

CA1967

動向レビュー

ワールドデジタルライブラリーの動向

ときぞねそういち
時実象一*
たかの あきひこ
高野明彦†
ふくばやしやすひろ
福林靖博‡

1. はじめに

筆者らは2019年3月21日から22日にかけてオランダのハーグ市にあるオランダ国立図書館(KB)で開催された会議「文化遺産オンライン: 全世界的な文脈でのワールドデジタルライブラリー」⁽¹⁾に参加した。ワールドデジタルライブラリー(WDL)⁽²⁾プロジェクトの事務局を担当してきた米国議会図書館(LC)が、2019年末で担当を終了する決定を受け、LCに代わって事務局を担当する研究図書館センター(CRL)が今後の方向性を議論するために開催したものである。本会議は、米国のニューヨーク・カーネギー財団及びホワイティング財団、カタール国立図書館(QNL)の支援と、図書館情報資源振興財団(CLIR)と欧州の図書館・博物館・文書館のコレクションをデジタル提供するヨーロピアーナ財団の協力を得て開催された。

この会議の参加者の一部は、カーネギー財団の資金により招待された。参加者は約70人であり、約半数がWDLに参加している国立図書館の代表であった。その他は米・スタンフォード大学図書館の技術者、ヨーロピアーナの職員などが目立った。



写真 WDL会議の様子(筆者(時実)撮影)

本稿では会議で行われた議論等を紹介するとともに、グローバルレベルでのデジタル文化資源の共有に係る今後の動向について論じることとする。

2. 背景

2.1. WDLとは

WDLとは、LCとUNESCO等が2005年から取り組んできたプロジェクトで、2009年4月21日にウェブサイトが公開された(E912参照)。

もともとこの構想は、2005年6月に、当時のLC館長のピリントン(James Billington)氏が米国UNESCO国内委員会総会の挨拶の中で提案したことに始まっている。このプロジェクトには立ち上げ当初、LCの他、エジプトのアレクサンドリア図書館、サウジアラビアの王立科学技術大学、ブラジル国立図書館、エジプト国立図書館・公文書館、ロシア国立図書館などが協力している⁽³⁾。2006年7月6日・7日に金沢工業大学で開かれた「図書館・情報科学に関する国際ラウンドテーブル会議」ではWDLがテーマとして取り上げられ⁽⁴⁾、LC准館長のマーカム(Deanna Marcum)氏がその構想について講演している。

WDLは「国際および異文化間の相互理解の促進、教育関係者への情報資源の提供」を目的とし、また「発展途上国における情報資源デジタル化の促進、インターネット上での非英語、非西洋世界の情報資源の拡大」に力点を置いている⁽⁵⁾。サービスの特徴は、(1) 世界中の重要な文化遺産をインターネットで公開し、(2) 各資料には解説が付され、英語、フランス語、スペイン語、ポルトガル語、ロシア語、中国語およびアラビア語の7か国語で利用できること、である。

2019年9月現在、WDLには、60か国・地域158機関が協力しており、1万9,147点のコンテンツが収容されている⁽⁶⁾。WDLはコンテンツの規模に比して閲覧数が多いことが特徴で、2017年度には710万人が訪問し、ページビュー数は3,110万件に上ったと報告されている⁽⁷⁾。

ヨーロピアーナ(CA1785、CA1863参照)、米国デジタル公共図書館(DPLA:CA1857参照)、あるいは2019年2月に試験版が公開されたジャパンサーチ(E2176参照)などの参加機関のメタデータのみを集約したポータル型のサービスと異なり、WDLはデジタル・コンテンツそのものを収容している点で、リポジトリの性格が強い。コンテンツの内訳は図1のとおりである⁽⁸⁾。

日本からは国立国会図書館(NDL)が2008年に参加契約を締結し、現在、累積で235点のコンテンツを提供している。

*東京大学大学院情報学環

†国立情報学研究所コンテンツ科学研究系

‡電子情報部電子情報企画課

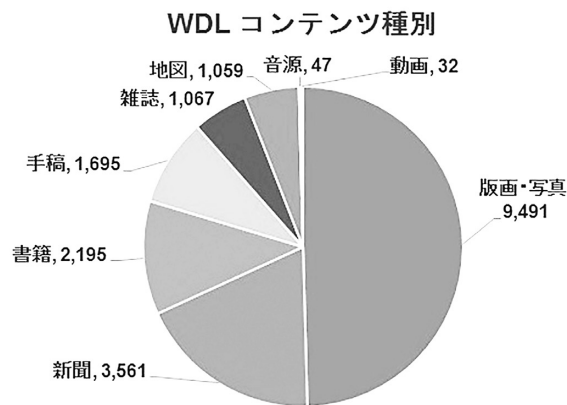


図1 WDLのコンテンツ種別(2019/8/20 現在)
出典: Collection Statisticsに基づき筆者作成
<https://www.wdl.org/en/statistics/>, (参照2019-08-20)。

