

## 実務者検討委員会（第 16 回） 議事概要

日時：令和 5 年 8 月 9 日（水）10 時 00 分～12 時 00 分

場所：オンライン

### 【議事】

- (1) 「デジタルアーカイブ活動」のためのガイドライン（最終案）及び、デジタルアーカイブアセスメントツール 2023 年版（最終案）について
- (2) 実務者検討委員会及びワーキンググループの議論の経緯等（最終案）について
- (3) ジャパンサーチの新機能について
- (4) その他

### (1) 「デジタルアーカイブ活動」のためのガイドライン（最終案）及び、デジタルアーカイブアセスメントツール 2023 年版（最終案）について

- 国立国会図書館より資料 1-1、資料 1-2、資料 1-3、資料 1-4、資料 1-5 について説明

#### 〈主な意見〉

- 表記の揺れや、同じ趣旨の記載の繰り返し等の修正、また言葉の統一を工夫されるとよい。
- CC0 と CC BY、どちらを基本としているのか、分かりやすく記載するとよい。
- 自然科学の標本情報においては、オープンデータの観点から CC BY が主流。また、政府が実施する生物系調査のデータの公開については、オープンガバメントの文脈に沿って、政府標準利用規約 2.0 版、CC BY 4.0 互換という形でのライセンスが基本。デジタルアーカイブの文脈における CC0 の推奨とどうすり合わせたらよいか検討しているところ。
- 先行して公開のルールづくりが進んでいるコミュニティについては、それらを否定するのではなく、CC0 が我々の目標であるということを伝えていくことで、個々のコミュニティが今よりも公開を緩和するアクションを取ろうとしたときの援護射撃になるのではないか。そのように当ガイドラインが活用されていく形が望ましい。ジャパンサーチでは CC0 を目指しているという方向性を示すことができればよいのではないか。
- 当ガイドラインは長文なので、セクションごとに分けて web 上に公開するなど、それぞれのセクションに応じて相互参照できるような構成にするとよいのではないか。
- 概要版を作成し、これを引用されやすい形で紹介する方向で検討したい。
- 当ガイドライン自体の URL の永続性を担保することは必要。
- 用語集の項目で、欧文の表記や読み方の表記について、表記基準を統一した方がよい。
- 表 4 の掲出位置がずれているのではないか。
- 用語集の中には、このプロジェクトの中でつくってきた用語もあれば、ごく一般に使われている用語もある。この用語集はガイドライン全体の品質や信頼性の評価にも関わるに使われると思われる

ので、ミスのないように留意する必要がある。

- 一般的な用語、専門用語、ガイドラインの中でだけ使われる方言のような用語が、混在し区別がつかないと、諸学者にとっては混乱を招く可能性もあるので、印をつけるなり、外部参照リンクを記載するなど、なにか工夫をしていきたい。

## (2) 実務者検討委員会及びワーキンググループの議論の経緯等（最終案）について

- 内閣府知的財産戦略推進事務局より資料 2 について説明

※特段、意見なし

## (3) ジャパンサーチの新機能について

- 国立国会図書館より資料 3 について説明

〈主な意見〉

- (SF プロトタイピングを活用してジャパンサーチの新機能を考えることについて) 大変面白いが、誰をメインの対象とするかということ想定しておいたほうがよい。学生、ベンダー、自治体など様々考えられるが、ある程度の絞り込みは必要ではないか。
- (画像検索のサムネイルのクラスタリング等、ジャパンサーチ上にラボのような実験ページを提供することについて) 地方機関では、所有コンテンツを可視化することについて、費用等の関係で、十分に対応できていないところが多い。ジャパンサーチ上にラボのような機能があれば、地方機関にとっては、可視化するシステムを自ら開発する必要がなくなるかもしれないので、早めの情報提供をお願いしたい。
- 地域アーカイブをジャパンサーチから見つけていくという観点は大切。類似画像検索と API を活用したギャラリー自動生成機能を組み合わせることで、地域にある類似画像の発見可能性が高まることはよいこと。地域の歴史とか文化を理解してもらう上でも結構重要な役割を果たすと思う。自動生成したものとの区別は必要。
- 地域由来のものをジャパンサーチ上に掲出していくことは有意義。技術的には自由自在にできる。
- 言葉による画像検索は非常に大きな可能性がある。ジャパンサーチにある大量の画像とメタデータを使って、日本版の CLIP<sup>※</sup>のモデルをつくるみたいな方向に行くと、とてもよいのではないか。
- ジャパンサーチの中で使うことができる画像とテキストを生成 AI の学習用データセットとして構築していく可能性はあるが、一方、その利用については、まだ議論の余地があるものと承知。

---

※ 言語と画像のマルチモーダル AI モデル。2021 年 2 月、OpenAI により公開。

- ジャパンサーチは、集約しているメタデータ等を活用していろいろなことができるということを、技術とともに示すことが、役割のひとつ。ラボの機能もそのひとつ。さらに各機関でローカライズして使ってもらうことで、さらに良いものになる。

#### **(4) その他**

- 国立国会図書館より、資料 4-1 ジャパンサーチ・アクションプランのアンケート結果について報告。

※特段、発言なし

- 知財事務局より、資料 4-2 デジタルアーカイブフェス 2023 の概要及び、デジタルアーカイブジャパン・アワード 2023 選考結果についてについて報告。

※特段、発言なし

以上