受けた議論が必要と理解した。

(事務局)フィードバックループを作るということも含めて検討する。

(委員)標準を作るのか、インフォーマティブな文章を作るのかを明確にすべき。例えば、規正文書は誰が読んでも理解できる、理解に齟齬がでない文書にすべき。人材育成の骨格(例えば、タイトルや対象など)は、今年度早いタイミングで決めたほうがよい。

(事務局)ガイドはあるが実運用できない環境を解決したい。

(委員)実際にデータ整備をしている立場からは、人材育成の分野では、まだまだデータリテラシーが低い。人材育成は重要と考える。

(事務局)データマチュリティと個人のリテラシーの強化が喫緊の課題と理解した。

(委員)データマチュリティの推進は省庁として可能なのか。

(事務局)海外では各省チェックリストなどが進められている。ヨーロッパでやっているようなレーダーチャートで公開するのもよいし、地理空間分野では、LIFO (Location Interoperability Framework Observatory) など各国でどこまで出来ているのかなどをアニュアルレポートとして公開している。そのような仕組みを日本でも導入してもよい。

(委員)業界毎のデータ辞書が多々あるが、データ辞書をつくる方法の一つとして、統合するデータ辞書を決めてまとめていく方法もある。しかし、理想を追求しすぎると、現場が混乱する。

(委員)農作業語彙を作った。その際には、参照するモデルがないので、農作業支援ソフト、農水省の用語集、NHK の用語集、公用語辞典などを調べてひたすら言葉を集めた。

(委員)体制の中で CSTI との関係はどのように考えればよいのか。

(事務局)AI 関連で内閣府科学技術・イノベーション推進事務局とは連携はしている。そこも含めてワンチームとして体制を検討したい。

(委員)用語の表記揺れについて、データそのものだけではなく、項目の用語の表記揺れ(例えば、時間と時刻)についてもルールを決めていくことから進めたい。

(事務局)ラベリングルールについても検討することも重要と理解した。

(委員)データモデルのフレームワークを定義したほうがよい。値を表記する時のルールまで踏み込むと互換性が取ることができる。そうすると、推奨データモデルと言えるのではないか

■8. クロージング

- ・ 議事要旨は1週間程度を目途に共有展開予定。承認後に公開する。(事務局)
- ・ 配布資料についても、懸念等あればメールでご指摘いただきたい。(事務局)

【決定事項】

- ・ 武田氏が、データ環境推進委員会の委員長に選出された。