Bib46W-01(私の発表論文等)　  
文化的資産の次世代への継承とAIを活用した利用に向けた活動履歴

中山正樹（元国立国会図書館　専門調査員）

目次

[(私の発表論文等)　 文化的資産の次世代への継承とAIを活用した利用に向けた活動履歴 1](#_Toc495070286)

[1. （2002年3月）国立国会図書館志望動機 29](#_Toc495070287)

[2. （2003年）e-JAPAN戦略を踏まえたNDLの活動の方向性（メモ） 30](#_Toc495070288)

[2.1. デジタル・アーカイブの推進に向けた申入れの骨子（ＮＤＬ関連部分） 30](#_Toc495070289)

[2.1.1. 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進 30](#_Toc495070290)

[2.1.2. 「ジャパン・ウェブ・アーカイブ」構想の推進 30](#_Toc495070291)

[2.1.3. その他 30](#_Toc495070292)

[2.2. e-JAPAN戦略で各機関が実施すべきとされる部分 31](#_Toc495070293)

[2.2.1. 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進 31](#_Toc495070294)

[2.2.2. 「ジャパン・ウェブ・アーカイブ」構想の推進 31](#_Toc495070295)

[2.3. 関係府省の実施項目との関係 32](#_Toc495070296)

[2.3.1. 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進 32](#_Toc495070297)

[2.3.2. 「ジャパン・ウェブ・アーカイブ」構想の推進 32](#_Toc495070298)

[2.4. NDLの基本姿勢 32](#_Toc495070299)

[2.5. 当面の作業内容 33](#_Toc495070300)

[2.6. 当面のスケジュール 33](#_Toc495070301)

[2.7. NDLが自ら実施すべき事項 33](#_Toc495070302)

[2.8. 要求機能要件 34](#_Toc495070303)

[2.9. 16年度予算要求に関して 34](#_Toc495070304)

[2.10. 実施に向けた体制整備 35](#_Toc495070305)

[3. （2004年）国立国会図書館「電子図書館中期計画2004」の実施に向けて 38](#_Toc495070306)

[3.1. 要約 38](#_Toc495070307)

[3.2. キーワード 38](#_Toc495070308)

[3.3. はじめに 38](#_Toc495070309)

[3.4. 電子図書館中期計画2004策定の背景 38](#_Toc495070310)

[3.5. 国立国会図書館電子図書館中期計画2004の骨子 38](#_Toc495070311)

[3.5.1. デジタルアーカイブの構築 38](#_Toc495070312)

[3.5.2. 情報資源に関する情報の充実 39](#_Toc495070313)

[3.5.3. デジタルアーカイブのポータル機能 39](#_Toc495070314)

[3.6. デジタルアーカイブ・ポータルの構築に当たっての考え方 39](#_Toc495070315)

[3.7. デジタルアーカイブ・ポータルが目指す方向 39](#_Toc495070316)

[3.7.1. 情報探索の行動パターンと提供すべきサービス 39](#_Toc495070317)

[3.7.2. 情報探索行動における図書館の方向性 40](#_Toc495070318)

[3.8. デジタルアーカイブ・ポータルの利用イメージ 40](#_Toc495070319)

[3.8.1. 情報資源に到達するための情報の充実 40](#_Toc495070320)

[3.8.2. デジタルアーカイブへのアクセス 40](#_Toc495070321)

[3.9. ポータルから見たデジタルアーカイブの要件 41](#_Toc495070322)

[3.9.1. コンテンツ仕様 41](#_Toc495070323)

[3.9.2. インターフェース仕様 41](#_Toc495070324)

[3.9.3. Webサービスとしての連携 43](#_Toc495070325)

[3.10. デジタルアーカイブ・ポータルの構築・運用に関しての考察 43](#_Toc495070326)

[3.10.1. コンテンツへの直接ナビゲーションが必要 43](#_Toc495070327)

[3.10.2. ナビゲーションすべき情報へは確実に 43](#_Toc495070328)

[3.10.3. 情報提供者の意志を尊重する 43](#_Toc495070329)

[3.10.4. 誰でも付加価値をつけて運営できるように 43](#_Toc495070330)

[3.10.5. データプロバイダは、情報収集ソフトからのアクセシビリティを確保 44](#_Toc495070331)

[3.10.6. 考察のまとめ 44](#_Toc495070332)

[3.11. デジタルアーカイブおよびポータル構築に向けて 44](#_Toc495070333)

[3.11.1. 公的機関のデジタルアーカイブ構築の推進 44](#_Toc495070334)

[3.11.2. デジタルアーカイブ・ポータルの推進 44](#_Toc495070335)

[3.11.3. ウェブ・アーカイブ構想の推進 45](#_Toc495070336)

[3.12. おわりに 45](#_Toc495070337)

[3.12.1. 参考文献 45](#_Toc495070338)

[4. （2004年）デジタルアーカイブの構築イメージとして想定したこと 47](#_Toc495070339)

[4.1. 2004年に想定したこと 47](#_Toc495070340)

[4.1.1. デジタルアーカイブ・ポータルの構築に当たっての考え方 47](#_Toc495070341)

[4.1.2. 情報探索行動における図書館の方向性 47](#_Toc495070342)

[4.1.3. デジタルアーカイブへのアクセス 47](#_Toc495070343)

[4.2. ポータルから見たデジタルアーカイブの要件 48](#_Toc495070344)

[4.2.1. コンテンツ仕様 48](#_Toc495070345)

[4.2.2. インターフェース仕様 49](#_Toc495070346)

[4.2.3. Webサービスとしての連携 50](#_Toc495070347)

[4.3. デジタルアーカイブ・ポータルの構築・運用に関しての考察 50](#_Toc495070348)

[4.3.1. コンテンツへの直接ナビゲーションが必要 50](#_Toc495070349)

[4.3.2. ナビゲーションすべき情報へは確実に 51](#_Toc495070350)

[4.3.3. 情報提供者の意志を尊重する 51](#_Toc495070351)

[4.3.4. 誰でも付加価値をつけて運営できるように 51](#_Toc495070352)

[4.3.5. データプロバイダは、情報収集ソフトからのアクセシビリティを確保 51](#_Toc495070353)

[4.3.6. 考察のまとめ 51](#_Toc495070354)

[4.4. デジタルアーカイブおよびポータル構築に向けて 51](#_Toc495070355)

[4.4.1. 公的機関のデジタルアーカイブ構築の推進 51](#_Toc495070356)

[4.4.2. デジタルアーカイブ・ポータルの推進 52](#_Toc495070357)

[4.4.3. ウェブ・アーカイブ構想の推進 52](#_Toc495070358)

[4.5. おわりに 52](#_Toc495070359)

[5. （2005年）各種デジタルアーカイブが提供するコンテンツの統合検索を目指して 54](#_Toc495070360)

[5.1. 背景 54](#_Toc495070361)

[5.2. 当館が目指すポータルのサービスは？ 54](#_Toc495070362)

[5.3. デジタルアーカイブのコンテンツをポータルから検索できるようにするために 55](#_Toc495070363)

[5.3.1. ポータル成功の鍵はメタデータの収集 55](#_Toc495070364)

[5.4. 16年度プロトタイプの概要 55](#_Toc495070365)

[5.4.1. どんなことができるのか 55](#_Toc495070366)

[5.4.2. 構築方針 55](#_Toc495070367)

[5.4.3. 適用技術 56](#_Toc495070368)

[5.5. デジタルアーカイブおよびポータル構築に向けて 56](#_Toc495070369)

[5.5.1. 公的機関のデジタルアーカイブ構築の推進 56](#_Toc495070370)

[5.5.2. デジタルアーカイブポータルの推進 56](#_Toc495070371)

[5.6. おわりに 57](#_Toc495070372)

[6. （2006年）Web2.0世代のデジタルアーカイブポータルの提供を目指して 58](#_Toc495070373)

[6.1. 要旨 58](#_Toc495070374)

[6.2. キーワード 58](#_Toc495070375)

[6.3. はじめに 58](#_Toc495070376)

[6.4. プロトタイプシステムの構築 59](#_Toc495070377)

[6.4.1. プロトタイプの各機能 60](#_Toc495070378)

[6.4.2. 統合検索機能の実装と「共通仕様」 61](#_Toc495070379)

[6.4.3. プロトタイプでの検証結果 62](#_Toc495070380)

[6.5. Web2.0世代のポータルを目指して 63](#_Toc495070381)

[6.5.1. メタデータの取得 63](#_Toc495070382)

[6.5.2. RSSの活用 63](#_Toc495070383)

[6.5.3. ソーシャルブックマーク 64](#_Toc495070384)

[6.5.4. 辞書の活用 64](#_Toc495070385)

[6.5.5. メタデータの自動生成，組織化，保存 64](#_Toc495070386)

[6.5.6. メタデータ提供とインタフェース 64](#_Toc495070387)

[6.5.7. ウェブサービスによる連携 65](#_Toc495070388)

[6.5.8. Blog, wiki 65](#_Toc495070389)

[6.5.9. システム基盤 65](#_Toc495070390)

[6.6. 終わりに 66](#_Toc495070391)

[7. （2006年）国立国会図書館における電子図書館サービスのゆくえ 68](#_Toc495070392)

[7.1. はじめに 68](#_Toc495070393)

[7.2. 資料のデジタル化と電子展示会の提供 68](#_Toc495070394)

[7.3. パッケージ系の収集・保存・提供 69](#_Toc495070395)

[7.4. ネットワーク系の収集・保存・提供 69](#_Toc495070396)

[7.5. デジタルアーカイブポータルの開発 70](#_Toc495070397)

[7.6. おわりに 70](#_Toc495070398)

[8. （2007年）デジタル情報も含めた統合検索サービスの提供を目指して 71](#_Toc495070399)

[9. （2008年）DA/PORTAの今後（将来像）【骨子】 72](#_Toc495070400)

[9.1. 概要 72](#_Toc495070401)

[9.1.1. 現状認識 72](#_Toc495070402)

[9.1.2. 何をすべきか（Web2.0時代の次を見据えて） 72](#_Toc495070403)

[9.1.3. どのように実施するか 72](#_Toc495070404)

[9.2. 何をすべきか（Web2.0時代の次を見据えて） 73](#_Toc495070405)

[9.2.1. データベース内の情報の「見える化」（可視化） 73](#_Toc495070406)

[9.2.2. 情報を全体で集合知化 73](#_Toc495070407)

[9.2.3. 集合知の永久保存 73](#_Toc495070408)

[9.2.4. 情報の組織化 74](#_Toc495070409)

[9.2.5. 利用者情報の組織化 74](#_Toc495070410)

[9.2.6. 情報と人の関係付け 75](#_Toc495070411)

[9.2.7. 情報検索手段（利用方法）の変革 75](#_Toc495070412)

[9.2.8. 利用のまとめ 76](#_Toc495070413)

[9.3. どのように実施するか 76](#_Toc495070414)

[9.3.1. NDL-DA、PORTAシステム開発における技術導入の考え方 76](#_Toc495070415)

[9.3.2. デジタルアーカイブシステムの構築に必要な技術の適用 77](#_Toc495070416)

[9.4. 国のデジタルコレクションの構築に向けた連携協力 79](#_Toc495070417)

[9.4.1. 主な連携協力先 79](#_Toc495070418)

[9.4.2. 政府の施策（2008年度） 81](#_Toc495070419)

[9.4.3. 公共図書館支援のイメージ 82](#_Toc495070420)

[9.4.4. 学術機関との連携協力 83](#_Toc495070421)

[9.4.5. 商用サイトとの連携協力 84](#_Toc495070422)

[9.4.6. Google,Yahoo等との連携の可能性 84](#_Toc495070423)

[9.4.7. 出版社、著作者等との連携協力 84](#_Toc495070424)

[9.4.8. MLA連携 85](#_Toc495070425)

[9.4.9. 歴史的音盤アーカイブの構築支援 86](#_Toc495070426)

[9.4.10. 日中韓の国立図書館との連携 87](#_Toc495070427)

[9.4.11. World Digital Library構想との連携 88](#_Toc495070428)

[9.5. PORTA 88](#_Toc495070429)

[9.5.1. PORTAのAPI提供準備状況 88](#_Toc495070430)

[9.5.2. PORTAの20FY開発の状況 89](#_Toc495070431)

[9.5.3. PORTAの21～22年度の方向性 89](#_Toc495070432)

[9.6. まとめ 91](#_Toc495070433)

[9.6.1. 5年程度までの方向性（クラウドコンピュータの世界の普及をイメージして） 91](#_Toc495070434)

[9.6.2. デジタルアーカイブの今後の構築計画（想定） 92](#_Toc495070435)

[10. （2008年）今後実施すべきことと連携協力（図書館総合展） 93](#_Toc495070436)

[10.1. 挨拶 93](#_Toc495070437)

[10.2. 本日お話しすること 93](#_Toc495070438)

[10.3. 現状の再認識 94](#_Toc495070439)

[10.4. 実施すべきこと 94](#_Toc495070440)

[10.5. 実施すべきことと、必要な機能の関連 94](#_Toc495070441)

[10.5.1. データベース内の情報の「見える化」（可視化） 95](#_Toc495070442)

[10.5.2. 情報を全体で集合知化 95](#_Toc495070443)

[10.5.3. 集合知の永久保存 96](#_Toc495070444)

[10.5.4. 情報の組織化、 96](#_Toc495070445)

[10.5.5. 利用者情報の組織化 97](#_Toc495070446)

[10.5.6. 情報と人の関係付け 98](#_Toc495070447)

[10.5.7. 情報検索手段（利用方法）の変革 98](#_Toc495070448)

[10.5.8. まとめ 99](#_Toc495070449)

[10.6. PORTAにおける「見える化」のイメージ 99](#_Toc495070450)

[10.6.1. NDL-DA、PORTAシステムの構築の考え方 99](#_Toc495070451)

[10.6.2. 情報の組織化の１つ。FRBRの概念での組織化による検索・閲覧サービスの実現（１ｍ） 100](#_Toc495070452)

[10.6.3. コンテンツ交換の方法の１つ（PeerToPeer）対等な関係でのデジタルアーカイブの構築 100](#_Toc495070453)

[10.7. デジタルアーカイブシステムの構築に向けた主な連携協力 101](#_Toc495070454)

[10.7.1. 産学官の研究機関との連携 101](#_Toc495070455)

[10.7.2. （連携協力）国際インターネット保存コンソーシアムとの連携 102](#_Toc495070456)

[10.7.3. 国のデジタルコレクションの構築に向けた主な連携協力 103](#_Toc495070457)

[10.7.4. 政府の施策（2008年度） 104](#_Toc495070458)

[10.7.5. 公共図書館支援のイメージ 106](#_Toc495070459)

[10.7.6. 関係機関拡大策の例公共図書館支援のスキーム 107](#_Toc495070460)

[10.7.7. 学術機関との連携協力 107](#_Toc495070461)

[10.7.8. 商用サイトとの連携協力 108](#_Toc495070462)

[10.7.9. Google,Yahoo等との連携の可能性 108](#_Toc495070463)

[10.7.10. 出版社、著作者等との連携協力 109](#_Toc495070464)

[10.7.11. MLA連携 109](#_Toc495070465)

[10.7.12. 歴史的音盤アーカイブの構築支援 112](#_Toc495070466)

[10.7.13. 日中韓の国立図書館との連携 112](#_Toc495070467)

[10.7.14. World Digital Library構想との連携 113](#_Toc495070468)

[10.8. まとめ 113](#_Toc495070469)

[10.8.1. まずは、連携による組織を越えた統合検索 113](#_Toc495070470)

[10.8.2. 次世代は、人と情報が関係付けられたサービスの連携 113](#_Toc495070471)

[10.8.3. 5年程度までの方向性（クラウドコンピュータの世界の普及をイメージして） 114](#_Toc495070472)

[10.9. デジタルアーカイブの今後の構築計画（想定） 115](#_Toc495070473)

[10.10. おわりに 116](#_Toc495070474)

[11. （2009年）「トータルな図書館システムの実現」に向けて【館内調整】 117](#_Toc495070475)

[11.1. アプローチ 117](#_Toc495070476)

[11.1.1. 当館の事業戦略 117](#_Toc495070477)

[11.1.2. 検討の柱 117](#_Toc495070478)

[11.1.3. 館全体システム 117](#_Toc495070479)

[11.2. 当面のアクション 119](#_Toc495070480)

[11.2.1. IT戦略本部「３カ年緊急プラン」、21年度補正予算等も踏まえて 119](#_Toc495070481)

[11.2.2. 全体 119](#_Toc495070482)

[11.2.3. 全館システム 120](#_Toc495070483)

[11.2.4. 共通インフラ 120](#_Toc495070484)

[11.2.5. 情報探索サービス 120](#_Toc495070485)

[11.2.6. 次期基盤 121](#_Toc495070486)

[11.2.7. 現行システム 121](#_Toc495070487)

[11.3. 論点 122](#_Toc495070488)

[11.3.1. 全体イメージ 122](#_Toc495070489)

[11.3.2. 紙とデジタルの業務を統合する時期は？ 122](#_Toc495070490)

[11.3.3. 情報探索サービスの部分は、市販のパッケージでいいのか？ 122](#_Toc495070491)

[11.3.4. 目録サービスの範囲の拡大は？ 122](#_Toc495070492)

[11.3.5. 「情報システムユーザスキル標準」で示されたようなスキルを具備する人材の確保方法 123](#_Toc495070493)

[11.3.6. 「次期基盤システムの開発方針」の見直しの前提 123](#_Toc495070494)

[11.4. 総合目録の実現方式の方向性を踏まえて 123](#_Toc495070495)

[11.4.1. 全体イメージの再確認 123](#_Toc495070496)

[11.4.2. 今後の検討 123](#_Toc495070497)

[11.4.3. パッケージの候補 123](#_Toc495070498)

[12. （2009年）次期基盤システムの開発方針について【館内調整】 125](#_Toc495070499)

[12.1. 背景 125](#_Toc495070500)

[12.2. 現行基盤システムの問題点 125](#_Toc495070501)

[12.3. 開発目的 125](#_Toc495070502)

[12.4. 基本方針 125](#_Toc495070503)

[12.5. 開発・運用予算の確保 127](#_Toc495070504)

[12.6. 次期基盤構築の体制確保 128](#_Toc495070505)

[12.7. 今年度中の作業 130](#_Toc495070506)

[12.8. 今後 130](#_Toc495070507)

[13. （2010年）国立国会図書館サーチと日韓自動翻訳の現況及び課題（日韓業務交流） 131](#_Toc495070508)

[13.1. はじめに 131](#_Toc495070509)

[13.2. 全体イメージ 131](#_Toc495070510)

[13.2.1. 国立国会図書館サーチ（NDL Search）とは 131](#_Toc495070511)

[13.2.2. 2012年1月でのサービスイメージ 131](#_Toc495070512)

[13.2.3. 2009年度開発版の到達点 132](#_Toc495070513)

[13.2.4. システム構成 133](#_Toc495070514)

[13.2.5. 2010年度、2011年度の開発予定 133](#_Toc495070515)

[13.3. 外部機関・サービスとの連携方針 133](#_Toc495070516)

[13.4. 実施中、実施予定の実証実験 134](#_Toc495070517)

[13.5. 日中韓翻訳実験 134](#_Toc495070518)

[13.5.1. 経緯 134](#_Toc495070519)

[13.5.2. 2009年度システムでの実装 134](#_Toc495070520)

[13.5.3. 2010年度システムでの実装 135](#_Toc495070521)

[13.5.4. 翻訳サービスの流れ（別紙PPT） 135](#_Toc495070522)

[13.5.5. 日中韓三か国による検索、翻訳の実験（提案） 136](#_Toc495070523)

[13.6. ポータルの相互運用性を高めるために 137](#_Toc495070524)

[13.7. 情報の利活用の促進を目指して（課題と今後） 137](#_Toc495070525)

[14. （2011年）電子図書館構想と日本の学術デジタルコミュニケーションの現状（e-Japanologyの構築に向けて） 138](#_Toc495070526)

[14.1. はじめに 138](#_Toc495070527)

[14.2. 電子図書館構想 138](#_Toc495070528)

[14.2.1. 当館蔵書のデジタル化 138](#_Toc495070529)

[14.2.2. 学位論文のデジタル化 139](#_Toc495070530)

[14.3. 知識インフラの構築と提供に向けて 139](#_Toc495070531)

[14.3.1. 知識インフラの必要性 139](#_Toc495070532)

[14.3.2. 近い将来に取り組むべき事項 139](#_Toc495070533)

[14.4. 知識インフラのノードとしての社会的な機能の展開 139](#_Toc495070534)

[14.4.1. NDL Searchとは 140](#_Toc495070535)

[14.4.2. NDL Searchが当面目指す方向性 140](#_Toc495070536)

[14.4.3. NDL Searchの将来像 140](#_Toc495070537)

[14.4.4. NDL Searchのサービスイメージ 140](#_Toc495070538)

[14.4.5. NDL Searchでの連携協力イメージ 141](#_Toc495070539)

[14.5. 知識の利活用の促進に向けて 141](#_Toc495070540)

[15. （2011年）知の共有を目指して（国立国会図書館におけるデジタルアーカイブ構築） 143](#_Toc495070541)

[15.1. 要約 143](#_Toc495070542)

[15.2. キーワード 143](#_Toc495070543)

[15.3. はじめに 143](#_Toc495070544)

[15.4. NDLの役割 144](#_Toc495070545)

[15.5. 電子図書館構築のあゆみ 144](#_Toc495070546)

[15.6. 最新の取組み 144](#_Toc495070547)

[15.6.1. 業務・システム最適化の推進 145](#_Toc495070548)

[15.6.2. デジタル情報資源の収集蓄積・利用の推進 145](#_Toc495070549)

[15.7. NDL Search 147](#_Toc495070550)

[15.7.1. NDL Searchが当面目指す方向性 147](#_Toc495070551)

[15.7.2. NDLSearchのサービスイメージ 147](#_Toc495070552)

[15.7.3. システム構成 148](#_Toc495070553)

[15.8. 知識インフラの構築 148](#_Toc495070554)

[15.8.1. 経緯 148](#_Toc495070555)

[15.8.2. 知識インフラとは 148](#_Toc495070556)

[15.8.3. 知識の集約と提供のための連携イメージ 149](#_Toc495070557)

[15.8.4. 次世代システム開発研究の概要 149](#_Toc495070558)

[15.8.5. NDLラボの設置 150](#_Toc495070559)

[15.9. 震災アーカイブの構築 150](#_Toc495070560)

[15.9.1. 震災アーカイブの構築に当たって 150](#_Toc495070561)

[15.9.2. 震災に関する知識インフラの構築 150](#_Toc495070562)

[15.10. 電子情報関連の組織再編 151](#_Toc495070563)

[15.11. おわりに 151](#_Toc495070564)

[15.11.1. 参考文献 151](#_Toc495070565)

[16. （2012年）国立国会図書館サーチのコンセプト・開発経緯と今後の展開 153](#_Toc495070566)

[16.1. はじめに 153](#_Toc495070567)

[16.2. 「知識はわれらを豊かにする」を情報処理システムの観点からみて 153](#_Toc495070568)

[16.3. 当館の位置づけの再認識（2004年） 153](#_Toc495070569)

[16.4. NDLサーチと業務基盤システム 154](#_Toc495070570)

[16.5. NDLSearchができるまで 154](#_Toc495070571)

[16.6. NDLSearchができるまで情報探索サービスの将来像(クラウドの世界でのサービスの連携)（2009年） 155](#_Toc495070572)

[16.7. 次期図書館システムの全体イメージ（2009年3月） 156](#_Toc495070573)

[16.8. 書誌・所蔵の視点でのコンテンツの体系的整理の概念（2009年3月） 156](#_Toc495070574)

[16.9. サービス要件定義・システム化要件定義作成（2010年） 157](#_Toc495070575)

[16.10. 適用すべき技術標準の指針（一覧） 157](#_Toc495070576)

[16.11. サービス構築の基本要件（2009年） 158](#_Toc495070577)

[16.12. NDLサーチのシステム化要件 159](#_Toc495070578)

[16.13. NDLシステムのリニューアル 160](#_Toc495070579)

[16.14. リニューアルシステム群の概要 160](#_Toc495070580)

[16.15. NDLサーチをサービスの起点に 161](#_Toc495070581)

[16.16. NDL蔵書の検索・申込システムとしてのNDL-OPACの刷新 161](#_Toc495070582)

[16.17. 館内利用者端末の多機能化 162](#_Toc495070583)

[16.18. NDLサーチでの当面の連携イメージ 162](#_Toc495070584)

[16.19. NDLサーチの機能改善（2012年度） 163](#_Toc495070585)

[16.20. 知識インフラと震災アーカイブ 163](#_Toc495070586)

[16.21. 新たな知識の創造と還流 164](#_Toc495070587)

[16.22. 知識インフラ構築に向けた有識者 164](#_Toc495070588)

[16.23. NDLラボ（仮称）の設置 164](#_Toc495070589)

[16.24. 現状のNDLサーチをベースに震災アーカイブとして機能拡張（想定） 165](#_Toc495070590)

[16.25. まとめ 165](#_Toc495070591)

[16.25.1. 知識インフラ構築に向けたNDLサーチの今後の展開 165](#_Toc495070592)

[16.25.2. NDLサーチは、知識インフラのポータルに 165](#_Toc495070593)

[17. （2012年）電子書籍等のディジタルコンテンツの長期保存と、将来にわたっての利用の保証－文化的資産の保存に向けた関係機関との連携協力－ 167](#_Toc495070594)

[17.1. タイトル 167](#_Toc495070595)

[17.2. 要約 167](#_Toc495070596)

[17.3. 目指すところ 167](#_Toc495070597)

[17.4. 使命・経緯 167](#_Toc495070598)

[17.4.1. 経緯 167](#_Toc495070599)

[17.4.2. NDLの使命と目標 168](#_Toc495070600)

[17.5. 関係機関との連携の観点 168](#_Toc495070601)

[17.6. 収集の観点 168](#_Toc495070602)

[17.6.1. 「インターネット資料」の許諾に基づく収集 168](#_Toc495070603)

[17.6.2. 公的機関の「インターネット資料」の制度的収集 168](#_Toc495070604)

[17.6.3. 民間の「オンライン資料」の制度的収集 168](#_Toc495070605)

[17.7. 保存の観点 168](#_Toc495070606)

[17.7.1. 所蔵資料のデジタル化 168](#_Toc495070607)

[17.7.2. 文化的資産の保存 169](#_Toc495070608)

[17.8. 組織化の観点 169](#_Toc495070609)

[17.8.1. 組織化の意義 169](#_Toc495070610)

[17.8.2. 組織化の連携協力 170](#_Toc495070611)

[17.9. 電子書籍出版社等との連携 170](#_Toc495070612)

[17.9.1. NDLデジタル化コンテンツの二次利用の促進 170](#_Toc495070613)

[17.9.2. 電子書籍サイト等、商用サイトへの案内の強化 170](#_Toc495070614)

[17.9.3. 電子書籍ビジネスのプラットフォーム整備に協力 170](#_Toc495070615)

[17.9.4. 電子書籍フォーマットの共通化 170](#_Toc495070616)

[17.9.5. 電子書籍に対する永続的識別子の付与 171](#_Toc495070617)

[17.9.6. 公共図書館での利用環境の共通化 171](#_Toc495070618)

[17.9.7. 著作権管理センターの構築・運用の協力 171](#_Toc495070619)

[17.10. 新たな取組 171](#_Toc495070620)

[17.10.1. 東日本大震災アーカイブ 171](#_Toc495070621)

[17.10.2. 知の共有化に向けた連携 171](#_Toc495070622)

[17.11. 政府の施策との連携 172](#_Toc495070623)

[17.12. 関係機関を繋ぐ役割を果たす 172](#_Toc495070624)

[18. （2012年）国立国会図書館が取り組むデジタルデータの保存と未来のデータ保存 173](#_Toc495070625)

[18.1. 企画内容 173](#_Toc495070626)

[18.2. 事前質問 173](#_Toc495070627)

[18.3. 回答概要 173](#_Toc495070628)

[18.4. 回答 174](#_Toc495070629)

[18.4.1. クラウドなどでデータの保存が消失する事件がおき、改めて「データの保存」が注目されています。「データを保存すること」の重要性についてお考えをお聞かせいただければと思います。 174](#_Toc495070630)

[18.4.2. 国会図書館が現在取り組んでいる資料の保存、デジタルデータの保存の取り組みについてお聞かせ願えればと思います。また以前の取り組みをどのように発展させていくかについてお考えをお聞かせください。 174](#_Toc495070631)

[18.4.3. 海外の図書館における資料保存・デジタルデータ保存についてはどういう面を参考にしたいとお考えでしょうか。 175](#_Toc495070632)

[18.4.4. 現在のデジタルデータの保存技術への要望を実際に資料保存の現場にいるお立場からお聞かせ願えればと思います。 176](#_Toc495070633)

[18.4.5. 注目されているデジタルデータの保存技術があればお聞かせください。 176](#_Toc495070634)

[18.4.6. 今後の国会図書館の資料保存、デジタルデータの保存の展望をお聞かせください。 177](#_Toc495070635)

[19. （2012年）国立国会図書館と出版界の連携（日仏シンポジウム講演内容） 178](#_Toc495070636)

[19.1. はじめに 178](#_Toc495070637)

[19.2. 電子書籍の保存 178](#_Toc495070638)

[19.3. 資料や資源への一元的なアクセス 178](#_Toc495070639)

[19.4. デジタル化したコンテンツの活用 178](#_Toc495070640)

[19.5. 公共図書館へのデータ提供環境の共通化 179](#_Toc495070641)

[19.6. 著作権情報の集中管理 179](#_Toc495070642)

[19.7. おわりに 179](#_Toc495070643)

[20. （2012年）資料組織化部会報告【将来構想会議】 181](#_Toc495070644)

[20.1. 基本方針 181](#_Toc495070645)

[20.2. 具体的方向性 181](#_Toc495070646)

[20.2.1. サービスの利便性向上 181](#_Toc495070647)

[20.2.2. 媒体に関わらずメタデータを一元的に扱えるフレームワークの構築 182](#_Toc495070648)

[20.2.3. メタデータの網羅性・信頼性の向上 183](#_Toc495070649)

[20.2.4. メタデータ作成作業の効率化及び迅速化 184](#_Toc495070650)

[20.2.5. 国立図書館としての貢献 185](#_Toc495070651)

[20.2.6. 体制と資源配分の見直しに向けて 185](#_Toc495070652)

[21. （2013年）OpenGLAMセッションでの手持ち資料（図書館総合展） 188](#_Toc495070653)

[21.1. 私の理念 188](#_Toc495070654)

[21.2. 連携が進まない理由 188](#_Toc495070655)

[21.3. 本日の議論の方向性 188](#_Toc495070656)

[21.4. OpenGLAMとNDL（概要） 189](#_Toc495070657)

[21.4.1. 知識インフラの構築に向けて 189](#_Toc495070658)

[21.4.2. 国としてのデジタルアーカイブの構築 189](#_Toc495070659)

[21.4.3. デジタル情報資源ラウンドテーブル 189](#_Toc495070660)

[21.4.4. MLA連携に関連したNDLの活動 189](#_Toc495070661)

[21.4.5. 大震災とMLA 189](#_Toc495070662)

[21.4.6. 国のオープンデータ戦略 190](#_Toc495070663)

[21.4.7. 国としてのアーカイブの構築 190](#_Toc495070664)

[21.4.8. MLA連携で想定する方向性 190](#_Toc495070665)

[21.4.9. NDLの役割 191](#_Toc495070666)

[21.5. 当館の姿勢 191](#_Toc495070667)

[21.6. 国としてのアーカイブの構築 191](#_Toc495070668)

[21.6.1. 方向性 191](#_Toc495070669)

[21.7. ナショナルアーカイブとは 191](#_Toc495070670)

[21.7.1. 国としての知識の貯蔵庫 191](#_Toc495070671)

[21.7.2. 現在の活動との関係 192](#_Toc495070672)

[21.7.3. 関係機関が協力してWinWinの関係で構築 192](#_Toc495070673)

[21.8. デジタル文化資源ラウンドテーブルの設置と成果 192](#_Toc495070674)

[21.8.1. 背景 192](#_Toc495070675)

[21.8.2. デジタル情報資源ラウンドテーブルの設置 192](#_Toc495070676)

[21.8.3. 目的 192](#_Toc495070677)

[21.8.4. 状況 192](#_Toc495070678)

[21.8.5. 提示された課題 193](#_Toc495070679)

[21.8.6. まとめ 193](#_Toc495070680)

[21.9. デジタル文化資源ラウンドテーブルの今後 193](#_Toc495070681)

[21.9.1. 現状 193](#_Toc495070682)

[21.9.2. 次のステップは 193](#_Toc495070683)

[21.10. 次の協議の場の設置 193](#_Toc495070684)

[21.10.1. 設置の目的 193](#_Toc495070685)

[21.10.2. 主な協議事項 193](#_Toc495070686)

[21.11. 今後のMLA連携の概要 193](#_Toc495070687)

[21.11.1. MLA連携の方向性 193](#_Toc495070688)

[21.11.2. ノウハウの共有 193](#_Toc495070689)

[21.11.3. 活動の連携 194](#_Toc495070690)

[21.11.4. 資源の共有化 194](#_Toc495070691)

[21.11.5. 人材の交流 194](#_Toc495070692)

[21.12. 今後のMLA連携の具体的項目 194](#_Toc495070693)

[21.12.1. 課題抽出 194](#_Toc495070694)

[21.12.2. メタデータ 194](#_Toc495070695)

[21.12.3. コンテンツ作成（デジタル化等）と利活用 194](#_Toc495070696)

[21.12.4. 法制度と予算 194](#_Toc495070697)

[21.12.5. 保存 194](#_Toc495070698)

[21.12.6. 人材育成と雇用保障 195](#_Toc495070699)

[21.13. 今後のMLA連携の留意点 195](#_Toc495070700)

[21.14. 知識インフラの構築 195](#_Toc495070701)

[21.15. 国のオープンデータ戦略 195](#_Toc495070702)

[21.15.1. オープンデータの取組み 195](#_Toc495070703)

[21.15.2. Europeanaの動きとの比較 195](#_Toc495070704)

[21.15.3. 府省のデータ公開に関する基本的考え方（ガイドライン） 196](#_Toc495070705)

[21.15.4. NDLが持つ文化資産 196](#_Toc495070706)

[21.16. NDLSearchの新たな挑戦　（１）目的 196](#_Toc495070707)

[21.16.1. 新しい検索サービスは、 196](#_Toc495070708)

[21.16.2. 第一に、当館資料の統合検索 197](#_Toc495070709)

[21.16.3. 第二に、関係機関の資料の統合検索 197](#_Toc495070710)

[21.16.4. 第三に、外部連携APIの提供 197](#_Toc495070711)

[21.17. NDLSearchの新たな挑戦　（2）現在有している機能 197](#_Toc495070712)

[21.17.1. まず、統合検索の対象は、 197](#_Toc495070713)

[21.17.2. 国内刊行図書については、 197](#_Toc495070714)

[21.17.3. さらに、共通プラットフォームの提供 197](#_Toc495070715)

[21.17.4. 一方、検索機能の充実 198](#_Toc495070716)

[21.18. NDLSearchの新たな挑戦　（3）今後の重点的な取組み 198](#_Toc495070717)

[21.18.1. 第一に、連携先の拡張による、統合検索対象の充実に取り組みます。 198](#_Toc495070718)

[21.18.2. 第二に、検索機能の向上に取り組みます。 198](#_Toc495070719)

[21.18.3. 第三に、メタデータの二次利用の促進に取り組みます。 198](#_Toc495070720)

[21.19. 大震災とMLA 198](#_Toc495070721)

[21.19.1. 大震災とMLA 198](#_Toc495070722)

[21.19.2. NDL 198](#_Toc495070723)

[21.19.3. MALUI連携 198](#_Toc495070724)

[21.20. ★知識インフラの構築に向けたNDLラボの活用 199](#_Toc495070725)

[22. （2013年）要件定義の必要性と人材育成（IT研修内容） 200](#_Toc495070726)

[22.1. 要件定義の必要性 200](#_Toc495070727)

[22.2. 知識情報基盤の構築 200](#_Toc495070728)

[22.3. 最適化計画の考え方 200](#_Toc495070729)

[22.4. 使命・目標達成に向けた考え方 200](#_Toc495070730)

[22.5. 館の情報化推進体制 201](#_Toc495070731)

[22.6. 情報化のプロセスと考慮点（サービス構築の流れに沿って）（流れ図参照） 201](#_Toc495070732)

[22.7. 過去の問題事例 202](#_Toc495070733)

[22.8. 要件定義の重要性と考慮点 203](#_Toc495070734)

[22.9. 各工程での成果物の妥当性評価 204](#_Toc495070735)

[22.10. 参考見積書の評価 205](#_Toc495070736)

[22.11. ステークホルダーとの関係での留意点 205](#_Toc495070737)

[22.12. まとめ 205](#_Toc495070738)

[22.13. 標準類 207](#_Toc495070739)

[22.14. 各種標準の概要 207](#_Toc495070740)

[22.14.1. 情報システム調達のための技術参照モデル（TRM）平成24年度版 207](#_Toc495070741)

[22.14.2. 情報システム調達のための技術参照モデル（TRM）に基づく工程とアウトプット 208](#_Toc495070742)

[22.15. 開発・運用の問題と解決に向けたアプローチ 209](#_Toc495070743)

[22.16. IT人材確保 210](#_Toc495070744)

[22.16.1. 国のIT戦略 210](#_Toc495070745)

[22.16.2. NDLにおけるIT活用の位置付け 210](#_Toc495070746)

[22.16.3. IT人材育成の必要性 210](#_Toc495070747)

[22.16.4. 必要なITリテラシー 211](#_Toc495070748)

[23. （2013年）我が国の知識インフラとしてのナショナル・アーカイブ構想（要約版） 213](#_Toc495070749)

[23.1. ナショナル・アーカイブ構築の意義と目標 213](#_Toc495070750)

[23.2. ナショナル・アーカイブの定義と社会的役割 213](#_Toc495070751)

[23.3. ナショナル・アーカイブを構成する主要機能 213](#_Toc495070752)

[23.3.1. コンテンツの生成機能 213](#_Toc495070753)

[23.3.2. コンテンツの収集・一時保管機能 213](#_Toc495070754)

[23.3.3. 保存機能 213](#_Toc495070755)

[23.3.4. 権利情報・管理情報の収集・管理機能 214](#_Toc495070756)

[23.3.5. 配信・流通機能 214](#_Toc495070757)

[23.4. 国立国会図書館の役割 214](#_Toc495070758)

[23.4.1. コンテンツの生成機能 214](#_Toc495070759)

[23.4.2. コンテンツの収集・一時保管機能 214](#_Toc495070760)

[23.4.3. 保存機能 215](#_Toc495070761)

[23.4.4. 権利情報・管理情報の収集・管理機能 215](#_Toc495070762)

[23.4.5. 配信・流通機能 215](#_Toc495070763)

[23.5. 実現に向けた取り組みの方向性 215](#_Toc495070764)

[23.5.1. 今後の取組課題 215](#_Toc495070765)

[23.5.2. 利用促進・啓発普及活動の必要性 216](#_Toc495070766)

[24. （2013年）我が国の知識インフラとしてのナショナル・アーカイブ構想（本文） 217](#_Toc495070767)

[24.1. ナショナル・アーカイブ構築の意義と目標 217](#_Toc495070768)

[24.2. ナショナル・アーカイブの定義と社会的役割 217](#_Toc495070769)

[24.3. ナショナル・アーカイブを構成する主要機能 218](#_Toc495070770)

[24.3.1. コンテンツの生成機能（図3） 219](#_Toc495070771)

[24.3.2. コンテンツの収集・一時保管機能（図4） 220](#_Toc495070772)

[24.3.3. 保存機能（図5） 221](#_Toc495070773)

[24.3.4. 権利情報・管理情報の収集・管理機能（図6） 222](#_Toc495070774)

[24.3.5. 配信・流通機能（図7） 223](#_Toc495070775)

[24.4. 国立国会図書館及び関係機関の役割と分担 224](#_Toc495070776)

[24.4.1. 関係諸機関の役割 224](#_Toc495070777)

[24.4.2. 国立国会図書館の役割 225](#_Toc495070778)

[24.5. 実現に向けた取り組みの方向性 227](#_Toc495070779)

[24.5.1. 今後の取組課題 227](#_Toc495070780)

[24.5.2. 利用促進・啓発普及活動の必要性 228](#_Toc495070781)

[25. （2014年）デジタル化資料送信サービスと今後の展開 229](#_Toc495070782)

[26. （2014年）第2回基本問題検討会議に関するメモ【本文】（私見） 230](#_Toc495070783)

[26.1. ナショナルアーカイブの考え方 230](#_Toc495070784)

[26.2. 第2回検討会議の議論のポイント 231](#_Toc495070785)

[26.3. 第2回検討会議資料への意見 232](#_Toc495070786)

[26.4. 事業の目標 233](#_Toc495070787)

[26.5. 検討の姿勢 233](#_Toc495070788)

[26.6. 問題検討にあたっての意識 233](#_Toc495070789)

[26.7. 個別の課題提起 236](#_Toc495070790)

[26.8. ナショナルアーカイブに関しての理念 246](#_Toc495070791)

[26.8.1. 姿勢（留意点） 246](#_Toc495070792)

[26.8.2. 文化情報保存センター構想 247](#_Toc495070793)

[26.8.3. ナショナルアーカイブ 247](#_Toc495070794)

[26.8.4. 著作権法の一部を改正する法律案の審議での説明・質疑 248](#_Toc495070795)

[27. （2014年）ナショナル・アーカイブ構想の検討状況と今後の方向性（スライド説明用手持ち原稿） 249](#_Toc495070796)

[27.1. ナショナル・アーカイブ構想の実現を目指す意義 249](#_Toc495070797)

[27.2. 問題意識と解決策の１つ 249](#_Toc495070798)

[27.3. ナショナル・アーカイブ関連の施策 250](#_Toc495070799)

[27.3.1. ナショナル・アーカイブ関連の各施策等 250](#_Toc495070800)

[27.3.2. ①電子書籍・②文化資産の両議員連盟の動き 251](#_Toc495070801)

[27.3.3. 「我が国の知識インフラとしてのナショナル・アーカイブ構想（案）」 252](#_Toc495070802)

[27.3.4. ナショナル・アーカイブと国立国会図書館 253](#_Toc495070803)

[27.3.5. ナショナル・アーカイブの機能情報関連図 254](#_Toc495070804)

[27.3.6. ③アーカイブに関するタスクフォース報告書 内閣官房知的財産戦略本部（知財本部） 254](#_Toc495070805)

[27.4. ④知的財産戦略調査会の提言とりまとめ（2014年5月27日自由民主党知的財産戦略調査会） 255](#_Toc495070806)

[27.4.1. ⑤知的財産推進計画２０１４（2014年6月20日知財本部決定） 255](#_Toc495070807)

[27.4.2. ⑥経済財政運営と改革の基本方針2014（2014年6月24日閣議決定） 256](#_Toc495070808)

[27.4.3. ⑧内閣官房高度情報通信ネットワーク社会推進戦略本部（IT戦略本部）の動き 256](#_Toc495070809)

[27.4.4. ⑦文化庁・内閣府防災の動き 257](#_Toc495070810)

[27.5. 国立デジタル文化情報保存センターのカバレージは？ 257](#_Toc495070811)

[27.6. 各種アーカイブ構築施策の一元化 257](#_Toc495070812)

[27.7. 日本の文化情報基盤の強化のための費用試算 257](#_Toc495070813)

[27.8. 国立デジタル文化資産振興センター 基盤整備の全体像 257](#_Toc495070814)

[27.9. 文化資産として関連付けて保存すべきインスタンス【源氏物語を例に】 257](#_Toc495070815)

[27.10. ナショナル・アーカイブの検討に当たっての考察 258](#_Toc495070816)

[27.11. 基本問題事項 258](#_Toc495070817)

[27.11.1. 基本問題事項としての論点 258](#_Toc495070818)

[27.11.2. 基本問題事項の検討の進め方 259](#_Toc495070819)

[28. （2014年）デジタル時代の国立国会図書館の活動 260](#_Toc495070820)

[28.1. はじめに 260](#_Toc495070821)

[28.2. この20年間の電子図書館の構築の歩み（1994年～2012年） 260](#_Toc495070822)

[28.2.1. 第1ステージ【1994～2002】 揺籃期・始動期 260](#_Toc495070823)

[28.2.2. 第2ステージ【2002～2012】 サービス離陸期・発展期 260](#_Toc495070824)

[28.2.3. 知識インフラの構築を目指して 261](#_Toc495070825)

[28.3. 2012年以降進めてきたこと 261](#_Toc495070826)

[28.3.1. 第3ステージ【2012～2014】 総括と再始動期、見直し期 261](#_Toc495070827)

[28.3.2. 2012年1月のシステムリニューアルの総括と、次期業務システム最適化計画（2013～2017）の策定 261](#_Toc495070828)

[28.3.3. 「私たちの使命・目標2012-2016」の策定 262](#_Toc495070829)

[28.3.4. 東日本大震災アーカイブ 262](#_Toc495070830)

[28.4. 今後10年で目指すところ（2015年～2024年） 263](#_Toc495070831)

[28.4.1. 第4ステージ【2015～2024】 本格的なデジタル情報の普及期、サービスの変革期 263](#_Toc495070832)

[28.4.2. 2020年から数年の予測と、国全体での対応 263](#_Toc495070833)

[28.4.3. アーカイブに関連した国の活動の方向性 264](#_Toc495070834)

[28.4.4. 必要な機能の想定 265](#_Toc495070835)

[28.4.5. 今後のNDLの活動の方向性 267](#_Toc495070836)

[28.4.6. 選択と協力 267](#_Toc495070837)

[28.5. 図書館における情報システムの役割 267](#_Toc495070838)

[28.6. おわりに 268](#_Toc495070839)

[29. （2015年）国立国会図書館のサービスシステムの歩みと今後の方向性 ―電子図書館事業20年を迎えた新たな方向性の模索― 270](#_Toc495070840)

[29.1. はじめに 270](#_Toc495070841)

[29.2. この20年間の電子図書館構築の歩み 270](#_Toc495070842)

[29.2.1. 第1ステージ【1994～2002】 揺籃期・始動期 270](#_Toc495070843)

[29.2.2. 第2ステージ【2002～2012】 サービス離陸期・発展期 270](#_Toc495070844)

[29.2.3. 知識インフラの構築を目指して 271](#_Toc495070845)

[29.2.4. 東日本大震災アーカイブ 272](#_Toc495070846)

[29.2.5. 第3ステージ【2012～2014】 総括と再始動期、見直し期 272](#_Toc495070847)

[29.3. 今後10年で目指すところ（2015年～2024年） 272](#_Toc495070848)

[29.3.1. 第4ステージ【2015～2024】 本格的なデジタル情報の普及期、サービスの変革期 272](#_Toc495070849)

[29.3.2. 2020年から数年の予測と、国全体での対応 272](#_Toc495070850)

[29.4. 国全体の活動の方向性 272](#_Toc495070851)

[29.4.1. アーカイブに関連した国の活動の方向性 272](#_Toc495070852)

[29.4.2. 恒久的保存基盤 273](#_Toc495070853)

[29.4.3. 知識創造基盤 273](#_Toc495070854)

[29.4.4. 情報発信基盤 273](#_Toc495070855)

[29.4.5. 運用基盤 273](#_Toc495070856)

[29.5. 今後のNDLの活動の方向性 274](#_Toc495070857)

[29.6. おわりに 274](#_Toc495070858)

[30. （2015年）【カット分戻し】国立国会図書館のサービスシステムの歩みと新たな方向性の模索 276](#_Toc495070859)

[30.1. はじめに 276](#_Toc495070860)

[30.2. この20年間の電子図書館構築の歩み 276](#_Toc495070861)

[30.2.1. 第1ステージ【1994～2002】 277](#_Toc495070862)

[30.2.2. 第2ステージ【2002～2008】 277](#_Toc495070863)

[30.2.3. 第3ステージ【2008～2012】 279](#_Toc495070864)

[30.2.4. 第4ステージ【2012～2014】 281](#_Toc495070865)

[30.3. 今後10年で目指すところ（2015年～2024年） 281](#_Toc495070866)

[30.3.1. 第5ステージ【2015～2024】 281](#_Toc495070867)

[30.4. 国全体の活動の方向性 282](#_Toc495070868)

[30.4.1. 今後10年の国全体の取り組み 282](#_Toc495070869)

[30.4.2. アーカイブに関連した国の活動の方向性 283](#_Toc495070870)

[30.5. 今後のNDLの活動の方向性 286](#_Toc495070871)

[30.6. おわりに 287](#_Toc495070872)

[30.7. 年表 289](#_Toc495070873)

[30.7.1. 第0ステージ 289](#_Toc495070874)

[30.7.2. 第1ステージ【1994～2002】揺籃期・始動期 289](#_Toc495070875)

[30.7.3. 第2ステージ【2002～2008】サービス離陸期 290](#_Toc495070876)

[30.7.4. 第3ステージ【2008～2012】 発展期 291](#_Toc495070877)

[30.7.5. 第4ステージ【2012～2014】　総括と再始動期、見直し期 291](#_Toc495070878)

[31. （2015年）文化的資産を将来に継承し、新たな知識が創造される社会を目指して【枠組みのみ】 293](#_Toc495070879)

[31.1. 骨子 293](#_Toc495070880)

[31.2. 中期計画で目指したこと 293](#_Toc495070881)

[31.3. 実施してきたこと 293](#_Toc495070882)

[31.4. 今後のナショナル・アーカイブの効率的・効果的な構築 293](#_Toc495070883)

[31.5. 文化的資産のアーカイブ 293](#_Toc495070884)

[31.6. ICTの活用によるイノベーションのためのアーカイブ 293](#_Toc495070885)

[31.7. デジタルネットワーク社会の進展に対応したサービスの構築・運用 293](#_Toc495070886)

[32. （2016年）電子図書館サービスからナショナルアーカイブの構築へ－LOD化によるデジタル文化財の利活用を目指して－ 294](#_Toc495070887)

[32.1. はじめに 295](#_Toc495070888)

[32.2. 電子図書館20年の歩み 295](#_Toc495070889)

[32.2.1. 電子図書館構想と実証実験 296](#_Toc495070890)

[32.2.2. 国立国会図書館中期計画2004 296](#_Toc495070891)

[32.2.2.1. アクションプランとしての3つの柱 297](#_Toc495070892)

[32.2.2.2. 電子図書館中期計画2004の実現イメージ 297](#_Toc495070893)

[32.2.3. デジタルアーカイブの構築 297](#_Toc495070894)

[32.2.3.1. デジタル情報の長期保存（OAIS、 情報パッケージ(SIP、 AIP、 DIP)） 298](#_Toc495070895)

[32.2.4. デジタルアーカイブポータルのプロトタイプ 298](#_Toc495070896)

[32.2.4.1. 開発当初のサービス要件とシステム化 298](#_Toc495070897)

[32.2.5. PORTAの構築 298](#_Toc495070898)

[32.2.5.1. システムアーキテクチャ 299](#_Toc495070899)

[32.2.5.2. 統合検索対象の拡大 299](#_Toc495070900)

[32.2.5.3. 統合検索のために苦労した点 299](#_Toc495070901)

[32.2.5.4. インキュベータとしてのPORTAの役割 299](#_Toc495070902)

[32.2.5.5. Web3.0の時代を見据えて適用すべき技術と進め方 299](#_Toc495070903)

[32.2.6. トータルな次期図書館システムを目指して 300](#_Toc495070904)

[32.2.7. 国立国会図書館サーチ（NDLサーチ） 300](#_Toc495070905)

[32.2.7.1. 機能の概要 301](#_Toc495070906)

[32.2.7.2. サービスの状況 301](#_Toc495070907)

[32.2.7.3. 連携方針・連携モデル 302](#_Toc495070908)

[32.2.8. 公共的書誌情報基盤 302](#_Toc495070909)

[32.2.8.1. 公共的書誌情報基盤の整備（2010年） 302](#_Toc495070910)

[32.2.8.2. 全国書誌情報の利活用に向けた超党派勉強会での検討（活字文化議員連盟　2015年9月） 303](#_Toc495070911)

[32.2.9. 知識インフラの構築を目指して 303](#_Toc495070912)

[32.2.9.1. 知識インフラとは 303](#_Toc495070913)

[32.2.9.2. 知識インフラの必要性 303](#_Toc495070914)

[32.2.9.3. 知識インフラの構築の目的 303](#_Toc495070915)

[32.2.9.4. 知識インフラの機能 303](#_Toc495070916)

[32.2.9.5. 知識情報基盤の構築モデル 304](#_Toc495070917)

[32.2.10. 東日本大震災アーカイブ 304](#_Toc495070918)

[32.2.10.1. 収集の基本的な考え方 304](#_Toc495070919)

[32.2.10.2. 東日本大震災アーカイブのシステムイメージ 305](#_Toc495070920)

[32.2.10.3. 「ひなぎく」での課題 305](#_Toc495070921)