LCが運用をやめる決定をした理由は示されていないが、年間約300万ドルとされる経費負担(システム運用と改良に加えて、メタデータの付与、7か国語への翻訳、開発途上国のデジタルセンターの技術指導、などを含む)が困難になったためと考えられる。

2.2. 米・スタンフォード大学図書館報告書

この会議に先立ち、米・スタンフォード大学図書館は、CRLからの委託で、WDLの現状分析を行って報告書をまとめている⁽⁹⁾。その要点は以下のとおりである。

まず、現行のWDLの強みとして、

- (1) 非常に高品質なデジタルコンテンツ
- (2) (点数は少ないが)トラフィックの多さ
- (3) デジタル化の能力の乏しい地域からのコンテンツを含む
- (4) すべての解説の7か国語での翻訳・提供(図2)
- (5) 典拠コントロールされた高品質のメタデータの提供
- (6) テキスト読み上げツール(ReadSpeaker)など豊富な機能
- (7) ネットワークでの世界へのコンテンツの配信
- (8) 統計や機能の充実
- (9) LCによるサポートの充実

等を指摘している。

一方で問題点としては、

- (1) 運用、技術、キュレーションなどLCへの過大な負担
- (2) 外部向けのデポジット、キュレーションといったツールの未整備(LCが単独で実施)
- (3) 寄せられるコンテンツの品質の大きなばらつき
- (4) LCの知的財産権に過度に敏感という性格が理

由ではじかれるコンテンツの多さ

- (5) 有用なものの高コストな多言語翻訳
- (6) より少数の機関から、より高品質のコンテンツを収集したほうが、量的にも質的にもよくなるというLC内部の意見
- (7) コンテンツとメタデータの収集において、ローカルなデータベースにコンテンツを分散してIIIFを使って閲覧するFast Trackと、WDLのリポジトリにキュレーションしたコンテンツを収めるSlow Trackに分けた方がよいというLC内部の意見
- (8) WDLの運用方式の、キュレーションや翻訳が完成してから公開する方式から、まず公開して後で内容を充実させていく方法への変更
- (9) 説明文、メタデータ、典拠コントロール、翻訳、などについては、センターで行うのではなく、クラウド方式、クラブ方式などで外部の力を借りる方式も考えられるという意見
- (10) より戦略的なコンテンツ収集の必要性

等をあげている。



図2. "South: The Story of Shackleton's Last Expedition, 1914-17" のロシア語ページ
<https://www.wdl.org/ru/item/7309/>, (accessed 2019-08-20)。

2.3. Pangia 構想

この会議で提案されたPangiaは、この会議に協力しているCLIRの会長であるヘンリー(Charles Henry)氏の構想である⁽¹⁰⁾。ヘンリー氏によれば、世界中の数十億のコンテンツをデジタルで提供することにより、市民の情報要求や教育者・研究者に役立つだけでなく、人工知能(AI)、ディープ・ラーニング、データマイニングなどのデジタル・イノベーションを促進することを期待したものである。ヘンリー氏の考えでは、Pangiaを構築するには、まず既存の選択されたプラットフォームの連携(federate)から始めることになる。その候補としては、ヨーロッパアーナ、DPLA、ニュージーランド国

立図書館のDigitalNZ⁽¹¹⁾、ブラジルの学術機関が運営するTainacan⁽¹²⁾、世界中に散らばる中東の文化遺産をデジタル化して統合的に活用することを目指す中東デジタルライブラリ(Digital Library of the Middle East)(開発中)⁽¹³⁾、インド・人材開発省が構築するインド国立デジタル図書館(National Digital Library of India)(開発中)⁽¹⁴⁾などをあげている。

3. 会議の概要

会議のプログラムの詳細はウェブサイトを参照されたい⁽¹⁵⁾。会議は概ね、WDLの背景に関する関係機関や各国の報告、WDLの現状と対策に関する報告、これを踏まえたグループ討論から構成されていた。本稿では、WDLの現状と対策に関する報告と議論についてまず述べ、後段で関係機関や各国の報告についても紹介する。

4. WDLの現状と対策に関する報告と討議

4.1. WDLの問題点と可能なモデルの提示

スタンフォード大学図書館のクラマー(Tom Crammer)氏から、WDLについて総括的な報告があった。報告の前半は先に紹介した報告書と同一である。この現状分析をベースとして、クラマー氏は次の3つのモデルを提案した。

(1) セントラル・ハブ型

従来型で、WDLが参加機関からすべてのコンテンツとメタデータを集める

(2) メタデータ・アグリゲータ型

ヨーロッパナ、DPLAのモデルで、WDLはメタデータのみを集める

(3) ネットワーク・ノードとスーパー・ノードの併用

WDLは大きな機関(スーパー・ノード)からはメタデータのみを集め、小さい機関からはコンテンツとメタデータの両方を集める。スーパー・ノードは他の機関からコンテンツを集めてWDLのアクセスに応じる

そのうえで、どのモデルをとるにせよ、次のような段階を踏んで進める必要があるとした。

- (1) 現在保有するメタデータ、コンテンツを新しい技術のもとにLC以外の場所に移動する
- (2) 新規コンテンツについては、提供とキュレーションのワークフローを見直す
- (3) 新機能と新規コンテンツでWDLの内容を強化する

4.2. 各国・地域等の取組

これらの問題提起を受け、参加者は3グループに分かれ、各グループが技術問題、ガバナンス問題、財政と持続可能性問題について順番に議論した。グループは、

大きく国立図書館系グループ(福林)、研究機関系グループ(高野)、その他(時実)、に分かれた(カッコ内はそのグループに参加した執筆者)。討議結果を各テーマのチューターがまとめて報告した。

4.3. 参加者の討議のまとめ

4.3.1. 技術問題

WDLの今後のオプションとして、セントラル・ハブ型の中央集権的なアプローチは支持されていないことが明らかになった。コストを減らすには、ワークフローの見直しが必要であり、それに技術の導入ができないかという議論があった。一方、WDLの価値としては、開発途上国に対する技術指導が大きいという意見であった。メタデータの自動付与や自動翻訳が考えられるが、人手が残るので効果は不明である。

WDLとPangiaの関係については、前者はリポジトリ、後者は便利なプラットフォームという分け方が述べられた。

また、現在行われている7か国語への翻訳についてはコストダウンが必要である。クラウド翻訳の可能性も提案されたが、品質管理の問題もある。

Pangiaのような大きなスケールになると、技術標準の管理も容易でない。たとえばヨーロッパナのデータモデルはメタデータの標準化に大きく役立っているが、参加者が増えるとこれをどうメンテナンスしていくか問題となる。

4.3.2. 財政問題と持続性問題

前述したように、WDLの運営には年300万ドルかかっているが、移行後は年350万ドル弱かかると考えられる。その内訳はAmazon Web Serviceに20万ドル、最低必要なプログラムに15万ドル、移行に10万ドル、編集、メタデータ作成、データの取込などの作業で300万ドルかかる。この費用を誰が出してくれるか。先進国の国立図書館、財団などが考えられる。

DPLAは年350万ドルの経費がかかっている。ヨーロッパナのフェアヴァイヤン(Harry Verwayen)氏によればヨーロッパナの予算は年500万ユーロである。しかし、アグリゲータに配分している費用を入れると合計1,000万ユーロに達する。これには各アグリゲータが得ている助成金は含まれない。また運営資金を財団に依存するのは、寄付が単年度単位であるなどのリスクがある。一方助成を新規技術の開発に使う場合は、完成した技術を皆で共有できるので、そのような資金の使い方が好ましい。

4.3.3. ガバナンス

WDLの独自の任務は何かという議論があった。

WDLは独自のデジタルライブラリーを持たない国にとってはその国のリポジトリとしての役割を果たす。一方で、各国独自のデジタルライブラリーと競合することはよくないとの意見があった。その他、デジタル化による、かつての植民地から略奪した文化遺産をデジタル化して公開することで旧植民地からも閲覧できるようにするヴァーチャルな返還はWDLの重要な任務であると考えられる。またWDLには個別のデジタルライブラリーのショーケースという役割がある。

現在のWDLのガバナンスでは、頻繁な会議、積極的でない会員など、いろいろ問題を抱えている。しかし、ガバナンス・ストラクチャについては、軽い方がよいとの意見であった。どのモデルを採用するかによって、ガバナンスは異なる。アグリゲータモデルではガバナンスは小さくなるが、アグリゲータとの調整が負担になる。

4.4. 今後の方向性

全体のまとめは、CRL代表(当時)のライリー(Bernard F. Reilly)氏が行った。技術・財政・ガバナンスといった大変大きな問題を議論したが、未来のシステムができるまでWDLを閉じることはできないので、移行期間が必要である。今後、会議主催者であるCRLが議論をまとめて公表するとのことであった。