[32.3. 今後10年で目指すところ（2015年～2024年） 306](#_Toc495070922)

[32.3.1. 「知的財産政策ビジョン」での今後10年の目標 306](#_Toc495070923)

[32.3.2. 電子書籍ナショナルアーカイブ 307](#_Toc495070924)

[32.3.2.1. 電子書籍分野のアーカイブの役割 307](#_Toc495070925)

[32.3.2.2. 電子書籍分野のアーカイブの機能モデル 307](#_Toc495070926)

[32.3.3. 国の文化資産のアーカイブ 308](#_Toc495070927)

[32.3.3.1. アーカイブに関連した国の活動の方向性 308](#_Toc495070928)

[32.3.3.2. 恒久的保存基盤 308](#_Toc495070929)

[32.3.3.3. 知識創造基盤 309](#_Toc495070930)

[32.3.3.4. 知識利活用基盤 309](#_Toc495070931)

[32.3.3.5. 運用基盤 309](#_Toc495070932)

[32.4. 文化財でのLOD化の目指すところ 309](#_Toc495070933)

[32.4.1. クラウドコンピューティングの世界でのサービスの連携（2007年想定） 309](#_Toc495070934)

[32.4.2. 文化財のアーカイブ構築のために何に留意するべきか 310](#_Toc495070935)

[32.5. まとめ 311](#_Toc495070936)

[32.6. おわりに 311](#_Toc495070937)

[≪討議報告≫ 314](#_Toc495070938)

[LOD化によるデジタル文化財の利活用を目指して 314](#_Toc495070939)

[33. （2016年）AIを活用した「知の共有化」システムの方向性―「未来の図書館を作るとは」の実現に向けて－ 316](#_Toc495070940)

[33.1. はじめに 316](#_Toc495070941)

[33.2. 「知の共有化」とは 317](#_Toc495070942)

[33.3. デジタルトランスフォーメーション(デジタル変革) 317](#_Toc495070943)

[33.4. 「未来の図書館を作るとは」【長尾先生】を読み返して 318](#_Toc495070944)

[33.5. 知的情報資源のビッグデータとしての活用の可能性 318](#_Toc495070945)

[33.6. 「知の共有化」システムの方向性 318](#_Toc495070946)

[33.6.1. 知識インフラの概念 319](#_Toc495070947)

[33.6.2. 文化情報資源のナショナルアーカイブ 319](#_Toc495070948)

[33.6.3. 恒久的保存基盤（あらゆる情報の恒久的保存） 320](#_Toc495070949)

[33.6.4. 知識創造基盤（新たな知識創造活動の場） 320](#_Toc495070950)

[33.6.4.1. 知識創造支援機能 320](#_Toc495070951)

[33.6.4.2. 知識創造活動 321](#_Toc495070952)

[33.6.4.3. 新たな知識創造の場としての図書館 321](#_Toc495070953)

[33.6.5. 知識利活用基盤（情報発信） 322](#_Toc495070954)

[33.6.5.1. 自分の知識の外部記憶装置 323](#_Toc495070955)

[33.7. 「知の共有化」システムの構築業務と、人材の資質とスキル 323](#_Toc495070956)

[33.7.1. 政府標準ガイドラインに沿ったシステム構築と運用 323](#_Toc495070957)

[33.7.2. 政府標準ガイドラインに沿った開発タスクとドキュメント 323](#_Toc495070958)

[33.7.2.1. システムで何ができるようにするのか（個別機能要件） 324](#_Toc495070959)

[33.7.2.2. どのようなデジタルコンテンツを用意するか（コンテンツ構築要件） 324](#_Toc495070960)

[33.7.3. iコンピテンシ・ディクショナリ（iCD） 325](#_Toc495070961)

[33.7.4. 業務遂行のタスクとスキル・知識の蓄積のスキーム 325](#_Toc495070962)

[33.7.5. iコンピテンシ・ディクショナリ（iCD）を活用した業務の遂行とスキル・知識の選択的習得 325](#_Toc495070963)

[33.8. まとめ 326](#_Toc495070964)

[参考文献 327](#_Toc495070965)

2002年4月に国立国会図書館に入館して、2015年3月に退館するまでの間に、私が雑誌投稿、外部講演等での原稿を取りまとめた。

入館した当時の目標と、13年経過した現在の「知の共有化を目指した知識情報の保存基盤と、新たな知識の創造を目指した利活用基盤の構築」は、表現こそ異なるが、今まで目指してきたところ、今後目指すところは、ほとんど同じであることを実感した。

過去の経緯を踏まえて、目標の達成に向けて進展していくことを願う。

# （2002年3月）国立国会図書館志望動機

平成7～10年、通産省高度情報化事業の一環として、教育や行政情報の電子化と共に、国立国会図書館と共同で電子図書館システムの実験と運用を行う機会を得ました。複数プロジエクトのシステムを運用する情報基盤センターにおいて、他のシステムとともに、電子図書館の実証実験・総合目録システムを実質的な責任者として運用してきました。

情報基盤センターでの電子図書館プロジェクトを通じて、**国会図書館が多くの書籍、情報を網羅的に収集し、過去や現在の文化資料を広く一般の関覧に供し、また、電子化においては最新の技術を駆使して保存し将来に残していくことや、様々な機関が分散して保管している資料の所在を一元的に把握できるようにして、一般の人にも利用を可能にすることは、文化の発展に大きく貢献し、大変有意義なことだと感じました。**私は、最新のシステム技術、セキュリティ技術に着日したシステム構築・運用の経験を、今後は、電子図書館や図書館システムの高度化、コンテンツの活用に生かしていきたいと思っています。

情報基盤センターの運用がほぼ終了し、現在は情報セキュリティに対する脅威と対策の啓発と対策技術の研究開発を企画・実施しています。2000年の中央省庁Web改ざん事件が発生後は、内閣官房に設置された情報セキュリティ対策推進室の専門調査チームに非常勤で、電子政府を構築する中央省庁の情報システムのセキュリティ対策への助言に当たってきました。これまでのシステム開発技術、ネットワークおよびデータベース構築・運用技術と、システムに対するセキュリティ対策技術の経験は、IPAにおいても国の情報化の促進に、微力ながら一翼を担えたと思っておりますし、また、国立国会図書館での電子図書館・情報システムの開発、運用、発展に多少は役に立てるのでないかと考えます。

# （2003年）e-JAPAN戦略を踏まえたNDLの活動の方向性（メモ）

平成15年7月31日

電子情報企画室　中山

## デジタル・アーカイブの推進に向けた申入れの骨子（ＮＤＬ関連部分）

### 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進

* 「国立デジタル・アーカイブ」は、関係府省等（国立国会図書館、独立行政法人を含む）が構築するデジタル・アーカイブと、それらのネットワーク上の統合ポータルサイトからなる
* 国は、「国立デジタルアーカイブ」構想を推進する
* 関係府省等は、公共的なコンテンツ・情報のデジタル・アーカイブ化を一層推進する
* 関係府省等は、「国立デジタル・アーカイブ」が有機的に機能するように、各アーカイブ間の互換性の確保・標準化への協力にしっかりと取り組む
* 国の取組み
  + 「国立デジタル・アーカイブ」は、各種ポータルサイト、地方自治体・民間のアーカイブと連携し、国民が必要とするあらゆるコンテンツへの道しるべとしての役割を果たすこと
  + 国立国会図書館は、関係府省等の協力のもと、「国立デジタル・アーカイブ」のポータルサイトを運営すること
  + 国立国会図書館は、国が保存し、国民に提供すべきコンテンツのアーカイブ化に努めること
  + 関係府省は、その際に生じる課題の検討と解決にも協力すること
  + デジタル・アーカイブに関する国際的な動向と整合性を図ること。諸外国との連携に努めること（とりわけ近隣アジア諸国）
  + 国は、デジタル・アーカイブとそのネットワーク化の推進に向けて、必要な研究開発を推進すること

### 「ジャパン・ウェブ・アーカイブ」構想の推進

* 「ジャパン・ウェブ・アーカイブ」は、ウェブページのアーカイブに関するネットワーク上の総合ポータル
* アーカイブ化について、関係府省等の協力のもと、国立国会図書館を中心に、公共図書館、ＮＰＯ等の参加を得て、「ジャパン・ウェブ・アーカイブ」構想を推進すること
* 国は、複数の主体によるウェブ・アーカイブが相互に連携し、１つのウェブ・アーカイブとして機能するよう、ウェブページ間の関連性・更新履歴を踏まえた情報解析等の研究開発、ウェブページの保存・検索等に要するメタデータ・フォーマットや自動情報収集等の技術の共通化・標準化・原本性認証・時間認証等の認証基盤の確立などを図ること
* 欧米の先進事例等を踏まえた世界最先端のものを目指すこと
* ウェブページの収集等に当たっては、既に収集を開始している欧米諸国等との連携に努めるとともに、ＯＤＡ等の活用も含め、近隣アジア諸国との連携・協力につとめること

### その他

* 統一的かつ整合的なデジタル・アーカイブ政策を推進するため、内閣官房を中心として、関係府省等の間の調整・連携を効率的に行うこと

## e-JAPAN戦略で各機関が実施すべきとされる部分

### 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進

* 内閣官房
  + 統一的かつ整合的なデジタル・アーカイブ政策を推進するため関係府省等の間の調整・連携を効率的に行うこと
* 国立国会図書館
  + 国立国会図書館は、国が保存し、国民に提供すべきコンテンツのアーカイブ化に努めること
  + 関係府省等の協力のもと、「国立デジタル・アーカイブ」のポータルサイトを運営すること
  + 各種ポータルサイト、地方自治体・民間のアーカイブと連携し、国民が必要とするあらゆるコンテンツへの道しるべとしての役割を果たすこと
* 関係府省等（国立国会図書館、独立行政法人を含む）
  + 公共的なコンテンツ・情報のデジタル・アーカイブ化を一層推進する
  + 「国立デジタル・アーカイブ」が有機的に機能するように、各アーカイブ間の互換性の確保・標準化への協力にしっかりと取り組む
  + 関係府省は、「国立デジタル・アーカイブ」構築・運用の際に生じる課題の検討と解決にも協力すること
  + デジタル・アーカイブに関する国際的な動向と整合性を図ること。諸外国との連携に努めること（とりわけ近隣アジア諸国）
  + デジタル・アーカイブとそのネットワーク化の推進に向けて、必要な研究開発を推進すること

### 「ジャパン・ウェブ・アーカイブ」構想の推進

* 内閣官房
  + 統一的かつ整合的なデジタル・アーカイブ政策を推進するため関係府省等の間の調整・連携を効率的に行うこと
* 国立国会図書館
  + アーカイブ化について、関係府省等の協力のもと、公共図書館、ＮＰＯ等の参加を得て、「ジャパン・ウェブ・アーカイブ」構想を推進すること
* 関係府省等（国立国会図書館、独立行政法人を含む）
  + アーカイブ化について協力すること
  + 複数の主体によるウェブ・アーカイブが相互に連携し、１つのウェブ・アーカイブとして機能するよう、ウェブページ間の関連性・更新履歴を踏まえた情報解析等の研究開発、ウェブページの保存・検索等に要するメタデータ・フォーマットや自動情報収集等の技術の共通化・標準化・原本性認証・時間認証等の認証基盤の確立などを図ること
  + 欧米の先進事例等を踏まえた世界最先端のものを目指すこと
  + ウェブページの収集等に当たっては、既に収集を開始している欧米諸国等との連携に努めるとともに、ＯＤＡ等の活用も含め、近隣アジア諸国との連携・協力に努めること

## 関係府省の実施項目との関係

### 「国立デジタル・アーカイブ（略称：Ｄアーカイブ）」構想の推進

* 公共的なコンテンツ・情報のデジタル・アーカイブ化を一層推進する
  + →各府省所管の公共情報のデジタル・アーカイブの構築
* 国立デジタル・アーカイブ」が有機的に機能するように、各アーカイブ間の互換性の確保・標準化への協力にしっかりと取り組む
  + →関係府省等全体での共通仕様の策定と適用
* 関係府省は、「国立デジタル・アーカイブ」構築・運用の際に生じる課題の検討と解決にも協力すること
  + →制度上、著作権上の問題解決
* デジタル・アーカイブに関する国際的な動向と整合性を図ること。諸外国との連携に努めること（とりわけ近隣アジア諸国）
  + →国際動向の調査及び共通仕様の国際標準の適用
* デジタル・アーカイブとそのネットワーク化の推進に向けて、必要な研究開発を推進すること
  + →技術的課題の解決と必要なシステム構築

### 「ジャパン・ウェブ・アーカイブ」構想の推進

* アーカイブ化について協力すること
  + →分担収集
* 複数の主体によるウェブ・アーカイブが相互に連携し、１つのウェブ・アーカイブとして機能するよう、ウェブページ間の関連性・更新履歴を踏まえた情報解析等の研究開発、ウェブページの保存・検索等に要するメタデータ・フォーマットや自動情報収集等の技術の共通化・標準化・原本性認証・時間認証等の認証基盤の確立などを図ること
  + →ウェブアーカイブを実施する関係府省等全体での共通仕様の策定と適用
* 欧米の先進事例等を踏まえた世界最先端のものを目指すこと
  + →国際的な実証実験プロジェクトの成果の活用
* ウェブページの収集等に当たっては、既に収集を開始している欧米諸国等との連携に努めるとともに、ＯＤＡ等の活用も含め、近隣アジア諸国との連携・協力に努めること
  + →大規模収集蓄積システムの開発及び運用

## NDLの基本姿勢

* 各府省に協力を求めて、各機関が実施できることは、そこに任せる
* 各府省の成果を活用する形で実施する
  + 各府省が実施するとしているものに関して、その成果を活用させてもらう
  + 将来的にNDLが国レベルのデジタルアーカイブで中心的な役割を果たすことを想定し、仕様の策定に参画する
  + ←各府省は、施策として提示したものは、関係機関を使って大規模な体制で実施する。「NDLが策定するものに協力してほしい」ではなく、各機関が仕様を決定する際に参画して、（統合もしくは横断による）国レベルのデータベース構築に必要な仕様になるように意見を提示する方向でなければ実施困難
  + ←
* 最終的にNDLが運用する際に必要とされる内容を「要求機能要件」として提示する
  + 他機関がe-JAPAN戦略で実施するとされた事項で、NDLが中心的な役割を果たすべきとされた事項は多岐に亘る。
* NDLに実行予算が付く形ではなく、NDLが必要とする事項を、各府省が実施する形とする←NDLからみて、各府省は外注先として想定
* NDLは、将来的にNDLが運用に関わると思われる部分に関して、各府省が独自システムとして構築されないように、コーディネータの役割を果たす

## 当面の作業内容

* 自民党e-JAPAN重点計画特命委員会から政府への申入れ（7月29日決定文書）
* NDLが提供すべきサービスの取りまとめ
  + 自民党デジタルアーカイブ小委員会まとめに記載されたNDLの役割の把握
  + e-JAPAN戦略のなかに位置付けられた場合のNDLの中長期計画骨子
* 各府省実施内容の把握→e-JAPAN重点計画2003
  + e-JAPAN戦略等で示される政府としての中長期の目標の把握
  + e-JAPAN戦略における他機関の実施内容の把握
* NDLが必要とする機能と各府省実施内容の関係の整理
  + 個別施策のうち、NDLが成果として活用する可能性のあるものを抽出
  + 個別施策へのNDLの関与の内容、方法の検討
* NDLと各府省の役割分担
* NDLが分担する部分において具体的に実施すべき内容の精査
  + 必要予算
  + 実施体制
  + 制度
* 各府省が実施する内容のコーディネーション

## 当面のスケジュール

* 自民党e-JAPAN重点計画特命委員会から政府への申入れ（7月29日決定文書）
* 中長期計画骨子
  + e-JAPAN基本戦略Ⅱを踏まえた将来イメージの作成と館内合意
  + NDLと他機関の役割分担案の作成と合意
  + 将来イメージと役割分担案の合意（内閣官房）

## NDLが自ら実施すべき事項

* 自民党デジタルアーカイブ小委員会まとめに記載されたNDLの役割の把握
  + 「国立デジタル・アーカイブ」（略称：Dアーカイブ）構想の推進
    - 関係府省と共に構想を推進する
    - 関係府省と共にデジタルアーカイブ構築と、アーカイブ間の互換性の確保・標準化への協力に取り組む
    - 国民が必要とするあらゆるコンテンツの道しるべとしての役割を果たす
    - NDLは、関係府省等の協力のもと、「国立デジタル・アーカイブ」の統合ポータルサイトを運営。国が保存し、国民に提供すべきコンテンツのアーカイブ化に努める。各府省は課題解決に協力する。
  + 民間デジタル・アーカイブの構築及びコンテンツ流通市場の形成促進
  + 「ジャパン・ウェブ・アーカイブ」構想の推進
    - NDLを中心に、関係府省等の協力のもと、統合ポータル・サイト「ジャパン・ウェブ・アーカイブ」構想を推進すること
    - 複数のウェブ・アーカイブが相互に連携が取れて構築されるよう、共通化・標準化等を図ること
* 開発
  + 他機関が実施する内容が、将来NDLが運用する際に必要とする機能を具備するようにコーディネーションできる体制を確立する
  + 実施に必要な制度的課題を提示し、NDLが実施しなければ解決できない部分に関しては自らが実施する
* 運用準備
  + 実業務として運用するために必要体制、手順を確立
* 予算

## 要求機能要件

* 共通事項
  + 「国立デジタル・アーカイブ」のイメージの具体化
  + 「ジャパン・ウェブ・アーカイブ」のイメージの具体化
* 個別事項

## 16年度予算要求に関して

* 短中期実施内容
  + 「国立デジタル・アーカイブ」
    - イメージの具体化
    - アーカイブ間の互換性の確保・標準化仕様の確立
      * 一次情報コンテンツ仕様
      * 一次情報データベース仕様
      * 二次情報仕様
      * 二次情報データベース仕様
      * 検索及び交換のためのインターフェース仕様
    - NDLが所蔵する資料のデジタルアーカイブ構築
      * デジタル化及びメタデータ作成
        1. 明治期
        2. 大正期
        3. ・・・
        4. 貴重書
      * データベース構築
      * 検索システム構築
    - 分散デジタルアーカイブコンテンツの統合ポータル
      * NDL内デジタルアーカイブの統合検索システムの構築
      * 分散デジタルアーカイブコンテンツのメタデータ交換もしくは横断検索システムの構築
  + 「ジャパン・ウェブ・アーカイブ」
    - Web情報の収集・組織化・保存仕様の確立
    - 収集システムの共同開発
    - 分担収集
      * NDL：政府系サイト？
        1. 政府系全体の収集に関して包括的許諾
        2. 実収集
      * NII：学術系サイト
      * ・・・

## 実施に向けた体制整備

* 業務遂行工数の確保→手続きの省力化により必要工数の捻出
  + 館内情報伝達の省力化
  + 館内合意形成の省力化
  + 決裁手続きの省力化
    - 文書決裁
    - 物品調達
* 業務遂行能力の確保
  + 政府、学界、産業界、利用者ニーズの動向把握
* **中長期計画への反映**
  + 使命
  + 長期的な目標
  + 中長期的な目標
    - 電子図書館として、情報の検索から一次情報の入手までのワンストップサービスを実現する。
      * 当館が保有する情報に限らず、公的機関が保有する情報のポータルサイトを構築し運用する
      * 各種ポータルサイト、地方自治体・民間のアーカイブと連携し、国民が必要とするあらゆるコンテンツへのポータルとしての役割を果たすこと
      * ポータルサイトを運営する組織として、当館が保有する情報に限らず、全ての情報に対するレファレンスサービスを実施する
    - 当館しか保有していない全ての資料をデジタル化し、また、公共機関が保有する情報を収集して、当館としてのデジタルアーカイブを構築する
      * 当館しか保有していない情報を、利用者ニーズの高いものから順次デジタル化する
      * 公共機関が保有する情報を、網羅的に収集・組織化し、アーカイブ化する
      * 当館のデジタルアーカイブは、他の組織が自由に利用できるようにする
    - 他機関と分担して日本のウェブアーカイブを構築する
      * 他機関と分担して日本のウェブ情報を収集する
      * 当館は、公共機関のウェブ情報を中心に収集する
      * 当館と他機関のウェブアーカイブをあわせたポータルサイトとして、日本のウェブアーカイブを提供する
  + アクションプラン
    - 国際的な相互運用性の確保
      * 各種仕様に関しては、国際的な標準に準拠させる
      * 各デジタル・アーカイブ組織は、情報の交換及び横断検索できるように、外部仕様を共通化させる
    - 実施に向けた課題解決
      * 所管の関係府省に解決を求める
        1. 制度的な課題は、各府省の施策により、著作権処理の工数を低減する制度の確立及びシステムの開発により解決を図る
        2. 技術的な課題は、各府省の研究開発の成果により解決を図る
    - 電子図書館サービス構築及び運用体制の確立
      * 電子図書館サービス体制の強化のために、管理業務の効率化
      * 企画、対外調整（政策的、技術的）、館内調整、システム開発、システム運用、電子化（一次情報、主題情報、レファレンス情報）、電子情報収集・組織化（電子納本、Webアーカイブ）、電子情報提供（ポータル運用を含む）、電子レファレンスの体制整備
    - システム開発
      * 基本的には、当館が独自にシステム開発しない
      * 稼動システムを構築するに当たっては、当館から要求仕様を反映した形で構築してもらい、成果を活用させてもらう。もしくは、標準仕様に準拠したパッケージソフトを導入する
  + 具体的な実施手順
    - 構築・運用イメージの確立
    - 関係機関に協力を求める
      * システム構築
        1. 関係府省に対して要求仕様を提示する
        2. 各種標準仕様の策定依頼
        3. 標準仕様を満足したシステム構築依頼
        4. システムの利用許諾を受ける
      * システム運用
        1. システムの運用支援
      * 制度解決
      * 保有している情報提供（電子納本、Web収集許諾）
    - ・・・・・

# （2004年）国立国会図書館「電子図書館中期計画2004」の実施に向けて

「情報の科学と技術 Vol. 54 (2004),　No.9」

特集＝デジタル情報資源のアーカイビング

## 要約

本稿は、国立国会図書館（NDL）が2004年2月に策定した「電子図書館中期計画2004」を紹介し、今後、国のデジタルアーカイブ・ポータルを有用なものとして活用されるために何をすべきか、個々のデジタルアーカイブを提供する組織、構築のための技術的課題を提供する組織に何を期待するか等について考察してみたい。

## キーワード

デジタルアーカイブ・ポータル、e-JAPAN、情報探索、メタデータ、XML、OAI-PMH、RSS、セマンティックWeb、Webサービス

## はじめに

　国立国会図書館（NDL）では、2004年2月に「電子図書館中期計画2004」を策定した。この計画はNDLが所蔵している図書のデジタル化、オンライン系情報資源の収集によるデジタルアーカイブの構築、インターネット上にある情報を一つの大きなデジタルアーカイブとして利用できるようにする「日本のデジタルアーカイブ・ポータル」の構築と提供を目指すものである。この計画の策定と時期を同じくして、政府は、「e-JAPAN戦略Ⅱ加速化パッケージ」 において、政府コンテンツの利用機会の拡大と保存を図るため、NDLにおける政府刊行物アーカイブ構築及び政府各機関ホームページの長期保存に関して、協力体制を確立することを示した。さらに、「e-JAPAN重点計画2004」 では、加速化パッケージの内容が、2005年の目標達成への施策の重点化の中の、「加速化5分野」として掲げられ、「コンテンツ政策の推進」の具体的施策の一つとして、「連絡会議の場において、国立国会図書館で検討しているアーカイブのポータルサイトとの連携のあり方についても検討する。」ということが決定された。

## 電子図書館中期計画2004策定の背景

NDLでは、これまでの事業を継続すると同時に、国会の図書館として、また、我が国唯一の国立図書館として、次の認識を重要視した。

第一は、情報通信ネットワークを活用することによって、時や場所に制約されることなく、当館のサービスの利用機会を格段に広げることができること。

第二は、国内外の動きや情報環境の変化の中で、デジタル情報の収集・組織化・保存・提供の重要性が高まっていること。

第三は、NDLの電子図書館サービスを充実するためには関係諸機関との連携協力が不可欠であり、当館の目指す方向を明示し、関係諸機関の理解を得ることが重要であること。

## 国立国会図書館電子図書館中期計画2004の骨子

NDLの目指す電子図書館サービスの目標は、デジタルコンテンツを広汎な利用者に提供するために、当館が国のデジタルアーカイブの重要な拠点となるということ、また国内外の多様な利用者層の需要に応じ、日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築することである。

### デジタルアーカイブの構築

NDL資料の利用における地域間格差を改善し、利用者の利便性を高めるために当館所蔵の資料の電子化を推進することと、オンライン系情報資源を広く収集し、消失を防ぐとともに、永続的な利用を確保することとした。

オンライン系情報資源の収集では、インターネット上のウェブ情報を、可能な限り発信時と同様な構造をもったウェブ・アーカイブとしてサイト単位に収集し、時系列的認識が可能な形で蓄積・保存・提供するものと、知的な著作単位で取り扱うべき情報資源を対象として、個別に収集・組織化・保存・提供するものがそれぞれある。

### 情報資源に関する情報の充実

NDLが集積し、構築するデジタルアーカイブを利用者が的確に利用できるようにするため、全文検索、知的概念検索等、新しい検索インターフェース、辞書等を用意し、また、必要な参考情報、レファレンスツールを充実させる。

### デジタルアーカイブのポータル機能

利用者の必要とする情報を入手できる窓口として、情報の探し方を利用者にオンラインで提供する機能、利用者が主題に沿って系統的に情報資源を発見できるよう案内する機能とともに、当館のデジタルアーカイブやOPACを含めて、国等の公的機関を中心とした電子的情報資源や情報提供サービスが提供している情報そのものに、サイトを渡り歩くことなく、一つの窓口で適切に案内する「日本のデジタルアーカイブ・ポータル」（仮称）の構築を目指す。その次の段階として、他の機関のウェブ・アーカイブ構築の動向を見つつ、日本全体のウェブ・アーカイブのポータル機能の構築も目標としている。

## デジタルアーカイブ・ポータルの構築に当たっての考え方

まず、自らが所蔵している膨大な図書のデジタル化およびオンライン系の情報資源の収集によって、デジタルアーカイブを構築する。その際には他機関のデジタルアーカイブとの連携を意識した仕様を適用する。次に、各デジタルアーカイブの提供機関に対しては、連携を意識した仕様の適用を求める。そして、連携が可能になったデジタルアーカイブを統合して利用できるようにしたデジタルアーカイブ・ポータルを構築する。様々な利用者層、利用形態に応じ、様々な機関が自らの優位性を生かして情報に到達するための情報、ナビゲーションの仕組みなどを付加価値として提供するポータルを構築する。NDLは、ポータルの1つとして、自らが保有する膨大な情報を、広く一般に提供するためのポータルの構築を目指し、様々なポータルとともに、日本のデジタルアーカイブ・ポータルを構築し、利用されることを目指す。

## デジタルアーカイブ・ポータルが目指す方向

### 情報探索の行動パターンと提供すべきサービス

デジタルアーカイブ・ポータルとしてどんな機能が必要かを考えるに当たって、一般的な利用者がどのような手順で情報探索を行うのかのモデルに対応してポータルのモデルを想定することが必要である。

まず第一に、利用者の情報探索の行動パターンは、おおよそ、6つの段階に分けて考えることができる。何かについて調べたいと思うことを課題として設定し「問題の定義、情報ニーズの識別」を行う。ポータルとしてはE-レファレンスやオンラインチュートリアルのサービスが必要となる。

次に、「情報探索戦略の策定」として、どんな情報源から調べるかを選び出す。「情報源の範囲及び利用順位の決定」を行うため、サブジェクトゲートウェイ的な機能により、情報を提供しているWebサイトやデータベースへのナビゲーションや情報源への拾い読み（ブラウジング）を支援するサービスが必要となる。

3番目に、「情報源の所在確認、情報源へのアクセス」として、選ばれた情報源に個別に実際に当たってみる行為で、として、メタデータ検索、シソーラス検索などのセマンティック検索を支援するサービスが必要となる。

4番目に、「情報の獲得（情報源の咀嚼、情報の抽出）」として、色々あたってみた情報を咀嚼して、有用そうな情報を選別することで、様々なデータベースを渡り歩くことなく統合検索し、あるデータベースの検索結果から他のデータベースの二次情報、一次情報への直接アクセスが一連の作業として行えるようにするサービスが求められる。

そして、選別された情報をあわせて自分が調べたかった内容として纏めることで、「情報の統合（情報の組織化、提示）」を行うために、検索結果を利用者の目的に応じて利用しやすい形でアウトプットする機能が必要となる。

最後に、今回調べた内容は正しそうだったか、調べ方は合理的だったかを改めて見直すことで、「評価（成果の評価、プロセスの評価）」を行うことで一連の情報探索行動を終えることとなる。

これらの段階を総合的に解決するポータルのイメージとしては、「あるテーマで何かを知りたいとき、色々なところから提供されている情報を、１つの窓口から入って、そこで解決できるようにしたもの」と想定する。

### 情報探索行動における図書館の方向性

図書館は、紙の資料の時代から「情報探索することを助けること」を仕事としてきた。しかし、現在のように多くの情報が氾濫し、かつ様々なジャンルでの情報探索のニーズに応えるためには、図書館自身の業務のやり方も変わっていかなければならない。今後図書館は、どんな方向に向かっていくべきかを考えてみる。

一つ目は、「個別の図書館から、壁のない図書館へ」ということ。「個別の図書館が、自館の所蔵資料に関する情報提供だけでは、利用者の情報収集行動を支援できない。」ということで「個別図書館サービスの横断的利用が可能になるようなサービスの提供を目指す」必要がある。

二つ目は、「図書館サービスの枠を越えて」ということで、「図書館だけでは、利用者の情報収集行動を支援できない。」ということ。「同じ分野、同じ利用者層をターゲットにした複数の専門情報サイトが連携して、利用者がワンストップで利用できるようにする」ことが大切と考える。

## デジタルアーカイブ・ポータルの利用イメージ

大きな流れとしては、まず、メタ情報がついたデジタルコンテンツをデジタルアーカイブに構築しておき、利用者の任意の条件からメタ情報を検索し、そのメタ情報の一覧から、必要なものを選択すると必要なデジタルコンテンツが直接見えるようになるというイメージを想定している。そしてそれは、一つの機関の情報についてではなく、他機関の情報を含めて統合的に利用できるようにすることにより、利用者は情報の所在に関係なく、利用できるようになるものである。（図１）

### 情報資源に到達するための情報の充実

必要な情報に辿り着くことを支援する仕組みと辞書類の構築と提供が必要で、それは自分の組織が提供できる情報には限らない。利用者が指定する任意の条件から、目録検索に必要な検索語を導出し、分かりやすい形式で検索結果情報を提示し解釈・処理を容易にするサービスが必要である。

### デジタルアーカイブへのアクセス

情報資源に到達するための情報から得た情報から、デジタルアーカイブ、冊子体等の所蔵情報へのアクセスできるようにすることである。

#### 二次情報の統合検索サービスの提供

図書館、学術ポータル、電子政府、民間ポータル、専門情報サイト等の目録、インデックスを、商用あるいはライセンス付きの電子資源に制限されることなく、異なる対象資源に対する統合的な探索及び情報検索をできるようにすることである。その際は、記述的メタデータの統合的な検索・情報探索を提供し、ポータルは、探索の設定・コントロール段階でユーザを支援し、探索結果が確実に再現されることを保証する。

#### 一次情報の統合閲覧サービスの提供

目録の検索結果から、自組織及び他機関が提供する一次情報を直接閲覧する機能を構築することであり、検索結果がインターネットで提供可能なデジタルコンテンツの場合はイメージ、フルテキストで閲覧できるようにし、デジタルコンテンツがなかった場合は、NDLの遠隔複写サービス等のドキュメントデリバリのサービスへつなげる。将来的には、有料コンテンツであった場合はそのコンテンツの注文サービス等へ誘導する。オンライン出版の各サイトや、Amazon.comのようなインターネット書店、いわゆる古本屋もWebサイトで注文を受け付けており、それらも誘導先の1つと想定する。

## ポータルから見たデジタルアーカイブの要件

デジタルアーカイブ・ポータルが提供すべきサービスを実現するためには、デジタルアーカイブは次の要件を持つ必要があると考える。

デジタルアーカイブとしてコンテンツを提供するデータプロバイダは、一次情報としてのコンテンツ、一次情報へのアクセス手段のために編集された二次情報を保有する。デジタルアーカイブ内のコンテンツに対して付加価値を付けたり、検索を支援するサービスプロバイダは、必要な検索語を導出するための辞書、案内情報、コンテンツに関する解題等の情報をデータベース化して保有する。

　データプロバイダが持っているコンテンツを統合的に利用するためには、それぞれのデータプロバイダが持っているコンテンツに関するメタデータが機械的に利用できる必要がある。またそのメタデータの利用に関しては、横断的な検索をしていく仕組みと、メタを収集しておいてそれを検索に利用していく仕組みがある。（図２）

### コンテンツ仕様

情報の内部形式は、完全に一つの形式で統一することは非現実的だが、統合利用のためには外部インターフェースの共通化が必要であり、そのためには、コンテンツやメタの内部形式も、ある程度共通化している必要がある。

コンテンツに関しては、将来的にも広く普及が見込め、利用が保証される形式を採用する必要がある。メタの記述内容に関しても、検索のヒット率をあげるためには、記述されるべき項目とその記述内容に関して、目録規則等を意識してある程度は調整された形でなければならない。また、人手で全てのコンテンツにメタを付与していくことは膨大な工数がかかるため、メタデータの自動付与も実現する必要があるが、そのためには、コンテンツからメタを自動生成できる要素がコンテンツそのものに存在しなければならない。将来的には、Webページやテキスト系文書は、HTML文書やPDF文書ではなく、文書に構造と意味を持たせたXML文書の形で公開されていくことが望まれる。また、画像・音声系の情報のファイル形式に関しては、多くの標準が存在しており統一していくことは困難であるが、少なくともそれぞれのファイルのプロパティには最低限のメタ情報が付与されていくことが必要である。

### インターフェース仕様

ポータルが対象とするデジタルアーカイブは、最終的には図書館界だけでなく、政府及び政府関係機関、公共機関、民間問わず広く世の中にあるWebサイトであり、様々な業種・業態を対象とする。そのようなサイトを対象とした連携のためには、広く普及が見込まれているインターフェース仕様を採用する必要がある。共通インターフェースのレベルとしては、個別に仕様の異なるデータベースのデータ操作言語レベル、仕様がある程度統一されたデータベースでの登録・更新・検索処理のアクセスメソッドレベル、それぞれがサービスとして構築されているシステムに対するリクエストとレスポンスの形での情報の受け渡しのレベルが考えられる。

#### データベースへのアクセスレベル

これは、それぞれのデータベースに対して、そのデータベースのスキーマに従ってSQL等のレベルで操作するもので、データベースの論理構造そのものを把握していなければならず、他機関との連携では、仕様変更の影響を受けやすく好ましくない。

#### データアーカイブへのアクセスレベル

　これは、OAI-PMH でのメタデータハーベスティングプロトコルのように、レポジトリの内容そのものでなくアクセスの際の論理データのレベルを規定することにより、レポジトリとプロトコルを用意したサイトから共通の仕様のデータを得られるものである。XMLでの受け渡しのため、リクエストとレスポンスのデータ項目の並びとか過不足は処理に影響しなく自由度が高い。

　メタデータをOAI-PMHでアクセス出来る形のレポジトリとしてサイト内において、サービスプロバイダ側が必要なメタデータで収集していく形である。小規模で変化の少ないサイトでは、データベースを構築せずに、静的レポジトリとしてXMLファイルで設置する方法もある。このようなレポジトリを用意することにより、メタデータのハーベスティングが可能になり、深層Webとされていたデータベースは表層Webと同様に扱うことが可能になる。

#### Webページへのアクセスレベル

また、Webページに関しては、RSS ファイルを設置して、Webサイトの見出しや要約などのメタデータを構造化して記述するXMLベースのファイルにより、Webサイトが持っている情報のメタを公開する方法も広く普及している。そのファイルそのものを収集することにより、メタデータの収集ができ、データ提供者、収集者に負荷が少なく、ニュースサイトや著名なサイトでは、更新情報を機械的に収集されることにより、その情報を使って利用者はアクセスできるため、情報提供に力を入れているサイトを中心に利用が広まっている。最近のWebページは、CMS により、プログラムにより動的なHTML文書が作られることが多くなっている。それにより、表層Webとして機械的収集が困難な深層Web化しているが、サイト内の情報をRSSファイル等で公開することにより、従来よりも機械的に認識可能な情報として扱えられるようになってきている。しかし、もともとはデータベースの形で論理構造を持った情報であるので、さらにもう一段進んで、近い将来的には、本格的なセマンティックWebの普及を期待する。

従来のWebページは、HTML言語で作成されているが、HTML文書は、ページの見た目を定義する言語であり情報を区別するための属性情報を持っておらず、機械的に内容を認識することを困難にさせている。

セマンティックWeb は、従来のHTML形式のように人が見て内容を理解する形のページを検索エンジンなどによって検索し、得られたデータや情報の意味を人間が判断する形ではなく、情報記述言語であるXML文書で、キーワードをキーワードとして認識できるメタデータを付与することによって人間の代わりにコンピュータが理解し、ソフトウェアで自動処理できるようにする新しいWebの形態である。Webページの個々の情報の意味情報であるメタデータを集約し、その知識を抽出して構造化することにより、インターネット上に散在するさまざまな情報全体をひとつの巨大な知識データベースと見なし、ソフトウェアの自動処理により効率的に処理し、コンピュータとの自然なやり取りで必要な情報が取得できるしくみが実現できる。また、XML文書をスタイルシートであるXSLTの定義に従ってHTML文書化する形であれば、スタイルシートを変えれば、印刷用、PC画面用、携帯電話用やWebコンテンツアクセシビリティのガイドラインを満たすWebページを動的に作成できシングルソース・マルチユースを実現できる。

### Webサービスとしての連携

WWW関連の技術を使い、ソフトウェアの機能をネットワークを通じて利用できるようにしたもので、コンポーネント化された複数のWebサービス同士をつなぎ合わせてアプリケーションを構築する形である。データ提供者側は、単なるデータプロバイダというよりもデータ提供システムプロバイダの位置付けになる。

この形は、検索キーを与えて検査結果を得るというデータベースアクセスというレベルではなく、複数のWebサービスサイトに処理要求のレベルでのリクエストを出して、各サイトの処理結果を、XML形式でのレスポンスとして返すものである。このシステム連携イメージは、B2B、B2Cでのシステム連携の基盤技術であり、それが、G２B2C での連携によるサービス提供につながる。同業種・業態でのそれぞれにWebサービスとしての連携仕様は確立しつつあるが、デジタルアーカイブとしての連携においては、図書館の枠を越えて様々な機関が提供するポータル、サービスプロバイダ、データプロバイダ間での共通の規約を整備することが必要である。しかし、この仕組みの技術基盤はすでに確立しており、属性名、属性値の必要最低限の記述規則（目録規則）を調整すれば、早期に実現は可能となっている。

## デジタルアーカイブ・ポータルの構築・運用に関しての考察

次に、「ポータルの構築・運用に関しての考察」として、検討のポイントを6つほど列挙してみる。

ポータルの構築にあたっては、情報提供者の意志と権利を十分に尊重しなければならない。また情報は、ある分野において最も影響のある情報が統合利用の対象外になっていては、真に有用なポータルとして使えない。

### コンテンツへの直接ナビゲーションが必要

ポータルは単なるサイトへのリンク集ではない。ポータルがポータルを指して、どこまでアクセスしても、一次情報に辿り着かないポータルは敬遠される。もし、リンク集的なものが必要だとすると、それは、システム連携、データ連携できるデジタルアーカイブをUDDIレジストリ に登録 して公開する形が望まれる。

### ナビゲーションすべき情報へは確実に

ポータルとして機能するためには、連携しやすいサイトではなく、連携すべきサイトであり、重要な情報を持つサイトが対象外であってはいけない。ロボット検索できるところでなく、含めるべきサイトは、ポータルから見えるようにしてもらうことが必要である。

実際には、提供者が広く利用されることを望んで意識しているサイトは、そういう配慮をしている。広く利用されるべきサイトに対しては、他から機械的に利用されるメリットを理解してもらって、検索エンジン、ニュースリーダ等が取りやすい仕組みを提示することが重要である。

### 情報提供者の意志を尊重する

しかしながら、機械的に情報を収集されたくないと思っているサイトや、第三者権利等のため収集されては困る情報を、無理して対象としてはいけない。提供者がポータルに登録されることを望むものと、ポータル側が見えるようにしたいものが一致できるようにすることが重要で、インデキシングされるメリットをきちんと理解してもらう必要がある。

### 誰でも付加価値をつけて運営できるように

ポータルは、ポータルとしてナビゲーションできる情報に付加価値をつけて情報を提供するもので、日本のポータルとして一つである必要はない。

多種多様な利用者ニーズに応えるポータルは、いくつあってもいい。それぞれがポータルのサービスコンポーネントを取捨選択して特色を持ったポータルを作れるようにすることが重要と考える。ポータルは既製品を押し付けるものでなく、利用者が自由にサービスコンポーネントを組み合わせて利用できるものでもある。既にあるポータルをリンクするのではなく、それぞれのポータルのサービスをワンストップで使える１つのコンポーネントとして組み込んだポータルが作れることが望ましいと考える。

### データプロバイダは、情報収集ソフトからのアクセシビリティを確保

Webページは、人間に対しての対話型提供のアクセシビリティだけでなく、機械可読型提供に対してのアクセシビリティに配慮してもらう必要がある。ポータル側から、データプロバイダ側に用意して欲しい機能を提示して、インプリメントしてもらう。その際は、将来的に普及する見込みのある技術を適用することが重要である。国の情報の場合は、パブリックドメイン化と利用促進のための仕組みを用意がさらに重要で、個別に利用許諾したり、独自仕様で提供して、ポータル側に個別のアクセスインターフェースを用意させるような形ではいけない。

### 考察のまとめ

そのような考察のもとで、日本のデジタルアーカイブ・ポータルの構築を目指すNDLは、「ポータル構築のインキュベーション役」なのかもしれない。全ての人を満足させられるポータルを運用することは困難であり、NDLは、日本のポータルを提供する一機関として、日本のポータル構築のために、情報を提供したい人が、情報を提供しやすくするための環境の普及を加速させることが、重要な役割と言えるかもしれない。

## デジタルアーカイブおよびポータル構築に向けて

デジタルアーカイブ及びデジタルアーカイブ・ポータルは、NDLが自ら構築・運用に向けて努力するとともに、関係機関との連携・協力を求めることなしに実現することはできない。ここでは、中長期的に協力して実施していくことにより、早期実現が図れると考える事項を列挙してみる。

### 公的機関のデジタルアーカイブ構築の推進

* 公共機関の情報を広く一般が簡便に利用できるようにするために、公共的な情報資源のデジタルコンテンツ化と、そのアーカイブを積極的に推進
* デジタルアーカイブの構築に当たっては、関係機関は、デジタルコンテンツの長期保存とアクセスの保障に必要な機能と標準的な仕様の研究開発・技術開発
* また、制度的な課題解決として、デジタルアーカイブとその利活用の推進に向けて、国有財産に属する電子情報の利用の自由化あるいは利用手続きの明確化及び簡素化を図ると共に、図書館等で所蔵する書籍等のデジタル化及びその提供に関し、著作権者の不明な著作物を利用するための手続の効率化

### デジタルアーカイブ・ポータルの推進

NDLは、標準仕様により構築されたデジタルアーカイブ全体へのナビゲーションと、それぞれのデジタルコンテンツへのワンストップアクセスができるポータルサイトを構築し運営することを目指す。そのために、関係機関には、デジタルアーカイブ内の情報を安定的かつ効果的に利用できるようにするための技術的課題の解決に協力を得る。

* 改廃等により消滅してしまうシステム環境の中で、電子情報格納フォーマットから情報を再現することが可能な長期保存技術
* メタデータの収集技術~
* 情報提供者によるメタデータ付与を容易にする技術
* 新しい概念の検索方式や共通検索技術
* また、デジタルアーカイブを構築する公共機関等には、デジタルコンテンツはデータ連携を可能にする標準的な仕様で提供していただくことが必要となる。。
* 標準的なフォーマット、インターフェース仕様の適用
* 標準的なメタデータの付与

各種ポータルサイトでは、デジタルアーカイブへの標準的なアクセスのための仕様を実装し、それぞれのポータルがシステム的に連携して、相互補完しあい、全体で相乗効果を発揮できるようになることが望ましいと考える。

### ウェブ・アーカイブ構想の推進

NDLは、公表されたインターネット情報資源については、インターネット情報資源が消滅しつつある状況において、個別にあるいはウェブ・サイトごとの収集・保存を早期に実施することが焦眉の課題となっており、実施を可能とするための法制度化等を納本制度審議会に諮りつつ、早急に行うこととする。

そのために、関係機関には、ウェブ上の情報を網羅的に収集するためシステム開発が必要である。また、ウェブ・アーカイブは1つの機関で全てを収集・蓄積することは困難である。国内・国際的な枠組みの中で、分担して収集することが望まれる。そのためには、国内外の複数の主体によるウェブ・アーカイブが相互に連携し、一つのウェブ・アーカイブとして機能するための技術的課題の解決が求められる。

* ウェブページの保存・検索等に要するメタデータ・フォーマットや自動情報収集等の技術の共通化・標準化
* その仕様を満たした収集システムの開発
* ウェブページ間の関連性・更新履歴を踏まえた情報解析等の研究開発

ウェブでデジタルコンテンツを提供している各機関には、NDLが行うウェブ・アーカイブ構築に関して、それぞれの機関が提供しているウェブページを、機械的に収集しやすくするための機能を実装する等の協力を得ることも重要である。

* ウェブページにある情報のメタデータファイルの設置等
* 第三者の権利を侵害する恐れのある情報の収集拒否設定等

## おわりに

デジタルコンテンツの利用を取り巻く環境は大きく変わりつつある。技術の進化により、提供可能なサービスが拡大し、そのサービスを受け入れた利用者からは、より高度なニーズが生まれてきている。

デジタルアーカイブ・ポータルは、図書館の枠を越えて、国、公共機関、学界に加えて民間、個人が保有する情報をワンストップで的確に閲覧利用できるようにするものである。

インターネット上にある膨大な情報を、「意味ある情報資源」として活用するための研究開発、技術開発を進めるとともに、情報の提供者はその技術を適用した情報提供することにより、巨大な知識ベースが構築できる。デジタル情報を日本の文化遺産として後世に残し、新たな知識を創出するための知識として、現在及び将来にわたって活用できるようにするために、データプロバイダ、サービスプロバイダのそれぞれの機関が「Win-Win」の関係で実施し発展していけることが重要であり、関係機関・関係各位の御協力を願いたい。

### 参考文献

高度情報通信ネットワーク社会推進戦略本部（IT戦略本部）, e-JAPAN戦略Ⅱ加速化パッケージ, http://www.kantei.go.jp/jp/singi/it2/kettei/040206ejapan.pdf

IT戦略本部, e-Japan重点計画-2004, http://www.kantei.go.jp/jp/singi/it2/kettei/040615honbun.pdf

三輪眞木子.情報検索のスキル, 東京, 中公新書, 2003, p.32-73.

Open Archives Initiative Protocol for Metadata Harvestingの略。OAI（Open Archives Initiative）が策定した、ウェブ上のメタデータを収集（メタデータ・ハーベスティング）するためのプロトコル

Rich Site SummaryもしくはRDF Site Summaryの略。Webサイトの見出しや要約などのメタデータを構造化して記述するXMLベースのフォーマットで主にサイトの更新情報を公開するのに使われている

コンテンツマネージメントシステム。Webコンテンツの作成、更新及び管理を行うシステムで、トップページも最新情報を含めて動的なページとして容易に配信することができる。

Government to Business to Consumer。国と企業、企業と国民で連携した付加価値サービス。

Webサービスの検索･照会システム

# （2004年）デジタルアーカイブの構築イメージとして想定したこと

平成26年9月1日

総務部　中山

今後のナショナルアーカイブ関連の議論のために、過去の経緯を総括する目的で、電子図書館中期計画2004の完成形をイメージする。

まず、「国のデジタル・アーカイブ・ポータルの構築 : 国立国会図書館電子図書館中期計画2004の実施に向けて」（情報の科学と技術 Vol. 54 (2004),　No.9）　<http://ci.nii.ac.jp/naid/110002826958>　から再確認したいポイントを抜粋する。

## 2004年に想定したこと

### デジタルアーカイブ・ポータルの構築に当たっての考え方

* まず、自らが所蔵している膨大な図書のデジタル化およびオンライン系の情報資源の収集によって、デジタルアーカイブを構築する。その際には**他機関のデジタルアーカイブとの連携を意識した仕様を適用**する。
* 次に、**各デジタルアーカイブの提供機関に対しては、連携を意識した仕様の適用を求める。**そして、連携が可能になったデジタルアーカイブを統合して利用できるようにしたデジタルアーカイブ・ポータルを構築する。
* **様々な利用者層、利用形態に応じ**、様々な機関が自らの優位性を生かして**情報に到達するための情報、ナビゲーションの仕組み**などを付加価値として提供するポータルを構築する。
  + **NDLは、ポータルの1つ**として、自らが保有する膨大な情報を、広く一般に提供するためのポータルの構築を目指し、様々なポータルとともに、日本のデジタルアーカイブ・ポータルを構築し、利用されることを目指す。

### 情報探索行動における図書館の方向性

* 図書館は、紙の資料の時代から「情報探索することを助けること」を仕事としてきた。しかし、現在のように多くの情報が氾濫し、かつ様々なジャンルでの情報探索のニーズに応えるためには、図書館自身の業務のやり方も変わっていかなければならない。
* 一つ目は、**「個別の図書館から、壁のない図書館へ」**
  + 個別の図書館が、自館の所蔵資料に関する情報提供だけでは、利用者の情報収集行動を支援できない。「個別図書館サービスの横断的利用が可能になるようなサービスの提供を目指す」必要がある。
* 二つ目は、**「図書館サービスの枠を越えて」**
  + 図書館だけでは、利用者の情報収集行動を支援できない。「同じ分野、同じ利用者層をターゲットにした複数の専門情報サイトが連携して、利用者がワンストップで利用できるようにする」ことが大切と考える。

### デジタルアーカイブへのアクセス

#### 二次情報の統合検索サービスの提供

* 図書館、学術ポータル、電子政府、民間ポータル、専門情報サイト等の目録、インデックスを、商用あるいはライセンス付きの電子資源に制限されることなく、異なる対象資源に対する統合的な探索及び情報検索をできるようにすることである。
* その際は、記述的メタデータの統合的な検索・情報探索を提供し、ポータルは、探索の設定・コントロール段階でユーザを支援し、探索結果が確実に再現されることを保証する。

#### 一次情報の統合閲覧サービスの提供

* 目録の検索結果から、自組織及び他機関が提供する一次情報を直接閲覧する機能を構築することであり、検索結果がインターネットで提供可能なデジタルコンテンツの場合はイメージ、フルテキストで閲覧できるようにし、デジタルコンテンツがなかった場合は、NDLの遠隔複写サービス等のドキュメントデリバリのサービスへつなげる。
* 将来的には、有料コンテンツであった場合はそのコンテンツの注文サービス等へ誘導する。**オンライン出版の各サイトや、Amazon.comのようなインターネット書店、いわゆる古本屋もWebサイトで注文を受け付けており、それらも誘導先の1つ**と想定する。

## ポータルから見たデジタルアーカイブの要件

* デジタルアーカイブ・ポータルが提供すべきサービスを実現するためには、デジタルアーカイブは次の要件を持つ必要があると考える。
* デジタルアーカイブとしてコンテンツを提供するデータプロバイダは、一次情報としてのコンテンツ、一次情報へのアクセス手段のために編集された二次情報を保有する。
* デジタルアーカイブ内のコンテンツに対して付加価値を付けたり、検索を支援する**サービスプロバイダは、必要な検索語を導出するための辞書、案内情報、コンテンツに関する解題等の情報をデータベース化して保有**する。
* データプロバイダが持っているコンテンツを統合的に利用するためには、それぞれのデータプロバイダが持っているコンテンツに関するメタデータが機械的に利用できる必要がある。またそのメタデータの利用に関しては、横断的な検索をしていく仕組みと、メタを収集しておいてそれを検索に利用していく仕組みがある。

### コンテンツ仕様

* 情報の内部形式は、**完全に一つの形式で統一することは非現実的だが、統合利用のためには外部インターフェースの共通化が必要**であり、そのためには、コンテンツやメタの内部形式も、ある程度共通化している必要がある。
* コンテンツに関しては、**将来的にも広く普及が見込め、利用が保証される形式を採用**する必要がある。メタの記述内容に関しても、検索のヒット率をあげるためには、**記述されるべき項目とその記述内容に関して、目録規則等を意識してある程度は統制された形**でなければならない。
  + また、人手で全てのコンテンツにメタを付与していくことは膨大な工数がかかるため、メタデータの自動付与も実現する必要があるが、
  + そのためには、**コンテンツからメタを自動生成できる要素がコンテンツそのものに存在**しなければならない。
* 将来的には、Webページやテキスト系文書は、HTML文書やPDF文書ではなく、**文書に構造と意味を持たせたXML文書の形で公開されていくことが望まれる**。また、画像・音声系の情報のファイル形式に関しては、多くの標準が存在しており統一していくことは困難であるが、少なくともそれぞれのファイルのプロパティには最低限のメタ情報が付与されていくことが必要である。

### インターフェース仕様

* ポータルが対象とするデジタルアーカイブは、最終的には**図書館界だけでなく**、政府及び政府関係機関、公共機関、民間問わず広く世の中にあるWebサイトであり、様々な業種・業態を対象とする。
* そのようなサイトを対象とした連携のためには、**広く普及が見込まれているインターフェース仕様を採用**する必要がある。
* 共通インターフェースのレベルとしては、
  + 個別に仕様の異なるデータベースのデータ操作言語レベル、仕様がある程度統一されたデータベースでの登録・更新・検索処理のアクセスメソッドレベル
  + それぞれがサービスとして構築されているシステムに対するリクエストとレスポンスの形での情報の受け渡しのレベルが考えられる。

#### データベースへのアクセスレベル

* これは、それぞれのデータベースに対して、その**データベースのスキーマに従ってSQL等のレベルで操作する**もので、データベースの論理構造そのものを把握していなければならず、他機関との連携では、仕様変更の影響を受けやすい。

#### データアーカイブへのアクセスレベル

* これは、OAI-PMH でのメタデータハーベスティングプロトコルのように、**レポジトリの内容そのものでなくアクセスの際の論理データのレベルを規定**することにより、レポジトリとプロトコルを用意したサイトから共通の仕様のデータを得られるものである。
* **XMLでの受け渡しのため、リクエストとレスポンスのデータ項目の並びとか過不足は処理に影響しなく自由度が高い**。
  + メタデータをOAI-PMHでアクセス出来る形のレポジトリとしてサイト内において、サービスプロバイダ側が必要なメタデータで収集していく形である。
  + 小規模で変化の少ないサイトでは、データベースを構築せずに、静的レポジトリとしてXMLファイルで設置する方法もある。このようなレポジトリを用意することにより、メタデータのハーベスティングが可能になり、深層Webとされていたデータベースは表層Webと同様に扱うことが可能になる。

#### Webページへのアクセスレベル

* また、Webページに関しては、**RSS ファイルを設置して、Webサイトの見出しや要約などのメタデータを構造化して記述する**XMLベースのファイルにより、Webサイトが持っている情報のメタを公開する方法も広く普及している。
  + そのファイルそのものを収集することにより、メタデータの収集ができ、データ提供者、収集者に負荷が少なく、ニュースサイトや著名なサイトでは、更新情報を機械的に収集されることにより、その情報を使って利用者はアクセスできるため、情報提供に力を入れているサイトを中心に利用が広まっている。
  + 最近のWebページは、**CMS により、プログラムにより動的なHTML文書が作られることが多くなっている**。それにより、表層Webとして機械的収集が困難な深層Web化しているが、**サイト内の情報をRSSファイル等で公開することにより、従来よりも機械的に認識可能な情報として扱えられる**ようになってきている。
  + しかし、もともとはデータベースの形で論理構造を持った情報であるので、さらにもう一段進んで、**近い将来的には、本格的なセマンティックWebの普及を期待**する。
* 従来のWebページは、HTML言語で作成されているが、**HTML文書は、ページの見た目を定義する言語であり情報を区別するための属性情報を持っておらず、機械的に内容を認識することを困難**にさせている。
* セマンティックWeb は、従来のHTML形式のように人が見て内容を理解する形のページを検索エンジンなどによって検索し、得られたデータや情報の意味を人間が判断する形ではなく、**情報記述言語であるXML文書で、キーワードをキーワードとして認識できるメタデータを付与することによって人間の代わりにコンピュータが理解し、ソフトウェアで自動処理できるようにする新しいWebの形態**である。
* **Webページの個々の情報の意味情報であるメタデータを集約し、その知識を抽出して構造化することにより、インターネット上に散在するさまざまな情報全体をひとつの巨大な知識データベースと見なし、ソフトウェアの自動処理により効率的に処理し、コンピュータとの自然なやり取りで必要な情報が取得できるしくみが実現できる。**
* また、XML文書をスタイルシートであるXSLTの定義に従ってHTML文書化する形であれば、**スタイルシートを変えれば、印刷用、PC画面用、携帯電話用やWebコンテンツアクセシビリティのガイドラインを満たすWebページを動的に作成できシングルソース・マルチユースを実現**できる。

### Webサービスとしての連携

* WWW関連の技術を使い、ソフトウェアの機能をネットワークを通じて利用できるようにしたもので、**コンポーネント化された複数のWebサービス同士をつなぎ合わせてアプリケーションを構築する形**である。
  + この形は、検索キーを与えて検査結果を得るというデータベースアクセスというレベルではなく、**複数のWebサービスサイトに処理要求のレベルでのリクエストを出して、各サイトの処理結果を、XML形式でのレスポンスとして返すもの**である。
  + このシステム連携イメージは、B2B、B2Cでのシステム連携の基盤技術であり、それが、G２B2C での連携によるサービス提供につながる。
  + 同業種・業態でのそれぞれにWebサービスとしての連携仕様は確立しつつあるが、**デジタルアーカイブとしての連携においては、図書館の枠を越えて様々な機関が提供するポータル、サービスプロバイダ、データプロバイダ間での共通の規約を整備することが必要**である。
  + しかし、**この仕組みの技術基盤はすでに確立**しており、属性名、属性値の必要最低限の記述規則（目録規則）を調整すれば、早期に実現は可能となっている。

## デジタルアーカイブ・ポータルの構築・運用に関しての考察

* ポータルの構築にあたっては、情報提供者の意志と権利を十分に尊重しなければならない。
* また情報は、ある分野において最も影響のある情報が統合利用の対象外になっていては、真に有用なポータルとして使えない。

### コンテンツへの直接ナビゲーションが必要

* **ポータルは単なるサイトへのリンク集ではない。**ポータルがポータルを指して、どこまでアクセスしても、一次情報に辿り着かないポータルは敬遠される。

### ナビゲーションすべき情報へは確実に

* ポータルとして機能するためには、**連携しやすいサイトではなく、連携すべきサイトであり、重要な情報を持つサイトが対象外であってはいけない。**ロボット検索できるところでなく、含めるべきサイトは、ポータルから見えるようにしてもらうことが必要である。

### 情報提供者の意志を尊重する

* しかしながら、機械的に情報を収集されたくないと思っているサイトや、第三者権利等のため収集されては困る情報を、無理して対象としてはいけない。**提供者がポータルに登録されることを望むものと、ポータル側が見えるようにしたいものが一致できるようにすることが重要**。

### 誰でも付加価値をつけて運営できるように

* ポータルは、**ポータルとしてナビゲーションできる情報に付加価値をつけて情報を提供する**もので、**日本のポータルとして一つである必要はない**。
* 多種多様な利用者ニーズに応えるポータルは、いくつあってもいい。**それぞれがポータルのサービスコンポーネントを取捨選択して特色を持ったポータルを作れるようにすることが重要**と考える。ポータルは既製品を押し付けるものでなく、**利用者が自由にサービスコンポーネントを組み合わせて利用できるもの**でもある。
* 既にあるポータルをリンクするのではなく、それぞれのポータルのサービスをワンストップで使える１つのコンポーネントとして組み込んだポータルが作れることが望ましいと考える。

### データプロバイダは、情報収集ソフトからのアクセシビリティを確保

* **Webページは、人間に対しての対話型提供のアクセシビリティだけでなく、機械可読型提供に対してのアクセシビリティに配慮してもらう必要がある**。
* ポータル側から、**データプロバイダ側に用意して欲しい機能を提示して、インプリメントしてもらう。**その際は、将来的に普及する見込みのある技術を適用することが重要である。
* **国の情報の場合は、パブリックドメイン化と利用促進のための仕組みを用意がさらに重要で、個別に利用許諾したり、独自仕様で提供して、ポータル側に個別のアクセスインターフェースを用意させるような形ではいけない。**

### 考察のまとめ

* そのような考察のもとで、日本のデジタルアーカイブ・ポータルの構築を目指すNDLは、「ポータル構築のインキュベーション役」なのかもしれない。**全ての人を満足させられるポータルを運用することは困難**であり、NDLは、日本のポータルを提供する一機関として、**日本のポータル構築のために、情報を提供したい人が、情報を提供しやすくするための環境の普及を加速させることが、重要な役割**と言えるかもしれない。

## デジタルアーカイブおよびポータル構築に向けて

### 公的機関のデジタルアーカイブ構築の推進

* **公共機関の情報を広く一般が簡便に利用できるようにするために、公共的な情報資源のデジタルコンテンツ化と、そのアーカイブを積極的に推進**
* デジタルアーカイブの構築に当たっては、関係機関は、**デジタルコンテンツの長期保存とアクセスの保障に必要な機能と標準的な仕様の研究開発・技術開発**
* また、制度的な課題解決として、デジタルアーカイブとその利活用の推進に向けて、**国有財産に属する電子情報の利用の自由化あるいは利用手続きの明確化及び簡素化を図ると共に、図書館等で所蔵する書籍等のデジタル化及びその提供に関し、著作権者の不明な著作物を利用するための手続の効率化**

### デジタルアーカイブ・ポータルの推進

関係機関には、デジタルアーカイブ内の情報を安定的かつ効果的に利用できるようにするための技術的課題の解決に協力を得る。

* 改廃等により消滅してしまうシステム環境の中で、**電子情報格納フォーマットから情報を再現することが可能な長期保存技術**
* **メタデータの収集技術**
* **情報提供者によるメタデータ付与を容易にする技術**
* **新しい概念の検索方式や共通検索技術**
* また、デジタルアーカイブを構築する**公共機関等には、デジタルコンテンツはデータ連携を可能にする標準的な仕様で提供していただくことが必要**となる。。
* **標準的なフォーマット、インターフェース仕様の適用**
* **標準的なメタデータの付与**

各種ポータルサイトでは、**デジタルアーカイブへの標準的なアクセスのための仕様を実装し、それぞれのポータルがシステム的に連携して、相互補完しあい、全体で相乗効果を発揮できるようになることが望ましい**と考える。

### ウェブ・アーカイブ構想の推進

国内外の複数の主体によるウェブ・アーカイブが相互に連携し、一つのウェブ・アーカイブとして機能するための技術的課題の解決が求められる。

* **ウェブページの保存・検索等に要するメタデータ・フォーマットや自動情報収集等の技術の共通化・標準化**
* その仕様を満たした収集システムの開発
* **ウェブページ間の関連性・更新履歴を踏まえた情報解析等の研究開発**

ウェブでデジタルコンテンツを提供している各機関には、NDLが行うウェブ・アーカイブ構築に関して、それぞれの機関が提供しているウェブページを、**機械的に収集しやすくするための機能を実装する等の協力を得ることも重要**である。

* ウェブページにある情報のメタデータファイルの設置等
* 第三者の権利を侵害する恐れのある情報の収集拒否設定等

## おわりに

* デジタルコンテンツの利用を取り巻く環境は大きく変わりつつある。技術の進化により、提供可能なサービスが拡大し、そのサービスを受け入れた利用者からは、より高度なニーズが生まれてきている。
* デジタルアーカイブ・ポータルは、**図書館の枠を越えて、国、公共機関、学界に加えて民間、個人が保有する情報をワンストップで的確に閲覧利用できるようにする**ものである。
* インターネット上にある**膨大な情報を、「意味ある情報資源」として活用するための研究開発、技術開発を進める**とともに、情報の提供者はその技術を適用した情報提供することにより、**巨大な知識ベースが構築**できる。**デジタル情報を日本の文化遺産として後世に残し、新たな知識を創出するための知識として、現在及び将来にわたって活用できるようにする**ために、**データプロバイダ、サービスプロバイダのそれぞれの機関が「Win-Win」の関係で実施し発展していけることが重要**であり、関係機関・関係各位の御協力を願いたい。

# （2005年）各種デジタルアーカイブが提供するコンテンツの統合検索を目指して

国立国会図書館における国のデジタルアーカイブポータルの取り組み

「行政とADP」（2005年）

## 背景

国立国会図書館（NDL）では、2004年2月に「電子図書館中期計画2004」[[1]](#endnote-2)を策定した。この計画はNDLが所蔵している図書のデジタル化、オンライン系情報資源の収集によるデジタルアーカイブの構築、インターネット上にある情報を一つの大きなデジタルアーカイブとして利用できるようにする「国のデジタルアーカイブポータル」の構築と提供を目指すものである。

デジタルアーカイブの構築は、NDL蔵書の利用における地域間格差を改善し、利用者の利便性を高めるために当館所蔵の資料の電子化を推進することと、インターネット上の情報資源を広く収集し、消失を防ぐとともに、永続的な利用を確保することを目指す。

インターネット上の情報資源の収集では、ウェブ情報を可能な限り発信時と同様な構造をもったウェブアーカイブとしてサイト単位に収集し、時系列的認識が可能な形で蓄積・保存・提供する。また、知的な著作単位で取り扱うべき情報資源を対象として、個別に収集・組織化・保存・提供する。

ポータル機能は，利用者の必要とする情報をワンストップで入手できる窓口を構築する。当館のデジタルアーカイブにとどまらず、国等の公的機関を中心とした電子的情報資源や情報提供サービスに利用者を適切に案内するものである。

## 当館が目指すポータルのサービスは？

NDLが構築するポータルは、NDLが保有している資料をデジタル化したデジタルアーカイブ、ウェブアーカイブ、図書館としてのナレッジ情報や、冊子体の資料の目録と、図書館に限らず他の機関が提供している情報を「いつでも、どこでも、どこにあっても」一つの窓口で、探し出し、閲覧もしくは入手できるようにする。

「いつでもどこでも」は、インターネットでのサービスとしてごく当たり前のことだが、NDLの開館時間に関係なく、また、遠路NDLまで足を運ばなくても資料を検索、閲覧できるようにすることであり、「どこにあっても」は、閲覧したい資料を、NDLが提供しているデジタルアーカイブを含めて、どこのデジタルアーカイブに格納されたものであっても所在場所を指定せずに、探し出せるようにするっことである。また、デジタルコンテンツが、ウェブページとして作成された、いわゆる表層ウェブにあっても、データベース化されて検索条件を入力しなければアクセスできない、いわゆる深層ウェブにあっても、１つの検索窓で探して、一覧表示できるようにする。（図１）

また、それに加えて、どんな形態であってもデジタル化されたコンテンツであれば、Webブラウザでその場で閲覧できるようにする。有料のデジタルコンテンツの場合は、そのコンテンツの閲覧申し込みページへ、またデジタル化されていないものは、その資料を所蔵している図書館等への遠隔複写申し込み、販売している書店へのオンライン注文ページなどへ案内する。

　さらに、数年後には、図書館業務として、専門調査業務やレファレンス業務で得られた知識をナレッジデータベースとして構築し、この情報を含めて検索できるようにする。このようなナレッジデータベースに関しても、デジタルアーカイブと同様に統合検索できる環境の提供を目指す。

そのために、NDLのデジタルアーカイブ内のコンテンツを含めて、インターネット上で色々な形で提供されているデジタルアーカイブを、１つの巨大なアーカイブとして、検索・閲覧できる基盤を構築して、様々な機関がその共通の基盤の上で、それぞれの対象とする利用者に対して、情報やサービスに付加価値を付けることにより、きめ細かなサービスを提供できる世界を目指す。

## デジタルアーカイブのコンテンツをポータルから検索できるようにするために

現状の大半のデジタルアーカイブは、個別コンテンツの特性に応じて、Webブラウザでの閲覧用システムでのみ提供している。利用者はデジタルアーカイブ毎の閲覧システムで個別に検索して利用しなければならない。また、Webブラウザでアクセスすることが目的のインターフェースであり、他のシステムとの機械的な連携機能ではないため、ポータルのような他のシステムからは容易にアクセスすることはできない。

NDLが目指す方法は、ポータルとのデータプロバイダの間のインターフェースとして共通仕様を策定して、その共通仕様を満たすインターフェースをそれぞれのデジタルアーカイブで実装する方法。共通仕様は、このポータルのための仕様ではなく、他のシステムからも利用できる国際標準、業界標準に従ったものを想定する。

デジタルアーカイブ側が連携機能を実装する形になれば、どのデジタルアーカイブでも、機械的なアクセスが容易になり、統合検索の対象が自然に広がる。このような形での実現を目指している。

### ポータル成功の鍵はメタデータの収集

ポータルが有用なものになるためには、多くのデジタルアーカイブ内のコンテンツが検索・閲覧できるようになることであり、その鍵は、いかにコンテンツのメタデータを収集できるかにある。

過渡期の技術としては、表層ウェブにあるコンテンツはRSS[[2]](#endnote-3)により収集することを想定する。ごく一般のサイトでも普及している仕組みであり、容易に実現が可能と思われる。RSS等の提供ができない場合は、ページ内容情報を機械的に判断してメタデータを作成するツールや、メタデータを自動生成するウェブサイトの利用を想定する。深層にあるコンテンツは、OAI-PMH[[3]](#endnote-4)、SRW[[4]](#endnote-5)（Webサービス化したZ39.50）のようなインターフェースでメタデータを収集することを想定する。

デジタルアーカイブ側できちんとメタデータを付与することは大変な労力が掛かるが、ウェブサイト側でコンテンツ管理システム（CMS）[[5]](#endnote-6)等の利用が進めば、一次情報、二次情報にメタタグが付与され、機械的に意味解析が容易になり、いわゆるセマンティックWeb化が意識せず進むと思われる。そうすると、メタデータの収集が容易になる。

## 16年度プロトタイプの概要

NDLでは、各機関が保有するデジタルアーカイブ内のコンテンツを統合検索するために必要な技術を実験的に適用したプロトタイプシステムを構築し、その試験的提供を通じて、適用する技術の妥当性及びサービスの有用性を検証している。

* URL：http://www.dap.ndl.go.jp/

### どんなことができるのか

複数のデジタルアーカイブ内の対象コンテンツを一つの検索窓から一元的に検索でき、検索結果が一覧で表示され、1次情報の提供画面へ辿り着ける。また、デジタルアーカイブ内のコンテンツまでは一元的に検索できないが、有用なコンテンツを持つウェブサイト及びデータベースへは、その入り口までナビゲーションする。

### 構築方針

プロトタイプの構築に当たっては、下記のような方針で実施した。

* 先進的かつ将来標準的な仕様となることが見込まれる技術の適用
* 適用事例が多いオープンソースの活用。可能な限り新規開発はしない。カスタマイズは必要最低限とする。
* 各々の機能は独立したWebサービス機能として、他システムからも容易に利用できるものを目指す。その仕様としては、図書館界のみならず、デジタルアーカイブの世界で標準となり得る仕様を採用

### 適用技術

　構築方針を踏まえ、以下の技術を採用した。

* フロントエンドとして、XOOPS[[6]](#endnote-7)の利用
* 日本語の分かち書きをするため、Chasenを採用
* データの通常の全検索にはNamazuを採用
* 連想検索にはGETA[[7]](#endnote-8)を採用
* メタデータの収集にはOAI-PMHを採用。横断検索では、SRWを採用。
* システム間連携はWebサービス化
* 書誌データの保存システムには、OAIS[[8]](#endnote-9)に準拠したDSpace[[9]](#endnote-10)及びDIAS[[10]](#endnote-11)を利用
* 画像のデジタル化フォーマットとして、JPEG2000[[11]](#endnote-12)を適用

## デジタルアーカイブおよびポータル構築に向けて

デジタルアーカイブ及びデジタルアーカイブポータルは、NDLが自ら構築・運用に向けて努力するとともに、関係機関との連携・協力を求めることなしに実現することはできない。

### 公的機関のデジタルアーカイブ構築の推進

* 公共機関の情報を広く一般が簡便に利用できるようにするために、公共的な情報資源のデジタルコンテンツ化と、そのアーカイブを積極的に推進することが望まれる。
* デジタルアーカイブの構築に当たっては、関係機関はデジタルコンテンツの長期保存とアクセスの保障に必要な機能と標準的な仕様の研究開発・技術開発に協力することが望まれる。
* また、制度的な課題解決として、デジタルアーカイブとその利活用の推進に向けて、国有財産に属する電子情報の利用の自由化あるいは利用手続きの明確化及び簡素化を図ると共に、図書館等で所蔵する書籍等のデジタル化及びその提供に関し、著作権者の不明な著作物を利用するための手続の効率化が望まれる。

### デジタルアーカイブポータルの推進

関係機関には、デジタルアーカイブ内の情報を安定的かつ効果的に利用できるようにするための以下の技術的課題の解決に協力を得ることを期待する。

* 改廃等により消滅してしまうシステム環境の中で、電子情報格納フォーマットから情報を再現することが可能な長期保存技術
* メタデータの収集技術
* 情報提供者によるメタデータ付与を容易にする技術
* 新しい概念の検索方式や共通検索技術

また、デジタルアーカイブを構築する公共機関等には、デジタルコンテンツはデータ連携を可能にする標準的な仕様で提供していただくことが必要となる。

* 標準的なフォーマット、インターフェース仕様の適用
* 標準的なメタデータの付与

各種ポータルサイトでは、デジタルアーカイブへの標準的なアクセスのための仕様を実装し、それぞれのポータルがシステム的に連携して、相互補完しあい、全体で相乗効果を発揮できるようになることが望ましいと考える。

## おわりに

デジタルコンテンツの利用を取り巻く環境は大きく変わりつつある。技術の進化により、提供可能なサービスが拡大し、そのサービスを受け入れた利用者からは、より高度なニーズが生まれてきている。

デジタルアーカイブポータルは、図書館の枠を越えて、国、公共機関、学界に加えて民間、個人が保有する情報をワンストップで的確に閲覧利用できるようにするものである。

インターネット上にある膨大な情報を、「意味ある情報資源」として活用するための研究開発、技術開発を進めるとともに、情報の提供者はその技術を適用した情報提供することにより、巨大な知識ベースが構築できる。デジタル情報を日本の文化遺産として後世に残し、新たな知識を創出するための知識として、現在及び将来にわたって活用できるようにするために、データプロバイダ、サービスプロバイダのそれぞれの機関がお互いにインセンティブをもって実施し発展していけることが重要であり、関係機関・関係各位の御協力を願いたい。

http://www.ndl.go.jp/jp/aboutus/elib\_plan2004.html

RSSとは、Rich Site SummaryもしくはRDF Site Summaryの略。Webサイトの見出しや要約などのメタデータを構造化して記述するXMLベースのフォーマットで主にサイトの更新情報を公開するのに使われている。

OAI-PMHとは、Open Archives Initiative Protocol for Metadata Harvestingの略。OAI（Open Archives Initiative）が策定した、ウェブ上のメタデータを収集（メタデータ・ハーベスティング）するためのプロトコル

SRWとは、Search/Retrieve Web Serviceの略。データの検索およびデータの取得に使用することを想定したWebサービスのインタフェース仕様

CMSとは、コンテンツマネージメントシステム。Webコンテンツの作成、更新及び管理を行うシステムで、トップページも最新情報を含めて動的なページとして容易に配信することができる。

XOOPSとは、コンテンツマネジメントシステムの１つで、オープンソースライセンスに基づいたソフト。コミュニティサイトを容易に構築できる。

GETA(Generic Engine for Transposable Association)とは、連想検索のエンジンで、入力した文章から「連想」されるような、文章と関連性の高いと思われるコンテンツを検索するもの

OAIS（Open Archival Information System）とは、情報を長期保存するために必要とされる機能要素と、そこで扱う情報の構成の規定がなされた技術標準

DSpaceとは、マサチューセッツ工科大学図書館とヒューレット・パッカード研究所が共同開発した,OAIS参照モデルに準拠したデジタル情報保存システム

DIAS(Digital Information Archiving System)とは、オランダ国立図書館とオランダIBMが共同開発した、OAIS参照モデルに準拠したデジタル情報保存システム

JPEG2000は、画像圧縮方式の一つ。国際標準規格であるJPEGを発展させた形で、高圧縮かつ高品質なデータ圧縮が可能

# （2006年）Web2.0世代のデジタルアーカイブポータルの提供を目指して

（国立国会図書館デジタルアーカイブポータルの現状と今後）

Current status and future of “NDL Digital Archive Portal”

-portal service of digital archives in the “Web 2.0” era-

2006年9月1日

久古聡美　1;　吉田曉　1 ;　中山正樹　2

KYUKO Satomi 1; YOSHIDA Satoru 1; NAKAYAMA Masaki 2

## 要旨

国立国会図書館が目指している「国立国会図書館デジタルアーカイブポータル」は，広く国のデジタルコンテンツ全体へのナビゲーションを行うことを目指すものである。平成19年度以降の稼働に先立ち，プロトタイプを構築し「統合検索」のために必要な機能の検証を行ってきた。本稿ではプロトタイプの概要を示し，さらに今後の方向性について述べる。デジタルアーカイブを保有する各機関にウェブサービスによるサービス提供機能の実装を呼びかけると共に，Web2.0時代にふさわしいポータルの構築を目指し，機械的な連携とデジタル情報の一層の利活用を図る。

The National Diet Library (NDL) is planning to construct “NDL Digital Archive Portal,” which aims at navigating its users to any digital information in Japan. Before establishing the system, which will happen at the beginning of fiscal year 2007, we have constructed its prototype system, and have verified the required functions. In this paper, we will show the outline of this prototype system and our plan. We not only construct the portal site according to ”Web 2.0,” but we also encourage digital archives of other institutions to implement the service functions using web services. These actions will increase the automatic coordination between systems and improve the usability of the digital information in Japan.

## キーワード

デジタルアーカイブ，ポータル，ワンストップナビゲーション，ダブリン・コア，辞書，Web2.0，OAI-PMH，SRW，RSS，CGM，ウェブサービス，SOA

Digital Archive, Portal, One-Stop Navigation, Dublin Core, Dictionary, Web2.0, OAI-PMH, SRW, RSS, CGM, Web Service, SOA

## はじめに

国立国会図書館（以下，「当館」という。）では，国のデジタル情報全体へのナビゲーションを行う総合的なポータルサイトとして，｢国立国会図書館デジタルアーカイブポータル｣（以下，「ポータル」という。）の構築を目指している。これは，当館の所蔵資料のみならず，当館が所蔵していない情報に関しても所蔵元へ案内する役割を果たすものであり，有用な各機関が提供しているデジタルアーカイブ内の著作物を，「いつでも，どこでも，どこにあっても」，ワンストップで検索，閲覧できるようにして，資料の利活用が促進されることを目指している。

当館では，当館を含む公共図書館の蔵書目録を一元的に検索できるようにする事業を実施しているが，デジタル情報に関しても同様の仕組みの提供が必要である。当館は，蔵書のデジタル化を進め，インターネットで提供している。また，各機関がインターネット上で提供している情報を収集し，長期保存している。しかし，リアルタイムに更新されるすべてのインターネット情報を当館が収集し，提供することは現実的ではない。

また，今までの検索エンジンでは，データベース内の著作物を直接検索することは困難であった。ポータルは，その各機関が保有するデジタルアーカイブ内のコンテンツを中心に「統合検索」し閲覧できるようにするものである。

「統合検索」とは，複数のデジタルアーカイブシステムのコンテンツを検索対象とし，次に挙げる点の実現を想定しているものである。

* 一つのキーワード入力画面(検索窓)から検索できること。
* 検索結果はデジタルアーカイブの所在や様態に依らず統合的に取り出すこと。
* 可能な限り検索結果から利用者が目的とする情報そのもの(一次情報，又は，紙媒体の資料はその入手元)までたどりつけるようにナビゲートすること。

さらに，ポータルサイトの機能として，利用者の情報行動を支援するための様々な機能や仕組みを充実させていく必要がある。一つの方向性として，Web2.0の概念を取り入れる可能性を考えている。

当館では，平成19年度以降の稼働を予定しているポータルの構築に先立って実験システムであるプロトタイプシステム（以下，「プロトタイプ」という。）を構築し，その試験提供を行ってきた。本稿では，まずプロトタイプ構築に関する概要と主な機能及びそこでの検証成果について紹介する。次に，Web2.0時代にふさわしいポータル機能の提供及び今後の方向性に関する考察を行う。

## プロトタイプシステムの構築

プロトタイプは，ポータルの構築に先立って，サービスの有用性及び適用技術の妥当性に関する検証を行うために構築し，試験提供を行うものである**注1)**。平成16年度から開発に着手し，平成17年7月には一般への試験公開を開始した。

公開後は，試用をしていただいた方々からご意見をいただくと共に，検索機能等について開発の立場から次々と課題等が洗い出されてきた。平成17年度は，これらの検証成果を踏まえて機能強化を進めた。新たに機能を開発するとともに，問題点の改善を行い，随時プロトタイプ上で追加公開していった。検索対象は，他機関のご協力を得ながら順次追加していき，現在は，次に挙げるデジタルアーカイブシステム全14種類が検索対象となっている。（[ ]内は，所蔵機関又はご協力いただいた機関）

●書籍，出版物，目録等

* 青空文庫 [青空文庫]
* 近代デジタルライブラリー [NDL]
* NDL所蔵貴重書サンプル（サンプル的にデジタル化したもの） [NDL]
* 貴重書画像データベース [NDL]
* 国立公文書館デジタルアーカイブ・システム [国立公文書館]
* 新書マップ [NPO法人連想出版]
* NDL蔵書目録 (和図書・和雑誌) [NDL]
* NDLプランゲ文庫（日本占領期検閲資料コレクション）雑誌・新聞目録 [NDL]
* NDL雑誌記事索引（2005年分のみ） [NDL]

●有用サイト，データベース情報

* 府省所管のデジタルアーカイブサイト情報 [内閣官房]
* Dnavi（データベース・ナビゲーション・サービス) [NDL]

●郷土情報

* デジタル岡山大百科 [岡山県立図書館]

●その他の参考情報

* カレントアウェアネス [NDL]
* レファレンス協同データベース [NDL]

### プロトタイプの各機能

統合検索では，キーワード検索に加えて自然文から検索する連想検索や分類を通じての検索など，複数の角度からデジタルアーカイブが検索可能となっている。各機能を用途に応じて使い分けることで，情報探索の幅を広げることが狙いである。

ここでは，プロトタイプで提供している統合検索の各機能について紹介する。

#### 統合検索(キーワード検索)

キーワード検索は，利用者が入力したキーワードに合致するデータを含む資料を検索する。

例えば，キーワード「吾輩　猫」を入力し，検索ボタンを押すと，検索結果が一覧表示される(図2参照)。ここで，「吾輩ハ猫デアル」が二つ得られるが，一つは「青空文庫」のものであり，もう一つは「近代デジタルライブラリー」のものである。通常この二つの結果を得るには，各々のサイトを訪れて別々に探す必要がある。しかし，統合検索では，このような手間をかけずに，一つの検索窓から複数のデジタルアーカイブシステムのメタデータを一度に検索し，結果を一覧表示する。

この検索結果一覧表示におけるタイトル部分のリンクからは，詳細情報表示画面へ移動するようになっている。ここでは「近代デジタルライブラリー」を例に見ることにする(図3参照)。

この画面で表示される情報は，デジタルアーカイブシステムからポータルが取得したメタデータで構成される。左側にはDublin Coreをベースにしたメタデータ要素が並ぶ。このうち「URL」は，デジタルアーカイブシステム（この場合「近代デジタルライブラリー」）での書誌情報画面へのリンクとなっている(図4参照）。

書誌情報画面には一次情報へのリンクがあり，利用者はこれを辿って一次情報を閲覧することができる(図5参照)。

このように，利用者はキーワードを入力し検索を行うところから，一次情報の閲覧までナビゲートされたことが分かる。「青空文庫」の場合もほぼ同様の画面遷移を経て，テキストファイル化された資料の本文を読むことができる(図6参照)。

キーワード検索には，辞書による検索支援機能がある。例えば，キーワード「コンピュータ」を入力すると，図7のような結果を得る。

統制語「コンピュータ」のデータが検索結果として得られ，そのシソーラスとして類義語，上位語，下位語，関連語が表示されている。利用者は表示された語群を検索のヒントとして得た上で，この中からより検索キーワードとして適切と思われる語を選択して再検索を行うことができる。

現在この辞書データには，当館の普通件名典拠，固有名件名典拠，著者名典拠を用いている。

#### 統合検索(連想検索)

連想検索は，入力したキーワードから「連想」されるようなコンテンツを検索するものである。文章などの自然文から関連する資料を検索できるのがその特徴の一つである。

例えば，Wikipediaの「源義経」の冒頭部分の文章を連想検索の検索窓にコピー＆ペーストして(図8参照)検索を行うと，義経に関連があると思われる資料が検索される(図9参照)。

#### 統合検索(分類による検索)

分類による検索では，日本十進分類法（NDC)に従い階層的に整理されたカテゴリのリンクをたどることで，資料を検索する。

例えば，浮世絵に関連する資料を見たいとする。

NDCの最上位の分類が表示されている画面で，浮世絵に最も関連があると思われる「芸術」を選択すると「芸術」の下位分類が表示される(図10参照)。

この中で浮世絵に最も関連があると思われる「絵画. 書. 書道」を選択する。すると第３分類が表示され，そのうち「日本画」を選択すると，浮世絵を含む「日本画」に関連する資料が一覧表示される(図11参照)。

この結果一覧表示から浮世絵の画像等を閲覧するまでの画面遷移は，他の検索の場合と同様である。また，第2分類以降は，選択した分類に属するデータに対してキーワードによる絞込み検索を行うことができるようになっている。

#### その他

この他，統合検索機能として「新着コンテンツ」「更新コンテンツ」「アクセスランキング」「キーワードランキング」がある。

また，利用者の情報入手全般を支援する機能を実験として提供している。RSS**注２)**を利用し，「デジタルアーカイブ関連」をテーマに外部のブログ等を含む他サイトが配信する情報を集めて表示している。また，プロトタイプからのお知らせや更新情報，新着コンテンツ，更新コンテンツを一覧表示し，それを個々の利用者がRSSリーダで受け取って閲覧できるようにしている。

### 統合検索機能の実装と「共通仕様」

図12はプロトタイプの実装イメージ図である。

上部の楕円は，検索対象又はその候補となる個々のデジタルアーカイブシステムを表している。

デジタルアーカイブシステムは，自身が保有するデジタル情報などを管理するためにそのメタデータを保有している。これらのメタデータを検索可能とするため，プロトタイプのサービスプロバイダでは，機械的に連携し，メタデータを事前に収集してメタデータデータベースに格納しておく方法（以下「ハーベスト」と呼ぶ。）又は利用者からの検索要求の都度にデジタルアーカイブシステムへ検索をかけて結果を受け取る方法（以下「横断検索」と呼ぶ。）のどちらかを採っている。

さらにポータルでは，ハーベストと複数の横断検索の結果を統合し表示している(図13参照)。これにより，利用者はデータの所在や検索方法の差異を意識することなく，一元的な結果の閲覧が可能となる。

現在，デジタルアーカイブシステムが保有するメタデータ要素は，各々のシステムに依存した定義により作成され，プロトコルも統一されていない状態である。ポータルでそのようなメタデータを統合的に扱うためには，個別に対応して他システムのメタデータ要素との差異を調整しなければならず，統合的に扱うことが難しいものとなっている。この問題は連携対象が増加するほど複雑になる。そこで，メタデータをハーベストもしくは横断検索するためのメタデータ要素，記述規則及び通信プロトコルの「共通仕様」を予め定めておくことが必要となる。各デジタルアーカイブシステム側でその仕様に沿ったインタフェースを実装してもらうことで，ポータルとしては連携対象をよりスムーズに拡大することができ，保有コンテンツが一層利活用されることになる。それによって個々のシステム間の連携も容易になり，また，各検索エンジンや各種ポータルがシステム内の情報を自由に検索でき，副次的に有用な連携が生まれることが期待される。（このような連携を行える機能を備えたデジタルアーカイブシステムを「データプロバイダ」と呼ぶことにする。）

プロトタイプ構築にあたっては，暫定的ながら「共通仕様」の案を定めた。今後は，他機関と調整を行いながら，案を確定し，普及していく予定である。

#### 「共通仕様」(メタデータ要素)

データプロバイダとポータルとのデータ交換に使用するメタデータ要素は，最低限Dublin Coreに定義されたものは可能としている。Dublin Coreは業種を超えてある程度普及しているという利点があるが，検索やその他実現したい機能のために必要と思われるメタデータ要素が定義されていない場合があった。そこでプロトタイプでは，Dublin Coreにいくつか必要と思われる要素や独自の限定子を追加して，「共通仕様」の案に取り入れている。

#### 「共通仕様」(通信プロトコル)

プロトタイプでは，次の通信プロトコルに関しても実験を行った(図14参照)。

・ハーベスト：OAI-PMH, SOAP, RSS

・横断検索：SRW, Z39.50

　「共通仕様」では上記の他，RSSやその他ウェブサービスによる横断検索も想定した。

### プロトタイプでの検証結果

検証の結果，複数のコンテンツを一元的に検索し，利用者を一次情報までワンストップでナビゲートすることが可能であり，有用だと確認できた。ただし，プロトタイプの機能やサービスはまだ実用としては十分でない点もあると認識しており，検証結果として問題点も多数洗い出されている。例えば，次のような点である。

* 検索機能の面では，ハーベストと横断検索の結果のマージに問題があり結果の並べ替え等に問題が生じている。
* ハーベストしたメタデータの分かち書きに問題があり，入力されたキーワードから適切に検索できないケースがある。
* 横断検索の手法では，横断検索先の検索機能に依存するので，他と結果の取り出しに相違がでてしまう可能性がある。
* 適用技術の検証部分に特に注力したため，利用者にとって大切な，操作性やサイトの見た目等への配慮が欠けている。

以上は，検証結果のほんの一部であるが，本格システムの構築においてはこれらの成果を十分活かしていく所存である。

## Web2.0世代のポータルを目指して

ポータルにおいて現在注目しているのは「Web2.0」と呼ばれる概念である。

Web2.0[[12]](#endnote-13)1)とは，このような名称の技術があるわけではなく，既存のウェブを「Web1.0」と総称し，これの対比として次世代のウェブに見られる特徴的な姿勢を示したものだと捉えることができる。Web2.0の性質を端的に表現する言葉として，次のようなものがあると言われている2)。

「ユーザの手で自由に分類」「リッチなユーザ経験」「ユーザ体験の蓄積をサービスに転化」「ロングテール」「ユーザ参加型開発」「ユーザ生成型コンテンツ」「知のオープンソース」「進歩的分散志向」

Web2.0で論じられる世界を(部分的に)実現する一形態として，既存サービスでは例えば，GoogleMap，Google Adsense, AmazonのユーザレビューやBlog, mixi, Wikipediaなどがよく言及されている。既存の技術も見直され，Ajax**注３)**が注目されている。

Web2.0の概念は，利用者の利便性を向上させるものであり，ポータルでも積極的に取り入れていくべきであると考える。本章では，Web2.0の概念に適合する要素を取り入れる可能性への考察を交え，デジタルアーカイブポータルの今後の展望を紹介する。

### メタデータの取得

ポータルでのナビゲーションは，各デジタルアーカイブシステムが持っているメタデータを収集，もしくは，ポータル側から横断検索できるようにしていただくことにより実現する。前述した「共通仕様」に従った提供機能があるデータプロバイダの場合，ポータルは円滑にメタデータを取得することが可能である。

このような提供機能がないシステムの場合，ポータルを含む他システムからアクセスするため，「共通仕様」に準拠した仕組みを提供してもらう。実装を促すために，当館が汎用的な提供機能のソフトウェアを開発して配布し，それを他システムに適用してもらうことも一法として考えている。

### RSSの活用

RSSはWeb2.0の「ユーザ生成型コンテンツ」の促進や「進歩的分散志向」を実現する技術といえるだろう。ウェブログ(Blog)におけるRSSを用いた更新情報の自動作成・配信機能，また，配信されたRSSデータの「RSSリーダ」を用いた受信はよく見られる。

またRSSはデータ構造が単純でありかつ拡張性があるという特徴を持つ**注4)**。この適用容易性ゆえに，ポータルではRSSの可能性に注目している。

ポータルのメタデータ収集場面での適用を考えると，他のサイトから配信されたRSSデータをサイト側の改修負担をかけることなく収集し，統合検索対象のコンテンツとして用いることもできるだろう。また，通信するデータフォーマットにRSS2.0を用いた検索「OpenSearch」3)を用いることも考えられる。これにより，一層RSS配信元のコンテンツが活用されることが期待できる。

### ソーシャルブックマーク

Web2.0の考え方では，利用者がコンテンツ作成に「参加」したり，「ユーザ体験の蓄積をサービスに転化」し他の利用者が利用可能としたりすることが行われている。

ポータルでは，データプロバイダ及びポータルでのメタデータ作成機能の一つとして，利用者が興味を持ったコンテンツやサイトをブックマークする機能を想定している。ブックマーク登録時に該当コンテンツのメタデータを取り出してデータベースに蓄積するが，その際利用者はフリーキーワードを含めたタグ(カテゴリ)を付与可能としておく。利用者が推薦したサイトを検索対象とし，さらに利用者の手で付与したカテゴリを用いた検索を行うことで，利用者の知識と体験を取り込むことができる。

### 辞書の活用

プロトタイプでは，当館の典拠データを辞書として用いたが，他の有用な辞書データを追加して共に検索可能とすることで，より幅広い支援を実現したいと考えている。辞書として有用なデータを蓄積したサーバ機能を用意し，外部システムと辞書検索のウェブサービスで連携することも想定している。例えば，WikipediaなどはWeb2.0の「ユーザ参加」により編集された，評価の高い「知のオープンソース」の一つの実現形態であり，このような辞書を取り入れることも考えられる。

辞書はそれ自体の検索サービスのみならず，プロトタイプの辞書による検索支援と同様の機能の実現やメタデータの自動付与及び統制等の様々な場面において活用が可能である。さらに，OWL(Web Ontology Language)を用いて辞書データを記述するなど，セマンティックWeb技術を活用すれば，機械的に意味解析可能となり，さらに活用の幅が広がると考えられる。

### メタデータの自動生成，組織化，保存

　ポータルで扱うメタデータは，先に示した「共通仕様」におけるメタデータ記述要素を用いることを基本としている。しかし，一般の組織では，図書館のように統制された書誌を作成することは難しいだろう。そもそも，多くのデジタルコンテンツに手作業できちんとしたメタデータを付与することは困難である。

メタデータ収集時や検索インデクス作成時には，辞書の活用や日本語構文解析技術等を活用して，他のメタデータ要素の値や一次情報のデータを解析し，メタデータ要素を推測して可能な限り機械的に付与することが考えられる。ただし，期待される精度が得られない場合のため，人手による一部要素の入力や，自動作成した結果の修正等（メタデータの半自動生成）を取ることを考慮する必要がある。

### メタデータ提供とインタフェース

ポータルがサービスを提供する際，そのインタフェースは大きく二つに分けられる。利用者に対するGUI（Graphical User Interface），システム間連携のためのAPI（Application Program Interface）である。

　GUIについては，従来のLook&Feelに優れたレイアウト・デザインの考慮は勿論だが，Web2.0に基づく「リッチなユーザ経験」を実現するために，利用者は特段の説明がなくとも直感的な操作が行えるGUIを備えていることが望まれる。Ajaxを用いることで，マウスポインタを当てた部分のヘルプを半透過表示したり，統合検索における結果表示において，特定のメタデータ情報の展開表示をしたりすることが可能となる。また，利用者自身により，表示するブロックを選択可能としたり，表示ブロックの配置をマウスのドラッグ操作により動かすなど，利用者毎に使い勝手の良いポータル画面を容易にカスタマイズできるパーソナライズ機能が実現可能だと考える。

　統合検索機能においては，収集したメタデータに対して辞書を活用することで，入力キーワードの推定表示等が可能と思われる。また，パーソナライズ機能として，利用者の志向に応じて検索対象を自動的に絞り込む機能を備えることも考えられる。

### ウェブサービスによる連携

Web2.0の「進歩的分散志向」に従うと，今後はそれぞれのデジタルアーカイブシステムが，ウェブ上で他システムからサービスとして利用可能な形での連携に対応が求められる。このようなウェブサービスによる連携機能が提供されると，コンテンツの検索サービス，閲覧サービス，貸出・複写申込みサービス，Amazonのように注文サービス等を自由に利用したサービスを提供することが可能になる。

ポータルでは，利用者がGUIを通じてメタデータ検索を行う機能だけでなく，統合検索などの各提供サービスについて，ウェブサービスのAPIとして提供することが好ましいと考えている。国等の公共機関に関しても，同様のサービスの提供を期待したい。例えば，公共図書館が「情報ハブ」として各地域のデジタルコンテンツ等を集約する枠組みが報告されているが，この集約した情報を提供するウェブサービスを提供する，すなわち蔵書目録検索サービス，複写申込みサービス等を独立したウェブサービス機能として構築しAPIを公開すれば，その情報が広く利活用されることになる。

各種ポータルサイト間でもシステム的に連携して相互補完しあい，全体で相乗効果を発揮できるようになることが望ましいと考える。

### Blog, wiki

　ポータルの説明文書や開発運用等の最新情報は，BlogやWikiを用いて編集し，利用者からもコメント付与や編集を可能とすることが考えられる。Web2.0の「ユーザ生成コンテンツ」やCGM（Consumer Generated Media）を実現する一つの方法である。Wikiはプロトタイプでも導入したが，Wikipediaのように，知識豊富な(不特定な)利用者の力を借りることで想像以上に優れた文書が完成することもあると考えられる。

### システム基盤

ポータルでは，多くの利用者及びシステムからアクセスがあると想定される。また，ポータルでの各種サービスに必要なデータは増加していくと予想できる。そこで，これらを考慮したシステム基盤の構築が不可欠である。

現在主流のアーキテクチャとして，システムをサービス単位で構築し，サービス間は疎結合な関係とするSOA （Service Oriented Architecture）が提案されている。SOAはサービスの改修の影響が他のサービスに及びにくい。また，グリッドコンピューティング技術を適用し，スケーラビリティ（拡張性）に富むシステムを構築することが可能である。

ポータルでもこのような技術を適用し，改修の負担の最小化，システム管理の省力化を図る想定である。

## 終わりに

インターネット上にある膨大な情報は，「意味ある情報資源」として活用するための技術開発を進めるとともに，情報の提供者はその技術を適用し提供することにより，巨大な知識ベースとなっていく。デジタル化されていく情報を日本の文化遺産として後世に残し，新たな知識を創出するための知識として現在及び将来にわたって活用できるようにするため，データ提供者，サービス提供者各々の機関が互いにインセンティブをもち，相互補完しあいながら発展していけることが重要である。当館は国のデジタルアーカイブポータルを構築し提供することにより，各機関のデジタルアーカイブが，他システムからの利用を意識したサービス構築のための仕組みを実装することの普及促進の役割を果たしたいと考えている。

利用者の日常の情報探索行動を支援するサービスとして，利用者のニーズに応えられる国のポータルを構築するためには，関係機関が持つコンテンツ，ウェブ上で提供しているサービス，および今後普及が見込まれる次世代ウェブに関連した有用な技術を活用させてもらうことが重要であり，関係機関と協力して実現していきたい。

注1)プロトタイプのURL： http://www.dap.ndl.go.jp/home/

注2) RDF Site Summary 又は Really Simple Syndicationの略。ウェブの見出しや要約を構造化して記述するフォーマットで，ニュースサイトやBlog等の更新情報の配信の用途によく用いられている。

注3) Asynchronous JavaScript + XMLの略。サーバと非同期に通信を行うことで，操作の結果をリロードすることなく随時に取得可能とする。これにより，サーバを意識しないスムーズな操作性が実現する。

注4) RSSの活用可能性を論じたものとして，次の文献がある。

林賢紀, 宮坂和孝「RSS(RDF Site Summary)を活用した新たな図書館サービスの展開」，情報管理Vol.49, No.1, April 2006.

1) Tim O'Reilly, “What Is Web 2.0”. (online), available from <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>, (accessed 2006-07-07).

2) 近江商人 JINBLOG. “Web2.0 とは　－7つの分類と要素MAP”. (オンライン), 入手先<http://ceonews.jp/archives/2005/10/web20\_7map.html>, (参照 2006-07-07).

3) “OpenSearch”. (online), available from <http://opensearch.a9.com/>, (accessed 2006-07-07).