議論では、特に非欧米から参加している国立図書館から、WDLは自国のコンテンツを発信し、またデジタル化やメタデータの技術を身に着ける上で重要であるとの意見があった。またヴァーチャル返還に意義を見出している様子であった。

今回の会議は、WDLからのLCの撤退による当面の措置(別機関と別システムへの移行)について参加国立図書館の理解を得ておくという意味が大きかったと思われる。

Pangiaについては、構想が漠然としていることもあり、参加した国立図書館の強い関心を集めるには至らなかった感がある。

5. 関連する報告

参加者の討議の前に、国際的、または各国の実践例について報告があった。大変参考になったので、以下に簡単に紹介する。

5.1. 挨拶

(1) クニベラー(E.J.B. Knibbeler)氏(オランダ国立図書館(KB))

KBでは文化遺産のオンライン提供を一貫して進めてきた。そのための仕組みとしてNetwork Digital Heritageが設立された。技術としてはLinked Open Dataを採用する考えである。

(2) ライリー氏(研究図書館センター(CRL))

この会議はWDLの将来の形を検討するものである。デジタル文化遺産プロジェクトは商業的であってはならない。またプロジェクトは透明であり、また内容は信頼できるもので、かつキュレーションされていることが好ましい。

5.2. 基調講演

バンダ(Fackson Banda)氏(UNESCO世界の記憶プログラム)

UNESCOでは「世界の記憶(Memory of the World)」プログラムを行っている⁽¹⁶⁾。

これは1992年に設立されたプログラムで、世界文化遺産に準じて文書遺産を登録している(2019年3月現在527件登録)。

5.3. IFLA Global Vision

ワイバー(Stephen Wyber)氏(国際図書館連盟(IFLA))

IFLAでは2018年にGlobal Visionを取りまとめたが(E2034、E2037参照)、その中で、期待される活動の9番目に「図書館は世界の記録遺産へのアクセスを最大限に広げる必要がある」としている⁽¹⁷⁾。

5.4. 個別報告

5.4.1. 国際的な活動

(1) フランス語圏デジタルネットワーク(Bibliothèque Réseau Francophone Numérique)

サニャ(Bruno Sagna)氏(フランス国立図書館(BnF))

フランス語圏デジタルネットワーク(Réseau francophone numérique (RFN:エレファントと発音))⁽¹⁸⁾はフランス語圏をターゲットとしたデジタルアーカイブである。27か国、その他の合計88か国でフランス語が用いられている。アフリカのフランス語圏は人口が増大しており、フランス語圏は拡大しつつある。RFNの運用はBnFがそのGallicaシステム(CA1905参照)を使って行っている。利用者の多くはフランス本国である。

(2) 中東デジタルライブラリ(Digital Library of the Middle East)

ヘンリー(Charles Henry)氏(図書館情報資源振興財団(CLIR))

世界中に散らばる中東の文化遺産をデジタル化して統合的に活用することを構想している。文化遺産のヴァーチャル返還の一環である。

(3) ヨーロピアーナ

フェアヴァイヤン氏(ヨーロピアーナ財団)

ヨーロピアーナの活動はコア・サービス(Horizon 1)、

コア・サービスの改善と拡張(Horizon 2)、新規なイノベーション(Horizon 3)の3方面で進めている。欧州では全コンテンツの10%しかデジタル化されておらず、そのうち36%がオンラインで提供されており、さらに7%(ヨーロッパのコンテンツの23%)しか再利用が認められていない。コア・サービスの改善としては、IIIFを推進している。その応用としては、Europeana Newspapers サービス(EI741参照)の改善がある。

5.4.2. 国別の活動

(4) インド国立デジタル図書館(National Digital Library of India)

ダス(Partha Pratim Das)氏(カラグブル・インド工科大学インド国立デジタル図書館プロジェクト)

インドは欧州より歴史が古く、さまざまな文化遺産がある。しかし多くが破壊されたり失われたりしている。これらをデジタル修復するところみを行っている。3Dによるバーチャル・ツアーも開発している。オーストラリアの大学の技術支援がある。

(5) イスラエル国立図書館デジタル図書館(National Library of Israel Digital Library)

ドイッチャー(Yaron Deutscher)氏(イスラエル国立図書館)

イスラエル国立図書館のデジタル図書館⁽¹⁹⁾はイスラエル国内だけでなく、世界中のユダヤ人を対象としている。現在新聞記事300万ページ、歴史的文書250万件、写真25万件、手稿6万件、書籍2万件、音楽4万時間、地図1万点などをデジタル化した。

(6) ウガンダ国立図書館デジタル図書館(National Library of Uganda Digital Library)