図1　プロトタイプのトップページ

図2　検索結果の一覧表示

図3　詳細情報表示画面

図4　URLのリンク先（近代デジタルライブラリー）

図5　近代デジタルライブラリーの本文

図6　青空文庫の本文

図7　辞書による検索支援の表示画面

図8　連想検索の検索窓

図9　連想検索の検索結果一覧

図10　分類（NDC）による検索　第2分類

図11　分類による検索の検索結果一覧

図12　平成17年度プロトタイプの実装イメージ

図13　統合検索のイメージ図

図14　適用した通信プロトコルの概念

# （2006年）国立国会図書館における電子図書館サービスのゆくえ

図書館雑誌平成18年11月号原稿

## はじめに

本稿では，電子図書館などに関する近年の国立国会図書館（以下「館」という。）の取組みについて、電子情報の保存と利用を中心に説明する。

最初に電子図書館事業の館における位置づけを明らかにしたい。館は，果たすべき使命・役割及び将来目指すべき方向性を明確にするため，平成16年に「国立国会図書館ビジョン2004」を定めた 。その中で次の3点を館の使命として挙げた。

・国民の知的活動の成果を，印刷物から電子情報にいたるまで広く収集し，国民共有の情報資源を構築する。

・国政課題に関する調査・分析及び情報の提供によって，国会の立法活動を補佐する。

・行政・司法各部門及び広く国民に対し図書館サービスを提供し，現在及び将来にわたり，情報資源へのアクセスを保障する。

つまり，電子情報を収集すること，当館資料を簡単に利用することができるようにすることは館の使命として位置づけられている。このビジョンを実現するため，4つの領域に重点を置いており，「デジタルアーカイブの構築」はそのひとつである。すなわち，「国民共有の情報資源として，電子情報を蓄積・提供するデジタルアーカイブを構築する」ことが目標である。同じく平成16年には「電子図書館中期計画2004」を定め，5年程度を目途として達成すべきことを示した 。

以下においては，資料のデジタル化と電子展示会，パッケージ系の収集・保存・提供，ネットワーク系の収集・保存・提供，デジタルアーカイブポータルの順で説明する。

## 資料のデジタル化と電子展示会の提供

館は，所蔵する資料のデジタル化を行ってきた。この結果，インターネットにつながっていれば，世界中どこからでも「近代デジタルライブラリ」によって明治期刊行図書約13万冊を閲覧することができるようになった 。提供開始以来の画像表示件数は800万件を超えている。著作権保護期間終了が不明の図書については文化庁長官裁定に基づき提供を行った。現在は，明治期に続いて大正期図書（約9万冊）のデジタル化の準備を行っているところである。

このほか，館は「児童書デジタルライブラリ」，「貴重書画像データベース」，「国会会議録」，「帝国議会会議録」をインターネット上で提供している。これらはレファレンスのツールとしても利用することができるであろう。また，これまでのデジタル化の経験を踏まえ，『資料デジタル化の手引き』を作成し，提供しているので，参考にしていただければと思う 。

また，平成10年の「デジタル貴重書展」以来，館は電子展示会を提供してきた。現在，「日本国憲法の誕生」，「史料にみる日本の近代」など9点を提供している。電子展示会は，館ホームページの中でも注目度の高いコンテンツであり，平成17年度末の時点で，それまでに提供されたコンテンツを累計して約1742万ページのアクセスがあった 。今後とも電子展示会を追加していく。

## パッケージ系の収集・保存・提供

電子出版物は大別してパッケージ系とネットワーク系に区別することができる。パッケージ系とは，CD-ROMやDVDのように有体物に固定された電子出版物のことである 。

館は，図書や雑誌などの図書館資料を国立国会図書館法の規定に基づき，できるだけ網羅的に収集してきたが ，パッケージ系の収集については，納本制度調査会の調査審議を経て，平成12年度に国立国会図書館法の改正が行われ，納本制度に組み入れられた。これにより，館は，図書や雑誌と同様に，パッケージ系の収集を開始した。収集されたパッケージ系は日本全国書誌やNDL- OPACにも掲載されており，また東京本館電子資料室において閲覧・利用することができる（ゲームやアプリケーションを除く。）。

課題は，こうして収集したパッケージ系を将来も利用することができるかどうかにある。例えば8インチのフロッピーディスクが保存されていたとしても，これを再生する装置はどこにでもあるわけではない。また，コンテンツを利用するのに必要なオペレーティングシステムなどのソフトウェアも絶えずバージョンアップがなされており，必ずしも完全に互換性があるとは限らない。こうした課題に対する対策としては，電子情報を他の媒体へ移行すること（マイグレーション）や昔の環境を擬似的に再現すること（エミュレーション）が考えられる。

館は平成15年度・平成16年度にかけて，パッケージ系の長期的保存とアクセス手段の確保をテーマに，委託調査を行った。これによれば，サンプル調査したパッケージ系（200点）の約7 割（138点）に利用上の問題があることが明らかになった。また，異種媒体への移行としてのマイグレーションは容易に実施できること，しかしエミュレーションやファイル形式の変換としてのマイグレーションは現時点では有効な技術的な解決手段とは言いがたいことが判明した 。

## ネットワーク系の収集・保存・提供

パッケージ系と異なり，ウェブ上の情報や電子ジャーナルといったネットワーク系は，納本制度に基づく収集対象に含まれていない。しかし，インターネット上の情報は1年で4割が消滅するといわれている。また，当館に納本されていた雑誌のうち，紙媒体での刊行をやめ，ウェブ上での提供に切り替えられたものが既に約250タイトル程度存在する。こうした事情を踏まえて，館は平成14年以来，「インターネット資源選択的蓄積実験事業」（WARP）によってウェブ情報の選択的な収集を行ってきたほか，技術的に収集が難しいデータベースについては「国立国会図書館データベース・ナビゲーション・サービス」（Dnavi）によってリンクしてきた 。

WARPでは，電子雑誌のほか，国の機関・都道府県，合併前の市町村，独立行政法人化前の政府関係機関，国立大学法人化前の国立大学，イベントのウェブサイトなどを収集対象とし，個別に許諾を得ながら，ロボットを用いて収集を行った。平成17年度末の収集コンテンツ累計は電子雑誌1490タイトル，ウェブサイト1898タイトルであり，収集した情報量は合わせて約3.1テラバイトに達する。平成18年からは全文検索機能を追加し，実験事業から本格事業へと移行した。

Dnaviは，インターネット上にあるデータベースへリンクして案内するサービスである。自治体が公開しているデータベース，独立行政法人及び特殊法人が公開しているデータベース，大学・学術機関が公開しているデータベース，貴重書・特別コレクションの一次情報・二次情報データベースを優先して収載した。平成17年度末時点での収載件数は9134件である。

しかし，既に述べたとおりネットワーク系は現時点では納本制度に含まれていないため，WARPは個別に発信者の許諾を得て収集を行っているが，この方法では収集可能な情報量が限定される。今後はネットワーク系を広く収集することを目指して法制度化を行いたいと考えている。

また，電子ジャーナルや電子学位論文など，ウェブ情報とは異質の情報を収集・保存するために，WARPとは別にデジタル・デポジットというシステムの開発に着手している。平成21年度にはこのデジタル・デポジット・システムも含め，デジタルアーカイブのシステムが整う予定である。

## デジタルアーカイブポータルの開発

電子情報を収集・保存するだけでは電子図書館事業として十分とはいえない。館のホームページにアクセスする利用者は，館に関心があるわけではなく，探している情報が館のページにありはしないかと期待してアクセスしていると考えられる。利用者のニーズを満たすためには，あちこちにある情報を統合して利用者に提供することが求められているといえる。

現在，館の各データベースサービスを利用するに際しては，個別の検索画面を用いる必要がある。このようにバラバラに案内していた館の情報を統合して提供するだけでなく，さらには館外のデジタルアーカイブの情報をも統合して提供することを目指して，デジタルアーカイブ・ポータル・サイトを構築することを企てている。このポータルでは，利用者による検索の結果が電子媒体であればただちに閲覧可能とし，紙媒体の場合には，例えば複写サービスへと誘導する。本格システム構築に先立って，プロトタイプシステムを構築しており，技術を検証する目的で試験的に提供中である 。ここでは，国立公文書館や岡山県立図書館と連携しており，近代デジタルライブラリ，NDL-OPACなどとともに，国立公文書館デジタルアーカイブやデジタル岡山大百科をも検索し，それらの一次・二次情報にジャンプする。

館がこのような仕組みを構築することは，館の利用者にとってだけでなく，他の図書館にとっても有用であると考えている。それぞれの図書館がデジタルアーカイブを構築するときに，共通のメタデータ発信機能を装備さえすれば，館のデジタルアーカイブポータルから統合的に検索することができるようになるためである。

また，利用者は図書館にある知識だけを求めているわけではないであろう。このため，図書館の壁を超えて，検索の範囲を公文書館，民間データベース，書店，古書店，博物館等へと広げたい。

## おわりに

館では，今後も電子図書館事業を計画的に進めていく。Google世代に対してもアピールする真の電子図書館の実現は一館のみではありえない。ぜひ日本中の図書館の御協力を希望したい。

# （2007年）デジタル情報も含めた統合検索サービスの提供を目指して

総合目録ネットワーク参加館フォーラム

2008年2月21日

国立国会図書館　中山正樹

私たちの日々の生活の中では、多くのデジタル情報が生まれ、ネットワーク上で流通している。しかし、それらは本などの印刷物と違って何の対策も講じなければ後世に残ることもなく、年月を経た後に利用できるとも限らない。このような状況において、当館は現実の状況への対応と、未来の利用者への対応に迫られている。

　当館は、日本における唯一の国立図書館であり、「国民の知的活動の成果を、印刷物からデジタル情報にいたるまで広く収集し、国民共有の情報資源を構築する」こと、そして「広く国民に対し図書館サービスを提供し、現在及び将来にわたり、情報資源へのアクセスを保障する」ことがその役割としてあると認識している。当館が納本制度の下で国内の刊行物を網羅的に収集していることから国民に期待されていることは、印刷物＝「物」に限らず「情報」を、そしてコレクションの完全性を目指し、時代を超えてこれを保存することではないだろうか。

昨今の「情報」は、誰でもインターネットで容易に提供できるようになってきており、その量は飛躍的に増大している。多くの電子書籍、電子雑誌が発行されている。「物」として刊行するのをやめた、いわゆる｢ボーンデジタル｣ものも多数ある。このように、情報は、冊子体、冊子体からデジタル化したコンテンツ、ボーンデジタルと様々な形態があり、さらに、デジタルコンテンツといっても、一つの情報源から利用形態に合わせて多くの形態が生成されている。

当館は、冊子体のみならず、デジタルコンテンツも収集・保存して提供しようとしているが、デジタルコンテンツの流通が飛躍的に増大している現在、当館のみでこれらを行っていくことは困難である。このことから、当館が収集し保存するデジタル情報だけでなく、様々な機関が提供しているデジタルコンテンツも併せて閲覧可能にすることも当館の役割だと考えている。

　近年、情報探索＝「検索エンジンを利用したインターネット検索」という行動パターンが定着してきている。しかし、検索エンジンでは、いわゆる深層ウェブと言われるデータベースの中にあるコンテンツを検索することができず、コンテンツがどこにあるのか見当を付けて、個別のデータベースを検索しなければならない。

　ここで求められるのは、情報の所在場所や、媒体、形態によらず、ワンストップで検索し、一次情報を入手できるサービスである。そこで、現在、当館では、1つの検索窓にキーワードを入力するだけで、各種データベース内の該当コンテンツの一覧を表示し、かつ、その一次情報へナビゲートするサービスの構築を進めている。

そのためには、情報を保有している機関がデジタルコンテンツを提供する際には、デジタルコンテンツの書誌であるメタデータの記述要素、記述規則を共通化すると共に、そのメタデータの授受のために標準仕様に準拠した通信規約の実装が求められる。

今後は、Web2.0及びその次の世代でのユーザニーズを踏まえて、単に統合検索するだけでなく、さらに利用者の情報探索行動、個別ニーズに合わせて、当館が保有しているか否かに関係なく、求める情報への迅速で的確なアクセスまたは案内を実現することを目指す。

# （2008年）DA/PORTAの今後（将来像）【骨子】

平成20年11月12日

関西館　中山

## 概要

### 現状認識

* + 1. 指数的に増大するデジタル情報
       1. 組織、マスメディアに加えて、個人が多くの情報を発信
       2. 情報の質がフラットに（専門家の情報も個人の情報も等価に）
    2. 利用者の情報探索行動の目的は、問題・課題の解決
    3. 検索エンジンで見えない情報はないも同然
    4. 有用な情報は、民間、個人のデータベース、ブログ等に多い
    5. インターネット上の情報は、すべてを１つの組織で収集・保存することは不可能
    6. 専門家の人海戦術では、情報の組織化は無理
    7. 個人の知は、専門家を凌ぐ
    8. 膨大な情報を検索できる時代に、従来の図書館的な指向で目録検索から入るような形では、利用されなくなる。

### 何をすべきか（Web2.0時代の次を見据えて）

* + 1. データベース内の情報の「見える化」（可視化）
    2. 情報を全体で集合知化
       1. 組織を超えて知識を集約
       2. 専門家の知識を集約
       3. 個人の知識を活用
    3. 集合知の永久保存
    4. 情報の組織化
       1. 情報に属性付与
       2. 情報と情報の関連付け
    5. 利用者情報の組織化
       1. 利用した人の属性把握
       2. 人と人の関係付け
    6. 情報と人の関係付け
    7. 的確な情報の検索（情報の選択範囲の拡大・絞込み）
    8. 的確な情報の閲覧（利用者属性・利用環境に応じた情報閲覧）

### どのように実施するか

* + 1. アーカイブシステム構築において、先進的な研究開発成果の適用
       1. 関係機関の技術開発成果の活用が必要
    2. コレクション構築において、図書館の枠を越えた連携協力
       1. すべてのデータプロバイダが連携できる仕組みとビジネスモデルが必要

以降、個別説明（デジタルアーカイブシステムの構築・国のデジタルコレクションの構築）

## 何をすべきか（Web2.0時代の次を見据えて）

### データベース内の情報の「見える化」（可視化）

* 1. 検索の入り口は、まずは検索エンジン。検索エンジンで見えない情報はないも同然。
  2. 多様な検索ニーズに答えるためには、多様な入り口（ポータル）が必要
  3. 各データベースが他のシステムから「見える化」すれば、様々なポータルサイトが実現できる。それは、データプロバイダが提供する検索サービスよりも使い勝手のいいものが実現される可能性も高い。

### 情報を全体で集合知化

* 1. 組織を超えて知識を集約
     1. 図書館、公文書館、博物館はそれぞれが書物を所蔵しているが、明確な分担保存の基準はない。
        1. GoogleやPORTAで「源氏物語」を検索して見ると、様々な機関、個人が情報を発信している。
        2. 原本を交換し合うのは難しいが、デジタルコンテンツなら、組織を超えて、必要なコンテンツを持ち合える。
  2. 専門家の知識を集約
     1. 図書館等の専門組織の強みは、司書、学芸員等が持つ知識。これらを集合して使えることが大切。
     2. ナレッジのデータベース化
        1. 専門家の知識は重要。
        2. 当館では公共図書館員による「レファレンス協同データベース」や「ナレッジデータベース」を構築しているが、専門家によるWikipediaのようなものは作れないか？
  3. 個人の知識を活用
     1. CGM(Consumer Generated Media)の活用
        1. ここ2～3年で、一般利用者が直接生成する情報は、膨大になってきた。それらは、断片であっても、集合すると有用な情報になっている。
        2. 多数の意見、類似嗜好の人の行動を情報として提供することにより、利用者の行動の参考となるようにする。
        3. Wikipedia、はてな、Yahoo知恵袋。。。。

### 集合知の永久保存

* 1. 今、そして将来も利用できる形で保存
     1. インターネット上の情報は消え行くもの。すぐに公開できなくても保存
     2. 現在の媒体、ファイルフォーマットは、将来読めなくなる。→国際標準、業界標準を策定した団体、その仕様に基づいて製造したベンダの協力のもと、マイグレーション、エミュレーションができるようにすることが必要
  2. インターネット上の情報は、すべてを１つの組織で保存することは不可能
     1. 複数のアーカイブ組織で分担保存
     2. 各アーカイブ組織が、自らの組織のタスクに応じて、必要なコンテンツを交換し合って保存
        1. ディザスタリカバリの面からも有効

### 情報の組織化

* 1. メタデータは一義的にはコンテンツの作成者が付与するのが望ましい
     1. blog、wiki、CMS（Contents Magage System）のような環境で構築されたコンテンツは、自然にメタデータが作られる。
     2. また、html文書に、マイクロフォーマットを使うことで、アプリケーションから使える意味ある情報をもたらすことが出来る。
  2. 情報に属性付与
     1. 情報のクラスタリング
        1. 著作の類似性（内容、レベル）等で著作物を把握
     2. FRBR、書誌ファミリーの概念での組織化
        1. 同一著作、原作から様々な媒体や形式に派生した著作物を書誌事項で体系的に把握
     3. CGM(Consumer Generated Media)の活用
        1. たとえば、フォークソノミーの１つとしての、ソーシャルタギング（はてなブックマーク等）
           1. 統制語彙のようにあらかじめ定められた言葉でものごとを分類するのではなく、各人が任意のタグ（自然語）を追加するもの 。ほかの利用者が目的のファイルを見つけられる確率が大幅に高まる。（＠ITより）
        2. ソーシャルタギングにおいて、専門家による統制語が使われていけば、より確実な分類になっていく。
     4. 自動組織化（自動メタデータ付与）
        1. 文書構造、文脈、画像・音声特徴の自動認識
        2. はてなブックマークでは、ブックマークすると、自動的にカテゴリ分類とキーワードが生成される。
        3. 主題分類も付与されれば。。。
  3. 情報と情報の関連付け
     1. セマンティックウェブ技術の活用

### 利用者情報の組織化

* 1. 利用した人の属性把握
     1. 趣味、嗜好があらかじめわかれば情報の絞込みが可能に
     2. 図書館の世界でほとんど利用されていない
     3. 実在する個人の行動を把握するためのものではない
  2. 利用情報の把握
     1. 利用者の利用傾向から、ユーザの特性・属性を推定
     2. その属性を元に、情報を推薦
     3. 注意することは、ユーザには、公的、私的で別の嗜好を持つ
        1. 例えば、情報処理の仕事をしているが、草花が好き。どちらも自分であるが、それを一人の興味・嗜好と認識すると推薦する情報は混乱する。
        2. ユーザの特別な行動は排除されなければならない。（人から頼まれて調べた。。。）
  3. 人と人の関係付け
     1. ユーザ同士の類似度を、同じアイテムにつけた評価の相関係数などで把握
     2. 類似性で、人と人のコミュニティは作られる。類似した人の情報は、自分にも有効。

### 情報と人の関係付け

* 1. 利用者が興味・嗜好を共有して、情報群を作っていく（あらかじめでなく）
  2. 利用者が主体のアーカイブ
     1. 利用者の観点で分類（クラスタリング）
        1. 図書館、公文書館、博物館、美術館、、、、ではない
        2. 一次、二次、三次情報ではない
        3. 一般人、児童、専門
        4. 学術、娯楽、、、、、
        5. 文学、科学、、、、
  3. 同じ嗜好を持つ人の利用傾向から自分にも役立つ情報の絞込みが可能に
  4. 他の人の利用実績から、こんな情報もあると提示されれば、ロングテールにもつながる

### 情報検索手段（利用方法）の変革

* 1. 情報の関連を活用した検索（情報の選択範囲の拡大・絞込み）
     1. 全文検索だけでは、的確に絞り込めない。また、検索語など、適切な検索条件を与えることは困難。
     2. 連想検索、あいまい検索
     3. 多角的な絞込み検索
        1. ファセット検索、クラスタリング検索
     4. 異業種、異文化圏の情報の検索
        1. オントロジー等の技術を活用した検索語の翻訳、検索結果の自動翻訳
     5. 組織を越えた統合検索と関連する情報への芋づる的なナビゲーション
        1. 組織を越えた大きな集合知の中から、多角的に絞込み、関連しそうな情報を例示
  2. 利用者属性、利用情報を活用した検索および閲覧（利用者属性・利用環境に応じた情報閲覧）
     1. パーソナライズ
        1. 利用者の背景、属性、行動履歴を参考に
        2. 利用場所、利用機器に応じて最適なコンテンツを自動選択
           1. 同一原本から、様々な媒体、形式に派生したコンテンツを、利用場所・利用機器により閲覧可能なコンテンツを選別

利用場所：図書館内、移動中、自宅等

利用機器：PC、携帯電話、PDA等

* + - * 1. 今や、インターネットはPCだけで見るものではない。携帯電話や携帯端末での利用は多い。携帯電話で近代デジタルライブラリーのようなイメージ画像は読みづらいが、青空文庫なら読める。また、朗読ならもっといい。
    1. 協調フィルタリング
       1. 他人の利用情報により推奨する検索
          1. 協調フィルタリング（Collaborative Filtering, CF）は、多くのユーザの嗜好情報を蓄積し、あるユーザと嗜好の類似した他のユーザの情報を用いて自動的に推論を行う方法論である。協調フィルタリングはリコメンデーション（推薦）やパーソナライゼーションに利用されている。実際にAmazon.comの「おすすめの商品」や、はてなアンテナの「おとなりアンテナ」などに応用されており、Web 2.0においてロングテールを支える要素技術として捉えることができる。（Wikipediaより）

### 利用のまとめ

* 1. 利用の仕方も、技術の進展とともに変化している。
     1. 書誌項目の完全一致、部分一致とかの検索ではニーズを満足しない
  2. ニーズとしての利用の仕方の調査研究も必要
     1. 技術の進展とともに、ニーズも変化している。新たなニーズに対応して、技術シーズも生まれてくる

## どのように実施するか

### NDL-DA、PORTAシステム開発における技術導入の考え方

* 1. サービス構築
     1. 基本的には作らない。OSSの組合せ、サービスのマッシュアップ。
     2. 技術開発、利用者ニーズに速やかに対応するためには、研究開発成果のタイムリーな実装が必要。
     3. 研究開発成果や、先進機能を使ってタイムリーにバージョンアップされていくOSSを組み合わせていくことが有効。（図書館用の商用パッケージでは、タイムリーに対応できない。）
     4. システムはWeb上にあるサービスを組み合わせて、マッシュアップで構築することが効率的。
        1. 各機関のサービスを活用しあうサービス連携
     5. Webサービス連携から、コンピュータ資源やサービスを共有しあうクラウドコンピュータの世界へ
  2. デジタルアーカイブ機能
     1. 対等な関係での分散デジタルアーカイブを指向
        1. 当館は国のデジタルアーカイブを構成するデータプロバイダの１つ。日本中のデジタルコンテンツを１つの機関ですべて収集・保存することは不可能。
        2. 分散デジタルアーカイブが連携して、1つの巨大なデジタルアーカイブとして利用できるようにすることを目指す
        3. メタデータ交換の統合検索連携は第一歩。次に、コンテンツの相互交換による連携。
           1. 必要なコンテンツを交換し合う方法として、セキュアなPeerToPeer（P2P）ネットワークでの連携が有望
  3. ポータル機能
     1. 一般の利用者にとって、データベースの所在を探すのも含めて、実質的な入り口は、GoogleやYahooのような検索サイト。
        1. Google、Yahoo等の表層ウェブを中心にした検索エンジンと、検索エンジンからはアクセスが困難な深層ウェブとのワンストップ検索サービスの融合が必要。
     2. 目的とするコンテンツを直接検索・閲覧できるサービスが求められている。たらいまわし的なナビゲーションは好まれない。
        1. 利用者層、直接統合検索できるコンテンツのカバレージなど、目的を明確にして、ユーザニーズに対応したポータル機能が必要。
     3. PORTAは？
        1. 当館が所蔵する資料をデジタル化したコンテンツに加えて、インターネット上から収集したコンテンツと、各機関が構築し提供しているコンテンツを検索できるようにしたデータプロバイダ。
           1. PORTAの統合検索は、統合検索先のメタデータを検索用データとしてハーベストして保持するリポジトリ機能と、検索要求の都度横断検索することにより実現。
        2. データプロバイダとして、検索機能をWebサービスのAPIとして、他のサービスプロバイダーに提供する。
           1. PORTAが持つGUI機能は、このAPIを利用して提供しているもので、デジタルコンテンツをアクセスするための入り口の一つ。デジタルアーカイブがWeb上で「見える化」されるのを加速させ、各機関のデジタルコンテンツの利活用が促進されることを目指すもの。

### デジタルアーカイブシステムの構築に必要な技術の適用

* 1. 技術
     1. 収集を容易にする技術
        1. *収集効率を高める技術（差分収集と再現技術を含む）*
        2. *収集品質を高める技術*
        3. *収集したサイトイメージから著作物を切り出す技術*
     2. 組織化（メタデータの付与等）を省力化する機能
        1. *可能な限り自動化*
        2. *セマンティックウェブ技術の活用*
        3. *クラスタリング技術*
        4. *FRBRの概念を適用した組織化も*
     3. 検索を容易にする技術
        1. *キーワードマッチングだけでなく、推定する技術も活用した検索技術*
        2. *検索、クラスタリング検索等*
     4. 長期保存を実現するストレージ技術　（次世代ストレージ）
  2. 進め方
     1. 国際標準、業界標準の積極的な適用
     2. 政府機関、民間を問わず、国内外の研究機関、研究者等と連携して調査研究
     3. 各機関の研究成果を組み合わせて機能を実現
     4. まずは、海外ではIIPC、国内では研究開発を行っているJST、NII、NICT、AIST、IPA等の政府機関、大学の研究室等との連携
     5. 有用な検索サービスを実施している商用ポータル機関（Google、Yahoo等）との連携
  3. 産官学の研究開発プロジェクトの成果の活用
     1. DA及びPORTAにおいて必要な技術であるが、技術的な課題がありパッケージやOSSが存在しない技術に関して、政府機関、大学等の研究機関に対して技術的なニーズを提示して、研究開発成果の技術移転を目指す。
     2. 縮減された予算の中でも機能強化を可能とする方策の一つ。
     3. 文明開化プロジェクト（総務省：21～23年度）
        1. 概要
           1. 「ICT 成長力強化プラン」に基づき、国立国会図書館、国立公文書館、他府省庁、地方公共団体、図書館・博物館・美術館、大学等との連携を図り、産学官を挙げてデジタル化を推進、日本中の知的財産を総デジタル化してつなぐ「デジタル文明開化プロジェクト」を実施する（重点計画2008（案）（6月11日）IT戦略本部）
           2. デジタルアーカイブ構築のために必要な技術開発

複数のデジタルアーカイブの統合に関する技術

分散型・大規模デジタルアーカイブの長期保管・利用に関する技術

デジタルアーカイブの多目的な利活用に関する技術

* + - * 1. 地域公共図書館等がデジタルアーカイブシステムを導入するための資金援助
      1. 状況
         1. 総務省での予算要求のために、技術開発が必要な分野、テーマ候補の提示、説明資料等の作成に協力
      2. 予算が確保された場合、来年度の実施に向けて、今年度末までに調達の準備をするとのこと
    1. 大航海プロジェクト（経済産業省）
       1. 経済産業省情報政策課に対して、今年度の「共同活用基盤を用いた実証実験」及び、来年度以降の「大規模図書館が将来的に活用可能なモデルサービス等の実証実験」として検討していただきたい研究開発テーマの一覧を提示。具体的なアクションにまでなっていない。
    2. 情報爆発IT基盤（文部科学省）
       1. 「情報爆発IT基盤」に対して、直接具体的な申入れはしていないが、このプロジェクトの個別研究開発テーマの研究者（木俵先生、黒橋先生）とはコンタクトを取っている。どの程度、当館で成果を活用できるかはわからない。
    3. 民間の技術開発、製品開発　（）
       1. 図書館パッケージベンダー　（図書館が利用する冊子体OPACからデジタルコンテンツの目録提供システムへ）
    4. 共通仕様・技術提供
       1. 日本規格協会　（メタデータ交換、サービス連携のための、通信プロトコル、メタデータ記述要素・記述規則）
       2. 電子情報技術産業協会（JEITA）　（製品開発者として、各種デジタルフォーマットの適用、長期的な利用を保証するためのエミュレーション、マイグレーション技術の提供、現在、オーディオアーカイブに関して仕様検討会が開催され、NDLはオブザーバ参加）
  1. IIPC
     1. 国際インターネット保存コンソーシアムへの加盟 (International Internet Preservation Consortium : IIPC)
        1. IIPCは、Webアーカイビングに資する相互運用可能なツールや技術の開発・標準化を推進し、国際的な利用を促進することを目的として結成された機関
        2. 現在、各国の国立図書館や公文書館など38の機関が参加
        3. IIPCを通じて国際的に貢献することはもちろん、相互運用可能なツールや技術を国内に還元することで、日本のWebアーカイブ進展に資する
     2. ワーキング
        1. スタンダード
           1. 短期的にはWARCの標準化、将来的にはメタデータ、評価基準の標準化等の標準化
        2. ハーベスティング
           1. Heritrixの開発、WARCフォーマットの開発とサポート、クローラに特化した要望、深層ウェブの収集、ビデオやストリーミングの収集
        3. アクセス
           1. アクセスツール（Wayback machine,WERA）改善、NutchWAXを使った全文検索のテスト、ユーザ認証、ユーザ権限、アクセスコントロール、アーカイブのコンテンツ解析のためのツール、アーカイブの構造解析のためのツール
        4. プリザベーション
           1. デジタル情報の保存をめぐる技術的課題への取り組み、ストレージの拡張性に関する研究、WARCフォーマットの保存機能の向上、ウェブをめぐる技術環境の変化の記録、再現できなくなりつつあるファイルフォーマットの特定

## 国のデジタルコレクションの構築に向けた連携協力

### 主な連携協力先

* 1. これは、連携協力先のリストです。予定も含んでいますが、コレクションの構築、システム構築、国際協力の観点で、様々な機関と連携協力のための調整を行っています。
  2. ナショナルデジタルコレクションの構築
     1. 内閣官房
        1. 知的財産戦略本部　（知財計画2008で示された、NDLのデジタルアーカイブ化、図書館資料の利用の促進、著作物のフェアユース規定の日本版、、、）
        2. IT戦略本部　（e-Japan重点計画I・重点計画2008？）
     2. 美術館、博物館、文書館、図書館連携（MLA連携）　（関係機関のサイトへのアクセス、来館利用が促進される形での連携・協力、デジタル文化資源ラウンドテーブルを設置予定）
        1. 国立公文書館 、東京国立博物館 、各美術館、博物館、公文書館、図書館
     3. 学術情報　（学術情報、科学技術情報の統合検索、コンテンツの長期保存、学位論文の保存と提供、NDLと大学図書館との連絡会議の下に学位論文電子化の諸問題に関するWGを設置して検討してきた）
        1. NII、JST 、大学図書館 、電子ジャーナル出版者
     4. 公共図書館（各図書館での郷土資料のデジタル化とアーカイブ構築を支援、公共図書館の総合目録ネットワークの便利さをデジタルの世界でも実現）
        1. 都道府県立図書館 政令指定都市立図書館 、市町村立図書館 、専門図書館
     5. 出版者等　（電子書籍作成支援、電子出版用に出版者へ提供、各出版者へのナビゲーション、出版者等のビジネスが拡大できる形での連携・協力）
        1. 日本電子出版協会（JEPA） 、日本印刷技術協会（JAGAT） 、冊子体書籍出版 、電子書籍出版者 、インターネット書籍販売者
     6. 民間デジタルアーカイブ　（色々あるがその一つとして）
     7. 歴史的音盤アーカイブ協議会（HiRAC）　（SPレコード3万タイトル6万曲、その他記録音盤、音源フォーマット、メタデータ仕様、デジタル化分担、統合検索、歴史的音盤を将来に亘って聞くことができるように）
     8. 商用ポータルサイト　（実質的なインターネットの入り口、色々考えられる）
        1. Google、Yahoo等　（デジタル化、深層ウェブの可視化、サービス連携、商用ポータルが保有する情報の利用、検索システム構築の技術トランスファー） 、ソーシャルブックマークサイト
     9. 政府情報
        1. 総務省行政管理局（e-GOV）　（政府サイト情報の統合検索、収集保存）
        2. 各府省支部図書館　（政府情報のデジタル化、長期保存、デジタルコンテンツの統合検索）
        3. 国立印刷局　（政府刊行物のデジタルフォーマットとしてXMLベースの原本作成）
  3. 国際協力
     1. World Digital Library　（歴史的文化遺産を世界中の国立図書館が協同でデジタル化して提供するもの、NDLは日本として、東アジアの一員として参画し、世界規模でのデジタルライブラリが構築されることを目指す）
        1. LC　LCが発行するガイドラインや施策は、今後の図書館を考える上で、とても参考になる。毎年、情報交換を行っている。
     2. IIPC　（ウェブアーカイブに関して、標準、収集、提供、長期保存の観点から、研究開発、NDLも4月より参加、構築中のデジタルアーカイブシステムでの適用と成果のフィードバックを目指す）
     3. 日中韓連携　（漢字文化圏の三国が協力してデジタルアーカイブを構築して、コンテンツの長期保存と提供を目指す。昨年度協力の合意をして、現在各国の現状をまとめている。まずはパイロットシステムを構築して、メタデータ交換、横断検索を試行する予定）

### 政府の施策（2008年度）

* 1. 今年度は、NDLのデジタルアーカイブ事業に直接関連する政府の施策として、このようなものがあります。
  2. まず一つ目として、「e-Japan重点計画特命委員会自民党デジタルアーカイブ小委員会」が、平成20年3月12日に公表した「デジタルアーカイブの推進に向けた申入れ」。
     1. 主な内容は、
        1. 平成16年6月の「国立デジタルアーカイブ構想」の提言に沿って、わが国のデジタルアーカイブの総合ポータルとなるPORTAが開設された。
        2. 国立国会図書館のウェブアーカイブの本格実施のための法制度の実現
        3. 全国図書館のデジタルアーカイブの統合化
           1. 全国の図書館が統合的なデジタルアーカイブの構築。
           2. 図書館間でのデジタル化された収集資料の相互利用、館内公衆端末での閲覧が可能となる著作権制度の創設。
           3. 著作権データベースの活用等著作権者が不明な場合の文化庁裁定制度の円滑な利用のための制度改善。
        4. 国立公文書館と国立国会図書館が協力
           1. 相互の役割の違いを尊重して、利用者の利便性の向上を図る観点から、連携・協力
           2. 全国の公文書館、図書館でのデジタルアーカイブ構築について、メタデータ付与等に係る技術面、ノウハウ面における支援、協力を行う。
           3. 国のデジタルアーカイブの総合ポータルの検索機能の充実等により、各分野でのデジタルアーカイブの利用の拡大を図る。
        5. 世界最先端のデジタルアーカイブ技術への対応
  3. 次に、内閣官房　知的財産戦略本部が6月18日に発表した「知財計画2008」
     1. 「国立国会図書館のデジタルアーカイブ化と図書館資料の利用を進める」ということで、
        1. 国立国会図書館において行われている貴重な図書等のデジタル化やインターネット情報資源等を収集保存し、ネット上で一般ユーザーの利用に供する取組''''について、その促進が図られるよう一層の連携を進める。
        2. このため、'権利者の経済的利益や出版ビジネスとの関係を考慮''しつつ、国立国会図書館における蔵書のデジタル化の推進に必要な法的措置を２００８年度中に講ずるとともに、国立国会図書館と他の図書館等との連携や図書館等利用者への資料提供の在り方については、関係者間の協議を促進し、２００８年度中に一定の結論を得る。（113/157ページ）
  4. さらに、IT戦略本部が、6月1１日に意見招請している''重点計画2008（案）
     1. 総務省が発表した、「ICT 成長力強化プラン」の中で記述した、「デジタル文明開化プロジェクト」の推進（総務省及び関係省庁）
        1. 国立国会図書館、国立公文書館、他府省庁、地方公共団体、図書館・博物館・美術館、大学等との連携を図り、産学官を挙げてデジタル化を推進、日本中の知的財産を総デジタル化してつなぐ「デジタル文明開化プロジェクト」を実施する。
     2. 具体的な内容の一つとして、地域公共図書館支援を想定しています。
  5. 「ICT 成長力強化プラン」の実施が盛り込まれた形で、「経済財政改革の基本方針」（いわゆる骨太方針2008）として、6月27日に決定されました。
  6. これに基づいて、スクラップ＆ビルドでの新規要求、既存経費の削減を含めて、来年度予算要求が行われています。

### 公共図書館支援のイメージ

* 1. これは、連携協力の例の１つとして、地域公共図書館支援のイメージです。
  2. 公共図書館が、地域情報ハブとして、郷土資料をデジタル化したり、県域の情報を収集して提供する役割を支援するものです。
     1. 公共図書館が、県内有用サイトの収集もしくは横断検索することを支援
     2. 公共図書館が、郷土資料をデジタル化することを支援
     3. 公共図書館が、収集コンテンツ及びデジタル化コンテンツのデジタルアーカイブを構築し、提供することを支援（必要に応じて当館が利用しているデジタルデポジットシステム等を貸与する）
     4. デジタルアーカイブを構築し運用することが困難な図書館に対しては、当館がASP的なサービスを行う。
     5. NDLは、各公共図書館のデジタルアーカイブを長期保存もしくはバックアップのために収集保存する。
     6. NDLは、各公共図書館のデジタルアーカイブのメタデータをハーベストもしくは横断検索して、利用者に対して、全国公共図書館が持つ郷土資料を統合的に検索し、各図書館サイトへナビゲートするポータル機能を提供する。
     7. また、NDLは、標準として適用すべき仕様のガイドラインや手引き類を提示し、実装を支援します。
        1. デジタル化及びシステム構築用の手引きの提供
           1. 「デジタル化の手引き」（公開中）
           2. 「デジタルアーカイブ構築の手引き」（計画中）
           3. 「共通仕様及び連携に関するガイドライン」（公開中）
           4. 「外部提供インタフェース仕様書」（公開中）
           5. 「メタデータスキーマガイドライン」、「メタデータ記述規則」（策定中）
        2. 各機関での地域情報収集用及びアーカイブ構築用の汎用ソフトウェアの提供も想定しています。
  3. 支援のスキーム
     1. PORTAでの連携協力先の拡大がなかなか進まない要因として、左の枠のように、各機関の対応が困難だったことがあげられます。
        1. 資料を、デジタル化する費用がない
        2. 自前で、サーバを立ち上げられない。運用できない
        3. 意義は理解したが、標準プロトコルを実装できない
        4. *ウェブページに貼り付けて、データベースとして検索できる形になっていない*
        5. *データベース化されていても、外部提供インタフェースを持っていない*
        6. *外部提供インタフェースを実装する費用がない*
        7. メタデータのマッピング調整に膨大な時間（工数・費用）がかかる
        8. *DCをベースにした記述要素を使っていても、使い方がそれぞれまちまち*
        9. *同じ要素を使っていても、記述規則が異なり、同じ内容として認識できない*
        10. さらに、統合検索できることのメリットが理解されていないということもあります。
        11. *アクセスが増えることが不安*
        12. *サービスが横取りされるのでは？*
     2. これらを解決することが、関係機関の拡大につながると考えています。
        1. 支援のスキームとして、
           1. 当館は、デジタルアーカイブシステムをオープンソースとして開発
           2. そのシステムをベースに、研究開発で、地域公共図書館用のソフトウェアを作成
           3. そのソフトウェアを各図書館が実装するために、国から何らかの資金援助
           4. 同時に、当館は、各種ガイドラインを提示
     3. このようなスキームで、各図書館の郷土資料が、デジタル化して公開され、それが、インターネットで発見されるようになって、資料の利活用が進み、さらに、サイトや来館での利用者が増えることを期待しています。

### 学術機関との連携協力

* 1. 学位論文のネットワーク形成国立国会図書館、国立情報学研究所、国公私立大学図書館の3者により提示された我が国の学位論文のデジタル化とネットワーク化の道筋（国会図書館と大学図書館の連絡会「学位論文電子化の諸問題に関するＷＧ中間報告」平成20年3月27日）に従う。
     1. 3者による協調の下、デジタル化を推進するため、学位論文をリポジトリに搭載し、JuNii+、国会のポータルサービスなどと相互にメタデータを交換するためのメタデータプロファイルを提案する。各機関はこのプロファイルにしたがい、リポジトリに搭載する学位論文の標準メタデータを作成するものとする
        1. 学位論文等の検索システムとメタデータの関係機関リポジトリでの学位論文メタデータプロファイルと、NDLが所蔵する紙媒体の学位論文のメタデータ、遡及デジタル化した学位論文のメタデータの関係を整理
        2. 電子ジャーナルとの相互連携も視野に
        3. DRF４において合意形成
           1. 大学において共通のメタデータプロファイルを策定中
           2. NDLにおいて、紙媒体、及び、紙から遡及デジタル化した資料のメタデータ案（DC-NDLベース）を策定中
           3. 両者の整合、クロスウォークを可能にすることにより、一元的なアクセスが可能になる。
  2. 国立情報学研究所（NII）のCiNii、JuNii（平成21年4月頃公開予定）
     1. CiNii、JuNiiのメタデータのハーベストは、両システムが改修中のため、新システム稼動後に実施することで内諾を得ている。
     2. CiNiiに関しては、OpenURLによる横断検索の準備は完了したが、CiNiiへの負荷が懸念されるため、メタデータをハーベストすることとした。
     3. JuNii+に関しては、OAI-PMHでのハーベストの準備は完了。
  3. 科学技術振興機構（JST）のJ-Stage、Jurnal@rchive（12月頃公開予定）
     1. J-Stage、Jurnal@rchiveに関しては、メタデータをファイル転送で受領することとして、準備は完了。現在、JSTから学会に対して、PORTAへのメタデータ提供及びPORTAから商用サイトを含めた機関への提供の許諾確認を行っている。

### 商用サイトとの連携協力

* 1. 電子書籍ポータルhon.jpの統合検索（11月公開予定）
     1. PORTAから、hon.jp Webサービス を利用して、hon.jpの電子書籍メタデータDBを統合検索する。冊子体資料として存在しない、「ケータイコミック」、「ケータイ小説」も検索可能となる。PORTAは、関連情報リンクで、既に商用サービスとの連携を実現しているが、統合検索先としての連携はこれが初めてのケースとなる。

### Google,Yahoo等との連携の可能性

* + 1. 商用ポータルは、実質的なインターネット情報の入り口であり、利用者の利便性向上、サービスの役割分担の上からも、重要連携先の１つ。特定商用ポータルとの排他的な連携ではなく、等距離での連携の実現を目指して、意見交換を行っている。
    2. Google
       1. GoogleBookSearchの検索結果から、PORTA（近代デジタルライブラリー、貴重書データベース等）、及び、総合目録ネットワークへのナビゲーション、ＮＤＬ蔵書のデジタル化、検索用テキスト化の可能性に関して、Google社ストラテジックパートナーディベロップメントマネージャと意見交換。（7月）。現在、日本Google社より、米国Google社に実施可能性について打診中。
       2. 今後、Googleが保有するデジタル化技術、検索技術等の技術移転の可能性についても意見交換することとしている。
    3. Yahoo
       1. Yahooカテゴリ分類ルール、登録されたURLを、有用なサイト・ページとして、PORTAからのナビゲーション先として、また、ウェブアーカイブでの有用なコンテンツの収集起点としての活用させていただくこと、Yahooでの検索キーワードを辞書として利用させていただく等、Yahoo検索事業部と意見交換（8月）

### 出版社、著作者等との連携協力

* 1. 資料の長期保存と利用の促進のために
     1. 当館はデジタルコンテンツを長期保存
        1. デジタルコンテンツの収集
           1. マルチユースコンテンツの原本、印刷用原本等、DRMのないもの
           2. 可能な限り詳細なメタデータの提供
     2. 資料のデジタル化
        1. 出版物を当館がデジタル化することの許諾
     3. 当館からのコンテンツ提供
        1. インターネット提供が許諾できない場合は、公共図書館内もしくは館内での閲覧提供まで
  2. 出版者のビジネス支援
     1. 各出版者の電子書籍作成支援
        1. 当館がデジタル化したコンテンツを電子出版用に提供
     2. 各出版者の電子書籍データベースの立上げ支援
        1. 当館デジタルアーカイブを利用して、出版者が販売することも検討
     3. 各出版者サイトへのナビゲーション
        1. 出版者の電子書籍データベースを、PORTAで統合検索
           1. 電子出版物の購読が促進されるように
        2. PORTAでの統合検索結果に、入手先として出版者へのリンクを表示
  3. 日本版「Book Rights Registry」実現の可能性
     + 1. 2008年10月、米国の著者団体と出版社団体は、Google社との和解条件の1つとして、著作権料徴収NPO「Book Rights Registry」の設立を発表した。日本国内でも似たようなシステムの整備を求める声は今後高まってくると思われるが、残念ながら現時点では、類似団体として音楽世界のJASRAC等があるが、雑誌や書籍出版業界にそのようなものはない。日本版「Book RightsRegistry」の実現に向けての制度面／法律面／技術面で問題点等を明確にしていく。

### MLA連携

* 1. 背景
     1. 文書館、博物館、図書館においては、これまで大量に蓄積されてきた過去の文化資源のデジタル化は、未だ進んでいない。
     2. また、デジタル化形式の標準化や、各組織・機関の作成するデータベース間の連携についても、その利活用において、多くの課題を抱えている。
  2. 目的
     1. 利用者は、所在場所を意識せず必要な資料を探せるように。
     2. 検索結果から各機関のDBへ案内することにより、各機関のDBの利活用が進むように。さらに来館して実物を見る利用者が増えるように。
  3. 状況
     1. 平成20年3月準備会開催
        1. 国立公文書館、東京国立博物館、国立情報学研究所、科学技術振興機構、慶応大学、奈良女子大、国立国会図書館
        2. 横断的アーカイブズ論研究会
     2. 平成20年度下期に、ラウンドテーブルを発足させる
        1. 日本国内の博物館、図書館、文書館が、館種を超えたデジタル化による仮想的なコレクションの構築を目指す。
        2. デジタル化の進展に係る共通の課題について、その解決に向けた協議・検討を行う。
     3. 連携の現状
        1. MLA連携のラウンドテーブルの開催前に一定の連携の実績を積んでおきたい。
        2. 国立美術館の「所蔵作品総合目録検索システム」（12月頃公開予定）
           1. 5国立美術館長会議（9月30日）にて、PORTAでの統合検索の承認が得られた。
           2. 国立美術館（東京国立美術館、京都国立近代美術館、国立西洋美術館、国立国際美術館）の総合目録のメタデータをハーベストする。現在、国立美術館において、メタデータ提供機能の実装作業中。（12月初旬、公開予定）
        3. 人間文化研究機構の「資源共有化システム」（未定）
           1. 人間文化研究機構内部の５研究機関（国立歴史民俗博物館、国文学研究資料館、国際日本文化研究センター、総合地球環境学研究所、国立民族学博物館）のデータベースを統合したシステムをPORTAで統合検索するもの。京都大学地域研究統合情報センター原正一郎教授を通じて、意見交換を行っている。
           2. 人間文化研究機構研究資源共有化事業委員会（8月）において、PORTAとの連携について、話し合われ、好意的な反応であり、人間文化研究機構側で調整を行っている。委員会において再度、議論され、原氏、東京大学史料編纂所の石上理事とで打合せを行うことになった。（10月28日、東京にて）
           3. OAI-PMHもしくはSRWで、DCベースのメタデータを受領可能と聞いているので、決定後、速やかに統合検索が実現できるものと思われる。
        4. 慶應義塾大学の「慶應義塾図書館稀覯書画像」、「慶應義塾写真データベース」、「慶應義塾図書館デジタルギャラリー」（実施時期未定）
           1. 各データベースは別々に構築されているものであるが、メタデータをハーベストする方法で統合検索が可能とのこと。現在、慶應義塾大学メディアセンター本部を窓口として、学内調整を行っていただいている。
        5. 早稲田大学の「古典籍総合データベース」（実施時期未定）
           1. メタデータをファイル転送していただくことで統合検索可能とのこと。現在、図書館情報管理課とメタデータのマッピングの調整を行っている。
        6. 東京大学の「東京大学総合研究博物館データベース」、「デジタル展示館」等（実施時期未定）
           1. 東京大学大学院情報学環　学環長は、前向きに検討したいとのことで、連絡あり。（10月7日）。
           2. 正式な依頼状を送ると共に、参加の仕方、プロセスを解説するような書類で説明する予定。それ以後、具体的な検討を開始する。（10月8日、説明資料案送付）

### 歴史的音盤アーカイブの構築支援

* 1. 歴史的音盤アーカイブ推進協議会（HiRAC)
     1. 趣意
        1. 文化的遺産であるSPレコード原盤が散逸・劣化し始めている
        2. これをデジタル化して保存し、広く国民に公開することによって、日本の文化と音楽産業界の発展に寄与する
     2. 対象
        1. SPレコード：3万タイトル（6万曲）、その他、商業用レコード以外の記録物
     3. HiRACメンバー
        1. 日本放送協会、日本音楽著作権協会（JASRAC）、日本芸能実演家団体協議会、日本伝統文化振興財団、映像産業振興機構、日本レコード協会
  2. 考えられる協力
     1. 方針
        1. HiRACへのオブザーバ参加
        2. HiRACによる音盤デジタル化の支援
        3. デジタル化音源の当館への納本
        4. NDLデジタルアーカイブからの公開　（館内、公共図書館、インターネット（？））
        5. オーディオアーカイブの標準規格の策定作業（JEITA）への参画
     2. アーカイブ及び公開用システムの構築
        1. デジタル化、統合検索等に関する研究開発の実装
        2. 著作権管理機能の研究成果の実装

### 日中韓の国立図書館との連携

* 1. メタデータ基準
     1. 保存のためのメタデータの形式は、各国の基準を尊重する。
     2. ・メタデータ交換は、DCベースとする。
     3. ・語彙等の違いは、オントロジー等による吸収を検討する
  2. 統合的な情報サービス
     1. コンテンツは3館で分散して持ち、統合検索できるようにする
     2. 機械翻訳は研究開発の状況をウオッチ
     3. 統合検索のプロトコルは、ハーベスト系（OAI-PMH、RSS）、横断検索系（SRU/SOAP、OpenSearch、OpenURL、Z39.50）のどの方法も3国とも可能
     4. ３国共通のポータルを立ち上げるのではなく、各国のポータルで見せていく
     5. まずは、歴史的資料から統合検索
  3. 長期保存
     1. DRセンターは、まず各国内で立ち上げ、その後、３国でのDRセンターの立上げを検討する。  
        その際は、CDNLAOでのシンガポールの提案のように国際的なＤＲの方向性で検討する。
     2. ＯＡＩＳはフレームワークであり、実装は各国の開発状況により異なる
     3. 情報パッケージの仕様は、3国とも手探り状態。まだ共通化する段階にない。
     4. まだ、コンテンツの交換の候補を提示する段階にない
  4. 情報交換
     1. 収集・組織化・保存及び提供機能に関する研究開発成果に関する情報
     2. デジタルアーカイブに関しての制度化に関する各国状況

### World Digital Library構想との連携

* 1. 米国議会図書館（LC）ビリントン館長が打ち出した構想
     1. 2005年6月、ユネスコ米国国内委員会のユネスコ米国国内委員会で
  2. 貴重書など一つしかない歴史的文化遺産を，世界各地の国立図書館等と協同でデジタル化し，無料公開しようとするもの
  3. 当面のアクション
     1. WDLサイトに、テーマを特定したコンテンツを提供
        1. 画像データ、メタデータ、解題

## PORTA

### PORTAのAPI提供準備状況

* 1. 「国立国会図書館デジタルアーカイブポータル（PORTA）外部提供インタフェースの提供について」（平成20年2月15日情報化推進委員会）に基づいて、平成20年3月末から、横断検索系の外部提供インタフェースは公開している。「国立国会図書館作成データベース提供方針」（国図総08071701号）により、ハーベスト系の外部提供インタフェースも含めて、商用サービス以外では、許諾手続きなしで、提供できるようになった。また、商用サービスへの提供は、許諾手続きによって行うこととなった。
  2. 国立情報学研究所Webcat-Plusでの近代デジタルライブラリーへのリンク（平成20年9月29日提供）
     1. PORTAのOAI-PMHにより、近代デジタルライブラリーのメタデータを提供することにより実現。
  3. 商用サイトであるhon.jpでのmobile.hon.jpによる書籍検索（一般利用者向け）及び、hon.jpターミナルによる電子書籍出版管理サービス（電子出版社向け）（平成20年10月末提供予定）
     1. PORTAのOpenSearchによりISBNを提供することにより実現。
  4. 国立情報学研究所連想情報学研究開発センターNPO法人連想出版の「想-イマジン」での連想検索（11月提供予定）
     1. PORTAのOAI-PMHにより、近代デジタルライブラリー、貴重書データベース、貴重書サンプルのメタデータを提供することにより実現。
  5. 検索結果のXML（DC-NDL）形式での出力、及び、OAI-PMHでの出力（11月末提供予定）
     1. XML出力に関しては、NDL蔵書（和図書、和雑誌）のメタデータのマッピングを修正し再投入中（平成20年11月末完了予定）。
     2. OAI-PMHに関しては、NDL蔵書（和図書、和雑誌）以外は、先行して提供する。（10月末提供予定）
  6. その他
     1. 広島市立図書館、神崎先生のサイト、Agropedia（農学情報資源システム）、、Ex Libris社の「MetaLib」（手続き未了）

### PORTAの20FY開発の状況

* + 1. 全体
       1. PORTAシステムの開発者のスキル不足等により、多くの不具合が発見されており、開発体制を大幅に入れ替えて、不具合の解消と、サービスとして必要不可欠な機能の実装を進めている。
       2. 現行サービスをDAインフラに移行
       3. ベンダー依存のミドルウェア（全文検索ソフト、運用管理ソフト）を含めて、全てをオープンスペックのツールにする←ベンダーロックイン状態であるため、今年度開発及び運用保守の調達では、適正な競争入札ができなかった。また、開発方式がベンダー製品の性能に依存しすぎているため
          1. 機能及び性能の不具合の改善
          2. SystemWalker→hinemos（OSS）
    2. ハーベスト機能及びメタデータ管理機能
       1. ハーベストもしくは横断検索して得られたメタデータを利用して、デジタルデポジットできるようにするための仕組みの実装（有用なコンテンツのブックマーク、RSSフィードを利用して、個別にメタデータを登録する機能等）
       2. DP登録機能の改善（標準仕様を実装したDPは容易に登録できるようにする）
       3. （メタデータの精緻なマッピングは求めない）
       4. PostgreSQL（現行のOSSのRDBMS）
       5. BizSearch→Solr(OSSの全文検索エンジン)
    3. 横断検索機能
       1. 汎用モジュール化
       2. 関連リンク情報に関して、汎用リンクリゾルバを適用する（商用サービス）
    4. API提供機能
       1. DP毎、プロトコル毎の提供可否設定
    5. GUI機能
       1. SunJavaPortalServer（ベンダー依存のOSSであるが、オープンコミュニティのOSSに変更することはできなかった。）
       2. 検索機能の改善（ファセット検索等）
       3. ブックマーク機能の改善（はてなブックマーク等の利用）

### PORTAの21～22年度の方向性

* 1. 全般
     1. 図書館、公文書館、博物館、美術館、大学、政府機関、民間等が分散して保有しているデジタルアーカイブを統合的に検索・閲覧できるようにして、国としてのデジタルアーカイブを構築する。
     2. 利用者の種別ごとに検索・閲覧方法は異なる。PORTAは統合検索の方法の１つとしてのGUIを提供するが、基本的には、各DBを統合的に検索できるようにして、その機能を利用した様々な利用者向けの「入り口」が用意されるようにAPIの提供を積極的に行う。
  2. ポータルの１つとしての「総合ポータル」
     1. 一般利用者の情報探索行動の段階に応じて、必要とする検索・閲覧が、ワンストップ（たらい回して的なナビゲーションではなく、一回で検索できる形を想定）で行える環境の構築と提供が促進されるようにする。
        1. Google等の検索サービスで有効な情報が得られなかったときに、デジタルアーカイブ内のコンテンツを統合検索することを想定
  3. データプロバイダ拡大策
     1. 様々な機関が分散して構築しているデジタルアーカイブが「可視化」されて、他のサービスプロバイダから統合的に利用できるように、標準APIが実装されるように働きかける。
        1. デジタル化、連携のための共通仕様のガイドラインの提示だけでなく、そのガイドラインに準拠したシステムモジュールを提供し、実装を支援していく。
        2. 成果物をOSS化して、公共図書館等へ提供していく
  4. サービスプロバイダ拡大策
     1. それぞれの機関が提供する情報資源をより多くの人に使っていただくために、アクセスの入り口を広く多様にする。
        1. 利用者ニーズに応じたサービスプロバイダ（ポータル等）が的確に開設されるように、PORTAでの統合検索のAPIを広く公開する。
  5. 機能強化策
     1. 国等の研究開発プロジェクトの成果を活用して、利用者ニーズ、技術動向を踏まえた機能を実装。
        1. メタデータ・コンテンツの収集・組織化・閲覧には、先進的な技術の適用が必須。それをNDLの資源で開発することは困難。
        2. 従来どおり、OSSをベースに開発する。
           1. 有用な機能モジュールが継続してリリースされ、研究開発成果も実装しやすいものをベースに構築する。
     2. 閲覧機能
        1. DAシステムの簡易検索機能はPORTAで実現（個々のコンテンツの詳細検索は、特性に応じた閲覧機能を構築する）
     3. 運用を効率化する機能
        1. DPプロバイダ追加を極力自動化したシステムの導入
  6. 普及啓発
     1. 有用なデジタルコンテンツを保有している機関に対して、広く連携を申し入れる
        1. 民間にしかないコンテンツ、サービスとの連携も積極的に行う
     2. 具体的な連携の打診内容
        1. 分散デジタルアーカイブを統合的に利用できるようにして国のデジタルアーカイブを構築する意義と、実施のための連携の必要性を説明
        2. まずは、ウェブアーカイブ/デジタルデポジットでのコンテンツを収集を打診、だめなら、PORTAでの統合検索、それもだめならDnaviでのサイトナビゲーションの実施の可能性を打診

## まとめ

### 5年程度までの方向性（クラウドコンピュータの世界の普及をイメージして）

* + 1. 過去２～３年の変化。
       1. Blog、wiki、SNS等で個人の情報が発信されるようになり、ソーシャルブックマーク機能で個人がタギングして情報を分類できるようになった。RSS1.0でメタデータの配信、RSS2.0やatomでコンテンツの配信まで行われるようになった。
    2. クラウドコンピュータの世界でのサービスのイメージ（図を作成）
       1. 5年後には、大きなパラダイムシフトがあると想定していい
       2. クラウドコンピュータは、ネットワークコンピューティング、グリッドコンピューティング、SaaS（Software as a Service、Webサービス等の技術を融合させた、データセンターのようなもの。
          1. クラウドコンピュータ自身は、コンピュータシステムの概念であるが、共通インフラの上で稼動することになるサービスは、共通のサービスインタフェースを持つことが必然となっていく。
          2. 個別の組織が、単独でインフラを抱えて、情報を抱えて、サービスを提供し、利用者が、個別のサービスを利用するような形ではない
       3. クラウドコンピュータの世界は、単に情報を共有する世界ではなく、サービスを共有する世界になっている
       4. NDL自身のサービスも、このパラダイムシフトに対応していかなければ利用されなくなる。
          1. Web2.0が2～3年で爆発的に普及したことを考えると、このような時代が、5年後に到来すると思われる。そのことから予測して、情報の管理の仕方、サービスのあり方を考えていく必要がある。
    3. クラウドコンピュータの世界で、情報と情報が情報群、サービスとサービスがサービス群、人と人がユーザ群として有機的に関連付けられていくものと思われる。
       1. 従来のように検索し、コンテンツを閲覧するような形から、クラウドの世界のリソースを、所在を意識せずにサービス群として利用
          1. Web2.0では、パーソナライズ、人と人の関係を活用して情報を選択してきた。また、サービスも先進的な機関が提供するWebサービスを個別に利用してきた。
          2. 次世代は、それに加えて、情報と情報、サービスとサービスの関係を組み合わせて、ユーザ群と情報群の多対多の関係で、より的確な情報の活用が可能になる。
          3. また、情報と情報、人と人、人と情報の関係を組み合わせれば、今まで見つけられなかった情報の選択が可能になる。
       2. このような時代には、図書館、文書館とか、博物館、美術館、民間、個人の区別なく集合が作られ、その中で、NDLは、複数のサービス群、情報群の１つとなっている。
          1. このような状況は、ウェブアーカイブ、ポータルサービスのあり方にも大きく影響する。当館のDAシステム、PORTAは、これらの情報発信の状況に対応したものでなければならない。
    4. 当館は、
       1. 国の情報資源を後世に残す役割と
       2. 現在の利用者のために、当館を収集・保存した情報資源を含めて、国の情報資源を的確に利用できるようにする役割を持つ
    5. 当館として、
       1. 情報を利用するというより、サービスを利用する時代に、どのようにして情報を収集するのか？サービスを保存することは意味がない
       2. また、当館のDAシステム、PORTAのシステム構築においても、それらの仕組みを取り入れていく必要がある。
          1. 5年後に当館を入り口としたサービスを想定しても、世の中の情報探索、情報入手の入り口として使われるものになるとは思えない。
    6. これらの変化に当館だけで対応することは不可能。システム構築、デジタルコレクション構築の双方ともに、関係機関との分担と協調なしでは達成し得ない。
       1. （当館がすべての情報アクセスの入り口になることはない）
       2. （有用な情報は必ずしも図書館に集約されない。図書館員が作れる有用な解題情報、レファレンス情報はほんのわずか。一般の人の集合知は、少数の専門家の知識に勝る。図書館を中心に考えた世界では利用者ニーズに応えられない）
    7. また、膨大な情報資源を保有する当館が、保存と利活用の両面で、先導的な役割を果たすためには、広く普及している技術やサービスを、後追いで適用する形であってはいけない。
       1. デジタルアーカイブシステムや統合検索システムの開発においても、当館からベンダーへの開発請負業務で行っていたり、図書館に閉じた世界でのパッケージ導入では、広く一般の利用者のニーズに応えたサービスは提供できない。
       2. 研究開発、技術開発の成果を活用させていただくことが重要で、それらを組み合わせて適用することが、効率的でありかつ効果的。（統合検索のAPIから、サービス連携のAPIによる連携を目指す）
    8. また、コレクション構築においても、当館が収集・保存できるデジタルコンテンツは、国の情報資源のごく一部。
       1. 公共図書館、大学図書館を含めて、MLAの各機関、民間、政府、さらに個人が持つ有用な情報を「見える化」して、各機関の持つ情報が分散データベース、分散サービスを統合して１つの巨大なデジタルアーカイブサービスとして利用できるようにしていかなければならない。
       2. そのためには、まず、関係機関のデータベースが相互に利用し合えるようにしていくことが重要。

### デジタルアーカイブの今後の構築計画（想定）

* + 1. 世界レベルでの電子図書館構築の構想の一翼を担う形で、日本としてのデジタルコレクションの構築を目指していきたい。
    2. そのために、この線表のように、スケジュール感で進めていきたいと考えている。

# （2008年）今後実施すべきことと連携協力（図書館総合展）

**2008年図書館総合展での発表内容**

平成20年11月25日

関西館　中山

## 挨拶

国立国会図書館の関西館　主任司書の中山です。デジタルアーカイブシステムを担当しています。

引き続いて、様々な機関や個人が提供するデジタルコンテンツを「見える化」して、統合的に利用できるようにすること、と

組織としても、公文書館、博物館・美術館、図書館等のいわゆるMLAが館種を超えた連携により、国全体のデジタルコレクションを構築していくための取り組みについて、「夢」を含めてお話しさせていただきます。

この図書館総合展のセッションには、専門家としての情報の組織化、レファレンスを重視している方が多く参加されていると思います。

私自身は、いわゆる図書館屋ではありません。そこで、別の観点から、一般ユーザ、情報処理技術者の視点で、今後実施すべきことをお話しさせていただきます。

## 本日お話しすること

今日のお話しは、次のような内容を用意しています。

何をすべきか、どのようにシステムを構築するか、どのようにしてコレクションを構築するか、をお話しし、最後に、5年程度の将来に向けた方向性として、想定していることをまとめたいと思います。

* 1. 何をすべきか
     1. *データベースの見える化*
     2. *情報を集合知化*
     3. *情報の組織化*
     4. *利用者の組織化*
     5. *情報と利用者の関連付け*
     6. *情報探索方法の変化*
  2. どのようにシステムを構築するか
     1. *当館システム構築の考え方*
     2. *デジタルアーカイブシステム構築に向けた連携協力*
  3. どのようにしてデジタルコレクションを構築するか
     1. *政府の施策（2008年）*
     2. *公共図書館との連携*
     3. *学術機関との連携*
     4. *商用サイトとの連携*
     5. *出版社、著作権者との連携*
     6. *MLA連携*
     7. *日中韓連携*
     8. *WDLとの連携*
  4. まとめ
     1. 最後に、5年程度の将来に向けた方向性として、想定していることをまとめたいと思います。

## 現状の再認識

* + まずは、現状の再確認です。
  + インターネット上では、指数的に、デジタル情報が増大しています。
    1. 従来のように組織、マスメディアからの発信に加えて、個人からの多くの情報が発信されています。
    2. 情報の質に関しても、個人からの情報も、専門家の情報に負けないくらい質の高いものも多く見られます。
  + 現在のユーザの、インターネットでの情報探索の行動の目的は、まさに問題・課題の解決であり、目録などで、情報がどこにあるのかを確認するのが目的ではありません。
  + 多くのユーザは、インターネットで問題解決をしており、Googleなどの検索エンジンで見えない情報はないも同然と言われています。
  + また、有用な情報は、民間、個人のデータベース、ブログなどに多く存在します。
  + それらのインターネット上の情報は膨大で、すべてを１つの組織で、収集・保存し、提供することは不可能です。
  + また、膨大な情報を、体系的に整理（紙資料の書誌に相当するメタデータ付与等の組織化）することは、専門家の人海戦術では、到底追いつきません。
  + 個人が持つ情報の集合知は、専門家をしのぎ、質的にも劣らない可能性があります。
    1. インターネット上の個人の知が集まれば、専門家が用意できる量に比べて格段多く、また質的にも劣らないものになる可能性があります。
  + 膨大な情報を検索できる時代に、従来の図書館的な考え方で、単独の図書館が所蔵資料の検索から入るようなサービスを提供していては、一般のユーザからは利用されなくなります。

## 実施すべきこと

* + これは、（Web2.0時代の次を見据えて）このような現状を踏まえて、何をすべきかを図示したものです。
    - 個々のデータベースのなかに隠れている情報を「見える化」すること。
    - それらを集合知化すること。
    - 情報と情報の関連付けをすること。
    - 人と人の関連付けをすること。
    - さらに、情報と人の関連付けをすること。
    - その上で、情報の選択の絞込みや、範囲の拡大、利用者属性に応じた情報閲覧環境を提供していくことです。

## 実施すべきことと、必要な機能の関連

実施すべきことを、具体的に実現するために必要な機能を、考えていきます。これらの機能をどのようにして開発していくかについて、連携のところでお話しさせていただきます。

### データベース内の情報の「見える化」（可視化）

* + まずは、データベース内の情報の「見える化」です。
    - 一般のユーザの情報探索行動として、検索の入り口は、まずは検索エンジンです。検索エンジンで見えない情報はないも同然です。データベースを公開していても、検索エンジンでは、データベースの中までは検索できません。
  + また、検索エンジンで見つかった無数のデータベースを、個別に確認していくことは困難です。
  + 個々に公開されているデータベースを、「見える化」すれば、検索エンジンや他のシステムからデータベースの中まで検索できるようになります。
  + データベースの中まで検索できるようになると、情報量は、また飛躍的に増加します。
  + ユーザの個々の検索ニーズに答えるためには、ニーズに応じて特化した多様な入り口（ポータル）が必要です。

### 情報を全体で集合知化

次に情報の集約に関してです。

* + 組織を超えて知識を集約すること、として
    - 現在、図書館、公文書館、博物館などのそれぞれの機関は、多くの書物を所蔵していますが、明確な基準に沿って、分担保存されていません。
      * 例えば、GoogleやPORTAで「源氏物語」を検索して見ると、様々な機関、個人が情報を発信しています。１つの機関だけを検索しても、全体から見ると一部に過ぎません。
      * 個々の機関が網羅的なコレクションを構築しようとしても、原本を交換し合うのは不可能です。しかし、デジタルコンテンツなら、組織を超えて、必要なコレクションを持ち合うことが可能です。
  + 専門家の知識を集約すること、として
    - 図書館、文書館等の専門組織の強みは、司書、学芸員などの職員が持つ知識です。その知識は、多岐に亘っていますが、それぞれが持つ知識は、単独では網羅的ではありません。これらの知識を集合して使えるようにすることが大切です。
    - その方法として、古くから、ナレッジデータベースの構築が叫ばれています。
      * 当館では公共図書館を中心とした司書による「レファレンス協同データベース」を運用し、また「ナレッジデータベース」の構築と提供も目指しているが、これらに参加する専門家が多くなればなるほど、利用価値の高いものになります。
      * 専門家によるWikipediaのようなものを作ってもいいのではないでしょうか？
  + 個人の知識を活用すること、として
    - CGM(Consumer Generated Media)の活用
      * ここ2～3年で、組織や専門家だけでなく、一般ユーザが直接生成する情報は、膨大になってきました。それらは、個々には断片であっても、集合すると有用な情報になっています。
      * はてな、Yahoo知恵袋、教えてgooなどでは、専門家のレファレンス情報を凌ぐ情報提供も多々あります。Wikipediaを辞書や解題として利用して、問題解決できる場合も多いと思います。
      * また、ユーザの利用行動の情報、コメントも、集約するとそれ自身も他のユーザにとって、有用な情報源となります。
        + *多数の意見、類似嗜好の人の行動を情報として提供することにより、利用者の行動の参考となるようにする。*

### 集合知の永久保存

* + 永久保存のために実施すべきことの１つとして
  + *組織、専門家、個人が生成しインターネットで公開した情報は、どれも現在のユーザにとって、重要なものです。*
  + *それらは、いつの時代も歴史的文献として古い資料を閲覧利用するのと同じで、将来のユーザにとっても重要なものです。そのために、将来に亘って利用できる形で保存しておくことが必要です。*
    - *インターネット上の情報は消え行くものです。各サイトから消えるまえにアーカイブ機関が保存できるようにする必要があります。*
      * *各サイトで利用できるうちはアーカイブからコンテンツを公開できなくても、消える前に保存する必要があります。*
    - 現在の媒体、ファイルフォーマットは、そのままでは、将来、読めなくなります。紙で残すのが一番いいという意見も聞かれますが、動画、音声を含んでいたり、動的に生成されるコンテンツは紙では保存できません。→国際標準、業界標準を策定した団体、その仕様に基づいて製造したベンダの協力のもと、マイグレーション、エミュレーションができるようにすることが必要です。
  + *インターネット上の情報は、すべてを１つの組織で保存することは不可能です。*
    - *複数のアーカイブ組織で分担して保存できるようにする必要があります。*
    - *各アーカイブ組織が、自らの組織のタスクに応じて、必要なコンテンツを交換し合って保存できるようにしていく必要があります。*
      * *永久保存の観点からのディザスタリカバリ対策の面からも有効なことです。*

### 情報の組織化、

* + 情報の組織化として、
  + 情報が膨大になるほど、情報の選択が必要になります。最近は、コンテンツの内容をキーワードで全文検索することで十分、という声もありますが、それでは本当に必要な情報が見逃される可能性も多くあります。
  + そのために情報に対する情報として、ある程度のメタデータが付与されることが有効です。メタデータは一義的には、コンテンツの作成者が付与するのが望ましいと考えます。
  + CGM(Consumer Generated Media)という概念がありますが、
    1. 現在、blog、wiki、SNSのようなCMS（Contents Management System）を利用した環境で生成されたコンテンツは、ユーザが意識しなくても、システムにより、自動的に、ある程度のメタデータが付与されます。
    2. また、htmlで書かれた文書に、システムにより、マイクロフォーマットのような形で、システムが判断できる意味ある情報をもたらすことも可能になってきています。
    3. また、フォークソノミーの１つとしてのソーシャルタギングとして、はてなブックマーク等のようにユーザが有用と思うページをソーシャルブックマークして、ユーザ自身の分類のためにタグを付与します。
       1. 統制語彙のようにあらかじめ定められた言葉でものごとを分類するのではなく、各人が任意のタグ（自然語）を追加するものです。ユーザが自ら情報群を作ることです。
  + より精度の高い検索を可能にするためには、情報に対して、タイトル、著者、作成日はもとより、多角的な検索のための各種分類を付与することが必要です。
  + 自動組織化（自動メタデータ付与）
    1. それには、人手を介さないで、コンテンツの文書構造、文脈、画像や音声の特徴を、システム的に自動で認識し、メタデータを生成する、自動メタデータ付与の仕組みが必要です。
       1. 現在利用されている例としては、はてなブックマークでは、ブックマークすると、自動的にカテゴリ分類とキーワードが候補として提示されています。
    2. 現在、また、今後も、様々な機関で研究開発や製品開発が進んでいくと思われます。
  + 情報と情報の関連付け、として
    1. 情報にある程度、メタデータが付与されれば、著作の内容、属性での類似性などで、グルーピング（いわゆるクラスタリング）して情報群を形成できます。
    2. あとで図示する、FRBR、書誌ファミリーの概念での組織化も一つの方法です。
       1. FRBRや書誌ファミリーの概念で、同一著作、原作から様々な媒体や形式に派生した著作物を書誌事項で体系的に把握することも有効です。
    3. CGMの活用例として、ソーシャルタギングなどにおいて、統制語が使われていけば、より確実な分類になって、ユーザの情報だけででも、精度の高いグルーピングが可能になっていきます。
    4. セマンティックウェブ技術の活用←分かりやすく解説
       1. 個々の情報に対して、ある程度、統制されたメタデータが付与されれば、情報の関連性も把握しやすくなり、また、オントロジーなどの仕組みにより、異なる文化で生成された情報のグルーピングが進みます。

### 利用者情報の組織化

* + 図書館では、従来、利用者情報、利用情報は、情報検索、閲覧には利用してきませんでした。
  + しかし、ユーザの利用傾向がわかれば、ユーザの趣味、嗜好などを推定でき、同じ検索手段でも、あらかじめ、提示する情報の機械的な絞込みが可能になります。
    1. 決して、実在する特定個人のプロフィールや、行動を監視するのが、目的ではありません。
       1. その意味では、実在する個人として登録されないと使えない申し込み機能と、情報探索のためのユーザ管理は別物として運用されることが好ましいかもしれません。
    2. また、推定に当たっては、ユーザには、公的、私的で別の嗜好を持つことも考慮する必要があります。
       1. 例えば、情報処理の仕事をしているが、草花が好きという人がいます。どちらも自分自身ですが、それを一人の興味・嗜好と認識すると、誤った推薦をしてしまいます。
       2. ユーザの特別な行動は排除されなければ、的確な情報の提示ができません。（人から頼まれて調べた。。。）
  + 人と人の関係付け、として、
    1. ユーザ個人の利用情報だけなく、複数のユーザ同士の利用傾向の類似性が把握できれば、他のユーザの利用傾向を、あたかも自らの利用傾向として認識でき、人と人のコミュニティが形成できます。そのコミュニティ内の誰かの行動は、自分の情報探索の行動にも参考になるものです。
    2. ユーザ同士の類似性で、が作られ、そのコミュニティに参加することにより、より大きなユーザ群としての利用行動として認識できるようになります。
    3. これは、ユーザ群が情報群を作ることです。

### 情報と人の関係付け

* + 情報と人との関係付けとして、
  + どういうユーザ群が、どういう情報群を利用したかという形で、ユーザ群と情報群が関係付けられれば、情報の選択がより一層容易になって、インターネット上の情報が膨大になっても、的確に情報に辿り着けるようにすることができます。また、ユーザ群から情報群の中にある情報を見つけ出すことにより、通常では見つけられなかったような情報も見えてきます。このような形で、ロングテールにもつながります。
    1. これは、協調フィルタリング（Collaborative Filtering, CF）の考え方です。
       1. 多くのユーザの嗜好情報を蓄積し、あるユーザと嗜好の類似した他のユーザの情報を用いて自動的に推論を行う方法論である。協調フィルタリングはリコメンデーション（推薦）やパーソナライゼーションに利用されている。実際にAmazon.comの「おすすめの商品」や、はてなアンテナの「おとなりアンテナ」などに応用されており、Web 2.0においてロングテールを支える要素技術として捉えることができる。（Wikipediaより）
  + ユーザが利用者の視点で分類する情報群は、
    1. 図書館、公文書館、博物館、美術館とかの所蔵館や館種の区分ではありません。
    2. また、一次、二次、参考情報といった区分でもありません。
    3. ユーザ群ごとに情報が検索しやすくなっていることが重要です。

### 情報検索手段（利用方法）の変革

このような状況において、情報の検索・閲覧環境として必要なサービスを想定します。

* + **情報の関連を活用した検索（情報の選択範囲の拡大・絞込み）**
    1. 全文検索だけでは、的確に絞り込めません。また、検索語など、適切な検索条件を、入力することは困難です。
    2. 連想検索、あいまい検索、さらに多角的な絞込み検索のために、ファセット検索、クラスタリング検索が効果的です。
    3. 異業種、異文化圏の情報の検索として
       1. オントロジー等の技術を活用した検索語の翻訳、検索結果の自動翻訳などの技術も期待しています。
    4. 組織を越えた統合検索と、関連する情報への芋づる的なナビゲーションができれば、
       1. 組織を越えた大きな集合知の中から、多角的に絞込み、関連しそうな情報を提示することが可能になります。
  + **利用者属性、利用情報を活用した検索および閲覧（利用者属性・利用環境に応じた情報閲覧）**
    1. いわゆるパーソナライズ機能として、
       1. ユーザの属性、行動履歴を参考に
       2. 利用場所、利用機器に応じて、最適なコンテンツを自動選択する仕組みを作ることができます。
          1. 同一原本から、様々な媒体、形式に派生したコンテンツを、利用場所・利用機器により閲覧可能なコンテンツを選別し、

利用場所：図書館内、移動中、自宅等

利用機器：PC、携帯電話、PDA等

最適なコンテンツ形態のものを利用できるようになります。

* + - * 1. 今や、インターネットはPCだけで見るものではありません。携帯電話や携帯端末での利用が増加しています。
        2. 携帯電話で近代デジタルライブラリーのようなイメージ画像を見ることは困難ですが、青空文庫なら容易に読めます。また、朗読の音声データがあれば満員電車の中でも聴けます。
    1. 協調フィルタリングの考え方に基づいて
       1. 他人の利用情報により推奨する検索が有効です。

### まとめ

* + 1. 情報の利用の仕方も、技術の進展とともに変化しています。
    2. ニーズとしての利用の仕方の調査研究も必要で、新たなニーズが見えてくれば、それに対応して、技術シーズも生まれてきます。

## PORTAにおける「見える化」のイメージ

NDL-DAのシステム構築の考え方については、割愛。

### NDL-DA、PORTAシステムの構築の考え方

*以上のようなニーズを満足するため、当館のサービス構築、システム開発では、次のような考え方で実装を推進していきます。*

* + *サービス構築*
    1. *基本的には作らない。OSSの組合せ、サービスのマッシュアップ。*
    2. *技術シーズ、利用者ニーズに速やかに対応するためには、研究開発成果のタイムリーな実装が必要。*
    3. *研究開発成果や、先進機能を使ってタイムリーにバージョンアップされていくOSSを組み合わせていくことが有効。（図書館用の商用パッケージでは、タイムリーに対応できない。）*
    4. *システムはWeb上にあるサービスを組み合わせて、マッシュアップで構築することが効率的。*
       1. *各機関のサービスを活用しあうサービス連携を想定。*
    5. *今後は、Webサービス連携から、コンピュータ資源やサービスを共有しあうクラウドコンピュータの世界への適応*
  + *デジタルアーカイブ機能*
    1. *対等な関係での分散デジタルアーカイブを指向*
       1. *当館は国のデジタルアーカイブを構成するデータプロバイダの１つ。日本中のデジタルコンテンツを１つの機関ですべて収集・保存することは不可能。*
       2. *分散デジタルアーカイブが連携して、1つの巨大なデジタルアーカイブとして利用できるようにすることを目指す*
       3. *メタデータ交換の統合検索連携は第一歩。次に、コンテンツの相互交換による連携。*
          1. *次期システムでのコンテンツを交換し合う方法として、セキュアなPeerToPeer（P2P）ネットワークでの連携も想定*
  + *ポータル機能*
    1. *一般の利用者にとって、データベースの所在を探すのも含めて、実質的な入り口は、GoogleやYahooのような検索サイト。*
       1. *Google、Yahoo等の表層ウェブを中心にした検索エンジンと、検索エンジンからはアクセスが困難な深層ウェブとのワンストップ検索サービスの融合が必要。*
    2. *目的とするコンテンツを直接検索・閲覧できるサービスが求められている。たらいまわし的なナビゲーションは好まれない。*
       1. *利用者層、直接統合検索できるコンテンツのカバレージなど、目的を明確にして、ユーザニーズに対応したポータル機能が必要。*
    3. *現状のPORTAは？*
       1. *当館が所蔵する資料をデジタル化したコンテンツに加えて、インターネット上から収集したコンテンツと、各機関が構築し提供しているコンテンツを検索できるようにしたデータプロバイダです。*
          1. *PORTAの統合検索は、統合検索先のメタデータを検索用データとしてハーベストして保持するリポジトリ機能と、検索要求の都度横断検索することにより実現しています。*
       2. *データプロバイダとして、検索機能をWebサービスのAPIの形で、他のサービスプロバイダーに提供します。*
          1. *PORTAが持つGUI機能は、このAPIを利用して提供しているもので、デジタルコンテンツをアクセスするための入り口の一つです。デジタルアーカイブがWeb上で「見える化」されるのを加速させ、各機関のデジタルコンテ*

### 情報の組織化の１つ。FRBRの概念での組織化による検索・閲覧サービスの実現（１ｍ）

これは参考です。情報の組織化の１つとして、書誌的体系に基づいた分類のモデルの１つです。

一つの著作物は、デジタル化されたコンテンツも含めると、この体現形の例示にありますように、様々な形態のコンテンツとして派生して流通しています。

体系的に分類されていると

利用者属性（知識レベル、嗜好等）、利用環境（PC、モバイル、アクセス場所等）を考慮して、コンテンツを的確に選択できるようにすることができます。

また、選択されたコンテンツを実際に入手する際には、個別資料として管理されている資料の中から、最も入手しやすいところで所蔵しているコンテンツを選択できるようにすることが容易になります。

### コンテンツ交換の方法の１つ（PeerToPeer）対等な関係でのデジタルアーカイブの構築

これも参考です。コンテンツ交換の方法の１つとして、組織が対等な関係でデジタルアーカイブを構築するイメージです。

**メ**タデータだけでなく、コンテンツ、さらに、コンピュータ資源も共有しあう仕組みが可能になり、コンテンツや、メタデータを、必要に応じて交換しあうことができるようになります。

当館は保存のためにコンテンツを保有し、また、各組織はそれぞれの利用者ニーズに応じて、必要なコンテンツを保有して、提供する形が見えてきます。

## デジタルアーカイブシステムの構築に向けた主な連携協力

「デジタルアーカイブシステムを構築」していくためには、これらの機関の協力なしには成しえません。

### 産学官の研究機関との連携

産学官の研究機関の成果をベースに構築していきたいと考えています。

* + 進め方、としては
    1. 国際標準、業界標準の積極的な適用
    2. 政府機関、民間を問わず、国内外の研究機関、研究者等と連携して調査研究
    3. 各機関の研究成果を組み合わせて機能を実現
    4. まずは、海外ではIIPC、国内では研究開発を行っているJST、NII、NICT、AIST、IPA等の政府機関、大学の研究室等との連携
    5. 有用な検索サービスを実施している商用ポータル機関（Google、Yahoo等）との連携
  + 産官学の研究開発プロジェクトの成果の活用

を基本的な方針としています。

デジタルアーカイブシステム及びPORTAの構築において、技術的な課題があり、パッケージやOSSが存在しない技術に関して、政府機関、大学等の研究開発成果の技術移転を期待してします。

* + 1. 縮減された予算の中でも機能強化を可能とする方策の一つ。
    2. 主な研究開発プロジェクトとして、総務省の「文明開化プロジェクト」、経済産業省の「情報大航海プロジェクト」、文部科学省の「情報爆発」などがあります。
    3. 文明開化プロジェクト（総務省：21～23年度）
       1. 概要
          1. 「ICT 成長力強化プラン」に基づき、国立国会図書館、国立公文書館、他府省庁、地方公共団体、図書館・博物館・美術館、大学等との連携を図り、産学官を挙げてデジタル化を推進、日本中の知的財産を総デジタル化してつなぐ「デジタル文明開化プロジェクト」を実施する（重点計画2008（案）（6月11日）IT戦略本部）
          2. デジタルアーカイブ構築のために必要な技術開発

複数のデジタルアーカイブの統合に関する技術

分散型・大規模デジタルアーカイブの長期保管・利用に関する技術

デジタルアーカイブの多目的な利活用に関する技術

* + - * 1. 地域公共図書館等がデジタルアーカイブシステムを導入するための資金援助
      1. 状況
         1. 総務省での予算要求のために、技術開発が必要な分野、テーマ候補の提示、説明資料等の作成に協力
      2. 予算が確保された場合、来年度の実施に向けて、今年度末までに調達の準備をするとのこと
    1. 大航海プロジェクト（経済産業省）
       1. 経済産業省情報政策課に対して、今年度の「共同活用基盤を用いた実証実験」及び、来年度以降の「大規模図書館が将来的に活用可能なモデルサービス等の実証実験」として検討していただきたい研究開発テーマの一覧を提示。具体的なアクションにまでなっていませんが、12月12日に、デジタルアーカイブWGが開催される予定です。
    2. 情報爆発IT基盤（文部科学省）
       1. 「情報爆発IT基盤」に対して、直接具体的な申入れはしていないが、このプロジェクトの個別研究開発テーマの研究者（木俵先生、黒橋先生）とはコンタクトを取っている。
    3. 民間の技術開発、製品開発　（）
       1. 図書館パッケージベンダー　（図書館が利用する冊子体OPACからデジタルコンテンツの目録提供システムへ）
    4. 共通仕様・技術提供
       1. 日本規格協会　（メタデータ交換、サービス連携のための、通信プロトコル、メタデータ記述要素・記述規則）
       2. 電子情報技術産業協会（JEITA）　（製品開発者として、各種デジタルフォーマットの適用、長期的な利用を保証するためのエミュレーション、マイグレーション技術の提供、現在、オーディオアーカイブに関して仕様検討会が開催され、NDLはオブザーバ参加）

### （連携協力）国際インターネット保存コンソーシアムとの連携

* + IIPC
    1. 概要
       1. 国際インターネット保存コンソーシアム (International Internet Preservation Consortium） 、略称IIPCは、Webアーカイビングに関する相互運用が可能なツールや技術の開発・標準化を推進し、国際的な利用を促進することを目的として結成された機関です
       2. 現在、各国の国立図書館や公文書館など38の機関が参加
       3. 当館は、今年4月に加盟しました。
       4. IIPCを通じて国際的に貢献することはもちろん、相互運用可能なツールや技術を国内に還元することで、日本のウェブアーカイブの進展にも貢献できると思っています。
       5. ワーキンググループとして、「スタンダード、ハーベスティング、アクセス、プリザベーション」の4つがありますが、当館は、現在、アクセスのWGに参加しています。来年度以降、プリザベーションのWGへの参加も検討しています。
    2. ワーキング
       1. スタンダード
          1. 短期的にはWARCの標準化、将来的にはメタデータ、評価基準の標準化等の標準化
       2. ハーベスティング
          1. Heritrixの開発、WARCフォーマットの開発とサポート、クローラに特化した要望、深層ウェブの収集、ビデオやストリーミングの収集
       3. アクセス
          1. アクセスツール（Wayback machine,WERA）改善、NutchWAXを使った全文検索のテスト、ユーザ認証、ユーザ権限、アクセスコントロール、アーカイブのコンテンツ解析のためのツール、アーカイブの構造解析のためのツール
       4. プリザベーション
          1. デジタル情報の保存をめぐる技術的課題への取り組み、ストレージの拡張性に関する研究、WARCフォーマットの保存機能の向上、ウェブをめぐる技術環境の変化の記録、再現できなくなりつつあるファイルフォーマットの特定

### 国のデジタルコレクションの構築に向けた主な連携協力

これは、国のコレクションを構築するための連携協力先のリストです。予定も含んでいますが、様々な機関と連携協力のための調整を行っています。 この中から２、３、簡単に説明します。

* + 1. 内閣官房。美術館、博物館、文書館、図書館連携（MLA連携）　。学術情報機関。公共図書館。出版者等。民間デジタルアーカイブ。歴史的音盤アーカイブ協議会。商用ポータルサイト。政府情報などがあります。
  + *ナショナルデジタルコレクションの構築*
    1. *内閣官房* 
       1. *知的財産戦略本部　（知財計画2008で示された、NDLのデジタルアーカイブ化、図書館資料の利用の促進、著作物のフェアユース規定の日本版、、、）*
       2. *IT戦略本部　（e-Japan重点計画I・重点計画2008？）*
    2. *美術館、博物館、文書館、図書館連携（MLA連携）　（関係機関のサイトへのアクセス、来館利用が促進される形での連携・協力、デジタル情報資源ラウンドテーブルを設置予定）* 
       1. *国立公文書館 、東京国立博物館 、各美術館、博物館、公文書館、図書館*
    3. *学術情報機関　（学術情報、科学技術情報の統合検索、コンテンツの長期保存、学位論文の保存と提供、NDLと大学図書館との連絡会議の下に学位論文電子化の諸問題に関するWGを設置して検討してきた）* 
       1. *NII、JST 、大学図書館 、電子ジャーナル出版者*
    4. *公共図書館（各図書館での郷土資料のデジタル化とアーカイブ構築を支援、公共図書館の総合目録ネットワークの便利さをデジタルの世界でも実現）* 
       1. *都道府県立図書館 政令指定都市立図書館 、市町村立図書館 、専門図書館*
    5. *出版者等　（電子書籍作成支援、電子出版用に出版者へ提供、各出版者へのナビゲーション、出版者等のビジネスが拡大できる形での連携・協力）* 
       1. *日本電子出版協会（JEPA） 、日本印刷技術協会（JAGAT） 、冊子体書籍出版 、電子書籍出版者 、インターネット書籍販売者*
    6. *民間デジタルアーカイブ* 
       1. *電子書籍ポータルhon.jpの統合検索（12月初旬公開予定）*
          1. *PORTAから、hon.jp Webサービス を利用して、hon.jpの電子書籍メタデータDBを統合検索する。冊子体資料として存在しない、「ケータイコミック」、「ケータイ小説」も検索可能となる。PORTAは、関連情報リンクで、既に商用サービスとの連携を実現しているが、統合検索先としての連携はこれが初めてのケースとなる。*
       2. *日経BP社の日経BPビズボードの統合検索（11月17日意見交換）*
          1. *日経BPビズボードで提供している雑誌記事、WebニュースをPORTAで統合検索。*
          2. *雑誌記事のメタデータはDCベース、プロトコルは、OAI-PMH。（実装に時間を要するようであればファイル転送も可）。WebニュースはRSS。*
    7. *歴史的音盤アーカイブ協議会（HiRAC）　（SPレコード3万タイトル6万曲、その他記録音盤、音源フォーマット、メタデータ仕様、デジタル化分担、統合検索、歴史的音盤を将来に亘って聞くことができるように）*
    8. *商用ポータルサイト　（実質的なインターネットの入り口、色々考えられる）* 
       1. *Google、Yahoo等　（デジタル化、深層ウェブの可視化、サービス連携、商用ポータルが保有する情報の利用、検索システム構築の技術トランスファー） 、ソーシャルブックマークサイト*
    9. *政府情報* 
       1. *総務省行政管理局（e-GOV）　（政府サイト情報の統合検索、収集保存）*
       2. *各府省支部図書館　（政府情報のデジタル化、長期保存、デジタルコンテンツの統合検索）*
       3. *国立印刷局　（政府刊行物のデジタルフォーマットとしてXMLベースの原本作成）*
  + *国際協力* 
    1. *World Digital Library　（歴史的文化遺産を世界中の国立図書館が協同でデジタル化して提供するもの、NDLは日本として、東アジアの一員として参画し、世界規模でのデジタルライブラリが構築されることを目指す）* 
       1. *LC　LCが発行するガイドラインや施策は、今後の図書館を考える上で、とても参考になる。毎年、情報交換を行っている。*
    2. *IIPC　（ウェブアーカイブに関して、標準、収集、提供、長期保存の観点から、研究開発、NDLも4月より参加、構築中のデジタルアーカイブシステムでの適用と成果のフィードバックを目指す）*
    3. *日中韓連携　（漢字文化圏の三国が協力してデジタルアーカイブを構築して、コンテンツの長期保存と提供を目指す。昨年度協力の合意をして、現在各国の現状をまとめている。まずはパイロットシステムを構築して、メタデータ交換、横断検索を試行する予定）*

### 政府の施策（2008年度）

* + まずは、政府機関の施策との関連です。
  + 今年度は、国のデジタルアーカイブ事業に直接関連する政府の施策として、このようなものがあります。
  + 一つ目として、「e-Japan重点計画特命委員会　自民党デジタルアーカイブ小委員会」が、平成20年3月12日に公表した「デジタルアーカイブの推進に向けた申入れ」。です。
    1. 主な内容として、
       1. 平成16年6月の「国立デジタルアーカイブ構想」の提言に沿って、わが国のデジタルアーカイブの総合ポータルとなるPORTAが開設された。
       2. 国立国会図書館のウェブアーカイブの本格実施のための法制度の実現
       3. 全国図書館のデジタルアーカイブの統合化
          1. 全国の図書館が統合的なデジタルアーカイブの構築。
          2. 図書館間でのデジタル化された収集資料の相互利用、館内公衆端末での閲覧が可能となる著作権制度の創設。
          3. 著作権データベースの活用等著作権者が不明な場合の文化庁裁定制度の円滑な利用のための制度改善。
       4. 国立公文書館と国立国会図書館が協力
          1. 相互の役割の違いを尊重して、利用者の利便性の向上を図る観点から、連携・協力
          2. 全国の公文書館、図書館でのデジタルアーカイブ構築について、メタデータ付与等に係る技術面、ノウハウ面における支援、協力を行う。
          3. 国のデジタルアーカイブの総合ポータルの検索機能の充実等により、各分野でのデジタルアーカイブの利用の拡大を図る。
       5. 世界最先端のデジタルアーカイブ技術への対応
       6. など、提言として政府に申し入れられました。
  + 次に、内閣官房　知財本部（知的財産戦略本部）が6月18日に発表した「知財計画2008」です。
    1. 「当館のデジタルアーカイブ化と図書館資料の利用を進める」ということで、
       1. 当館において行われている貴重な図書等のデジタル化や、インターネット情報資源等を収集保存し、ネット上で一般ユーザーの利用に供する取組'について、その促進が図られるよう一層の連携を進める。と、うたわれています。
       2. このため、'権利者の経済的利益や出版ビジネスとの関係を考慮''しつつ、国立国会図書館における蔵書のデジタル化の推進に必要な法的措置を２００８年度中に講ずるとともに、国立国会図書館と他の図書館等との連携や図書館等利用者への資料提供の在り方については、関係者間の協議を促進し、２００８年度中に一定の結論を得る。（113/157ページ）
  + 次に、同じく内閣官房のIT戦略本部が8月20日に発表した'重点計画2008の中では、
    1. 総務省の「ICT 成長力強化プラン」の中で記述されている、「デジタル文明開化プロジェクト」の推進（総務省及び関係省庁） があります。
       1. 国立国会図書館、国立公文書館、他府省庁、地方公共団体、図書館・博物館・美術館、大学等との連携を図り、産学官を挙げてデジタル化を推進、日本中の知的財産を総デジタル化してつなぐ「デジタル文明開化プロジェクト」を実施する。とされています。
    2. 具体的な内容の一つとして、地域公共図書館支援を想定しています。
  + 「ICT 成長力強化プラン」の実施が盛り込まれた形で、「経済財政改革の基本方針」（いわゆる骨太方針2008）として、6月27日に決定されました。
    1. これに基づいて、各省庁において、来年度予算要求が行われています。

### 公共図書館支援のイメージ

* + これは、連携協力の例の１つとして、公共図書館支援のイメージです。
  + 公共図書館が、地域情報ハブとして、郷土資料をデジタル化したり、県域の情報を収集して提供する役割を支援するものです。
    1. *公共図書館が、県内有用サイトの収集もしくは横断検索することを支援*
    2. *公共図書館が、郷土資料をデジタル化することを支援*
    3. *公共図書館が、収集コンテンツ及びデジタル化コンテンツのデジタルアーカイブを構築し、提供することを支援（必要に応じて当館が利用しているデジタルデポジットシステム等を貸与する）*
    4. *デジタルアーカイブを構築し運用することが困難な図書館に対しては、当館がASP的なサービスを行う。*
    5. *NDLは、各公共図書館のデジタルアーカイブを長期保存もしくはバックアップのために収集保存する。*
    6. *NDLは、各公共図書館のデジタルアーカイブのメタデータをハーベストもしくは横断検索して、利用者に対して、全国公共図書館が持つ郷土資料を統合的に検索し、各図書館サイトへナビゲートするポータル機能を提供する。*
    7. *また、NDLは、標準として適用すべき仕様のガイドラインや手引き類を提示し、実装を支援します。* 
       1. *デジタル化及びシステム構築用の手引きの提供* 
          1. *「デジタル化の手引き」（公開中）*
          2. *「デジタルアーカイブ構築の手引き」（計画中）*
          3. *「共通仕様及び連携に関するガイドライン」（公開中）*
          4. *「外部提供インタフェース仕様書」（公開中）*
          5. *「メタデータスキーマガイドライン」、「メタデータ記述規則」（策定中）*
       2. *各機関での地域情報収集用及びアーカイブ構築用の汎用ソフトウェアの提供も想定しています。*
  + PORTAでの連携協力先の拡大がなかなか進まなかった要因として、左の枠のように、各機関の対応が困難だったことがあげられます。
    1. 資料を、デジタル化する費用がない
    2. 自前で、サーバを立ち上げられない。運用できない
    3. 意義は理解したが、標準プロトコルを実装できない
    4. ウェブページに貼り付けて、データベースとして検索できる形になっていない
    5. データベース化されていても、外部提供インタフェースを持っていない
    6. 外部提供インタフェースを実装する費用がない
    7. メタデータのマッピング調整に膨大な時間（工数・費用）がかかる
    8. DCをベースにした記述要素を使っていても、使い方がそれぞれまちまち
    9. 同じ要素を使っていても、記述規則が異なり、同じ内容として認識できない
    10. さらに、統合検索できることのメリットが理解されていないということもあります。
    11. アクセスが増えることが不安
    12. サービスが横取りされるのでは？
    13. これらを解決することが、関係機関の拡大につながると考えています。

### 関係機関拡大策の例公共図書館支援のスキーム

* + 支援のスキームとして、
    1. 当館は、デジタルアーカイブシステムをオープンソースとして開発し、
    2. そのシステムをベースに、研究開発により、公共図書館用のソフトウェアを作成。
    3. そのソフトウェアを各図書館が実装するために、国は、何らかの資金援助を行い、
    4. 同時に、当館は、構築のためのガイドラインを提示して支援するものです。
  + これにより、各図書館の郷土資料が、デジタル化して公開され、それが、インターネットで発見されるようになって、資料の利活用が進み、さらに、サイトや来館での利用者が増えることを期待しています。

### 学術機関との連携協力

* + これは、大学機関リポジトリと、当館が遡及デジタル化するコンテンツを含めて、学位論文のネットワークを形成するイメージです。
  + 学位論文のネットワーク形成
    1. 当館、大学図書館、国立情報学研究所の連絡会において、「学位論文電子化の諸問題に関するＷＧ」が設置され、平成20年3月27日に中間報告を公表しました。この内容に従って、国の学位論文のデジタル化とネットワーク化の実施に向けて、活動をしていきます。
    2. 三者による協調の下、デジタル化を推進するため、学位論文を、大学機関リポジトリに搭載し、リポジトリ、国立情報学研究所、当館などのシステムで、相互にメタデータを交換するためのメタデータプロファイルを検討しています。各機関は、このプロファイルにしたがい、リポジトリに搭載する学位論文の標準メタデータを作成します。
    3. 学位論文等の検索システムとメタデータの関係として
       1. 機関リポジトリでの学位論文メタデータプロファイルと、当館が所蔵する紙媒体の学位論文、当館が遡及デジタル化する学位論文のメタデータの関係を整理しています。
    4. 電子ジャーナルとの相互連携も視野に検討
    5. 明日のセッションDRF４において議論される予定です。
       1. 大学において共通のメタデータプロファイルを策定中
       2. NDLにおいて、紙媒体、及び、紙から遡及デジタル化した資料のメタデータ案（DC-NDLベース）を策定中
    6. 三者で整合をとり、クロスウォークを可能にすることにより、一元的なアクセスが可能になると考えています。
  + また、学術、科学技術関係での連携の実現を目指して、
    1. 国立情報学研究所（NII）のCiNii、JuNii＋との統合検索（平成21年4月頃公開予定）
       1. *CiNii、JuNiiのメタデータのハーベストは、両システムが改修中のため、新システム稼動後に実施することで内諾を得ている。*
       2. *CiNiiに関しては、OpenURLによる横断検索の準備は完了したが、CiNiiへの負荷が懸念されるため、メタデータをハーベストすることとした。*
       3. *JuNii+に関しては、OAI-PMHでのハーベストの準備は完了。*
    2. 科学技術振興機構（JST）のJ-Stage、Jurnal@rchiveとの統合検索の準備を行っています。（12月頃公開予定）
       1. *J-Stage、Jurnal@rchiveに関しては、メタデータをファイル転送で受領することとして、準備は完了。現在、JSTから学会に対して、PORTAへのメタデータ提供及びPORTAから商用サイトを含めた機関への提供の許諾確認を行っている。*

### 商用サイトとの連携協力

* + 商用サイトとの連携協力として、まず、電子書籍ポータルhon.jpの統合検索の公開準備を進めています。12月に公開予定です。
    1. PORTAから、hon.jp のWebサービスAPI を利用して、hon.jpの電子書籍データベースを統合検索するものです。冊子体資料として存在しない、「ケータイコミック」、「ケータイ小説」も検索可能となります。
       1. PORTAでは、関連情報リンクで、既に商用サービスとの連携を実現していますが、統合検索先としての連携はこれが初めてのケースとなります。
  + また、日経BP社との連携の検討も行っています。（11月17日意見交換）
    1. 日経BPビズボードで提供している雑誌記事、Webニュースを、PORTAで統合検索できるようにするものです。現在、意見交換を進めており、比較的早く実現できる可能性があります。
    2. 雑誌記事のメタデータはDCベース、プロトコルは、OAI-PMH。（実装に時間を要するようであればファイル転送も可）。WebニュースはRSS。

### Google,Yahoo等との連携の可能性

* + 商用ポータルは、実質的なインターネット情報の入り口であり、利用者の利便性向上、サービスの役割分担の上からも、重要な連携先の１つです。*特定商用ポータルとの排他的な連携ではなく、等距離での*連携の実現を目指して、意見交換を行っています。
  + 現在インターネットで提供されている各機関の表層ウェブ、データベースなどの深層ウェブ、それに加えて過去のウェブで公開されていた表層ウェブをあわせて検索できるようになれば、網羅的な検索利用が可能になります。
  + *Google*
    1. *Googleとの連携として、GoogleBookSearchの検索結果から、PORTA（近代デジタルライブラリー、貴重書データベース等）、及び、総合目録ネットワークへのナビゲーション*
    2. *蔵書のデジタル化、検索用テキスト化の可能性に関して、意見交換を行いました。（7月）*
    3. *Google社ストラテジックパートナーディベロップメントマネージャと意見交換。。現在、日本Google社より、米国Google社に実施可能性について打診中。*
    4. *今後、Googleが保有するデジタル化技術、検索技術等の技術移転の可能性についても意見交換することとしています。*
  + *Yahoo*
    1. *また、Yahooとの連携として、Yahooカテゴリの分類ルール、登録されたURLを、有用なサイト・ページとして、PORTAからのナビゲーション先として検索、*
    2. *また、ウェブアーカイブでの有用なコンテンツの収集起点としての活用させていただくこと、*
    3. *Yahooでの検索キーワードを辞書として利用させていただく等、Yahoo検索事業部と意見交換（8月*）

### 出版社、著作者等との連携協力

* + 出版者、著作権者との連携協力ということで、
  + 資料の長期保存と利用の促進のために
    1. 現在、協議会を設置して、デジタル化とコンテンツ提供に関して、制度面での課題解決に向けた協議を進めています。
  + 合意が形成されたのち、実施に向けた具体的な検討を進めます。
    1. まず、当館は出版物のデジタルコンテンツを長期保存します。
       1. デジタルコンテンツの収集
          1. マルチユースコンテンツの原本、印刷用原本等、DRMのないもの
          2. 可能な限り詳細なメタデータの提供
       2. 資料のデジタル化
          1. 出版物を当館がデジタル化することの許諾
       3. 当館からのコンテンツ提供
          1. インターネット提供が許諾できない場合は、公共図書館内もしくは館内での閲覧提供まで
  + 出版者のビジネス支援として、
    1. 各出版者の電子書籍作成支援、データベースの立上げ支援、各出版者サイトへのナビゲーションなどを想定しています。
    2. 各出版者の電子書籍作成支援
       1. 当館がデジタル化したコンテンツを電子出版用に提供
    3. 各出版者の電子書籍データベースの立上げ支援
       1. 当館デジタルアーカイブを利用して、出版者が販売することも検討
    4. 各出版者サイトへのナビゲーション
       1. 出版者の電子書籍データベースを、PORTAで統合検索
          1. 電子出版物の購読が促進されるように
       2. PORTAでの統合検索結果に、入手先として出版者へのリンクを表示
  + *日本版「Book Rights Registry」実現の可能性*
    - 1. *2008年10月、米国の著者団体と出版社団体は、Google社との和解条件の1つとして、著作権料徴収NPO「Book Rights Registry」の設立を発表した。日本国内でも似たようなシステムの整備を求める声は今後高まってくると思われるが、残念ながら現時点では、類似団体として音楽世界のJASRAC等があるが、雑誌や書籍出版業界にそのようなものはない。日本版「Book RightsRegistry」の実現に向けての制度面／法律面／技術面で問題点等を明確にしていく。*

### MLA連携

* + 背景
    1. 文書館、博物館、図書館においては、大量の歴史的資料を所蔵していますが、デジタル化は十分には進捗していません、
    2. また、デジタル化形式の標準化や、各組織・機関の作成するデータベース間の連携についても、その利活用において、多くの課題を抱えています。
  + 目的
    1. 各機関の所蔵資料のデジタル化が進んで、利用者が、所在場所を意識せず、必要な資料を探せるように。
    2. 統合検索結果から各機関のデータベースへ案内できるようにすることにより、各機関のデータベースの利活用が進むように。
    3. さらに来館して実物を見る利用者が増えるようにする。ことを目指します。
    4. 主な検討項目は、
       1. デジタル化における標準・ガイドラインの検討
       2. 地域におけるデジタルアーカイブ構築支援
       3. 一元的な閲覧環境の構築
       4. コンテンツの長期保存

などです。

* + 状況
    1. 平成20年3月準備会開催しました。
       1. 国立公文書館、東京国立博物館、国立情報学研究所、科学技術振興機構、慶応大学、奈良女子大、国立国会図書館
       2. 横断的アーカイブズ論研究会
    2. 平成20年度中に、ラウンドテーブルを発足させる予定です。
       1. 日本国内の博物館、図書館、文書館が、館種を超えたデジタル化による仮想的なコレクションの構築を目指す。
       2. デジタル化の進展に係る共通の課題について、その解決に向けた協議・検討を行う。
    3. 連携の現状
       1. MLAに関連する個別の連携の準備状況として、
       2. まず、国立美術館との連携があります。（12月初旬、公開予定）
          1. *5国立美術館長会議（9月30日）にて、PORTAでの統合検索の承認が得られた。*
          2. 国立美術館の４つの機関（東京国立美術館、京都国立近代美術館、国立西洋美術館、国立国際美術館）の総合目録である「所蔵作品総合目録検索システム」をPORTAから統合検索できるようにする予定です。現在、メタデータの交換機能の実装作業中で、12月中の公開を目指しています。
       3. 次が、人間文化研究機構との連携です。（未定）
          1. 人間文化研究機構内の５つの研究機関（国立歴史民俗博物館、国文学研究資料館、国際日本文化研究センター、総合地球環境学研究所、国立民族学博物館）のデータベースを統合した「資源共有化システム」をPORTAで統合検索するものです。*京都大学地域研究統合情報センター原正一郎教授を通じて、*まだ、意見交換の段階ですが、歴史的資料として、図書館を含めて分散所蔵されている資料が統合的に利用できることが加速されることを期待しています。28日に意見交換を行う予定です。
          2. *人間文化研究機構研究資源共有化事業委員会（8月）において、PORTAとの連携について、話し合われ、好意的な反応であり、人間文化研究機構側で調整を行っている。委員会において再度、議論され、原氏、東京大学史料編纂所の石上理事とで打合せを行うことになった。（10月28日、東京にて）*
          3. OAI-PMHもしくはSRWで、DCベースのメタデータを受領可能と聞いているので、決定後、速やかに統合検索が実現できるものと思われる。
       4. さらに大学図書館の歴史的資料のコレクションとの統合検索を目指して、各大学と個別に調整を行っています。
       5. 慶應義塾大学の「慶應義塾図書館稀覯書画像」、「慶應義塾写真データベース」、「慶應義塾図書館デジタルギャラリー」（実施時期未定）
          1. 各データベースは別々に構築されているものであるが、メタデータをハーベストする方法で統合検索が可能とのこと。*現在、慶應義塾大学メディアセンター本部を窓口として、学内調整を行っていただいている。*
       6. 早稲田大学の「古典籍総合データベース」（実施時期未定）
          1. メタデータをファイル転送していただくことで統合検索可能とのこと。*現在、図書館情報管理課と*メタデータのマッピングの調整を行っている。
       7. 東京大学の「東京大学総合研究博物館データベース」、「デジタル展示館」等（実施時期未定）
          1. *東京大学大学院情報学環　学環長は、前向きに検討したいとのことで、連絡あり。（10月7日）。*
          2. *東京大学大学院情報学環　学環長に正式な依頼文書を送付済み（11月）*
          3. 今後、参加の仕方、プロセスを説明するとともに、具体的な検討を開始する。
       8. 奈良女子大学附属図書館（奈良女子大学所蔵資料電子画像集）（10月27日公開）
          1. 奈良女子大学の歴史や、奈良女子大学の所蔵する貴重資料のほか、社寺等が所蔵する奈良地域に関連する貴重資料
       9. 茨城県立図書館デジタルライブラリー（11月4日公開）
          1. 茨城県の貴重な郷土資料
       10. 神戸大学付属図書館（11月25日公開予定）
           1. 「神戸大学経済経営研究所によって作成された明治末から昭和45年までの新聞切抜資料を、デジタル化したアーカイブ」
       11. 筑波大学付属図書館（12月末公開予定）
           1. マッピングの調整終了（11月20日）。現在、実装作業中。