ムヨムバ(Hellen Agnes Muyomba)氏(ウガンダ国立図書館(NLU))

2009年にWDLプロジェクトの一環として、NLUとLCの合意により、NLUにデジタル化センターを設置した。このプロジェクトにより、ウガンダの各地、各機関に散らばっている文書などが集められ、保存が図られた。ウガンダは政情が不安定であり、これが問題である。

(7) カタールデジタル図書館(Qatar Digital Library)

ライリー(Susan Reilly)氏(カタール国立図書館(QNL))

カタールは2012年に教育研究図書館を発展させて国立図書館とした。すでに手稿や地図をデジタル化しており、一部をWDLにも提供している。さらに英国外務省の18世紀から20世紀のアラビア湾地域に関する文書、インド省の文書や英国図書館(BL)のイスラム関係の手稿の電子化を進めている⁽²⁰⁾。

6. おわりに

日本でもジャパンサーチがようやく試行の段階に入り、デジタルアーカイブのコンテンツ充実と利活用のためのネットワーク構築をどのように推進すべきかの議論が始まっている。WDLの今後の行方については不確かなところも多いが、世界規模での貴重な試行例であり、今後の進め方を含めて学ぶべきことが多い。引き続き注視していきたい。

- (1) Cultural Heritage Online: The World Digital Library in a Global Context. Center for Research Libraries. <https://www.crl.edu/wdl-planning>. (accessed 2019-09-30).
- (2) World Digital Library. <https://www.wdl.org/en/>. (accessed 2019-09-30).
- (3) 田中久徳. ワールドデジタルライブラリー—文化遺産を一望する. 国立国会図書館月報. 2009, (580), p. 4-9. <https://doi.org/10.11501/1001547>. (参照 2019-11-14).
- (4) 「情報管理」事務局. 集会報告: 図書館・情報科学に関する国際ラウンドテーブル会議: ワールド・デジタル・ライブラリー—その挑戦と目標—. 情報管理. 2006, 49(6), p. 348. <https://doi.org/10.1241/johokanri.49.348>. (参照 2019-11-14).
- (5) 田中. 前掲.
- (6) “Collection Statistics”. World Digital Library. <https://www.wdl.org/en/statistics/>. (accessed 2019-09-30).
- (7) Annual Report of the LIBRARIAN OF CONGRESS For the Fiscal Year Ending September 30, 2017. LC. 2018, p. 24-25. <http://www.loc.gov/static/portals/about/reports-and-budgets/documents/annual-reports/fy2017.pdf>. (accessed 2019-09-30).
- (8) “Collection Statistics”. World Digital Library. <https://www.wdl.org/en/statistics/>. (accessed 2019-09-30).
- (9) “World Digital Library: A Technical Assessment by Stanford Libraries”. Center for Research Libraries. 2019-03. https://www.crl.edu/sites/default/files/attachments/pages/WDL_Tech_analysis.pdf. (accessed 2019-08-20).
- (10) Alisa Alering. The infinite library. ScienceNode. 2018-05-21. https://sciencenode.org/feature/The_infinite_library.php. (accessed 2019-09-30).
- (11) DigitalNZ. <https://digitalnz.org/>. (accessed 2019-11-14).
- (12) Tainacan. <https://tainacan.org/>. (accessed 2019-11-14).
- (13) Digital Library of the Middle East. <https://dlme.clir.org/>. (accessed 2019-11-14).
- (14) National Digital Library of India. <https://ndliitkgp.ac.in/>. (accessed 2019-11-14).
- (15) Cultural Heritage Online: The World Digital Library in a Global Context. Center for Research Libraries. <https://www.crl.edu/wdl-planning>. (accessed 2019-09-30).
- (16) Memory of the World. UNESCO. <https://en.unesco.org/programme/mow>. (accessed 2019-08-20).
- (17) IFLA Global Vision. IFLA. <https://www.ifla.org/node/11900>. (accessed 2019-08-20).
- (18) Réseau francophone numérique. <http://rfnum-bibliotheque.org/rfn/>. (accessed 2019-08-20).
- (19) Digital Library. National Library of Israel. <http://web.nli.org.il/sites/NLI/English/digitallibrary/Pages/default.aspx>. (accessed 2019-08-20).
- (20) Claudia Lux. Qatar National Library — Architecture as innovation in the Arab world. IFLA Journal. 2014, 40(3), p. 174-181. <https://doi.org/10.1177/0340035214546984>. (accessed 2019-11-14).

[受理:2019-11-18]

Tokizane Soichi

Takano Akihiko

Fukubayashi Yasuhiro

Where will the World Digital Library Head?