### 歴史的音盤アーカイブの構築支援

* + 歴史的音盤アーカイブとは、SPレコードなどに記録された音源の保存と利活用を目指すものです。
  + 歴史的音盤アーカイブ推進協議会（HiRAC)が、設置され、当館もオブザーバ参加しています。
    1. 趣意としては、
       1. 文化的遺産であるSPレコード原盤が散逸・劣化し始めており、
       2. これをデジタル化して保存し、広く国民に公開することによって、日本の文化と音楽産業界の発展に寄与するものです。
    2. 対象としては
       1. SPレコード：3万タイトル（6万曲）、その他、商業用レコード以外の記録物をデジタル化して提供できるようにすることを目指しています。
    3. HiRACメンバー
       1. 日本放送協会、日本音楽著作権協会（JASRAC）、日本芸能実演家団体協議会、日本伝統文化振興財団、映像産業振興機構、日本レコード協会
  + *考えられる協力*
    1. *方針*
       1. *HiRACへのオブザーバ参加*
       2. *HiRACによる音盤デジタル化の支援*
       3. *デジタル化音源の当館への納本*
       4. *NDLデジタルアーカイブからの公開　（館内、公共図書館、インターネット（？））*
       5. *オーディオアーカイブの標準規格の策定作業（JEITA）への参画*
    2. *アーカイブ及び公開用システムの構築*
       1. *デジタル化、統合検索等に関する研究開発の実装*
       2. *著作権管理機能の研究成果の実装*

### 日中韓の国立図書館との連携

* + 10月（23～24日）には、日中韓の三国交流を開催し、デジタルアーカイブの構築と提供に関する協力として、メタデータ基準、統合的な情報サービス、長期保存に関して、意見交換しました。
    1. 交換用メタデータ仕様、交換用プロトコルに関しては、各国で大きな差はなく、統合検索サービスは、早ければ、来年中（5月）に実現できる可能性があります。
    2. また、今後の情報交換として、収集・組織化・保存及び提供機能に関する研究開発成果の情報に加えて、デジタルアーカイブに関しての制度化に関する各国の状況も情報交換することになりました。
  + メタデータ基準
    1. 保存のためのメタデータの形式は、各国の基準を尊重する。
    2. ・メタデータ交換は、DCベースとする。
    3. ・語彙等の違いは、オントロジー等による吸収を検討する
  + 統合的な情報サービス
    1. コンテンツは3館で分散して持ち、統合検索できるようにする
    2. 機械翻訳は研究開発の状況をウオッチ
    3. 統合検索のプロトコルは、ハーベスト系（OAI-PMH、RSS）、横断検索系（SRU/SOAP、OpenSearch、OpenURL、Z39.50）のどの方法も3国とも可能
    4. ３国共通のポータルを立ち上げるのではなく、各国のポータルで見せていく
       1. 早ければ、来年の5月に統合検索が実現
    5. まずは、歴史的資料から統合検索
  + 長期保存
    1. DRセンターは、まず各国内で立ち上げ、その後、３国でのDRセンターの立上げを検討する。  
       その際は、CDNLAOでのシンガポールの提案のように国際的なＤＲの方向性で検討する。
    2. ＯＡＩＳはフレームワークであり、実装は各国の開発状況により異なる
    3. 情報パッケージの仕様は、3国とも手探り状態。まだ共通化する段階にない。
    4. まだ、コンテンツの交換の候補を提示する段階にない
  + 情報交換
    1. 収集・組織化・保存及び提供機能に関する研究開発成果に関する情報
    2. デジタルアーカイブに関しての制度化に関する各国状況

### World Digital Library構想との連携

* + World Digital Libraryは、
    1. 米国議会図書館（LC）のビリントン館長が2005年6月に打ち出した｢世界電子図書館構想｣です。
       1. 2005年6月、ユネスコ米国国内委員会のユネスコ米国国内委員会で
  + 貴重書など一つしかない歴史的文化遺産を，世界各地の国立図書館等と協同でデジタル化し，公開しようとするものです。
  + 当面のアクション
    1. 現在のアクションは、World Digital Libraryのサイトに、各国が、テーマを特定したコンテンツを提供することで、来年、公開される予定です。
    2. WDLサイトに、各国がテーマを特定したコンテンツを提供
       1. 画像データ、メタデータ、解題

## まとめ

### まずは、連携による組織を越えた統合検索

　まとめとして、まずは、各機関が持っている情報を可視化して、MLA、学術、公共図書館連携、国際連携など、様々な形での連携により、各機関にデータベースが分かれていても、検索できるようにすることです。

　当館が唯一の入り口になることを目指したものではなく、様々な連携の入り口から統合的に検索できるようにすることを目指します。

### 次世代は、人と情報が関係付けられたサービスの連携

　次は、人と情報が関連付けられて、精緻な検索の絞込みと、ニーズに応じた閲覧ができるサービスを提供することです。

### 5年程度までの方向性（クラウドコンピュータの世界の普及をイメージして）

* + 過去２～３年の変化。
    1. Blog、wiki、SNS等で個人の情報が発信されるようになり、ソーシャルブックマーク機能で個人がタギングして情報を分類できるようになりました。*RSS1.0でメタデータの配信、RSS2.0やatomでコンテンツの配信まで行われるようになった。*
  + クラウドコンピュータの世界でのサービスのイメージ（図を作成）
    1. 5年後には、大きなパラダイムシフトがあると想定されます。その一つとして、クラウドコンピュータの世界があります。
    2. クラウドコンピュータは、ネットワークコンピューティング、グリッドコンピューティング、SaaS（Software as a Service）、Webサービス等の技術を融合させた、データセンターのようなものです。
       1. クラウドコンピュータ自身は、コンピュータシステムの概念ですが、システムの共通基盤（インフラ）の上で稼動することになる各サービスは、共通のサービスインタフェースを持つことが必然となっていきます。
       2. 個別の組織が、単独でコンピュータ資源（ハードウェア、ソフトウェア）や、情報を抱えて、サービスを提供し、ユーザが、個別のサービスを利用するような形ではありません。
          1. クラウドコンピュータの普及する時代には、単に情報を共有する世界ではなく、サービスを共有する世界になっていると思われます。
       3. Web2.0が2～3年で爆発的に普及したことを考えると、このような時代が、5年後には到来すると思われます。
          1. 各機関、当館のサービスも、このパラダイムシフトに対応していかなければ、利用されなくなります。
          2. そういう時代になることを予測して、情報の管理の仕方、サービスのあり方を考えていく必要があります。
  + 今後ますます、情報と情報が情報群、サービスとサービスがサービス群、人と人がユーザ群として有機的に関連付けられていくものと思われます。
    1. 従来のように検索し、コンテンツを閲覧するような形から、クラウドコンピュータの世界のリソースを、所在を意識せずに、また、インタフェースの違いを意識せずサービス群として利用していくことになると思われます。
       1. Web2.0では、パーソナライズ、人と人の関係を活用して情報を選択してきました。また、サービスも先進的な機関が提供するWebサービスを個別に利用してきました。
       2. 次世代は、それに加えて、情報と情報、サービスとサービスの関係を組み合わせて、ユーザ群と情報群の多対多の関係で、より的確な情報の活用が可能になっていきます。
    2. このような時代には、所蔵機関、専門家・一般個人の区別なく集合が作られていきます。その中では、当館のサービスは、複数のサービス群、情報群の１つにすぎません。
  + 当館は、
    1. 国の情報資源を後世に残す役割と
    2. 現在のユーザのために、当館を収集・保存した情報資源を含めて、国の情報資源を的確に利用できるようにする役割を持ちます。
  + しかし、
    1. インターネットの世界での情報発信の状況が変革すると、当館が進めているウェブアーカイブ、ポータルのサービスのあり方にも大きく影響します。*（もし対応していかなければ、利用されない孤立したサービスになってしまいます。）*
    2. 今後のウェブの世界では、情報を提供するというより、サービスを提供する時代に、どのようにして情報を収集していくのかを考える必要があります。サービスの形をそのまま保存することは困難です。
    3. また、当館のシステム構築においても、それらの仕組みを取り入れていく必要があります。
       1. *5年後に当館を入り口としたサービスを想定しても、世の中の情報探索、情報入手の入り口として使われるものになるとは思えない。*
  + 膨大な情報資源を保有する当館が、これらの変化に対応して、情報の保存と利活用の両面で、先導的な役割を果たすべきと考えますが、当館の資源（人・物・金）だけで対応することは不可能です。当館におけるシステム構築、デジタルコレクション構築の双方ともに、関係機関の協力なしでは対応できません。
    1. *（当館がすべての情報アクセスの入り口になることはない）*
    2. *（有用な情報は必ずしも図書館に集約されない。図書館員が作れる有用な解題情報、レファレンス情報はほんのわずか。一般の人の集合知は、少数の専門家の知識に勝る。図書館を中心に考えた世界では利用者ニーズに応えられない）*
    3. デジタルアーカイブシステムや統合検索システムの開発においても、当館からベンダーへの開発請負業務で行っていたり、図書館に閉じた世界でのパッケージ導入では、先導的な役割は果たせず、広く一般の利用者のニーズに応えたサービスは提供できません。
    4. システム開発においては、研究開発、技術開発の成果を活用させていただくことが重要で、それらを組み合わせて適用することが、効率的でありかつ効果的です。*（統合検索のAPIから、サービス連携のAPIによる連携を目指す）*
  + また、コレクション構築においても、当館が収集・保存できるデジタルコンテンツは、国の情報資源のごく一部です。
    1. 公共図書館、大学図書館を含めて、MLAの各機関、民間、政府、さらに個人が持つ有用な情報を「見える化」して、クラウドコンピュータの上で、各機関の持つ分散データベース、分散サービスが統合して、１つの巨大なデジタルアーカイブサービスとして利用できるようにしていかなければなりません。
    2. そのためには、まず、関係機関のサービスが相互に補完して、一つのサービスとして利用し合えるようにしていくことが重要で、それの実現に向けて、連携協力を進めていきたいと考えています。

## デジタルアーカイブの今後の構築計画（想定）

* + これは、デジタルアーカイブの今後の構築計画です。
  + 世界レベルでの電子図書館構築の構想の一翼を担うことを目指して、まずは、日本としてのデジタルコレクションを構築していきたいと思います。
    1. 関係機関の協力のもとで、人と情報が有機的に関係付けられて利用できるシステムを構築する。
    2. 当館としてのコレクションを構築する。
    3. 関係機関と協力して、国のデジタルコレクションの構築を目指す
    4. それを、関係機関と分担して、国のコレクションとして後世に残していく
  + 以上を、この線表のようなスケジュール感で進めていきたいと考えています。

## おわりに

当館は、デジタル化されていく情報を、日本の文化遺産として後世に残し、新たな知識を創出するための知識として、現在及び将来に亘って活用できるようにする役割をもっています。

この役割を果たすためには、技術の進展やニーズの変化に対応して、ユーザへの新たな形態のサービスを提供していかなければなりません。

このような取り組みに関して、今後とも、ご協力をお願いします。

以上で、ご説明を終わらせていただきます。

ご清聴ありがとうございました。

# （2009年）「トータルな図書館システムの実現」に向けて【館内調整】

平成21年3月18日

情報システム課　中山

将来の「トータルな図書館システムの実現」を目指して、数年先のスコープを見据えて次世代の図書館システムを想定し、そこまでのステップとして、次期の図書館システムを構築していく

## アプローチ

### 当館の事業戦略

* ビジョンに沿った事業戦略の策定と、資源の再配分
* 日本の情報資源をすべて収集・保存・蓄積し将来にわたって利用を保証させていくという、NDLの使命を、NDL単独で実現することは不可能。様々なレベルにおいて、関係機関との連携が必要。
* 関係機関と連携して実現に向けた全体イメージと実現に向けた実施体制、スケジュール案を策定する。

### 検討の柱

#### 当館インターネットサービスの改善

* 新技術の取り込み
* 共同研究（研究素材・課題の提供、成果の活用）

#### 商用サイトのサービスとの差別化と共存

* Googleを含めて、市場に対して影響力のある商用サービス機関との差別化と共存の方向付け

#### 関係機関と連携したコンテンツ構築

* 電子納本の促進
  + デジタル化推進戦略（全体計画・方針）の策定
  + 電子出版物流通センター構想の具体化
  + 出版者の電子書籍作成支援
  + 出版物の当館によるデジタル化
  + 出版者の電子書籍データベース立ち上げ支援
  + 当館サービスから各出版者サイトへのナビゲーション
* 媒体の種別、所蔵場所によらないトータルな統合検索・閲覧サービスの構築
  + MLA連携、日中韓連携の具体的アクション
  + レファレンス協同データベース、PORTA、総合目録の統合利用環境の構築

### 館全体システム

#### トータルな図書館サービス

* サービスの概念
  + 基本サービス
  + 付加価値サービス
* 外部連携
  + 紙とデジタルをシームレスに利用できるサービス
    - 電子図書館課と協力して推進
  + 関係機関と連携したサービスの提供
    - 公的機関だけでなく、民間企業、商用ポータルも含めて。
  + 関係機関と分担による保存
    - 各国の国立図書館、IA社、国内でのアーカイブ機関と連携。

#### 情報探索サービス

* 要求要件定義
  + 新利用者サービス2009及び書誌サービスの新展開（2009）から、情報探索に係る業務・サービス要件を洗い出し、要件を多角的に分析する。
    - 類別
      * サービスの種別、利用者層、
  + 市場動向（外部のユーザ・組織の要望・状況、商用を含めた外部のサービスの動向（トレンド））を分析する。（ニーズ）
    - 調査内容
      * インターネットサービスの5年後程度のロードマップ
      * 先進的なサービスを提供するサイトの今後の動向
      * Amazon、Google、Yahoo、YouTube、Wikipedia、はてな等
  + ニーズと技術を総合的に判断して、要件定義書を作成する。
  + 作成に当たっては、外部の有識者の参画を求め、外部からの目線での意見を尊重する。
* システム化要件定義
  + システムイメージ及びリリース計画案の作成
  + 新技術、研究開発成果の適用
    - 国の施策としての実施、研究機関と共同で。
  + 段階的なサービスの提供
    - まず、既存のサービスを順次拡充していく。
    - 並行して、拡張性のある新しいシステムとしてOSS、パッケージを利用して構築を進める。
    - 実証実験のテストベッドとなることも想定
  + 民間のサービスとの連携
    - Amazon、Google、Yahoo、YouTube、Wikipedia、はてな等のサービスの活用
  + 調査内容（実装技術）
    - 現在適用されている技術
    - 普及が見込まれる技術
    - 先進技術を駆使したパッケージ、OSSの普及見込み（ロードマップ）
    - 技術要素
      * 情報の可視化技術
      * 情報・知識の集合知化技術
      * （集合知の永久保存技術）
      * 情報の組織化・情報と人の関係づけをする技術
      * 膨大な情報から的確な情報を選択する技術
      * 利用者属性・利用環境に応じた閲覧提供技術

#### 次期基盤システム

* 従来型の図書館サービスを継承
  + 紙資料の業務システム（収集・組織化・提供（検索）、閲覧複写申込機能）としての役割の継続
  + 従来型OPACに慣れ親しんだユーザ向けの検索機能
* 追加機能
  + 検索結果から、デジタルコンテンツがある場合のナビゲーション
  + データプロバイダとして、他システムから、検索及び複写申込みができるようにするためのWebAPIの追加

#### 共通インフラ・標準化

* 世界レベルでの資源の共有を目指すための共通仕様
* 他機関サービスとの連携技術
  + クラウドコンピューティングの世界でのWebサービスAPI
    - 通信プロトコル
* 書誌（メタデータを含む）記述要素、記述規則
  + 管理用、検索用
* デジタルコンテンツ仕様
* 館内システムでの資源の共有
  + ハードウェア、ネットワーク資源の共有
  + 分散拠点でのディザスタリカバリ対策

## 当面のアクション

### IT戦略本部「３カ年緊急プラン」、21年度補正予算等も踏まえて

* 21年度補正予算候補
  + 候補の前提
    - 22年度以降に実施するものの中で、前倒しで実施できるものがないか？
    - 21年度に実施し、繰り越さないもの。後年度負担増にならないもの
  + 留意点
    - 21年度執行計画にあるものを単純に増額することはしない。

### 全体

* 電企室検討
  + 新技術の取り込み
  + 共同研究（研究素材・課題の提供、成果の活用）
  + 出版物の版下データの蓄積・保存の仕組み作り
* その他

### 全館システム

#### 資料デジタル化

* 媒体変換計画の実施計画の前倒し
* 博士論文のデジタル化　：　～100億円
* 貴重書のデジタル化 ：　～35億円
* 当館所蔵図書のデジタル化 ：　～300億円
* 当館所蔵雑誌・新聞 ：　～70億円

#### 公共図書館所蔵資料等のデジタル化

* 公共図書館所蔵資料のデジタル化支援
* 公共図書館デジタルアーカイブ及び提供システム構築支援
  + DAシステムを公共図書館用デジタルアーカイブシステムとしてOSSパッケージ化
  + ASPサービスも考慮

#### デジタルデポジット

* コンテンツ種別に特化した利便性の高い閲覧システム
  + 歴史的音源再生システム（試聴機能付）（サービスプロバイダシステムの購入）
  + 高精細画像閲覧システム

#### Webアーカイブ

* 過去、IPAが支援して作ったウェブページの収集組織化

#### 保存システム

* 集合知の永久保存
* 次世代保存メディアの利用の実証実験

### 共通インフラ

* 新規技術導入・外部システム連携強化に向けた、技術アドバイザー会議体の設置

### 情報探索サービス

#### 調査及びプロトタイプ構築

* OSS、ベンチャー企業パッケージのトライアルユース
  + 21FYにα版構築、22FYにβ版、正式版、以降エンハンス

#### 自動翻訳Webサービス

* 検索語、検索結果等を翻訳

#### 高度な検索機能及び付加価値サービス

* Google等のサービスと連携すると、より便利になるサービス
* 電子出版社サイトの統合検索

#### 検索GUIのプロトタイピング（実証実験としてβ版を開発し公開実験）　：　数千万～

* 情報の選択範囲の拡大
  + 膨大な情報から的確な情報を選択する機能
  + 多角的な絞り込み検索機能
  + レコメンデーション機能
* 情報の絞り込み
  + 知識レベル、専門領域、趣味嗜好に応じた検索機能
  + 利用機器に応じた検索機能
  + 利用場所に応じた閲覧環境

#### 組織化、API提供（実証実験としてβ版を開発し公開実験）　：　数千万～

* 情報の可視化
* 情報を全体で集合知化
* 情報の組織化、情報と人の集合知化

### 次期基盤

* 自動書誌作成支援ツール
  + 前例、目録規則による自動書誌作成
  + 自動書誌作成内容のチェック機能
* 各種書誌・メタデータ相互変換Webサービス
  + 各データ変換プログラムの作成と、汎用Webサービス化　：　数千万
* 現行システムからのデータ移行プログラム（データコンバータ）　：　数千万
* 最適化計画検証等作業：　　　　数千万円　（21年度執行計画記載案件）
  + 平成20年度に策定した次期基盤システムの将来体系について、業務・サービスに関する実現可能性検証およびその結果に基づく業務等の検討を行う。

### 現行システム

* 22FY以降の費用削減につながるもの
  + 22FY以降の調達準備
    - ドキュメント整備　：　数千万
    - 調達支援、プロジェクト管理支援の調達　：　数千万
* 来館者管理システム
  + 来館者システムとの連携機能
    - ILSパッケージのAPIの確定が先
* メールシステム設定改修　：　百万程度？
  + 霞が関WAN及び総合行政ネットワーク（LGWAN）とのメール経路を柔軟に設定できるようにするもの

## 論点

### 全体イメージ

#### 次世代図書館システムのイメージ

* 「Project Next-L」の挑戦は参考になるのではないか？
  + 世の中の進歩は、図書館よりも確実に早い
  + 今の利用者は、新しいサービスに驚かない
    - 検索エンジンとは違うインタフェースがかえって驚かれる。
    - たらい回しは嫌われる。（ナビゲーションもその一つ）
  + ALEPHは、従来型の図書館業務システムのパッケージ→新しいサービスの提供のための機能を期待してはいけない。あくまで、コストダウンのために利用する。
* 新サービスは、「情報探索サービス」の世界でトータルなサービスとして。

#### 次世代図書館システムの全体イメージ（あり方、具体的内容）を検討する組織はPMO？

#### 情報探索サービスのサービス及び実現形を検討する組織体は「情報探索サービスPJMO」？事務局は、サービス部門？それとも企画課、情報システム課？

* サービスの要求要件定義はサービス部門、システム化要件、システム開発・システム運用は、情報システム課、電子図書館課？

#### 3月末までにシステム導入計画を作れるか？

* 世の中の動向とNDLサービスの提供のスケジュール感は?

### 紙とデジタルの業務を統合する時期は？

* DAと次期基盤は、ある意味、バックオフィスの業務系システムであるが、紙とデジタルの業務の統合は、いつを目指すか？
* 業務の統合時期には、システム的な整合性も確保しておくべきと考えるが、24年1月にリリースする次期基盤システムは、28年頃まで使われることになる。早期の統合は困難になる。
* DSpaceのようなソフトでは、基盤は作れないのか？

### 情報探索サービスの部分は、市販のパッケージでいいのか？

* この部分は、NDLとして図書館界にリーダシップを発揮すべき部分と考える
  + OSSに機能を付加して、公共図書館向けに提供したいが、それは可能か？
* 研究開発成果がたくさんあり、それらの成果が順次容易に取り込めるような拡張性が必要だが、「Project Next-L」のようなOSS等は利用できないのか？
* **FRBR的な概念の実現方法は？組織化は無理。検索機能でどこまでできるか？**
  + **目録管理とPORTAと総合目録の合体を早く考えられないか？→紙とデジタルの総合目録を早期に実現（FRBR図の実現形部分）**

### 目録サービスの範囲の拡大は？

* 目録サービスとして、紙だけでなく、デジタルも含めたものにするのはいつから？
  + 全国書誌の範囲の拡大等
  + メタデータ付与の所管は、電図課？収集書誌部？
* 目録検索機能及び外部API等は、レファ協、PORTA、総目のAPI機能と合体して考えたほうがいいのではないか。

### 「情報システムユーザスキル標準」で示されたようなスキルを具備する人材の確保方法

* 体系的にスキルを習得できる場の用意。（研修、OJT）→単に情シスを経験するだけではスキルは身に付かない。身につけられるように情シスが変わらなければならない。
* 組織体制の見直し
  + 今後の業務量の試算
* 仕様調整、見積もり精査の実施姿勢
  + 開発、検収プロセスの見直し

### 「次期基盤システムの開発方針」の見直しの前提

* 何故この方針が出されたのか？
  + 今までの検討が、館全体のシステムとの整合に明確でなかった。
  + 今までの検討では、膨大な費用がかかりそうだった。
* 今回の全体イメージ及びスケジュールは24年1月に基盤システムが、パッケージを導入した次期基盤としてリリースすることを前提としたもの。
  + 28年度くらいまでこの形で運用することになるが、それでいいか？
  + どこまで仕切り直しができるか？
  + 世の中がどんどん変わり、利用者ニーズはどんどん高度化する。次期基盤は24年1月でいいのか

## 総合目録の実現方式の方向性を踏まえて

### 全体イメージの再確認

* 実現方式の概要
  + ILSパッケージとは別に、統合検索用の簡略書誌のDBとインデックスを持つ
  + NDL所蔵情報はハーベストもしくは横断検索
  + 参加館所蔵情報は、ハーベストと横断検索の両方を用意
  + 簡略書誌として、DC-NDLの項目とかなり似たものになると思われる

### 今後の検討

* バックオフィスとしてのPORTA、レファ協の統合DB、ハーベスト機能、横断検索機能の統合の可能性
* フロントエンドとしてのPORTAと総合目録を含む次世代OPACの統合の可能性⇒情報探索サービスの中で検討

### パッケージの候補

* PRIMOと同等のパッケージはOSSを含めて数多くある。
* 主要な要件は？
  + 拡張性
    - 検索関連の技術開発、研究開発は活発。それらの成果を容易に実装できること
  + ....

# （2009年）次期基盤システムの開発方針について【館内調整】

平成21年2月25日　加筆

平成21年2月20日　情推

情報システム課　中山

## 背景

* 次期基盤システムについては、平成20年10月30日の館議においてILSパッケージと次世代OPACとの組み合わせで実現することとし、パッケージとしてはAlephを採用するとの報告を行ったところであるが、現行の機能をそのまま移行するには多大な費用が必要であるとの第1次見積りが提示された（平成20年12月19日情報化推進委員会）。
* さらに、当館のシステム全体のイメージを館内で共有すべきとの指摘を受けて、平成21年1月8日の館議懇談会で「次世代図書館システムの全体イメージについて」を議題として懇談が行われ、利用者からみたサービスのあり方をより明確にすべきとの意見が出された。
* 一方、利用者サービスの側からは、新たなサービス像として「新・利用者サービス2009（仮称）」が2月5日の館議・庶務担当課長会議合同懇談会で提示されている。
* また、2月12日の館議において平成21年度の推進事業の一つとして、「トータルな図書館システムの実現」について館長により決定されたことが報告された。
* 上記の状況を踏まえ、新しい利用者サービスのあり方とシステムとの関わりを明確にし、良質で安価なシステムの構築・移行を実現するため、館全体の将来システムを見通しつつ次期基盤システムの開発方針について、今回とりまとめた。

## 現行基盤システムの問題点

「次期基盤システム導入準備作業実施計画書（1.0版）」（平成20年8月　電子図書館基盤システムPJMO）によると、現状基盤システムの問題点として次の点が挙げられている。

* 運用・保守作業や機能拡張に係る経費が大きい。予算的な観点から経費の削減が必要だが、費用削減の余地が少ない。
* 複雑な業務要件を具現化して複雑に入り組んだシステムになっているため、機能拡張への対応力が低い（業務ニーズをシステムに組み込みすぎて複雑で大きなシステムになっている ）
* **要約すると、現行基盤システムの運用コストが高くて、また、改善費用も高いので、改善できない。**

## 開発目的

運用性、柔軟性の高い次期基盤システムを構築することにより、これらの問題点を解決し、さらに、削減された資源を活用して、社会のニーズの変化に対応した新たな利用者サービスを展開できるようにする。

* **作り直して、運用コストを下げて、下げたコストで新しい利用者サービスを構築しようということ。**

## 基本方針

#### 次期基盤システムは、当館全体システムを構成する次期図書館システムの一機能として検討する。

* 実施に当たっては、館長により決定のあった「平成21年度推進事業」（平成21年2月12日館議報告）に掲げられた優先的かつ重点的に取り組む事業を各部局が中心に推進することと密接に調整を図りつつ、トータルな次期図書館システムの構築と提供の実現を目指す。
  + *「国立国会図書館業務システム最適化計画」に沿って、「業務・システム最適化の実施」を行い、「次期基盤システムの開発予算獲得」等により、確保した実行予算の中で、現行基盤システムの問題点を解消できるように「次期電子図書館基盤システムの導入」を目指す。*
  + *「書誌データの作成・提供の方針（2008）」等の方向性に沿って、「書誌サービスの改善、効果的なデータ整備の実施」を目指す。*
  + *2月5日の館議・庶務担合同懇談会で提示された「新・利用者サービス2009（仮称）」で示された今後の検討の推移を見つつ「新しい利用者サービスの実現」を目指す。*

#### 次期基盤システムは、ILSパッケージと次世代OPACで構成されるとしてきたが、付加価値サービスを提供する情報探索サービスについては、利用者サービスを提供する部署、データ作成部署が中心となった別途の体制を設置し、そこで検討することとする。

* 図１の緑の横楕円の範疇で、「情報探索サービス」として検討する。

#### 次期基盤システムは、

* ILSパッケージを核として、総合目録システムの要求機能要件も含めて検討する。
  + 実現方式及びサービスのリリース時期は今後の検討による。
* 特殊要因等での予算要求を前提とせず、当館総予算の中で捻出可能な範囲に抑える。
  + 次期基盤システムの運用経費は、現行システムの運用経費を大幅に下回ることを目指す。
* **図１の赤の縦楕円の範疇で、「次期基盤システム」として検討する。**

#### 情報探索サービスについては、

* 図書館ポータル部会の中で調整する準備を進めていたが、現在、「新・利用者サービス2009」の検討が進められており、図書館ポータル、次世代OPACとして範囲を限定せず、情報探索サービスとして体系的に整理されたサービスの構築を目指す。
* 当館資料と利用者をつなぐ書誌については、利用者の視点に立って今後の方針を明確にし、「新・利用者サービス2009」と並行して検討を進め、新たな情報探索サービスにおいて実現させる。
* **検討組織としては、図書館ポータル部会（情シス、電図課、参企）を拡大して、図協、電企室を想定していたが、新しい利用者サービスとして、主題情報部、資料提供部、関西館、書誌サービスの改善として、収集・書誌部が入り、全体システムとの整合のために企画課、情報システム課が入った組織を想定する。**
  + *一般利用者向けのサービスとして、紙、デジタルの区別なく、また当館が保有しているか否かを問わず、統合的に検索・閲覧できるサービスを提供する。*
  + *システム構築に当たっては、利用者視点で、様々な検索方法等、付加価値を付けたサービスが容易に実装できるものとする。*
* **要約すると、従来の基盤システムに総合目録管理機能を追加して、パッケージを導入して、業務改善する。**
* **新たなサービスとして検討している次世代OPACとか図書館ポータルは、紙、デジタルを問わない統合的なサービスとして検討しようということ。**
* **【参考】の図は、利用者が得たい資料を入手するためのどのように体系的に整理されていればいいかを概念で示したもの。このように組織化すべきということではなく、このように使えるような目録システムがあればいいという例の１つ。（利用者視点では、主題を用いたものなど多角的なアプローチがある）**
  + **「情報探索サービス」の目録検索機能の部分は、緑の楕円の「実現形」の部分を整理し**
  + **個別資料として、紙資料の「NDL蔵書」、「XX図書館の蔵書」は、「次期基盤システム」がサービスする部分**
  + デジタルの「NDLがデジタル化」、「NDLが収集」は、デジタルアーカイブシステム。それに加えて、「XXで公開」、、を含めて、PORTA

#### 「NDL業務・システム最適化計画」（平成20年3月）及び「情報システムに係る予算概算要求・調達実施ガイドライン」（平成21年1月21日改訂）の趣旨、方向性に則る。

## 開発・運用予算の確保

#### 現行システムの運用経費の削減

* 22年度以降の現行基盤システム運用の契約内容、作業内容、単価の再評価（センター、ハード、ミドルウェア、アプリケーション保守、ハードウェア保守、システム運用）
  + 現在の契約内容の妥当性を評価するとともに、SLAを下げるところまで削減する必要があれば、館としてサービスのあり方を合意し、運用業者と調整する。
    - *従来からの業者の総合試験によるバグ出しに職員が参画し、調達コストも職員の工数も多大になるにも関わらず、品質が上がらない開発及び改修プロセスの見直し。*

#### 次期基盤システムの導入経費の削減

* システム化要件を明確にし、業者見積もりの妥当性を的確に判断する。
  + *曖昧な要件定義では、安全を見込んで工数を上積みされる。*
  + *業者見積もりをベースに実施項目を削減するのではなく、業者見積もり工数の妥当性を的確に判断して予定工数を削減させる。*
    - *受注者は、発注者がきちんと評価することができないと思われると工数を容易に下げない。業者見積もりがそのまま関係者に提示されるような状況を改善しなければならない。*
    - *そういうことができるスキルを持った人材を確保する。職員で確保できないのであれば、支援業者に評価を求める。また、CIO補佐官、外部等に第三者評価を求める。*
  + *その上で、競争原理が働く調達にする。*
    - *特定ベンダーしか実施できない要件での随意契約は前提としない。競争原理が働かない状態での業者見積もりを削減させるには限度がある。時間的余裕、ノウハウの蓄積を理由に挙げるケースが多いが、そのために高くてもいいことにはならない。いかなる場合でも見積もり工数を精緻に評価する。（その時間を考えると入札のほうが効率的な場合もある）*
* 競争原理が働く調達を行う。
  + *特定ベンダーしか提案できないRFPにはしない。*
    - *ALEPH, Millenium だけでなく、国内ベンダー製品、OSSパッケージの組み合わせでも提案が可能な形とする。*
* 新規開発を極力なくす。
  + 必須な機能ではなくて、パッケージで実現できない業務は、業務の内容や流れを見直す。
  + 必須な機能で、何らかの開発が必要な機能は、パッケージベンダでしかできない仕様を極力なくす。
  + 必須な機能で、パッケージの設定で容易に実現できそうもないものは、他のパッケージの適用等も検討する。
    - ただ、高いからといって機能を削るのではなく、安くできる方法を考える
      * *他システムとのデータ交換用プログラムも含めて、モジュールを切り出して、利用できるようにする。（既存システム開発者に対して、コストパフォーマンス等、実現可能性を調査させる）*
      * *雑誌記事索引等、１つのパッケージで統合的に構築することが困難な場合は、他のパッケージ等の適用も考慮する。  
        （DAでは、デジタルデポジットシステムで、デジタル化もしくはポーンデジタルの雑誌記事の管理と、検索・閲覧システムを開発している。紙の雑誌記事も同様に扱えないか？）*

## 次期基盤構築の体制確保

#### 全館体制

* 21年度推進事業の中に複数部局から次期基盤に関連する推進事業があり、館全体として、連携して推進する必要がある。
  + 現行ＰＭＯもしくはＰＭＯの機能を強化した組織が核となって、「トータルな図書館システムの実現」を、推進できるようにする。
  + CIO,企画、情シスだけでなく？？？
  + 方向性、方針の決定に当たっては、館議（情報システム関連）等の場で合意を形成する。
  + 次期基盤関連の推進事業の主管部局課及び各業務の主管課によるPJMO体制とする。
  + 全体システムのあり方、その具体的内容を決める組織は、どこか？どんな体制を構築するか？
  + 全体を見渡して、遅れているところ、見検討の部分はないか。
  + これに基づいたシステム導入計画（スケジュール）は？

#### 関係部局課

* 21年度は、関係各部局課において、パッケージのトライアルユースによる機能把握、及び、要求要件を満たせることの確認をする体制の確保が必要となる。
  + *21年度当初まで、要求要件をまとめる体制が必要。*
  + *21年度、パッケージソフトのトライアルユースにおいては、業務毎の各原課が主体となって、要求要件を満足した運用が可能かどうか検証していくことになる。*
* 21年度を含め、今後、次期基盤システムの導入完了までは、業務・システムを検討するために関係部局による継続的な体制が必要である。
  + *検討の単位*
    - *現在の主管課・導入課体制、次期基盤における検討領域、及び導入パッケージシステムの機能単位を考慮した検討単位を決定し、その検討単位ごとに代表する主管課課長を長として、次の検討を行う。その配下に作業グループを設置し、実作業を行う。*
  + *業務の検討*
    - *次期基盤システムで実現する業務の検討においては、検討単位ごとに、館内合意された方向性・方針に基づく検討を行い、PMO（？）が全体調整や取りまとめを行う。*
  + *システムの検討*
    - *次期基盤システムで実現するシステムの検討においては、検討単位ごとに、館内合意された方向性・方針に基づく検討を行い、情報システム課が全体調整や取りまとめを行う。*

#### 情報システム課

* 課内の業務実施体制を再編して、要員を確保するとともに、若干名の増員を図り体制強化する。
* 業務毎の担当割
  + 全体管理　（補佐1名）
  + 情報探索サービスの開発　（主査1名、係員２名）
    - NDL統合情報検索画面（GUI）
    - 目録検索機能、外部インタフェース（API）
  + 次期基盤システムの開発　（係長級1名、係員4名）
    - 収集領域、書誌領域、利用提供領域（NDL統合情報検索と兼ねる）、物品及び資料管理領域、総合目録領域
    - データ移行・システム連携（来館者システム等）

#### PJMO支援等の調達による体制の確保

* 適正な仕様、費用での構築できるように、PJMO支援等の調達により、「情報システムユーザスキル標準」（平成18年経済産業省策定）等のスキルを具備する人材を確保する。
  + 最適な業務・システムを、効率的、効果的に構築するためには、組織内に「情報システムユーザスキル標準」（平成18年経済産業省策定）等で示されるスキルをもって実施することが有効であるが、そういうスキルをすでに持っている人材を確保することは困難である。
* **各要求部門とアウトソーシング先との橋渡し。情報システム課から見れば、ユーザは各部局課及びその職員。各部局課から情報システム課へアウトソーシングする形。**
* **スキルを持った人材を確保することは困難。しかし、スキルを持った人と一緒にOJTで実施することにより、スキルを持つことが可能。PJMO支援や非常勤職員等の登用により、そういうスキルを持った人を参画させることが有効。**
* **また、やみくもにスキルを身につけることは困難であり、「情報システムユーザースキル標準」等のガイドラインとして、ハイレベル、ミドルレベル、エントリーレベルに見合った業務を行って、ステップアップすることは可能。**
* **「情報システムユーザースキル標準」は、経済産業省が平成18年度に策定したもの。情報システム部門に限らず、組織において、ISに関わる全ての部署、人材を対象とする。**
  + **組織力強化のために利用**
  + **システムベンダー企業への発注時に利用**
  + **ISに関わる組織及び人材の役割・キャリアパスの理解のために利用**
    - **組織の観点：所属メンバーの現状のスキル把握、強化すべきスキルの把握などに活用**
    - **キャリアアップやキャリアチェンジを図るための参照モデルとして活用**
  + **組織への適用**
    - **IS機能の社内配置状況**
    - **IS機能の個人への割り当て状況**
    - **業界、業態、企業固有の知識・スキル**
  + **タスク概要**
    - **共通業務**
      * **IS機能をアウトソーシングした場合、発注者側の責務と役割として、受け入れ検査もしくはこれと同等の作業が発生する。**
      * **契約管理**

## 今年度中の作業

* 21年度PJMO支援調達の準備
* 総合目録機能の要求要件定義
  + ILSパッケージを利用した将来体系に、総合目録の機能も含める。
* 現在作成中の「次期基盤システム業務システム最適化計画」（以下、「計画書」）の案の取りまとめ
* 全体サービスイメージとの整合
  + ＰＭＯ、館議懇談会等での全館的な認識の共有化を図る。
  + 関係部局課の要求との整合性を確認する。
  + 「計画書」へ共通認識を反映させる。
* 「計画書」の第三者評価及び見直し
* 「計画書」の「将来体系」をベースにRFP作成準備
  + RFPの範囲と、要求要件の記述レベルを確認する。
* ALEPHのトライアルユースの調達及び実施準備
  + *関係部局課によるALEPHのトライアルユースのための調達及び、汎用図書館パッケージの機能を把握し、要求要件を満たせることを確認するための実施準備。*

## 今後

* 今後、本文書の方針に沿って、具体的な内容を詰める。
* なお、想定スケジュールを含めた全体イメージは、具体的な個別システムの連携、統合の実現方式、実現時期等の検討内容に応じて、より具体的な内容としてブラッシュアップする。
* 今後の方向性、方針の検討及び見直しに当たっては、館議懇談会（情報システム関連）等の場で合意を形成する。

# （2010年）国立国会図書館サーチと日韓自動翻訳の現況及び課題（日韓業務交流）

第13回日韓業務交流

業務交流Ⅱ

2010年10月14日

国立国会図書館　総務部 情報システム課長

中山　正樹

## はじめに

国立国会図書館　総務部 情報システム課長の中山です。当館の基幹の情報システムの構築及び運用を担当しています。

本日は、「国立国会図書館サーチと日韓自動翻訳の現況及び課題」と題して、2010年8月17日にプロトタイプ版を試験公開した「国立国会図書館サーチ」（通称：NDL Search）と、その中で実現しようとしている「日中韓翻訳サービス機能」の現状と課題について、お話しします。なお、このシステムは2012年1月の正式公開を目指して構築中のものであり、今後ユーザ等の評価を受けながら継続的に開発し、機能追加を行っていきます。それらの機能強化の一つとして、日韓翻訳機能の実装を進めています。

## 全体イメージ

### 国立国会図書館サーチ（NDL Search）とは

国立国会図書館サーチNDL Searchは、国内の各機関が持つ豊富な「知」をご活用いただくためのアクセスポイントとなることを目指した検索サービスです。「当館が保有しているか否かを問わず、冊子体に加えて、デジタル化された画像、テキスト、音声等の様々な形態の情報を、いつでも、どこでも、利用者が求める形で、迅速かつ的確に、アクセスまたは案内できるようにすること」を目的としています。

現在、当館が所蔵する図書、雑誌等の資料を探すことができるほか、都道府県立図書館、政令指定都市の市立図書館の蔵書、国立国会図書館デジタルアーカイブポータル（PORTA）が収録している各種のデジタル情報も探すことができます。

### 2012年1月でのサービスイメージ

国立国会図書館サーチは、従来の図書館利用者、図書館員の方だけではなく、広く一般の利用者、各種~~の~~Webサービスを提供している個人、企業・団体の方などの利用者も含め、幅広い範囲の方々に利用していただくサービスの提供を目指しています。NDL Searchの用途として、次の３つのようなケースを想定しています。

表題だけ読み上げますが、GUIを利用して一般的な検索・閲覧利用をすること。APIを利用して情報を再利用すること。そして、このシステムのサービスや機能を他の機関のシステムで再利用することです。

* 一般的な検索・閲覧（GUIでの利用）
  + 当館が所蔵する各種資料を対象とした検索をWeb上で行う（NDL-OPAC）
  + 当館が所蔵する雑誌記事を対象とした検索をWeb上で行う（雑誌記事索引）
  + 日本中の図書館が所蔵する図書を対象とした検索をWeb上で行う（公共図書館総合目録）
  + 当館および提携機関が所蔵するデジタル資料やレファレンス記録などを、図書や雑誌記事と併せてWeb上で検索する（統合検索サービス）

APIを利用した情報の再利用すること。

* 情報の再利用（APIでの利用）
  + 国立国会図書館サーチの検索結果を、利用者自身のWebサービス上で利用する（検索結果API提供機能）
  + 国立国会図書館サーチの収録データを、まとめて入手して利用する（メタデータダウンロード機能）

このシステムのサービスや機能を再利用すること。

* サービス・機能の再利用
  + 国立国会図書館サーチで開発されたシステムを使って、各機関・企業が作成したデータと併せて提供する（マッシュアップによるサービス提供支援）
  + 国立国会図書館サーチのシステムとデータを研究・開発に利用する（テストベッド環境の提供）
  + 国立国会図書館サーチのために開発されたオープンソース・ソフトウェアを利用する（図書館システムとしてオープンソフトウェア（OSS）で提供）

### 2009年度開発版の到達点

2009年度開発版システムを2010年8月に試用公開しました。公開当初から予想外に多くのアクセスがあり、レスポンスの低下を招きましたが、各種マスメディア、Twitter、ブログ、ソーシャルブックマーク等でも多く言及され、概ね好意的な評価をいただきました。検索システムとして、基本的な機能は実装されており、また、書誌情報の定期的な登録更新も行っていますので、レスポンス等が改善されれば、十分実利用に耐えうるものと認識しています。提供している主な機能とシステムのイメージは次の通りです。

主なサービスとして、これも内容は読み上げませんが、ここにあるような、基本機能、検索支援機能、グルーピング機能、絞り込み機能、再検索機能、ブックマーク機能、外部サービス連携機能があります。

* 基本機能
  + 現在、34個のデータベースから収集した約5,500万件の文献情報等を検索できます。（検索対象は、<http://iss.ndl.go.jp/information/target/> を参照）
  + 全文テキスト化された資料に関しては、書誌情報だけでなく、本文の全文検索ができます。
  + 統合検索の結果について、可能な限り入手手段を案内します。（近くの図書館、Amazon、GoogleBookSearch等へもナビゲート）
* 検索支援機能
  + 連想検索、類義語・同義語検索等を用いて検索を支援します。（あいまい検索機能）
  + 「日本語 ⇔ 中国語」「日本語 ⇔ 英語」の翻訳検索・翻訳表示ができます。（翻訳機能）
* 検索結果のグルーピング機能
  + 複数の機関で所蔵している同一の資料をまとめて表示します。（書誌同定機能）
  + 形態を異にする同一著作を隣接表示します。（著作単位でのグルーピング）
  + 検索結果は、適合度順（検索語に対する各資料の関連性が高いもの順）で排列します。
* 絞り込み機能と再検索機能
  + 資料種別、所蔵館等から絞り込み検索を行うことができます。（ファセット検索）
  + 関連キーワード等から再検索を行うことができます。（シソーラス検索、連想検索等）
* ブックマーク機能
  + 書誌情報の固定URL表示、Twitterへの投稿機能、検索結果一覧の動的RSSの配信機能等、検索結果を活用するための様々な付帯機能を利用できます。
* 外部サービス連携機能
  + このサービスを他のシステムから利用するための各種標準的なAPIが利用できます。

### システム構成

括弧２のシステムイメージとしては、別紙システム構成にありますように、

国立国会図書館サーチ（開発版）では、オープンソースの統合図書館システムであるNext-L Enju[[13]](#footnote-1)をコアに、Heritrix、Hadoop、GETAssoc[[14]](#footnote-2)、WordPressといったOSSを活用してシステムを構築しています。開発した部分のソフトウェアは、将来的にはOSSとして公開し、公共図書館等での利用に供することを想定しています。（別紙システム構成図を参照）

### 2010年度、2011年度の開発予定

試用公開したシステムについては、ユーザの意見を反映し、より使いやすいシステムに改善すると共に、今後、インターネットで普及しているサービス、実証実験等で実用化が検証された技術を積極的に採用し、順次、機能強化を行う予定です。現在想定している主な機能強化項目は以下の通りです。

* + 2010年度機能拡張開発は

検索、収集・組織化、ナビゲーション、利用者付加価値機能等、本システムの中核となる利用者共通的な機能の強化、性能改善を行っています。

* + 2011年度機能拡張開発は、

国会関係者・児童・障害者・来館者等、独自のインターフェースを必要とする利用者のための機能を開発する予定です。

## 外部機関・サービスとの連携方針

デジタルネットワーク時代に、利用者に求められるサービスと機能を持ったシステムを構築し提供するためには、外部の機関との連携協力が必須であり、当館は積極的に連携協力を行っていきます。その連携の姿勢として、次のような４つの方針を掲げています。

* + メタデータの収集または横断検索等による統合検索サービスの提供

外部機関・サービスが提供するコンテンツのメタデータを当該機関・サービスの許諾を得て収集、もしくは横断検索します。

* + 外部のウェブサービスとの連携によるサービスの提供（マッシュアップサービス）

外部で提供されている連想検索サービスや機械翻訳サービス等のウェブサービスを有機的に組み合わせて、付加価値の高い検索サービスを実現します。また、外部の情報サービスへの効果的なナビゲーションを実現することにより、利用者の情報探索を支援します。

* + コンテンツの統合利用促進のための環境整備

有用なコンテンツを保有しているにもかかわらず、データベースの構築や検索サービスの提供ができない機関に対して、データベースの構築やAPI実装等を支援します。

* + そして、研究開発、技術開発における連携として、

　利便性の高いシステム構築のためには、現状で確立した技術のみでは実現が困難です。大学の研究室、官民の研究機関、ベンチャー企業等による各種の情報技術に係る研究開発を支援するために、当館の情報資源を利用した実用化・実証実験を行うことができるよう、テストベッドの場を提供します。

## 実施中、実施予定の実証実験

研究開発、技術開発の成果を活用した実用化実証実験として、現在実施中、実施予定のものは以下の通りです。

* 県域の市町村立図書館蔵書目録の検索
* 書誌同定・集約表示の精度向上
* シングルサインオンの実現方式の検証
* 全文テキスト化・全文検索
* 日中韓翻訳機能

## 日中韓翻訳実験

### 経緯

NDL Searchの検索対象の多くは日本語の資料ですが、当館のアジア言語OPACをはじめ、統合検索対象には、韓国語、中国語の文献も含まれます。利用者が日本語で検索しても、他の言語の資料は検索・閲覧できず、その言語で検索しなければなりません。そこで、日本語を入力して~~も~~他の言語に翻訳して検索し、その検索結果も日本語に翻訳できれば、他の言語の資料の~~も~~利用も容易になると考えました。

### 2009年度システムでの実装

2009年度開発システムにおいては、情報通信研究機構(NICT)の「日中・中日言語処理技術の開発研究」(<http://www2.nict.go.jp/x/x161/project.html>)の実証実験システムを、WebAPI経由で利用させていただき、検索キーワードを日本語から中国語に変換して、中国語文献等の検索を行えるようにしました。この実証実験システムは、日中の汎用的な用例翻訳手法および科学技術文献の翻訳・情報検索用辞書を半自動作成する手法を利用したものです。また、現在のNICTの実証実験システムでは、科学技術分野以外の自然文翻訳は困難なため、書誌詳細画面での翻訳は、Googleの翻訳サービスを利用して、書誌項目を中国語から日本語に翻訳できるようにしました。

### 2010年度システムでの実装

2009年度のシステムでは、中国語翻訳機能しか実装できませんでしたが、日中韓の言語の壁を越えることは重要なことであり、韓国語に関しても同様に進めたいと考え、後になってしまいましたが、韓国語翻訳の実装も進めています。

NDL Searchでは、今後の拡張性として、複数の翻訳サービスを自由に選択できるように開発しており、現在、韓国語も翻訳できる翻訳サービスとして実績のある（株）高電社の翻訳ソフトウェア（<http://www.kodensha.jp/>）を採用することとして、追加実装を行っています。

（株）高電社は、1979年に設立され、1990年より日韓翻訳システムの開発に着手しており、現在、韓国語、中国語、英語と日本語の双方向のテキスト翻訳機能として（単語及び自然文翻訳）及びウェブページ翻訳機能として（ウェブページ内全文テキスト翻訳）を、パッケージソフトウェア及びウェブサービス（ASP）の形で提供しています。構文解析においては、ルールベースを利用して、助詞、助動詞に関わる手法など韓国語の特性に合わせた言語処理と、意味解析辞書による訳語の最適化処理が特徴となっています。基本辞書としては、韓国語は、日韓35万語、韓日27万語、中国語は、日中35万語、中日29万語を持ち、ユーザ辞書も追加可能になっており、日中翻訳で80％以上、日韓翻訳で90％以上の認識率を持つということです。地方自治体の日本語ホームページを英語・韓国語・中国語に自動的に翻訳したり、大手のポータルサイトへの翻訳サーバの導入実績があると聞いています。

これにより、韓国語や中国語の文献を日本語で検索し、検索結果を日本語に翻訳することも可能になります。また、日本語の文献を、韓国語や中国語で検索することも可能になります。

### 翻訳サービスの流れ（別紙PPT）

翻訳サービスの流れのイメージは別紙スライドのとおりです。

*NDL Searchの検索画面、書誌詳細画面、情報保有サイトの表示（目録、本文、ウェブページ）（実装準備中）において、翻訳機能を実装しています。①検索画面においてキーワードを入力して、「日本語→韓国語」「日本語→中国語」、を選択して、検索実行をします。②キーワードが、テキスト翻訳サーバにより、翻訳されて、③その翻訳されたキーワードにより、文献DBを横断検索します。④検索先から検索結果の書誌情報（原文）を得て、利用者画面に表示されます。ここで、⑤「翻訳ボタン」を押下することにより、書誌情報（原文）がテキスト翻訳サーバで翻訳され、利用者画面に書誌情報（翻訳結果）が表示されます。さらに、書誌情報（翻訳結果）にある文献ページへの「翻訳リンクボタン」を押下することにより、そのページの全文テキストを翻訳して表示されます。*

12日に、実装したばかりの日韓翻訳機能を、デモンストレーションさせていただきます。（デモ操作）

* + - 1. 検索トップページ

簡易検索窓にキーワードを入力し、詳細条件で「キーワードを韓国語に翻訳(高電社)」（赤枠部分①）をクリックして検索する。

* + - 1. 検索結果一覧画面

韓国語に翻訳されたキーワードで検索した結果が一覧に表示される。入力したキーワード(原文)とそのキーワードの翻訳結果は、検索結果一覧画面の右上で確認できる（赤枠部分②）。検索結果一覧に表示された書誌を日本語で表示させたい場合は、「日本語に翻訳」というボタンをクリックする（赤枠部分③）。

「日本語に翻訳」というボタンをクリックすると、日本語に翻訳され、「原文を表示」というボタンに切り替わる（赤枠部分④）。原文(韓国語)に戻したい場合は、「原文を表示」というボタンをクリックする。

* + - 1. 書誌詳細画面

書誌詳細画面を開くと、原文(韓国語)のまま書誌情報が表示される。「日本語に翻訳」というボタンをクリックすると、日本語に翻訳される(赤枠部分⑤)。

* + - 1. 提供元データベースの書誌詳細画面

「見る・借りる」に提供元データベースの書誌詳細画面へのリンクが表示される。提供元データベース名をクリックすると、原文のままの提供元データベースの書誌詳細画面が表示されるが、（日本語翻訳を通して閲覧）というリンクをクリックすると、日本語に翻訳された提供元データベースの書誌詳細画面が表示される(赤線部分⑦)。

### 日中韓三か国による検索、翻訳の実験（提案）

8月に締結された「日中韓電子図書館イニシアチブ協定」で掲げられた目標の実現策として、

* NLC、NDL及びNLKは、メタデータ・スキーマの標準化、情報サービス（ポータル）の統合及び電子情報への長期のアクセスを可能にするための共同開発を促進する。
* 各図書館が有するそれぞれのポータルの相互運用性を高めることにより、三者の統合的情報サービスの第一段階を実現する。

ここに記載のように掲げられましたが、この合意に基づく最初の共同実験として、日本語、中国語、韓国語の言語の壁を超えた実験を提案させていただいています。

NDL Searchでは、現在、日本国内の統合検索先が保有している韓国語、中国語の資料の翻訳検索を行なえるようにしようとしているところですが、今後、双方向での検索・閲覧を実現することを視野に、まずは、翻訳機能の実用性の検証のために、NLK、NLCの蔵書目録等を検索・閲覧対象としてさせていただきたいと考え、下記のような提案をさせていただいております。

* NDL Searchから、NLK、NLCの蔵書目録を横断検索させていただきたい。
* そのために、利用可能な通信プロトコル（SRU/SRWまたはZ39.50等）の種別と、アクセスのためのURL等の情報を教えていただきたい。

6月から、NLC、NLKに対して、それぞれ担当者レベルで具体的な調整をさせていただいています。

次のステップとしては、韓国、中国で実験的に利用可能な翻訳システムがあれば同様に実装させていただくこと、また、翻訳精度をより高めるために、三か国協力して、用例集、用語の語彙の充実に取り組みたいと思っています。

この成果を各国のポータルシステムに実装することにより、言語の壁を越えて、各国が保有している情報を相互に検索・閲覧できることを目指したいと思います。

## ポータルの相互運用性を高めるために

NDL Searchは、まずは、日本国内の様々な機関が保有する情報資源を、「いつでも、どこでも、誰でも、所蔵機関や媒体の形態を問わず、「情報」を閲覧もしくはナビゲーションする」ことを目指しています。そこに、翻訳機能が実用化されれば、「言語を問わず」も目指すことができます。

今後はさらに、各図書館の資料が同一もしくは類似の資料と判断できるように、交換されるメタデータの記述要素、記述規則の共通化や、類義語を把握するシソーラス、語彙の違いを吸収するオントロジー等の言語や表現の差異を吸収する技術を相互交換して適用することにより、より的確な検索が可能になるようにしたいと考えています。これにより、「三者の統合的情報サービスの第一段階の実現」に近づけるものと考えています。

## 情報の利活用の促進を目指して（課題と今後）

最後に、利用者の情報探索の目的は、問題・課題の解決であり、回答が掲載された資料の所在ではなく、回答そのものを知識として得ることです。現在では、多くの利用者がインターネットで閲覧できる情報だけで問題を解決しつつあり、Googleなどの検索エンジンで見えない情報はないも同然と言われています。こうした情報環境の中で、知識・情報の利活用を促進するためには、紙の形態で流通している資料もデジタル化し、内容をインターネット上で検索・閲覧できるようにする（可視化）ことが必要です。

今後の課題として、従来の単なる情報探索から、事実としての知識検索へ進化させ、知識の再利用による新たな知識の創造が求められています。それを実現するためには、単に資料の内容を可視化して集積するだけでなく、個別の情報に意味的にタグ付け（自動組織化）し、知識として相互に関連付けて（自動集合知化）、利用者が求める知識として、より的確に取り出せるようにすること、また、知識として有効に活用するために、情報の信頼性を確保することが必要です。

「可視化」において翻訳により言語差異を吸収し、また、「組織化」、「集合知化」等において情報の記述要素、記述規則等の差異を吸収することができれば、「国の文化・科学遺産への容易なインターネット・アクセスを提供すること、豊かな多文化・多言語コンテンツを一般の人々に享受してもらうこと、学術世界へ貢献することを目的とする。」という日中韓電子図書館イニシアチブ協定の目標の達成の一翼を担えると考えています。

# （2011年）電子図書館構想と日本の学術デジタルコミュニケーションの現状（e-Japanologyの構築に向けて）

東京外国語大学国際日本研究センター

日本語・日本学研究 Vol1 (2011)

国立国会図書館　総務部 情報システム課長

中山　正樹

## はじめに

「電子図書館構想と日本の学術デジタルコミュニケーションの現状」と題して、当館の電子図書館のあゆみと、学術関係を含めた知識インフラの構築と提供に向けた状況を報告する。

まずはNDLの役割について、納本制度の下で国内の刊行物を網羅的に収集する当館が、国民に期待されていることは、印刷物＝「物」としてではなく、「情報」として、網羅的に収集し、時代を超えて保存し、提供していくことではないか。当館の使命は、利用者が必要とする情報を確実に提供することであって、当館が保有している情報だけを提供すればいいというものではないと考える。

当館が保有している情報は利用しやすく、また、保有していない情報へは、所在場所へ確実に案内することが求められている。このような使命を果たすために、様々な施策を講じている。

## 電子図書館構想

当館は、1994年ころから電子図書館構築に向けた活動を行ってきている。1995年にはパイロット電子図書館プロジェクトでの実証実験を実施した。2002年、電子図書館サービスを中心的な機能に置く関西館が開館し、本格的な電子図書館事業を開始し、以降、「近代デジタルライブラリー」の公開、インターネット資源の選択的収集事業（WARP）、データベースナビゲーション事業（Dnavi）、各種の電子展示会の開催など各種の新規サービスを公開・提供し、全国の各種デジタル情報資源を一元的に検索し、ナビゲーションを可能とするポータルシステム（PORTA）に着手するなど、日本全国のデジタル資源の連携を視野に入れたサービスを始めた。

現在は、さらに飛躍する時期であると考える。その方向性として、国内においては館種を問わない全国の図書館との連携の強化と、図書館に限らず広く博物館や文書館などのデジタルの文化・歴史遺産をもつ機関との連携の強化が考えられる。また、海外に対しても、東アジアの日中韓3カ国の連携や、世界各国とのグローバルな協力の推進も視野に収める。また、デジタル化対象コンテンツについては、これまで主として資料の画像イメージを作成し蓄積してきたが、画像データベースは継続して強化していくとともに、これまで不十分であったテキスト化した資料のデータベース化も、重要な課題となる。さらに、音楽・映像データ等、文書記録のデジタル化に留まらない別のカテゴリーの資料のデジタル化も、視野に入ってくるものと考えられる。

### 当館蔵書のデジタル化

蔵書をデジタル化してインターネットで提供できる資料を増やすために努力してきたが、2009年以降、新たな枠組みとして、利用による資料損傷を避けるために保存を目的とするデジタル化を開始した。なお、2010年1月から施行された新著作権法では、当館が資料保存の目的でデジタル化することを、法律上明確にするように法改正が行われた。また、2009年に国の経済対策の一環として127億円の補正予算が計上され、また2010年にも10億円が計上された。これにより、2011年3月末の想定で、当館所蔵資料915万冊のうち、1/4にあたる230万冊がデジタル化される。

### 学位論文のデジタル化

学位論文電子化の諸問題に関するワーキンググループ中間報告（2008年3月）に基づいて、国公私立大学図書館協力委員会との協議により、学位授与日が2001年3月31日以前の学位論文を、当館が遡及してデジタル化し、デジタル化する学位論文は、著作権処理を行った上でインターネット公開することが確認された。

## 知識インフラの構築と提供に向けて

日本の学術デジタルコミュニケーションの現状については、学術情報を含めた知識インフラとして、その構築に向けた状況を報告する。

### 知識インフラの必要性

我が国の第4期科学技術基本計画に係る「科学技術基本政策策定の基本方針」（2010年7月16日総合科学技術会議報告）でも、文献等研究情報のデジタル化、オープンアクセスの推進等とともに、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」とされている。 知識インフラは、科学技術研究活動の実践を根本で支え、科学、技術、学術、文化活動によって生み出される多様なデータ、情報、文献、知識を開放し、それらへの迅速で適切なアクセスを可能にすることで、次の研究、開発、教育、その他の社会的・文化的実践へとつなげる動的サイクルを形成することを目的としている。つまり、情報の生産→流通→アクセス→再生産という知識の循環を促進するネットワーク、プラットフォームとなることを目指すものである。 これにより、組織や個別学術分野を越えた知識の融合を可能とし、学際的な新しい知識やイノベーションの創造を容易にする。

### 近い将来に取り組むべき事項

近い将来に取り組むべき事項として、７つ掲げている。 （ここの内容はスライドを参照。）

* + - 1. 国内学術出版物のデジタル化と電子情報資源の収集
      2. デジタル化のための環境整備
      3. 電子情報資源の管理・保存
      4. 電子情報資源の利活用の促進
      5. 従来の所蔵資料・サービスと電子情報資源との有機的連携
      6. 利用情報の解析と利活用
      7. 知識インフラのノードとしての社会的な機能の展開

## 知識インフラのノードとしての社会的な機能の展開

知識インフラの課題に対して取り組むべき事項の実現形の１つが、利用者の窓口として想定している「国立国会図書館サーチ（開発版）」（NDL Search）である。

### NDL Searchとは

NDL Searchは、当館をはじめ、全国の公共図書館、公文書館、美術館や学術研究機関等が持つ豊富な「知」をご活用いただくためのアクセスポイントとなることを目指している。2010年8月に試行公開し、2012年1月から本格システムとして稼働する。

「当館が保有しているか否かを問わず、冊子体に加えて、デジタル化された画像、テキスト、音声等の様々な形態の情報を、いつでも、どこでも、利用者が求める形で、迅速かつ的確に、アクセスまたは案内できるようにすること」を目的としている。

### NDL Searchが当面目指す方向性

NDL Searchが当面目指す方向性としては、冊子体資料、デジタルコンテンツ、レファレンス情報の３つのシステムが統合されるイメージで、機能としては、全ての情報の統合し、先進的な検索機能を提供し、利用目的に応じてあらかじめ情報を選別できるようにする。また、直接サービスと、APIを利用した間接サービスサービスが提供されることを目指す。

### NDL Searchの将来像

NDL Searchの将来像は、クラウドの世界で各組織体のサービスが連携して利用できるようになることである。コレクション構築、サービスの構築、サービスの提供のどれも、当館単独だけではなし得ず、関係機関の協力と連携が不可欠である。 また、今後のサービストレンドを踏まえたサービスの構築は、関係機関による研究開発、技術開発の成果を活用させていただかなければ、実現が困難である。

コレクション構築においては、2009年７月に国立国会図書館法が改正になり、国・地方自治体、国立大学、独立行政法人等の機関のウェブを収集・保存し提供できるようになったが、それだけでも、すべてを当館だけで保存することは困難である。同種のコレクションを保有する公共図書館、大学図書館を含めて、博物館、美術館、公文書館等の各機関、政府機関、民間、さらに個人が持つ有用な情報が、クラウドの世界でサービス連携して、１つの巨大なデータベースとして利用できるようにすることが必要である。

### NDL Searchのサービスイメージ

2012年1月正式稼働時には、以下のようなサービスを実現する。

* + - 1. 所蔵機関、情報種別を問わない統合検索機能の提供

所蔵機関、情報種別を問わず、データベースを統合検索（書誌、目次及び全文テキストからの検索）し、同一著作物、同一資料をグルーピング表示する。また、日中韓の国立図書館の統合検索（翻訳検索・翻訳表示）できるようにする。

* + - 1. コンテンツの閲覧及びナビゲーションを容易にする機能の提供

情報の入手手段へナビゲーション（デジタルコンテンツ、所蔵機関、オンライン書店、近くの図書館へ案内）する。また、携帯電話、スマートフォン、タブレット端末等、閲覧デバイスに最適化したユーザインタフェースを提供する。

* + - 1. ユーザビリティを向上させた検索機能の提供

あいまい検索（表記ゆれ、キーワードサジェスト）、絞り込み検索（ファセット検索）及び再検索機能（連想キーワード検索、シソーラス検索、キーワードレコメンド）を可能にする。

* + - 1. 情報及びサービスの再利用のための機能の提供

納本された新着資料の書誌情報の早期提供（MARC及びDCメタデータ形式）し、また、検索結果のRSS情報の提供、ブックマーク機能、Twitter、書棚サイトへの投稿も容易に行えるようにする。

さらに、外部システムが本サービスを利用するための連携機能の提供（ハーベスト及び横断検索API）を実装する。

### NDL Searchでの連携協力イメージ

デジタルネットワーク時代に、利用者に求められるサービスと機能を持ったシステムを構築し提供するためには、外部の機関との連携協力が必須であり、当館は積極的に連携協力を行っていく。その連携の姿勢として、次のような方針を掲げている。

* + - 1. メタデータの収集または横断検索等による統合検索サービスの提供

外部機関・サービスが提供するコンテンツのメタデータを当該機関・サービスの許諾を得て収集、もしくは横断検索する。

* + - 1. 外部のウェブサービスとの連携によるサービスの提供（マッシュアップサービス）

外部で提供されている連想検索サービスや機械翻訳サービス等のウェブサービスを有機的に組み合わせて、付加価値の高い検索サービスを実現する。また、外部の情報サービスへの効果的なナビゲーションを実現することにより、利用者の情報探索を支援する。

* + - 1. 研究開発、技術開発における連携

利便性の高いシステム構築のためには、現状で確立した技術のみでは実現が困難である。大学の研究室、官民の研究機関、ベンチャー企業等による各種の情報技術に係る研究開発を支援するために、当館の情報資源を利用した実用化・実証実験を行うことができるよう、テストベッドの場を提供する。

* + - 1. コンテンツの統合利用促進のための環境整備

有用なコンテンツを保有しているにもかかわらず、データベースの構築や検索サービスの提供ができない機関に対して、データベースの構築やAPI実装を支援する。

## 知識の利活用の促進に向けて

利用者の情報探索の目的は、問題・課題の解決であり、回答が掲載された資料の所在ではなく、回答そのものを知識として得ることである。

現在では、多くの利用者がインターネットで閲覧できる情報だけで問題を解決しつつあり、Googleなどの検索エンジンで見えない情報はないも同然と言われている。

こうした情報環境の中で、知識・情報の利活用を促進するためには、紙の形態で流通している資料もデジタル化し、内容をインターネット上で検索・閲覧できるようにする（可視化）ことが必要である。

今後の課題として、従来の単なる情報探索から、事実としての知識検索へ進化させ、知識の再利用による新たな知識の創造が求められている。それを実現するためには、単に資料の内容を可視化して集積するだけでなく、個別の情報に意味的にタグ付け（自動組織化）し、知識として相互に関連付けて（自動集合知化）、利用者が求める知識として、より的確に取り出せるようにすること、また、知識として有効に活用するために、情報の信頼性を確保することが必要である。

「可視化」において翻訳により言語差異を吸収し、また、「組織化」、「集合知化」等において情報の記述要素、記述規則等の差異を吸収することができれば、各機関が持つ知識が意味的に関連を持って利用でき、知識の利活用の促進を目指した「知識インフラ」の構築という目標の達成の一翼を担えると考えている。

# （2011年）知の共有を目指して（国立国会図書館におけるデジタルアーカイブ構築）

「情報管理54(11), 715-724」（科学技術振興機構2012年2月）をベースに追加訂正。

http://dx.doi.org/10.1241/johokanri.54.715

## 要約

本稿では、国立国会図書館（NDL）のデジタルアーカイブの構築の現状と、「知識の共有化」が目指す「新たな知識の創造と還流」に向けた活動の方向性について述べる。NDLは、納本図書館として、冊子体資料だけでなく、政府系インターネット情報等のデジタルコンテンツを含めて収集保存する責務を持っており、それらをいつでもどこでも利用できるようにすることが望ましい。NDLは、あらゆる資料や情報を可能な限り収集し保存し、NDLデジタルアーカイブを構築する。しかしながら全てを収集することは不可能であるので、他の機関と併せて網羅的な知識の蓄積を図り、分散デジタルアーカイブを構築する。それらNDLが収集できていないものも含めて、分散したデジタルアーカイブの情報を一元的にナビゲートし、かつ、意味的に関連付けて知識として利用できるようデータプロバイダの役割を果たすNDLSearchを構築する。このような既存の情報を知識として再利用して新たな知識の創造を可能にする知識インフラを構築する。

This article shows the current situation of construction of the National Diet Library(NDL)'s digital archive and the direction of activities for “creating and reproducing new knowledge” which “knowledge sharing” aims at. The NDL、 as a deposit library has a responsibility to acquire and preserve not only paper materials but also digital contents、 and needs to make them accessible anytime and anywhere. The NDL acquires and preserves digital contents to a maximum extent to build the NDL digital archive. However、 it is impossible to collect all of them、 so the NDL plans to accumulate all knowledge in cooperation with other institutions to build distributed digital archives. The NDL plays a data provider role to navigate all the distributed digital archival information and make it available as semantically-related knowledge to build “NDL Search”. The NDL will make it possible to create new knowledge reusing the existing information as knowledge to build up the knowledge infrastructure.

## キーワード

国立国会図書館、デジタルアーカイブ、知識インフラ、国立国会図書館サーチ、Next-L Enju、Hadoop、メタデータ、ダブリン・コア、Webサービス、セマンティックウェブ

National Diet Library、 digital archive、 knowledge infrastructure、 NDLSearch、 Next-L Enju、 Hadoop、 meta data、 Dublin Core、 Web service、 Semantic Web

## はじめに

NDLは、日本の唯一の国立図書館であり、国の中央図書館として位置づけられている。

NDLは、納本制度の下で、国内の刊行物を網羅的に収集して提供する使命を持っているが、刊行物が、冊子体からデジタル情報資源にシフトしていく中で、すべての情報を収集することは不可能である。また、歴史的な刊行物は、必ずしも図書館にはなく、公文書館、博物館、図書館等で、分散所蔵している。

このような現実で、利用者はNDLが全ての図書を所蔵することを期待している。その期待に応えるためにも、民間データベース、大学・研究機関等が持つ学術情報、データベース、美術館・博物館・図書館・文書館の連携（MLA連携）、海外の図書館等とのサービス連携が必須な時代になっている。また、高度化する利用者ニーズに対応するためには、先進技術を活用して、サービスを提供する使命があり、そのためには、国や民間が持つ先進の技術の移転が必要となる。

NDLは、保存図書館として、情報資源を保存する立場での管理的指向から、利用者指向でサービスを提供する方向へ移行していかなければならない。

従来の国の機関の枠に捉われないサービスの提供に当たっては、効率的・効果的に実現するために、業務・システムの構築・運用も効率的に行われなければならない。

システムの開発に当たっても、サービスの提供に当たっても、機械的な作業は可能な限り省力化し、人は、より知的な活動へシフトしていく。そして、デジタル情報と情報技術の活用によって、より創造的な社会が築かれることを目指す。

## NDLの役割

NDLは、「知識はわれらを豊かにする」というビジョン1）の中で、「日本の知的活動の所産を網羅的に収集し、国民の共有資源として保存する」「利用者が求める情報への迅速で的確なアクセスまたは案内をできるようにする」「来館者がどこにいても、来館者と同様のサービスが受けられるように努める」という柱を掲げている。

NDLの使命は、利用者が必要とする情報を確実に提供することであって、NDLが保有している情報だけを提供すればいいというものではない。NDLが保有している情報は利用しやすく、また、保有していない情報へは、所在場所へ確実に案内することが求められている。このような使命を果たすべく、資料のデジタル化及び情報の収集・保存そして提供を行うサービスを構築している。

## 電子図書館構築のあゆみ

NDLは、1994年ころから電子図書館構築に向けた活動を行ってきている。1995年にはパイロット電子図書館プロジェクトでの実証実験を実施した。2002年、電子図書館サービスを中心的な機能とする関西館が開館し、本格的な電子図書館事業を開始した。以降、「近代デジタルライブラリー」の公開、インターネット資源の選択的収集事業（WARP）、データベースナビゲーション事業（Dnavi）、各種の電子展示会の開催など各種の新規サービスを公開・提供している。2007年に、全国の各種デジタル情報資源を一元的に検索し、ナビゲーションを可能とする国立国会図書館デジタルアーカイブポータル（PORTA）に着手するなど、日本全国のデジタル資源の連携を視野に入れたサービスを始め、館種を問わない全国の図書館との連携の強化と、図書館に限らず広く博物館や文書館などのデジタルの文化・歴史遺産をもつ機関との連携の強化を目指した。

2009年より、広範囲にデジタルコンテンツを収集、所蔵資料のデジタル化を行い、ウェブ情報の制度的収集も始まった。また、冊子体資料もデジタル情報も併せて網羅的に検索・案内する情報探索サービスとして、国立国会図書館サーチ（NDLSearch）の開発と提供を目指した。

## 最新の取組み

NDLが新しい利用者サービスを実現するために取り組んでいることは、大きく2つある。1つ目は「業務・システム最適化の推進」であり、2つ目は「デジタル情報資源の収集・蓄積・利用の推進」である。従来型のシステムの機能強化・運用に係る膨大な人物金の資源を削減して、削減した資源で新たなサービスシステムを構築し提供するということである。

### 業務・システム最適化の推進

NDLはサービスの改善を目指して、トータルな図書館システムと新しいサービスの実現に向けて活動している。2008年に「国立国会図書館業務・システム最適化計画」2）を策定し、利用者本位のサービスと効率的な業務遂行を可能とするよう、費用対効果の高い情報システムの実現に取り組んでいる。 また、情報システム資源の無駄を省き、利用者の利便性を高めるため、インタフェースやデータ構造の標準化、ハードウェア、ソフトウェア資源の共有化、さらには類似の機能やサービスを提供する情報システムの統廃合を進め、併せて業務の見直しを行うこととしている。サービスは、システムとそれを使った業務により実現するもので、情報システムは、サービスを効率的・効果的に実現するためのものと位置づけている。

情報システムは、サービスを効率的に遂行することを支援するもので、情報システムでできないことを、業務により人が行うものとしている。システムの構築に当たっては、個別サービスの最適化ではなく、国全体、NDL全体の最適化の観点で実施すべき要件を明確化し、他機関と連携分担して構築する。

#### 業務基盤システムの構築

現行の電子図書館基盤システムの次期システムとして、本年1月にNDL所蔵資料の統合書誌データベースシステムとNDL-OPACの機能を担う業務基盤システムが運用を開始した。

現行の電子図書館基盤システムは、NDLの独自仕様で開発・運用しているもので、運用コストや、柔軟性と相互運用性の面で課題となっていた。次期システムではこれらの課題を解決するため、業務を見直して、図書館パッケージソフトを導入し、最低限の仕様変更や外付け開発を行った。

#### デジタルアーカイブの構築

NDLのもう1つの基幹システムであるデジタルアーカイブシステムは、インターネット情報をサイト単位で収集・保存するシステムと、デジタルコンテンツを著作別で収集・保存するシステムと、大規模なファイルシステムを保有する電子書庫3つで構成される。NDLが個別データベースとして公開していた、貴重書データベース、近代デジタルライブラリー、児童書デジタルアーカイブを統合して、一元的に管理できるようにして、効率化を進めている。

#### 情報探索サービス（NDLSearch）の提供

3つ目の新サービスである情報探索サービスシステムは、業務基盤システムの稼働に合わせて、本年1月に本格稼働した。PORTA、国立国会図書館総合目録ネットワーク（ゆにかねっと）を統合し、NDL-OPACを統合検索することにより、システムの最適化を図り、かつ、従来のOPACでは提供できていない次世代OPAC機能を実装して、利用者サービスの向上を図った。NDLSearchの概要は、後掲する。

### デジタル情報資源の収集蓄積・利用の推進

NDLの現在の主な取組みの2つ目は、「デジタル情報資源の収集・蓄積・利用の推進」である。

#### インターネット資料の制度的収集

2009年7月に国立国会図書館法と著作権法が改正され3)、国、地方公共団体、国立大学等の公的機関が発信するインターネット資料について、著作権者の許諾なく収集できるようになった。 この法改正に基づき、NDLは、2010年4月から、公的機関のウェブサイトの網羅的な収集を開始しました。

#### 民間の「オンライン資料」の収集制度化

2010年6月に納本制度審議会から提出された答申4）では、電子書籍、電子雑誌など従来の図書、雑誌に相当し、ネットワーク上を流通する情報を「オンライン資料」と定義し、NDLが法的に収集できる仕組みを整備すべきとされた。この答申を受けて、現在、「オンライン資料」の収集に係る制度設計の検討を行っている。利用者が書籍に触れる機会が増えることによりニーズを喚起し、出版市場が拡大することにより、出版界が発展する枠組みを目指す。出版界、著作権者等の関係団体・機関との協議を経て、2012年以内には制度化を実現したいと考えている。

#### 所蔵資料のデジタル化

2009年に著作権法が改正5）され、2010年1月に施行されたことに伴い、制度上重要な変更があった。

1つは、国立国会図書館においては原本保存のため、著作権法第31条第2項の規定により著作権者の許諾を得ないで所蔵資料をデジタル化することが認められたことである。2009年度補正予算で、NDL所蔵資料のデジタル化経費として約127億円が計上され、デジタル化作業を実施した。貴重書を含む古典籍資料約7万冊、和図書約90万冊（1968年以前刊行分）、和雑誌約114万冊（2000年受入分まで）、博士論文約14万冊（1991年度～2000年度受入分）を対象とした。デジタル化は、画像データのみだが、目次のあるものは目次のテキストデータを作成している。 デジタル化した資料の利用については、NDL内での提供が基本だが、和図書のうち、戦前期刊行のもの及び博士論文は、著作権処理を行った上で、インターネットで提供している。

改正著作権法に基づきデジタル化した資料の原本は、利用に供さないことが条件のため、従来行っている公共図書館等への図書館間貸出ができないことになる。現在、デジタルデータを用いた代替案について、出版社、著作権者の関係団体等と協議している。

2つ目は、障害者の情報利用の機会確保のため、国立国会図書館、公共図書館、大学図書館、学校図書館等において、著作権者の許諾を得ないでテキストデータ、拡大図書などを作成し、障害者に提供できるようになったことである。関係機関等から障害者サービスにデジタル化資料を活用することが求められている。方法としては、作成した画像データからOCR等により全文テキストデータを作成のうえ、NDLの館内、他の公共図書館もしくは障害者の使用する専用端末に配信し、読み上げソフト等を用いて提供する方式が想定される。

#### 障害者向け及び本文検索のためのテキスト化

また、2011年4月に、出版関係者や有識者でつくる文化庁の「電子書籍の利用と流通の円滑化に関する検討会議」6）は、NDLがデジタル化した蔵書のうち、市場で入手困難な出版物について、全国の公立図書館や大学図書館などに画像データを配信し、閲覧できる仕組みを整備すべきであるとしている。　　また、さらなる利用者の利便性の向上を図るために、本文検索サービスの提供が必要であり、検索サービスとして利用するためのテキスト化であれば、権利者の利益を不当に害さないとの中間報告がなされた。

本文テキスト化に当たっては、視覚障害者向けの読上げに対応すること、本文の記述部分の章・節・項単位で的確に検索できるようにするために全文インデックスが作れること、さらに、電子出版社等が、DRM等を施して著作者の権利を保護した形で電子書籍ビューア等で閲覧できるようにすることが必要である。

視覚障害者向けのドキュメントフォーマットとして、DAISY4に準拠し、テキスト部分に関してはEPUB3.0の仕様に従った日本語書籍基本テンプレート（JBasic）7）というマークアップ指針も策定が進んでいる。

論文等に関しては、科学技術振興機構（JST）が運営する科学技術情報発信・流通総合システム（J-STAGE）8）で、論文データを多言語対応の論文用XML DTDを適用することになっている。

閲覧用コンテンツに関しては、縦長の冊子体形式を、そのまま閲覧できるタブレット端末の普及が見込まれる状況で、PC、スマートフォン、タブレット端末等で快適に検索・閲覧できる仕様への対応を進めている。

## NDL Search

NDLSearchは、NDLが利用者とコンテンツを結ぶ基本動線であり、本年1月から正式に稼働した。

### NDL Searchが当面目指す方向性

NDL Searchは、国内の各機関が持つ豊富な「知」を活用するためのアクセスポイントとなることを目指した検索サービスである。「国立国会図書館が保有しているか否かを問わず、冊子体に加えて、デジタル化された画像、テキスト、音声等の様々な形態の情報を、いつでも、どこでも、利用者が求める形で、迅速かつ的確に、アクセス又は案内できるようにすること」を目的としている。

書誌情報だけでなく、資料の全文も検索の対象とするなど、より掘り下げた検索を可能としている。最終的には、利用者が必要とする事実や情報・知識そのものを検索できるようにすることを目指している。

### NDLSearchのサービスイメージ

NDLSearch は、以下のようなサービスを実現する。9)

#### 所蔵機関、情報種別を問わない統合検索機能の提供

所蔵機関、情報種別を問わず、データベースを統合検索（書誌、目次及び全文テキストからの検索）し、同一著作物、同一資料をグルーピング表示する。また、日中韓の国立図書館の統合検索（翻訳検索・翻訳表示）できるようにする。

#### コンテンツの閲覧及びナビゲーションを容易にする機能の提供

情報の入手手段へナビゲーション（デジタルコンテンツ、所蔵機関、オンライン書店、近くの図書館へ案内）する。また、携帯電話、スマートフォン、タブレット端末等、閲覧デバイスに最適化したユーザインタフェースを提供する。

#### ユーザビリティを向上させた検索機能の提供

あいまい検索（表記ゆれ、キーワードサジェスト）、絞り込み検索（ファセット検索）及び再検索機能（連想キーワード検索、シソーラス検索、キーワードレコメンド）を可能にする。

#### 情報及びサービスの再利用のための機能の提供

納本された新着資料の書誌情報の早期提供（MARC及びDCメタデータ形式）し、また、検索結果のRSS情報の提供、ブックマーク機能、Twitter、本棚サイトへの投稿も容易に行えるようにする。

#### 他機関サービス向けのAPIの提供

他機関のシステムが、このサービスをデータプロバイダとして利用するために、国際標準等として普及しているハーベスト及び横断検索用の各種APIによる連携機能を実装している。

### システム構成

NDLSearchは、オープンソースソフトウェア（OSS）や外部のウェブサービスを積極的に活用して、開発することを基本方針としている。ベースとなるシステムには、日本で開発されたオープンソースの統合図書館システムであるNext-L Enjuを採用し、そのほか、ウェブの収集ロボットであるHeritrix、分散ファイルシステムであるHadoop、連想検索エンジンであるGETAssoc、CMSであるWordPressといったOSSを活用してシステムを構築している。開発したソフトウェアも、将来的にはOSSとして公開し、公共図書館等での利用に供することを想定している。

## 知識インフラの構築

NDLSearch、NDLデジタルアーカイブの発展形として、今後取り組もうとしている計画は、「知識インフラの構築」である。

### 経緯

国の総合科学技術会議が決定した「第4期科学技術基本計画」（2011年8月19日閣議決定）10）においては、研究情報基盤の整備を推進することとし、推進方策が示された。また、NDLでは、これに先立ち、「国立国会図書館における今後の科学技術情報整備の基本方針に関する提言」（第52回科学技術関係資料整備審議会（2011年1月19日））を受けて「第三期科学技術情報整備基本計画」11)を定め、「知識インフラ」の構築を進めることとした。

### 知識インフラとは

知識インフラとは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を知識として集約し、新たな知識の創造を促進させるもの。その知識をさらに集積・流通・活用と創造するサイクルの構築が必要である。

新たな知識の創造のためには、分野を越えた知識の関連付けが必要であり、日本中に散在するコンテンツの所在を集中管理し、そこに検索をかければ、関連する全ての必要なコンテンツが得られるようにする。知識は関連するものが有機的に結合され、ネットワーク的に統合化されたものであり、単に情報を集めたものではない。日本中にある芸術を含んだあらゆる学問・研究のコンテンツ、研究ツール、社会状況データ等が知識の形に組織化され、これらの知識・情報が公開され、全ての人が共有できることが大切である12)。

＊所在場所、媒体形式を意識にした「書籍、文献単位の提供」から、所在場所、媒体形式を問わず、「情報・知識・物語の単位での提供」へ。

＊メタデータの収集、探索、提供ではなく、情報の内容で、情報を関連付けし、提供する。

＊人間頭脳の持つ知識とその活用の機能にできるだけ近いサービスの提供。

＊ユーザの意図した情報がより的確に取り出せる「知識検索サービス」。

＊より安心して利用できる検索システム。キーワード、曖昧検索を駆使した「情報探索」から、信頼性を評価した「事実検索」へ。

### 知識の集約と提供のための連携イメージ

情報を集約し提供するシステムとして、図1のように、NDLSearchを核として想定している。まず、情報を集約し、それを提供する。利用者に求められるサービスと機能を持ったシステムを構築し提供するためには、外部の機関との連携協力が必須であり、NDLは積極的に連携協力を行っていく。その連携の姿勢として、次のような方針を掲げている。

#### メタデータの収集または横断検索等による統合検索サービスの提供

外部機関・サービスが提供するコンテンツのメタデータを当該機関・サービスの許諾を得て収集、もしくは横断検索する。実施に当たっては、共通的なメタデータ仕様、メタデータ交換プロトコルの実装が必要であり、また、個々のコンテンツに永続的識別子が付与されることが重要である。

メタデータに関しては、NDLは図書・雑誌についてMARC21、冊子体とデジタル化資料の検索のために必要な記述規則として、ダブリンコア（Dublin Core: DC）をベースに拡張したメタデータとしてDC-NDL（国立国会図書館ダブリンコア記述規則）を定義している。

永続的識別子については、冊子体資料では、図書にはISBNが付与され、NDLに納本された刊行物には独自の書誌IDが、さらに国内刊行物には全国書誌番号(JP番号)が付与されている。学術・科学技術分野において、JSTが、ジャパンリンクセンター（JaLC）を設置し、DOIをベースにした永続的識別子を付与し、論文等へのリンクを保証するサービスを準備している13)。出版界でも書籍と電子書籍を関連付ける識別子の検討が進められている。これらを統合的に関連付けるサービスの提供を目指す。

#### 外部のウェブサービスとの連携によるサービスの提供（マッシュアップサービス）

外部で提供されている連想検索サービスや機械翻訳サービス等のウェブサービスを有機的に組み合わせて、付加価値の高い検索サービスを実現する。また、外部の情報サービスへの効果的なナビゲーションを実現することにより、利用者の情報探索を支援する。実施に当たっては、各サービスが、単なるメタデータ交換ではなく、ウェブサービスの連携に必要な共通的なAPIを持ち、必要に応じて自由に組み合わせられるようにプロトコルの共通化を図っていく。

#### )研究開発、技術開発における連携

利便性の高いシステム構築は、現状で確立した技術のみでは実現が困難である。大学の研究室、官民の研究機関、ベンチャー企業等による各種の情報技術に係る研究開発を支援するために、NDLの情報資源を利用した実用化・実証実験を行うことができるよう、実験用プラットフォームを提供する。

#### コンテンツの統合利用促進のための環境整備

有用なコンテンツを保有しているにもかかわらず、データベースの構築や検索サービスの提供ができない機関に対して、データベースの構築やAPI実装を支援する。大学等においては、機関リポジトリの構築が進んでいるが、公共図書館、出版社等では、デジタル化及びデータベース化の実施が先進的な組織に留まっている。NDLは、標準的な仕様を実装したアーカイブシステム、検索・閲覧システムをOSSとして提供し、各期間でのデジタルアーカイブの構築を支援していく。

### 次世代システム開発研究の概要

NDLは、次世代技術の研究開発成果を活用して、これまでの単なる「情報検索」から、事実としての「知識検索」へ進化させ、知識の再利用による新たな知識の創造に寄与することを目指す。図2のようなアプローチを想定している。

これを実現するために、NDLラボの設置を想定している。

### NDLラボの設置

NDLラボは、次世代サービスの研究開発と実用化を促進するために、実用化実証実験の場と成果の利用のための研究者が集える場を想定している。概念としては、図3のように、NDLが保有しているコンテンツ、システムを研究者に提供する。研究者は、それらの資源を活用して実用化システムを開発する。その成果を、NDLのシステムに実装して次世代のサービスを提供する。このスキームで2010年度に、情報探索サービスシステムの環境をテストベッドとして利用して、全文テキスト化実証実験を行った。

## 震災アーカイブの構築

2011年3月11日の東日本大震災で、日本は未曾有の被害を受けた。その復興のために「復興への提言」（災害の記録と伝承）が示され、関係する記録、資料等の収集・保存を行い、一元的なアクセスが可能な仕組みとして、東日本大震災アーカイブ（仮称）の構築を進めるとされた。「第4期科学技術基本計画」の基本認識と理念の中でも「震災からの復興、再生の実現」が示され、2011年度三次補正予算で震災アーカイブを具体的に構築することとなった。

東日本大震災に関する記録、資料等は多種多様である。震災時の状況、被害の実態、復旧・復興の状況、支援の記録などを将来に残していく必要がある。

これらを1つの機関が収集することは困難であり、関係府省の各機関が分担して収集・保存を行い、各機関に保存された記録、資料等に対して一元的なアクセスと永続的な保存を保障するための「震災アーカイブポータル」の構築と提供を目指す。

### 震災アーカイブの構築に当たって

NDLでは、知識インフラの構築の一環として、震災アーカイブ及び震災アーカイブポータルを構築することとした。構築に当たっては、効率的、効果的に進められるように、次世代技術の研究成果を積極的に活用する。

### 震災に関する知識インフラの構築

アーカイブでは、クローラによる収集のみならず各機関が収集したアーカイブを将来的には長期保存のために一括して受け取る仕組みを構築する。また、画像・映像なども的確に検索ができるように、明確なメタデータが付与されていない情報にも可能な限りメタデータを自動付与する仕組み、本文も含めたテキストの全文インデキシング等を行う。さらに、今後、復興の記録なども含めていくことを想定して、ストレージ容量や大量の情報の組織化のための処理能力も必要に応じて増強できる仕組みを構築する。また、震災アーカイブ自身が、災害で消失してしまわないように、ディザスタリカバリも考慮する必要がある。それらの要件を満たすシステムとして、必要に応じて、処理能力、ストレージ容量が段階的に増やせる、PaaS（Platform as a service）、分散処理サーバ、分散ファイルシステムの導入を検討している。

このようなシステムを構築するためには、知識インフラの構築で目指す「次世代技術の研究開発成果の活用」のスキームで行うこと有効と考え、それによって、図4に示す「新たな知識の創造と還流」により、「震災に関連する知識のインフラ」の構築を目指したいと考えている。

## 電子情報関連の組織再編

NDLでは、以上のような次世代に向けた電子情報に関する事業を効率的、効果的に実施するために、2011年10月に電子情報部を設置した。15)

設立趣旨は、NDL全体の電子情報、情報システムの企画立案が効率的に行えるようにして、統合的に情報システム基盤の構築・運用を図る。分散して行っていたシステム関連業務を一元的に行い、現行システム・サービスを効率的に再構築・運用する。将来的な展望を持って、トータルな図書館システムを実現し、図書館の枠を超えて利用者サービスを向上させる。

2012月1月の業務・システムのリニューアルの次の目標は、知識インフラの構築であり、国立図書館として、知識情報資源のアーカイブ基盤の構築やデジタルコンテンツの利用促進等の情報流通基盤の整備を推進し、次世代の図書館サービスを提供することである。

これらの実現に向けて、先進サービス動向、技術を把握してサービス要件、システム化要件を取りまとめ、構築・運用するために高いマネジメント能力を持った人材育成・確保を進める。それに伴い、外部の有識者の実践的な助言・提案をいただくために、有識者が集まれる場として、前掲のNDLラボの運営を想定している。そこでのテーマは多岐にわたる。たとえば、技術要素では、パターンマッチング、パターン認識、マルチメディア技術、画像、映像処理技術、クラスタリング、キーワード抽出、シソーラス、テキストマイニング、関連性検出、文章解析、対話システム、情報圧縮・要約技術、情報分析技術、機械翻訳技術等がある。

## おわりに

利用者が情報を知識として活用するための情報探索を行う目的は、問題・課題の解決であり、回答が掲載された資料の所在ではなく、回答そのものを知識・情報として得ることである。今後は、従来の単なる情報としての検索でなく、事実としての知識検索へ進化させ、知識の再利用による新たな知識の創造が求められている。それを実現するための前提として、資料をデジタル化し、インターネット上に内容を可視化することが必要。加えて、個別の情報に意味的にタグ付けし、知識として相互に関連付けて、利用者が求める知識として、より的確に取り出せるようにすることが必要である。「可視化」においては、自動翻訳等により言語差異を吸収し、また、交換される情報に付与されたメタデータの記述要素、記述規則を共通化し、用語の類義語を把握するシソーラス、語彙の違いを吸収するオントロジー等の言語や表現の差異を吸収する仕組みが出来上がれば、言語、業種、業態の壁を超えて、網羅的な情報資源の中からより的確な情報を取り出すことができる。

NDLは、各国の国立図書館と協力して、世界レベルでの電子図書館構想の一翼を担い、国内外の関係機関と連携して、知の共有と提供の基盤を構築することを目指す。

### 参考文献

1)国立国会図書館. “国立国会図書館60周年を迎えるに当たってのビジョン（長尾ビジョン）”. 国立国会図書館. http://www.ndl.go.jp/jp/aboutus/vision\_60th.html、 (accessed 2011-11-07).

2)国立国会図書館. “国立国会図書館業務・システム最適化計画”. 国立国会図書館. http://www.ndl.go.jp/jp/aboutus/pdf/optimization.pdf、 (accessed 2011-11-07).

3)国立国会図書館. “改正国立国会図書館法によるインターネット資料の収集について”. 国立国会図書館. http://www.ndl.go.jp/jp/library/news/1188809\_1484.html、 (accessed 2011-11-11).

4)国立国会図書館. “納本制度審議会答申「オンライン資料の収集に関する制度の在り方について」について”. 国立国会図書館. http://www.ndl.go.jp/jp/news/fy2010/1189274\_1531.html、 (accessed 2011-11-19).

5)文化庁. “平成21年通常国会　著作権法改正等について”. 文化庁. http://www.bunka.go.jp/chosakuken/21\_houkaisei.html、 (accessed 2011-11-11).

6)文化庁. “電子書籍の流通と利用の円滑化に関する検討会議”. 文化庁. http://www.bunka.go.jp/bunkashingikai/kondankaitou/denshishoseki/index.html、 (accessed 2011-11-19).

7)電子出版環境整備事業. “JBasicもうすぐ08”. 電子出版環境整備事業. http://www.epubcafe.jp/jbasic、 (accessed 2011-11-19).

8)科学技術振興機構. “J-STAGE次期バージョン（J-STAGE3）開発について”. J-STAGE（科学技術情報発信・流通総合システム). http://info.jstage.jst.go.jp/society/development/index.html、 (accessed 2011-11-11).

9)国立国会図書館. “機能概要 ≪ 国立国会図書館サーチについて”. 国立国会図書館サーチ. http://iss.ndl.go.jp/information/function/、 (accessed 2011-11-30).

10)内閣府 科学技術政策・イノベーション担当. “第4期科学技術基本計画（平成23年8月19日　閣議決定）”. 内閣府. http://www8.cao.go.jp/cstp/kihonkeikaku/4honbun.pdf、 (accessed 2011-11-11).

11)国立国会図書館. “第三期科学技術情報整備基本計画”. 国立国会図書館. http://www.ndl.go.jp/jp/aboutus/pdf/basic\_plan03.pdf、 (accessed 2011-11-11).

12)長尾真. “新しい時代の情報図書館学へ向けて（九州大学創立百周年記念講演）”. 九州大学. https://qir.kyushu-u.ac.jp/dspace/bitstream/2324/19701/1/20110611.pdf、 (accessed 2011-11-11).

13)科学技術振興機構. ジャパンリンクセンターの開発について. J-STAGEニュース. 2011、 no. 28. http://www.jstage.jst.go.jp/jnews/J-STAGE\_NEWS\_NO28.pdf、 (accessed 2011-11-11).

14)内閣官房 東日本大震災復興構想会議. 復興への提言～悲惨のなかの希望～. 2011-06-25、 第12回東日本大震災復興構想会議. 57p. http://www.cas.go.jp/jp/fukkou/pdf/kousou12/teigen.pdf (accessed 2011-11-11).

15)国立国会図書館. 国立国会図書館月報. 2011、 no. 607. http://www.ndl.go.jp/jp/publication/geppo/pdf/geppo1110.pdf、 (accessed 2011-11-11).

# （2012年）国立国会図書館サーチのコンセプト・開発経緯と今後の展開

平成24年3月24日

三田図書館・情報学会　月例研究会

## はじめに

国立国会図書館（通称NDL）、電子情報部の中山です。

本日は、お話しする機会を設けていただき、ありがとうございます。

3年前に、関西館から戻ってPORTAのお話しをさせていただいた頃から、当館のシステムの全面リプレースを担当してきました。

この1月にPORTAの後継となるNDLサーチを含めて、当館の基幹システムがリニューアルオープンしました。

本日は、「国立国会図書館サーチのコンセプト・開発経緯と今後の展開」と題して、NDLサーチを開発するに当たって考慮したこと、どのようにデザインしてきたか、また、今後何を目指しているのかを、お話しさせていただきたいと思います。

## 「知識はわれらを豊かにする」を情報処理システムの観点からみて

まずはNDLの役割についてです。前回も最初のスライドに使わせていただきました。

* 当館は、日本における唯一の国立図書館として、「知識はわれらを豊かにする」というビジョンの中で、
* （２）「日本の知的活動の所産を網羅的に収集し、国民の共有資源として保存する。」、
* （３）「利用者が求める情報への迅速で的確なアクセスまたは案内できるようにする。」、
* （４）「来館者がどこにいても、来館者と同様のサービスが受けられるように努める。」という柱を掲げています。
* 納本制度の下で国内の刊行物を網羅的に収集する当館が、国民に期待されていることは、印刷物＝「物」としてではなく、「情報」として、網羅的に収集し、時代を超えて保存し、提供していくことと考えています。
* 当館の使命は、利用者が必要とする情報を確実に提供することであって、当館が保有している情報だけを提供すればいいというものではありません。
* 当館が保有している**情報は利用しやすく、**また、保有していない情報へは、**所在場所へ確実に案内すること**が求められています。

このような使命を果たすために、**資料のデジタル化及び情報の収集・保存そして提供するシステムを構築しています。**

## 当館の位置づけの再認識（2004年）

当館の位置付けを再認識してみます。

* 当館の位置付け
  + 世界の国立図書館の一つ
  + 日本の唯一の国立図書館
  + 日本の公共図書館の中央館
* 使命と現実
  + 納本制度のもとで、国内の刊行物を網羅的に収集して提供する使命を持っている  
    →しかしながらすべてを収集することは不可能
  + 歴史的資料は、公文書館、博物館、図書館等で、分散所蔵している
  + 当館単独で、すべての利用者ニーズに応えることはできない  
    →民間DB、学術情報、MLA、日中韓、WDLとのサービス連携を視野に
  + サービスを提供する使命があっても、当館には技術はない  
    →政府や民間の技術の移転が必要

図書館も変わらないと、世の中で役に立つ組織として存立し続けることが困難になる危機感があった。組織の枠を越えて、時代のニーズにあったサービスの提供が必要です。

* + 国の機関であっても、民間レベルをサービスを
    - 国だからできないという先入観念を持たない
      * 「国だからこんな程度」と言われないように
    - 「立法府は政府と違う」とは言わない
      * 政府でできて立法府ができないということはない。
* 当館もまだまだ変革の途上です。
* **これからも従来の枠に捉われないサービスを模索中**
* 計画倒れにならないように、できるところから進めていきたいと思っています。

## NDLサーチと業務基盤システム

それでは本題に入ります。まずは、リニューアルしたシステムの中核のシステムとして、「検索サービスとしてのＮＤＬサーチと、従来型図書館の延長の業務基盤システム」の開発経緯と目指したところを、紹介します。

## NDLSearchができるまで

デジタルアーカイブポータル、PORTAの目指したのは、デジタルコンテンツの統合検索

NDLサーチが目指したのは紙・デジタルコンテンツの統合検索、当館の情報提供サービスの入り口になることです。

2004年3月

電子図書館中期計画2004が策定され

デジタルアーカイブの構築

情報資源に関する情報の充実

デジタルアーカイブのポータル機能

の3本柱を立てました。その直後から、デジタルアーカイブポータルのプロトタイプを構築を始め

2005年7月

デジタルアーカイブポータルプロトタイプ公開

近代デジタルライブラリと青空文庫の統合検索

その後、プロトタイプでの意見を踏まえて、正式版の開発に着手。

2007年10月

PORTA正式版公開

8機関、20アーカイブ

約800万件を対象

3年前に、システムのリニューアルを担当してから、次期システムの全体イメージを見直して、当館サービスの窓口となる「情報探索サービス」を含めた次期システムを計画

2009年4月

NDLSearchの開発を決定

当館コンテンツのみへ案内するポータルは構築しない

PORTAを取り込む

ゆにかねっとを取り込む

OPACサービスを拡張する次世代OPAC機能は、OPACではなく全てのコンテンツを検索するNDLサーチに実装する

2010年8月

NDLSearch試験公開

2012年1月

NDLSearch正式公開

NDL-OPACのリニューアル、加えて来館者管理システム、館内サービスシステムもリニューアル

NDLデジタルアーカイブも部分的ですがリニューアルされた

加えて、2月にはHPもリニューアルした

**このように、基本のコンセプトは2004年のままで、より利便性の高いサービスを追求してきました**。

## NDLSearchができるまで情報探索サービスの将来像(クラウドの世界でのサービスの連携)（2009年）

これは、3年前に、館長に提示した情報探索サービスの将来像です。

それぞれの分野、カテゴリで、サービスがクラウドの世界で連携して、網羅性のある情報として利用可能になることをイメージしています。

このなかでは、当館はデータ及びサービスを提供するプロバイダの１つとして位置づけています。

ただ、当館は巨大なデータプロバイダ

・巨大なデータプロバイダとして、中核的となる

　当館が唯一の入り口になることを目指したものではなく、**様々な連携の入り口から、クラウド上のサービスを自由に利用できるようにすることを目指す**ことを想定しました。

～～～～～～～～～～～～～～～～～～～

1. *コレクション構築、サービスの構築、サービスの提供のどれも、当館の力だけではなし得ず、関係機関の協力と連携が不可欠です。*
2. *今後のサービストレンドを踏まえたサービスの構築は、関係機関による研究開発、技術開発の成果を活用させていただかなければ、実現が困難です。*
3. *また、コレクション構築においても、７月に国立国会図書館法が改正になり、国・地方自治体、国立大学、独立行政法人等の機関のウェブを収集・保存し提供できるようになりましたが、それだけでも、すべてを当館だけで保存することは困難です。*
4. *同種のコレクションを保有する公共図書館、大学図書館を含めて、博物館、美術館、公文書館等の各機関、政府機関、民間、さらに個人が持つ有用な情報が１つの巨大なデータベースとして利用できるようにすることが必要です。*
   1. *そのためには、１つのデータベースに集約するのではなく、個別の分散データベースとして提供され、他のサービスからナビゲートできるように「見える化」していただくことが重要となります。*
   2. *「見える化」された各機関の持つ分散データベース、分散サービスが、相互に補完しあって、より高度なサービスを提供できるようになることを期待しています。*

*それぞれの分野、カテゴリで、サービスが連携して、網羅性のある情報として利用可能になることをイメージしています。*

*このなかでは、当館はデータ及びサービスを提供するプロバイダの１つです。*

*当館が唯一の入り口になることを目指したものではなく、様々な連携の入り口から統合的に検索できるようにすることを目指します。*

## 次期図書館システムの全体イメージ（2009年3月）

これは、1月に公開したシステムの当初のシステムイメージです。

次期システム構築の背景には、経費節減がありました。

(1) 現行システムの運用経費の削減

(2) 次期基盤システムの導入経費の削減

開発目的

運用性、柔軟性の高い次期基盤システムを構築することにより、これらの問題点を解決し、さらに、削減された資源を活用して、社会のニーズの変化に対応した新たな利用者サービスを展開できるようにする。

基本方針

当初は、OPAC、PORTA、リサーチナビ、レファレンス協同DBを合わせて検索できるポータル機能の構築を想定した。

調整の結果

(1) 次期基盤システムは、 ILSパッケージを核として、当館全体システムを構成する次期図書館システムの一機能として検討する。

(2) 図書館ポータル機能、付加価値サービスを提供する次世代OPAC機能は、情報探索サービスで実現する。

(3)さらに、総合目録機能、PORTAを統合した実現イメージを想定する

**それぞれのシステムが独自に検索機能を持っていたが、共通な機能の部分をNDLサーチに集約させる形での構築に変更された。**

## 書誌・所蔵の視点でのコンテンツの体系的整理の概念（2009年3月）

これは、書誌的体系に基づいた分類のモデルの一つとして

『*書誌的記録の機能要件（Functional Requirements for Bibliographic Records：*FRBR）』モデルに基づいた資源の管理のイメージです。

一つの著作物は、デジタル化されたコンテンツも含めると、この体現形の例示にありますように、様々な形態のコンテンツとして派生して流通しています。

体系的に分類されていると

利用者属性（知識レベル、嗜好等）、利用環境（PC、モバイル、アクセス場所等）を考慮して、コンテンツを的確に選択できるようにすることができます。

また、選択されたコンテンツを実際に入手する際には、個別資料として管理されている資料の中から、最も入手しやすいところで所蔵しているコンテンツを選択できるようにすることが容易になります。

**NDLサーチは、体現形と言われる、緑色の丸の部分をグルーピングして、かたまりとしてリストアップ**できることを目指しました。

**オレンジ色の丸の部分は、従来からの公共図書館の総合目録の部分で、同じ資料を複数の機関で保有しているものを、体現形のレベルで同定**しています

これがNDLサーチの中の情報のモデルとなっています。

## サービス要件定義・システム化要件定義作成（2010年）

サービス要件定義書・システム化要件定義書作成までの流れ

システムの構築に当たっては、まず提供すべきサービスの要件を取りまとめました。

1. 長尾ビジョン→各種実施計画書→サービス要件として、どんなサービスを提供するかをリストアップ
2. このサービス要件に対して、「利用者ニーズ調査」、「外部サービス動向調査」を踏まえて、「サービス要件定義書」を策定
3. 「利用者ニーズ調査」では、フォーカスグループインタビュー、一般ユーザアンケート調査を行ってきました

この「サービス要件定義書」をもとに、システムとして実現するための要件を取りまとめました。

その際には、「技術・製品開発及び適用動向調査」及び、「技術標準適用指針・各技術標準適用ガイドライン」を参考しました。

この**システム化要件に沿って開発の調達仕様書をまとめ、プロトタイプシステムを構築し、公開実験を行い、反響とか意見をいただいてました**。

その意見等を反映させながら、22年度（2010）から23年度（2011）までに、機能強化して、正式版として公開しました。

## 適用すべき技術標準の指針（一覧）

システム化の要件を定義するに当たって、適用すべき技術の選択基準として、「技術標準適用指針」を参考にしました。

これは、個々の技術項目について館の標準を定めるうえでの基本的な**方針**して、各種の技術や標準を評価するための**原則**です。

**これに沿って、より効率的・効果的にシステムを構築できる技術を積極的に適用してきた。**

* 利用者の利便性向上に資する技術の積極的な採用
  + 一般利用者向けのサービスにおいては、直感的に理解できるユーザー・インターフェースや、一元的な情報の検索・参照を実現するための機能など、**利用者の利便性向上に貢献する技術を積極的に活用する**。
* オープンな標準に基づいた技術・仕様の採用
  + 特別な事情のない限り、国際標準やデファクト・スタンダード等のオープンな規約に準拠した技術・仕様を採用する。また、**館内および館外のシステムとの連携インターフェース仕様は、国際標準やオープンな規約に準拠したものを採用**する。
* 技術・仕様の共通化
  + システム資源の共有化とサービスの統合化を実現するため、**特別な理由のない限り、採用する技術・仕様は共通化**する。
* システムの特性に応じた成熟度を持つ技術の採用
  + システムの重要度や先進性などの特性に相応する**成熟性や先進性を備えた技術・仕様を採用**する。
* パッケージ・ソフトウェアやオープンソース・ソフトウェアの活用
  + 機能要件の実現手段として、**パッケージ・ソフトウェア**を活用する。さらに、同様な機能を有する**オープンソース・ソフトウェア(OSS)**が利用可能な場合は積極的に採用する。
* 資源の共同利用および柔軟な配分・拡張に資する技術の採用
  + ハードウェア、ソフトウェア、ネットワーク等のシステム基盤資源を共通システム基盤として**複数アプリケーション間で共有する技術**や、その共有資源の配分や拡張を**利用状況に応じて柔軟に行うための技術を積極的に採用**する。
  + 業務アプリケーションのシステム化に際しては、業務固有の要件に基づいて、共通システム基盤の採否を個別に判断する。
* システムの重要度に応じた障害対策技術の選択
  + システムの障害が社会に与える**影響の程度に応じて、障害対策に係る技術を選択**する。
* 情報セキュリティを考慮した技術の選択
  + システム資源や情報の保護の観点から、適切な情報セキュリティ対策が施された技術を採用する。
* 運用・保守業務の集約化・共通化に資する技術の採用
  + 情報システムの運用・保守に係るノウハウの共有化と要員の活用を推進するため、**運用・保守業務の全館的な集約化・共通化**、ならびに運用・保守**作業の負担軽減に資する技術**を積極的に採用する。

## サービス構築の基本要件（2009年）

提供するサービスの構築に当たっては、このような基本要件を示しました。

**これにより、より利用者ニーズに適合したサービスの提供を目指した。**

* **ＮＤＬの新しい利用者サービスの方向性を打ち出す**

「網羅性、確実性、信頼性」が保証された良質な情報資源へと利用者をナビゲートする。

* **利用者オリエンテッドでユーザビリティを追求する**

利用者の検索プロセスを考慮したユーザインターフェースの提供

旧来のOPACのような検索表示にこだわらない。

* **利用者をターゲッティングする**

当館サービスをあまりりようしたことのない「一般ユーザ」。（来館での従来型OPAC利用者向けでなく）

* + 探す目的が明確かつ専門的知識の獲得のためのものであり、目的とするコンテンツがＮＤＬサイトのどこかにあるはずだという確証をもった形での利用
  + どこからどうやって探せばよいかわからず試行錯誤しながら探しているうちに、Googleやリンク集、Blogや記事など、ＮＤＬ以外のサイトで紹介・引用されたものをきっかけにたまたまＮＤＬサイトに行き当たった場合の利用
* **検索エンジン経由で訪れるユーザを重視する**

ＮＤＬのサイトという認識なしに訪問したユーザを適切にナビゲートすることを重視

* **デザイン・操作性を磨き上げる**

他の優れたアイデアや工夫は積極的に取り込み、無理なオリジナリティは追及しない。

利用者の反応・評価を適時に反映させながらアジャイル的に開発を行い、プロトタイプの期間中を通じて継続的に評価・改善を重ねることで完成度を高めていく。

* **「いつでも、どこでも」を実現する**

携帯端末利用者にも、PCの場合と遜色ない高い操作性とデザインのGUIを提供する。場所と時を選ばず、情報にアクセスすることを可能とする。

* **新しい付加価値を生み出す**

民間企業や非営利団体、個人が提供しているサービスとの連携や複数の異質なサービスの組み合わせ

従来の図書館の枠にとらわれない自由な発想による付加価値創造の仕掛けを作る。

## NDLサーチのシステム化要件

**サービス要件を実現するシステムの要件としては、このようなものを提示しています。**

* **情報の収集**

書籍、ジャーナル、雑誌、地図、画像、映像、音楽等のコンテンツのメタデータを収集する。

総合目録ネットワーク「ゆにかねっと」が提供している総合目録の情報もこの中に含まれる。

ただし、当館ホームページ上のテキスト情報については、全文検索の対象とすることから、収集の対象とする。

* **情報の組織化**

収集したメタデータについて、インデックスを作成して組織化し、DC-NDL形式により体系化された形でデータベースに保管する。なお、関連資料をグルーピングするなど、情報の構造的な見せ方も可能となるよう所要の対応を行う。

* **データ管理**

情報探索サービス用に収集・組織化したメタデータを最新の状態で管理する。

* **情報の検索**

簡易検索、詳細検索その他様々な方法で検索を行う。その際、利用者は必要に応じ、当館が用意するサジェスト機能やレファレンス情報、外部機関が提供する連想検索機能等のナビゲーションサービスを活用する。

* **付加価値サービスの提供**

検索機能のほか、RSS配信やブックマーク機能など情報探索に役立つ検索以外のサービスも利用者に提供する。

また、公共図書館等に対し、情報提供・収集用のAPIを強化する。

民間企業や非営利団体、個人等と連携しながら、保有する情報資源を活用した様々なサービスを提供することを目指す。

## NDLシステムのリニューアル

このような要件をもとに開発、試行公開を経て、NDL-OPAC等も含めて公開となった。

24年1月には、新規に構築していたＮＤＬサーチのほか、旧システムをリニューアルし、大きく４つのシステムに統合した。

1. 資料デジタル化の進展を踏まえ、紙資料とデジタルコンテンツの一元的な利用環境を整備、デジタルコンテンツの更なる活用を促進
2. 内外の情報資源への統合的なアクセス
3. システムの最適化、運用コストの合理化

**要件整理の段階から目指していた、ここにあるような目標は、達成できたと考えている**

## リニューアルシステム群の概要

システムごとに、ここに記載のサービスを提供している

おもなサービスは個別にスライドで説明します。

* NDLサーチ
  + 当館のみならず多様なデジタルアーカイブ機関との連携により、一次資料の入手手段までナビゲート
  + Next-L Enju をコアに、Heritrix、Hadoop、GETAssoc 、WordPressといったOSSを活用してシステムを構築
  + GETAssoc(連想キーワード）、J-Global（科学技術専門用語）連携により再検索機能を提供
  + 公共図書館の総合目録（ゆにかねっと）機能を統合
  + メタデータ同定、グループ化に分散処理（Hadoop）を採用
  + 各種連携APIを公開
    - 検索用API
      * 国立国会図書館サーチの検索ができます。（検索対象範囲は、使用するAPIのプロトコルによって異なります。）
      * SRU、SRW、OpenSearch、OpenURL、Z39.50、RSSに対応しています。
    - ハーベスト用API
      * 国立国会図書館サーチに格納されたメタデータのダウンロードができます。
      * OAI-PMHで提供しています。RSSでも可能です。
  + ShibbolethでのSSOとして、IDPを同システム環境に実装
* 業務基盤システム（NDLｰOPAC）

・ILSパッケージ（ExLibris社Aleph）を採用、一部改修＆外部開発

・収集、整理（雑誌記事索引作成も含む）、蔵書管理（出納・複写）、逐次刊行物管理、OPAC（アジア言語OPACも統合）

・MARC21準拠、書誌情報提供サービス、遠隔複写サービス

・Shibboleth　SSOにはAlephパッケージ内のPDS（Patron　Directory　Service）にて対応

* 館内サービスシステム

・スクラッチ開発。来館利用者への一元的な情報提供、各種申込（閲覧、複写、デジタルコンテンツプリントアウト、状況照会）、東京本館・関西館・国際子ども図書館で稼働

・従来のコンテンツ/サービス別に細分化されていた端末環境を統合

・利用者カード認証（端末ログイン）時、業務基盤システムPDS経由でログイン（SSOには未対応）

* 来館者管理システム

・ゲートとカード発行機のH/W、管理システム(スクラッチ）

・業務基盤システムとMQ＆REST通信にて連携し、ステータス管理

・東京本館、関西館で稼働

* Web NDL Authorities

・NDL典拠データを検索、提供するサービス

・典拠詳細画面からNDLサーチ検索が可能

・SPARQL APIを提供

・RDF、JSON形式でダウンロード可能

* 児童向けOPAC

・国際子ども図書館の開架資料検索システム。NDLサーチのサブシステムとして開発。児童用UI

## NDLサーチをサービスの起点に

**今回のリニューアルでは、NDLサーチをサービスの起点に位置づけている**

* 内外の情報資源を統合検索、一次情報の入手手段までナビゲート
* 連携対象（82DB、7千万件のメタデータを検索可能）
* JPO近刊情報センターとの連携開始
  + 出版前情報から、作成中書誌、完成書誌まで一貫した提供が可能に（出版前、納本後3日程度、貸出提供直前）
* CiNii Booksとの連携により、大学図書館蔵書も検索可能に
  + OpenURLによる書誌間連携を実現、横断検索について調整中
* 韓国及び中国との連携（中国国家図書館との連携が課題）
* シングルサインオン
  + Shibbolethにより、OPACとのシングルサインオンを実現、今後対象拡大
* 震災アーカイブ
  + NDLサーチをベースとして

## NDL蔵書の検索・申込システムとしてのNDL-OPACの刷新

NDL蔵書の検索・申込システムとしてのNDL-OPACの刷新

* ILSパッケージ（ExLibris社Aleph）を採用、一部改修・外部開発
* 検索範囲の拡大
  + アジア言語資料、雑誌記事索引、電子ジャーナル等を統合検索
* デジタル化資料へのナビゲート
  + 近代デジタルライブラリーに加え、「国立国会図書館のデジタル化資料」も
* 書誌データの変更
  + MARC21形式、ユニコードの採用
* 書誌データのダウンロード機能
* 書誌情報の迅速な提供
  + 作成中書誌（インプロセスデータ）の提供
* 新機能（マイリスト、検索履歴等）

## 館内利用者端末の多機能化

* 紙資料とデジタルコンテンツの統合利用環境の整備による利便性向上
  + コンテンツ／サービス別に細分化されていた端末環境を統合
* デジタル化したコンテンツの閲覧・複写環境の整備
  + 利用条件に対応したアクセス制御、複写制御
* 各種申込／確認
  + 閲覧・複写・デジタルコンテンツプリントアウトの申込、状況確認

リニューアルオープン後、ICカードになったこと、到着モニター表示がなくなったこと、検索システムが変わったこと、遅くなったこと、デジタル複写の印字品質が悪いことなど、改善要望が多く寄せられている。

また、3回のサービス停止もあり、来館者の評価は決して高くない状況。

**システムの安定化と性能改善は来年度の優先事項として取り組むことになっている**

## NDLサーチでの当面の連携イメージ

NDLサーチを核とした、他機関サービスとの連携イメージです。

まず、情報を集約

*国立国会図書館【蔵書目録、デジタルコンテンツ（当館デジタル化、収集コンテンツ）】、公共図書館、大学図書館【デジタルアーカイブ】、専門図書館*

そして、それを提供する

*提供方法として、NDLが直接利用者に届けるサービス*

*関係機関と相互補完して利用者に届けるサービス*

*サービスプロバイダを通じて利用者に届けるサービス*

NDLSearchでは、

デジタルネットワーク時代に、**利用者に求められるサービスと機能を持ったシステムを構築し提供するためには、外部の機関との連携協力が必須**であり、当館は積極的に連携協力を行っていきます。その連携の姿勢として、次のような方針を掲げています。

* 統合検索サービスの提供

外部機関・サービスが提供するコンテンツのメタデータを当該機関・サービスの許諾を得て収集、もしくは横断検索します。

* 外部Webサービスとの連携

外部で提供されている連想検索サービスや機械翻訳サービス等のウェブサービスを有機的に組み合わせて、付加価値の高い検索サービスを実現します。

また、外部の情報サービスへの効果的なナビゲーションを実現することにより、利用者の情報探索を支援します。

* 研究開発における連携

利便性の高いシステム構築のためには、現状で確立した技術のみでは実現が困難です。

大学の研究室、官民の研究機関、ベンチャー企業等による各種の情報技術に係る研究開発を支援するために、当館の情報資源を利用した実用化・実証実験を行うことができるよう、テストベッドの場を提供します。

* 統合利用促進のための環境整備

有用なコンテンツを保有しているにもかかわらず、データベースの構築や検索サービスの提供ができない機関に対して、データベースの構築やAPI実装等を支援します。

このように、**他の機関との連携により、補完しあいながら、利用者が必要とする情報を利用できるようにすることを目指しています。**

## NDLサーチの機能改善（2012年度）

1月に正式公開しましたが、来年度は、ここにあげられたような改善を想定しています。

* 統合検索サービスの強化

検索性能の向上

組織化精度の向上

異体字、表記ゆれの吸収

* 外部サービスとの連携強化

NDL-OPACとの連携改善

SSO利用者情報のリアルタイム更新

点字総目、占領期資料のNDL-OPACへの遷移

APIでの提供機能の充実

情報提供機能の強化

全国書誌の提供内容の充実（公共図書館にしかない郷土資料、ボーンデジタルも含めたものを全国書誌として）

* 研究開発・技術開発成果の活用

NDLラボでの提供機能の充実

Next－L Enjuの最新機能の取り込み

組織化精度の向上

* コンテンツ統合利用のための環境整備

NDLSearchのOSS化

## 知識インフラと震災アーカイブ

NDLSearch，NDLデジタルアーカイブの発展形として，今後取り組もうとしている計画は，「知識インフラの構築」である。

経緯

国の総合科学技術会議が決定した「第4期科学技術基本計画」（2011年8月19日閣議決定）においては，研究情報基盤の整備を推進することとし，推進方策が示された。また，NDLでは，これに先立ち，「国立国会図書館における今後の科学技術情報整備の基本方針に関する提言」（第52回科学技術関係資料整備審議会（2011年1月19日））を受けて「第三期科学技術情報整備基本計画」11)を定め，「知識インフラ」の構築を進めることとした。

知識インフラとは

知識インフラとは，情報資源を統合して検索，抽出することが可能な基盤で，国内の各機関が保有する情報を知識として集約し，新たな知識の創造を促進させるもの。その知識をさらに集積・流通・活用と創造するサイクルの構築が必要である。

知識インフラ構築の一環で、先行する分野の1つとして、震災アーカイブの構築を目指す

## 新たな知識の創造と還流

これは、国としての「知識インフラの構築」で、

「新たな知識の創造と還流」の概念をイメージしたもの。

様々な分野の情報が、様々な関係機関によって発信され、

それらが、連携機関によって、分担して所蔵している

その情報が、一元的にアクセスでき、ナビゲーションするものを「知識インフラ」として位置付けている

そこでは、単なる「情報の検索」でなく、事実としての「知識検索」へ進化させ、知識の再利用による新たな知識の創造に寄与することを目指す。

その**「知識インフラ」を利用して、新たな知識として創造され、蓄積されることをイメージする**

＊所在場所，媒体形式を意識にした「書籍，文献単位の提供」から，所在場所，媒体形式を問わず，「情報・知識・物語の単位での提供」へ。

＊メタデータの収集，探索，提供ではなく，情報の内容で，情報を関連付けし，提供する。

＊人間頭脳の持つ知識とその活用の機能にできるだけ近いサービスの提供。

＊ユーザの意図した情報がより的確に取り出せる「知識検索サービス」。

＊より安心して利用できる検索システム。キーワード，曖昧検索を駆使した「情報探索」から，信頼性を評価した「事実検索」へ。

## 知識インフラ構築に向けた有識者

そのような知識インフラの構築には、同種の情報が意味的に連携できる形になっていかなければならない。

そのために、関連する情報のコンテンツホルダーが知恵を出しあう場として、それぞれのWGを組織し、また、それらの情報を知識として活用できるシステムを構築するために、**「NDLサーチでの当面の連携イメージ図」の「③研究機関における連携」の実施母体として、IT技術関連の有識者会議を想定した。**

これらのWGが集まった組織体として、「デジタル情報ラウンドテーブル」を想定している。

システム技術ＷＧは、次世代技術を活用して、「知識インフラを構築」するために支援・助言をいただく場であるが、その具体的な活動として、震災アーカイブの構築があり、当館と総務省で共催する「東日本大震災アーカイブラウンドテーブル」の下位の「システム技術WG」としても位置付ける

## NDLラボ（仮称）の設置

これは、次世代サービスの**研究開発と実用化を促進**するために設置する**実用化実証実験の場の提供と成果の利用**のための研究者が集える場の想定です。

* 研究開発を促進するために、当館は、保有コンテンツ、保有システムを、研究機関に提供する
* 研究機関等の皆さんに、それらの資源を活用して、実用化システムを開発していただく
* その、成果を、当館のシステムに実装して、次世代のサービスを提供する

この具体的な実施として、震災アーカイブの構築を位置付け、ＩＴ技術ＷＧのメンバーを始めとする研究者が、震災アーカイブの構築環境をテストベッドとして利用していくことを想定している

## 現状のNDLサーチをベースに震災アーカイブとして機能拡張（想定）

これは、現状のNDLサーチ、デジタルアーカイブシステム、電子書庫をベースに震災アーカイブとして機能拡張（想定）

来年度末には、震災アーカイブを構築して公開すると発表しています。

短期間に、効率的・効果的に構築するためにも、ベースは既存のシステムを活用することとしています。

ベースは、今回リニューアルしたシステムです。

**ピンクの吹き出しは、来年度の機能拡張を想定している部分**で、さまざまな研究機関の成果を活用していきたいと考えています。

## まとめ

### 知識インフラ構築に向けたNDLサーチの今後の展開

目指すところは、「知識インフラの構築」であるが、知識インフラ構築の一環で、先行する分野の1つとして、来年度から再来年度にかけて、震災アーカイブの構築を目指す

震災アーカイブは、この１月にリリースした、NDLSearchとNDLデジタルアーカイブをベースに機能拡張して構築することを想定している。

震災アーカイブで実証された技術・スキームで、2014年度頃から、本格的な知識インフラの構築と提供を目指していきたい

### NDLサーチは、知識インフラのポータルに

次世代図書館サービスで想定しているイメージは、知識インフラと≒です。

* 知識インフラの提供を目指して

仮想的な図書館：図書館同士の壁を越えて

知識を蓄積し提供する場所

物理的な組織の壁を越えて

MLA（美術館、図書館、博物館）の館種、学術（大学、独立行政法人）、民間（企業の研究所等）の業務業態を越えて

* 情報の内容で関連付け

現在のNDLサーチの到達点は、メタデータレベルでの統合検索

これからは、1つの社会で共有する中立的な知識構造、知識システムとして、

資料、記事が、分解されて、さまざまな観点から情報同士がリンクされた巨大なネットワーク構造

関連付け

情報に固有の、共通的な知識の体系化、構造化

人それぞれの観点での、利用者の要求に応じて知識を関連付けてられるもの

* NDLSearchは、知識インフラの入り口

情報そのものの組織化と、取出し機能（発見・閲覧）を担う

そのために情報を組織化

*メタデータの自動付与*

*類似性のある情報の検出と分類（自動リンク付け）*

コンテンツの内容での検索

*マルチメディア情報の内容検索*

* 知識インフラの目標イメージ

他の知識との関係がチェックされ、新しい関係のリンクが付けられていく仕組み

*情報が単に記録されていくのではなく、*

自分のほしい情報そのものが出てくる仕組み

**自分のほしい情報の書かれている本を取り出して、書かれている個所を探すのではなく、自分のほしい知識そのものが出てくる仕組み**

**このようなサービスを提供する入り口になることを目指しています。**

# （2012年）電子書籍等のディジタルコンテンツの長期保存と、将来にわたっての利用の保証－文化的資産の保存に向けた関係機関との連携協力－

「情報処理学会誌（特集：電子書籍の未来）」（2012年8月情報処理学会）

## タイトル

電子書籍等のデジタルコンテンツの長期保存と、将来にわたっての利用保証

―文化的資産の保存に向けた関係機関との連携協力―

Long-term Preservation of Digital Contents (eBooks etc.) and Ensuring Access to them in the Future

―Cooperation and Coordination with the Relevant Organizations for Preserving Cultural Heritage

## 要約

国立国会図書館では、電子書籍等のデジタルコンテンツの長期保存と、将来にわたっての利用の保証に取り組んでいる。これを実現するためには、収集、組織化、保存、提供のシステム構築・運用、業務実施のあらゆる局面での関係機関との連携協力が不可欠である。

The National Diet Library is engaged in the long-term preservation of digital contents (eBooks etc.) and ensuring access to them in the future. To achieve these goals, it is essential to cooperate and coordinate with the relevant organizations throughout all phases of acquisition, organization, preservation, establishment and operation of providing system, and service implementation.

## 目指すところ

国立国会図書館（NDL）は、納本図書館として、印刷出版物を網羅的に収集保存し、将来にわたって利用を保証する責務を持っている。しかしながら印刷出版物として発行されないボーンデジタルの電子情報を含めて、その全てを収集することは不可能である。 NDLは、NDL及び他機関が分散して収集・保存している資料・情報を、意味的に関連付け、一元的に検索・ナビゲートできるようにする。また、情報を文化的資産として、将来にわたって利活用できるようにすることを目指している。

## 使命・経緯

### 経緯

NDLは、1994年頃から電子図書館構築に向けた活動を行ってきている。1995年にはパイロット電子図書館プロジェクトでの実証実験を実施し、2002年、電子図書館サービスの拠点として中心的な機能を持つ関西館が開館し、電子図書館事業を開始した。2004年には、「電子図書館中期計画2004」を策定し、デジタルアーカイブの構築、レファレンス情報等に到達するための仕組みの充実、さらに、所蔵場所に依らず一元的にアクセスできるポータルの構築に取り組んできた。

「電子図書館中期計画2004」が策定されてからの電子情報に対する取組みは、国の高度情報通信ネットワーク社会推進本部（IT戦略本部）が策定したe-Japan戦略等の取組としても掲げられ、関係府省の協力のもと、デジタルアーカイブ構築事業として進めてきた。その実現形が、2012年1月にリニューアルオープンした国立国会図書館サービスシステムの基本的な形であり、印刷刊行物等の物理的な資料の収集整理と閲覧提供を管理する業務基盤システム、電子情報を統合的に収集保存するデジタルアーカイブシステム（DAシステム）、NDLおよび他の機関が保有する印刷刊行物、電子情報等を一元的に検索・ナビゲートする国立国会図書館サーチ（NDLサーチ）等で構成する。[[15]](#endnote-14)

### NDLの使命と目標

NDLでは、2012年7月末に、今後5年程度の活動の基本となる使命として「私たちの使命・目標2012-2016」を策定し、文化的資産としての収集・保存、迅速かつ的確なアクセス環境・手段の整備、国内外の関係機関と連携して、知識・文化の基盤を一層豊かにすること、さらに、東日本大震災に関しては、出版物に限らず全ての記録を後世に残すことを、柱になる目標として掲げた。

## 関係機関との連携の観点

様々な資料・情報を文化的資産として保存し、利活用を促進するためには、収集、組織化、保存、提供のシステム構築・運用、業務実施のあらゆる局面での関係機関との連携協力が不可欠である。連携協力の実施に当たっては、姿勢として合意形成に留まらず、具体的なアクションとして、実施していくことが重要である。

## 収集の観点

### 「インターネット資料」の許諾に基づく収集

ウェブサイトでは日々刻々と新しい情報が発信され、同時に消されていく。時に、機関・組織の改廃・合併により、サイトそのものが消失する。消されていく情報には、後世に残すべき文化的資産も多く含まれ、印刷出版物の形態を取っていないものもある。NDLでは、2002年から、各機関の協力のもと、個別の許諾に基づいて、ウェブサイトを収集し、時間軸で再現できるように保存している。

### 公的機関の「インターネット資料」の制度的収集

2009年7月に国立国会図書館法と著作権法が改正され、国、地方公共団体、国立大学等の公的機関が発信するインターネット資料について、個別著作権者の許諾なく収集できるようになった。 この法改正に基づき、NDLは、2010年4月から、公的機関のウェブサイトの網羅的な収集を開始した。公的機関の協力により、順調に収集が行われている。なお、国公立大学の機関リポジトリ等は、早期に消失されることなく保存・提供が保証されているとみなされるので、当面は制度的収集対象とはせず、国立国会図書館サーチで、資料の所在場所にナビゲートして、閲覧利用を保証している。

### 民間の「オンライン資料」の制度的収集

電子書籍、電子雑誌など従来の図書、雑誌に相当しネットワーク上を流通する電子情報を「オンライン資料」と定義している。2012年6月に、国立国会図書館法と著作権法が改正され、オンライン資料の収集が許諾なしにできるようになった。この法律は、2013年7月に施行される。しかしながら、有償で提供されているオンライン資料に関しては、条件の整備等、時間を要するため、当分の間、無償で提供され、閲覧制限機能が実装されていないものを収集することとし、2013年7月からの収集に向けて、運用の準備を行っている。これにより、無償の電子書籍、電子雑誌は、ネットワーク上から消えていっても、将来にわたって利用が可能になる。

## 保存の観点

### 所蔵資料のデジタル化

NDLが収集保存している印刷出版物は、経年劣化が進むとともに、閲覧・複写提供により劣化が加速される。原本保存のために、2009年に著作権法が改正され、著作権者の許諾を得ないで所蔵資料をデジタル化することが認められた。 NDLでは、約150億円で、和図書約90万冊（1968年以前刊行分）、和雑誌約114万冊（2000年受入分まで）、博士論文約14万冊（1991年度～2000年度受入分）等、所蔵資料の1/4程度をデジタル化したが、引き続きデジタル化を進めていく必要がある。しかしながら、今後も継続的にデジタル化を行うための予算確保には課題がある。

### 文化的資産の保存

NDLが収集したパッケージ系電子出版物、インターネット資料、もしくはデジタル化した電子情報も、国の知識・文化の基盤となる資料・情報であり、データを失うことはあってはならない。

電子書庫としてのストレージは、東日本大震災アーカイブのためのものも含めると、 現在においても2PB（ペタバイト）の容量となる。今後、更に増加する電子情報の利用を保証するためには、大きく2つの観点がある。一つは、物理的に読めなくならないように保存（物理保存）すること、もう一つは、ファイルの内容が読めなくならないように保存（論理保存）することに留意する必要がある。

物理保存について、現時点において半永久的に保存できる記録媒体は実用化されていないので、膨大なデータを物理的に読めるようにしていく仕組みの確立が課題である。現在、 GlusterFSというシステムの適用を試行している。これは、寿命が5～10年で数TB（テラバイト）程度の容量の磁気ディスクを備えたPCを並列に配置した拡張容易な大容量分散ファイルシステムで、順次容量の大きな磁気ディスクに置き換えることで、少しずつ媒体変換を進め、かつ、必要な容量を確保できる仕組みである。また、このシステムにより、大規模災害に備えたディザスタ・リカバリー対策として、複数の分散したセンターで同期する仕組みの実装も想定している。近い将来には、クラウドサービスを活用し、複数の民間クラウドサービスを組み合わせて、相互に同期させることで災害時も含めたデータの保存を図ることも想定する。

論理保存については、ITの発展とともに、様々な国際標準、業界標準のフォーマット仕様を適用した電子情報が存在するが、将来にわたって読めるようにすることは大きな課題である。

この問題は、保存の使命を持つNDLだけでは解決が不可能である。新しい媒体、フォーマット仕様を開発してきた技術者・研究者の方々、国際標準・業界標準を策定してきた機関、その仕様を適用したアプリケーションやコンテンツを開発、販売してきた企業などの協力が不可欠である。様々な関係者に対して、現在の利用者のみならず、文化的資産として後世においても利用できるようにするために、仕様の共通化、マイグレーション等への協力を働き掛けていきたい。

## 組織化の観点

### 組織化の意義

組織化とは、利用者が迅速、的確かつ容易に検索できるように、メタデータ（書誌データを含む）を付与して整理することである。

一つの著作物が、単行本として出版され、のちに文庫本となり、また様々な形態の電子書籍として、派生して流通しているが、体系的に整理されたメタデータが付与されていると、利用者属性（知識レベル、嗜好等）、利用環境（PC、モバイル、アクセス場所等）を考慮して、コンテンツを的確に選択できるようにすることが容易になる。

外形的な情報によるメタデータのみならず、本文テキスト等からの組織化も行う技術（例えばセマンティックウェブ技術等）を駆使することにより、利便性の高い検索サービスが実現できる。

### 組織化の連携協力

図書・雑誌の出版者、博物館、文書館、図書館等のいわゆるMLA機関、その他著作物を提供する全ての機関が、語彙の違いを吸収できる共通のメタデータ記述規則を適用し、意味的に関連付けられることが重要である。関係機関で協力してメタデータの相互交換の仕組みを構築する必要がある。

また、爆発的に増加する電子情報には、従来の印刷刊行物のように人海戦術的な精緻なメタデータの付与は困難である。自動的にメタデータを付与する技術、本文情報も含めて組織化する技術等、大量のデータを構造化・意味的情報を取り出す技術（例えばデータマイニング技術等）の研究開発とその成果の実用化が期待されている。

## 電子書籍出版社等との連携

電子書籍は、印刷出版物の延長にあるものであり、文化的資産の１つの形態である。

現在、電子書籍出版は、ビジネスとして立ち上がろうとしている。NDLは、電子書籍によって読者人口が増えて、出版全体の市場が拡大し、出版ビジネスが加速されるように支援するとともに、電子書籍の将来に亘る利用を保証することが役割と考える。そのためにも、民間の市場経済活動を阻害することなく、市場拡大のために、出版界と下記のような様々な連携協力が今後の課題である。

・NDLデジタル化コンテンツの二次利用の促進

・電子書籍サイト等、商用サイトへの案内の強化

・電子書籍ビジネスのプラットフォーム整備

・電子書籍フォーマットの共通化

・電子書籍に対する永続的識別子の付与

・公共図書館での利用環境の共通化

・著作権管理センターの構築・運用の協力

### NDLデジタル化コンテンツの二次利用の促進

国のオープンガバナンスの方向性に沿って、NDL保有の資産で、第三者の権利を侵害しないものは、積極的に二次利用を促進させたい。電子書籍出版社に、画像データをとして提供し、二次利用によって、電子書籍を作成してビジネスが行えるように支援することも想定する。

### 電子書籍サイト等、商用サイトへの案内の強化

NDLサーチは、紙・デジタル、有償・無償、商用サイト・公的機関等に関わらず、ロングテールで容易に資料の存在を確認することを目的としている。利用者が最も迅速に入手し閲覧可能な入手先へ利用者をナビゲートすることが目的である。NDLは今後利用者の資料の有力な入手手段となる電子書籍サイト等への案内を強化することを想定している。

### 電子書籍ビジネスのプラットフォーム整備に協力

NDLが構築しているデジタルアーカイブ機能とポータル機能は、スケーラブルな書籍データベース及び配信サービスシステムであり、検索及び閲覧のAPIを公開している。今後ソフトウェア資産もオープンプラットフォームとして公開することにより、他の公的機関サイト、商用サイトでも適用もしくは連携が容易になる。

### 電子書籍フォーマットの共通化

現在の電子書籍フォーマットは多種多様であり、かつ、電子出版サイト毎にビューアが異なる状態である。これは、電子書籍の流通の拡大、NDL等による長期保存の観点からみると障害になっている。電子書籍フォーマットの国際標準、業界標準の策定を支援して、共通フォーマットの普及を促進させたい。

### 電子書籍に対する永続的識別子の付与

出版に先立って販売促進のために作られた出版前情報、出版情報は、NDLで蔵書として管理するための書誌情報には活用されておらず、また関連付けもされていない。出版情報はONIXで、書誌情報はMARCで、電子情報はDCベースでというように、書誌的事項の記述規則も共通化されず、再利用もされていないため、検索時に同一のものと認識することも困難な状態になっている。

著者が作品を作成した時点で、永続的識別子を付与し、販売のために作成された出版情報と、図書館での書誌情報をリンクさせる形で相互連携できるようにしていきたい。また、実際に永続的識別子を付与する手段として、JaLCを活用したDOI付与も想定している。

### 公共図書館での利用環境の共通化

今後、電子書籍が、公共図書館等でも電子書籍サイトからそれぞれのビューアを利用する形で提供されることが予想される。NDLからの提供は、別の著作権保護方式で別のビューアを利用する形では、利用者にとって利便性が悪い。電子書籍サイトとNDLとで、共通の著作権保護機能とビューアで提供できるように、システムの共通化を図っていくことが、市場の拡大に繋がると考える。今後、NDLでの資料のデジタル化状況、出版社でまだ電子書籍化されていない資料の著作権状況など、出版界とNDLで協力して、著作権管理データベースを構築すべきと考える。

### 著作権管理センターの構築・運用の協力

今後、NDLでの資料のデジタル化状況、出版社でまだ電子書籍化されていない資料の著作権状況など、出版界とNDLで協力して、著作権管理データベースを構築すべきと考える。

## 新たな取組

### 東日本大震災アーカイブ

東日本大震災アーカイブは、大震災に関連する、災害現象そのもの、災害前・災害直後・復興の過程、災害時の対応、他地域・次世代への教訓等を記録として網羅的に収集し、後世に残すものである。大震災の記録は、従来からの収集対象である印刷刊行物に留まらず、ビラ類、写真、動画、音声はもとより観測記録等、多種多様である。また、記録を保有している機関も様々であり、早期に収集保存に着手しなければ、散逸の恐れがある。記録を保有もしくは集約している関係府省、博物館・美術館、図書館、文書館、企業及び先行して震災アーカイブを構築・運営している組織と協力して、分担収集し一元的なアクセスを保証したい。

大震災アーカイブポータルに関しては、既存のデジタルアーカイブシステムをベースに、分散アーカイブを構築し、また、統合的に利用できるポータルを構築している。また、見せ方も記録の日時、場所も意識した閲覧機能の実装を目指している。

### 知の共有化に向けた連携

国の第4期科学技術基本計画で示された「知識インフラ」は、知の共有化を目指す分野を問わないモデルであり、2012年1月にリニューアルしたサービスシステムの延長線上にあるものである。

東日本大震災アーカイブは、コンテンツ、システムともに、分野を特定した「知識インフラ」の実現形であり、既存のサービスをベースに、必要な機能を実装する。このアプローチは、国の施策としての「ビッグデータの利活用」「知の共有化」に繋がる。

## 政府の施策との連携

知的財産推進計画2012 （2012年5月知的財産戦略本部）[[16]](#endnote-15)[2]の「戦略2:日本を元気にするコンテンツ総合戦略」では、「電子書籍の本格的な市場形成」及び「コンテンツのアーカイブ化とその活用促進」において、著作物のデジタル化、コンテンツ流通の一層の促進や、アーカイブに関する博物館、図書館及び公文書館の連携の取組、東日本大震災のデジタル・データを一元的に検索・利用できるポータルサイトを構築等、関係府省とNDLが協力して取り組むべき事項の内容とスケジュールが示されている。

「電子行政オープンデータ戦略」（2012年7月IT戦略本部）[[17]](#endnote-16)[3]では、オープンガバナンスの方向性として、国民共有財産である公共データを積極的に公開すること、機械判読可能な形式で公開し、営利目的、非営利目的を問わず活用を促進することが示されている。

また、内閣府総合科学技術会議の科学技術イノベーション政策推進専門調査会[[18]](#endnote-17)[4]において、第4期科学技術基本計画の重点化課題「新たな産業基盤の創出」の重点的取組として「大規模情報（ビッグデータ）の利活用の基盤技術の開発・標準化・普及促進」が明確化され、ビッグデータの収集・蓄積・分析等の研究開発及び国際標準化を進めるとされている。NDLのデジタルアーカイブ構築、東日本大震災アーカイブポータル構築は、まさにビッグデータを扱うシステムであり、これらの施策での技術開発、実証実験の成果を活用していきたい。

## 関係機関を繋ぐ役割を果たす

NDLは、「私たちの使命・目標2012-2016」を策定し、具体的な実施計画の策定作業を進めている

NDLは、関係機関との連携により、国として資料・情報を収集・組織化し、ビッグデータとして利活用できることを目指していく。

今後、様々な業種・業態で情報を発信者している機関同士、それらの情報を発信している機関と情報の利用者同士、また、膨大な情報を高度に処理・活用するための研究開発・技術開発を行っている組織同士を繋ぐ役割を果たしたいと考えている。

繋ぐに当たっては、 関係機関間の利害調整ではなく、未来志向でより創造性を持って、資料・情報の権利保持者の権利を尊重し、将来的な利活用の拡大を目指して、共存共栄で協力・分担して進めることが大切である。

# （2012年）国立国会図書館が取り組むデジタルデータの保存と未来のデータ保存

**ハミングヘッズ取材対応**

平成24年8月7日

電子情報部　中山

## 企画内容

* 1. 『HH News & Reports』
     + 『Monthly Interview』の9 月企画
     + 参考：「国立国会図書館の蔵書電子化と未来の図書館の姿」
       - <http://www.hummingheads.co.jp/reports/interview/n0101206/interview51_01.html>
  2. **「国立国会図書館が取り組むデジタルデータの保存と未来のデータ保存」（仮）**

クラウドなどでデータの保存が消失する事件がおき、改めて「データの保存」が注目されている。その意義を再確認しつつ、国立国会図書館の取り組んでいるデジタルデータをはじめとした資料の保存の取り組みについて紹介。今後の展望についても話をうかがう

## 事前質問

* + クラウドなどでデータの保存が消失する事件がおき、改めて「データの保存」が注目されています。「データを保存すること」の重要性についてお考えをお聞かせいただければと思います。
  + 国会図書館が現在取り組んでいる資料の保存、デジタルデータの保存の取り組みについてお聞かせ願えればと思います。また以前の取り組みをどのように発展させていくかについてお考えをお聞かせください。
  + 海外の図書館における資料保存・デジタルデータ保存についてはどういう面を参考にしたいとお考えでしょうか。
  + 現在のデジタルデータの保存技術への要望を実際に資料保存の現場にいるお立場からお聞かせ願えればと思います
  + 注目されているデジタルデータの保存技術があればお聞かせください。
  + 今後の国会図書館の資料保存、デジタルデータの保存の展望をお聞かせください。

## 回答概要

* 電子図書館課がこれまで取り組んきた、電子情報の長期的な保存と利用
  + <http://www.ndl.go.jp/jp/aboutus/preservation.html>
* 個別
  + 短期的なデータの機密性・完全性・可用性の確保（たとえばe-文書法）、長期保存の必要性
  + 情報資源を後世に残すことの意義、現状の長期保存のための取組み、ウェブアーカイブ、デジタルデポジット、制度的収集、保存のためのデジタル化
  + DRAMBORAの意義、オランダ王立図書館e-Deposit、IA社、JIIMAでの取組み。。。
  + ④⑤長期保存メディア、長期保存システム、必要に応じて順次拡張できるストレージ技術、分散ファイルシステム、ディザスタリカバリ技術、クラウドコンピューティングとか、膨大な情報を検索・閲覧できるようにする技術にも触れていいかも
  + ⑥当館は、今後増え続けるデジタルコンテンツをどのようにして、後世に残し、将来的なアクセスを保障していくか

## 回答

### クラウドなどでデータの保存が消失する事件がおき、改めて「データの保存」が注目されています。「データを保存すること」の重要性についてお考えをお聞かせいただければと思います。

* 「私たちの使命・目標2012-2016」
  + 国立国会図書館は、出版物を中心に国内外の資料・情報を広く収集し、保存して、知識・文化の基盤となり、国会の立法活動を補佐するとともに、行政・司法及び国民に図書館サービスを提供することを通じ、国民の創造的な活動に貢献し、民主主義の発展に寄与することを使命としている。
  + 確実な保存
    - 電子情報も従来の紙資料と同じく、国の知識・文化の基盤となる資料であり、当館の使命を果たすためにも、データを失うことはあってはならない。
* クラウドでのデータ消失に関して
  + ただ、例の件を「クラウド」という言葉で一般化するのは、やや正確さを欠くようにも思う。
  + 分散保存が必要
  + 東日本大震災では、バックアップデータを同じ施設内に置いておくだけでは、データの安全を確保するには足りないということが広く認識された。
  + その意味で、クラウドの重要性は今後むしろ高まるのではないか。
* データを有効活用
  + 同時に、震災時の対応やその後の復興に向けての検討においても、データを有効活用することが、社会にとっていかに重要かということも再認識されている。
  + 図書館にとって、電子情報が知識・文化の基盤であることはもちろんだが、広く電子データ一般として社会の基盤となっているということもできるだろう。
  + データを保存することは、社会の基礎・基盤を守る活動の一部だと考えている。

### 国会図書館が現在取り組んでいる資料の保存、デジタルデータの保存の取り組みについてお聞かせ願えればと思います。また以前の取り組みをどのように発展させていくかについてお考えをお聞かせください。

* デジタルコンテンツの保存の取組み
  + 電子資料の保存については、NDLデジタルアーカイブシステムの構築に取り組んできた。当館所蔵資料のデジタル化、インターネット資料（ウェブサイト）の収集を行ってきた。
  + また、今年からは東日本大震災アーカイブ（仮称）の構築にも着手している。来年7月からは、国立国会図書館法の改正により、オンライン資料の収集も可能となる。
  + デジタルデータの保存には、媒体やストレージなどの「物理保存」の取り組みと、記録されたデータの内容を人が理解できる形式で再生できるように維持する「論理保存」の取り組み２つの保存への取り組みが必要である。
  + 当館ではまだ、論理保存の取り組みを本格的には行えていないが、平成18年度から5カ年をかけて調査を行った。詳しくは、
    - 電子情報の長期利用保証に関する調査研究
    - http://www.ndl.go.jp/jp/aboutus/preservation\_02.html
    - の下にある各報告書を見てもらいたい。
* 紙資料などの保存の取組み
  + 電子に限らない、紙資料などの資料保存一般としても、当館は、国際図書館連盟（IFLA）の資料保存コア活動（PAC: Core Activity on Preservation and Conservation）のアジア地域センター（<http://www.ndl.go.jp/jp/aboutus/iflapac.html>）を担うなど、様々な活動に取り組んでいる。
  + 詳細については、担当部署に聞いていただいた方がよいだろう。
* 今までの取り組みに関しては、
  + 例えば、2009年度補正予算で行った当館所蔵資料のデジタル化では約200万冊以上の資料がデジタル化されたが、著作権処理が終わっていないものは、当館内でしか利用できない。
  + しかし、2012年6月に著作権法が改正され（6月15日衆議院で修正可決、6月20日参議院で可決（第180回国会）。6月27日付公布。平成25年1月1日施行。）、来年1月以降、絶版に相当する資料については国立国会図書館から公共図書館、大学図書館等への送信が可能となる。
    - 現在、具体的な送信方法について関係者と協議しつつ検討を行っている所である。
  + また、資料のデジタル化については、財政は厳しいが、今後も、必要な予算を確保し、電子化を進めていきたい。
* 情報技術を活用して社会全体の知的活動の基盤を構築し、日本における知的活動をさらに活発にすることの重要性を常に指摘されていた。（いわゆる知識インフラの構築）
  + 国立国会図書館だけで実現できることではないが、様々な関係機関と協力・連携しながら、日本の知的活動を支える仕組みを、今後も充実させていきたい。

### 海外の図書館における資料保存・デジタルデータ保存についてはどういう面を参考にしたいとお考えでしょうか。

* 海外では、とくに欧米で、デジタルデータの保存について盛んに研究がおこなわれている。
  + とくに日本ではあまり行われていない、論理保存の取り組みが盛んである。
  + 今後、電子書籍が普及すれば、フォーマットの問題など、論理保存がきちんとできなれば、将来、その電子書籍を利用することができなくなってしまう。当館も、今後、論理保存の取り組みを強化していきたい。
* さらに、海外ではDRAMBORA[[19]](#footnote-3)のような、デジタルデータ保存のリスク評価を行う活動や、LIFE[[20]](#footnote-4)などのデジタルデータの保存コストに着目した研究も行われている。これらについても、参考としたい。
* 海外では、技術的な側面だけではなく、保存するための活動を行う機関・組織の政策や、機関・組織自体の長期的な安定性についても、議論が進められていると聞いている。
* 紙の本であれば、たまたま倉庫の隅に残っていた、というものでも、読むことができるが、8インチのフロッピーディスクが発見されても、機器やソフトウェアが維持されていなければ、読めるわけではない。
* 保存には継続的な取り組みが必要であり、関係する機関・組織が果たす役割は大きい。
* その一方で、個人や家族の記録がデジタルに移行している状況を反映して、私的なデータ保存活動に関する啓蒙などに取り組んでいる例もあるようだ。
* 様々な点で、海外の事例に学ぶべき点はまだまだ多いと考えている。

### 現在のデジタルデータの保存技術への要望を実際に資料保存の現場にいるお立場からお聞かせ願えればと思います。

* ストレージについては、当館がデジタルアーカイブの構築に取り組み始めた以降も随分安くなった。しかし、運用費用も含めるとまだまだ安いとは言えない。
* かといって、クラウドでのデータ消失事故のようなことはあってはならない。
* データを消失しないことを前提としながらも、運用コストを含めた妥当なコストでデータを保存する技術の確立を望みたい。
* ビット列の保存のための物理保存については、
  + 今のところ、長期に安定して保存できるメディアによる解決と、HDDのようにある程度壊れることを前提としつつ、モニタリングと適宜の交換等の運用で解決する方策が考えられるが、どちらも適材適所で使えるよう、さらに発展していくことを期待している。
* 再生可能にしておくための論理保存については、
  + ファイルフォーマットを標準化し、技術仕様の入手を容易にすることで、再生のための技術を維持する方法などがある。
  + エミュレーションという方法もありうるが、知的財産権の問題を解決しなければ、過去のOSですら正規には入手困難（例えば、Windowsの古いバージョンは、Javaがらみの特許侵害問題で、新規の購入ができない）な状況でもあり、制度的な側面も含めて、より議論が深まることを期待したい。

### 注目されているデジタルデータの保存技術があればお聞かせください。

* ビックデータを扱う（並列）分散処理システム、分散ファイルシステムやバックアップ技術
  + 分散ストレージは導入しているので、今後実際に運用しながら評価していくことになる。東日本大震災をきっかけに大規模災害時のディザスタ・リカバリが話題になっているが、当館で扱うようなPBを超えるデータを遠隔地にリアルタイムで保存できる技術が必要となるだろう。
* 複数のクラウドの組み合わせ
  + 中小規模のデータについては、複数の民間クラウドサービスを組み合わせて、相互に同期させることでデータの保存を図るサービスなども、海外では出てきているようだ
  + （ 例えば、Duracloud。Amazon、Microsoft等のクラウドサービスをストレージとして使いつつ、同じデータを複数クラウドに保存することで、安全性を確保しているようです。）。国内でも、様々な技術やサービスが登場してきてほしい。

### 今後の国会図書館の資料保存、デジタルデータの保存の展望をお聞かせください。

* 今後も増え続ける膨大なデジタルデータをどのように保存し、利用可能としていくかが課題となる。
* ストレージのコストは今後も安くなるだろうが、膨大な情報を検索・閲覧できるようにするアプリケーションとそれを動かすサーバも必要であり、システム全体のコストを抑える必要がある。
* また、論理保存の技術をどうやって確立するかも課題となる。
* これまで当館が扱ってきた、デジタルデータは、デジタル化資料（画像データ）やインターネット資料が主だったが、オンライン資料（いわゆる電子書籍・電子雑誌に相当するもの）の収集が始まり、その何年か後には、再生機器、再生アプリケーションの問題が発生する可能性がある。そのときにどのようにして、デジタルデータを利用可能とするかが課題となるだろう。
* この問題は、国立国会図書館だけで解決できるものではない。
* 技術者・研究者の方々や、国際標準を策定してきた機関、ソフトウェアを開発、販売し、業界標準を推進させてきた企業など、様々な関係者が、現在の利用者、将来の利用者の利用を想定して、協力して取り組んでいくことが必要だ。
* そのためには、まずは課題がある、ということをもっと広く知っていただく必要があるだろう。御社にも引き続き、協力をお願いしたい。

# （2012年）国立国会図書館と出版界の連携（日仏シンポジウム講演内容）

2012年11月18日　日仏シンポジウム「デジタル時代の本のゆくえ」第2部デジタル書籍と出版業界について（ずぼん19号ポット出版）

## はじめに

　国立国会図書館電子情報部長の中山です。私からは、国立国会図書館が民間と連携することによって、電子書籍の普及や出版産業の発展にどのように寄与できるか、現在の取り組みを中心にお話ししたいと思います。

　まず前提として、現在、膨大な量のデジタルな出版物がインターネット上に流通しているという事実があります。それらを網羅的に収集・保存して、将来にわたって利用を保障していくのが、国立国会図書館の役割です。しかし今の時代、それを私たちだけで実現することはもはや不可能です。今後、電子書籍の普及はもとより、デジタル化資料の利活用を促進するためには、国立国会図書館以外にも国のさまざまな機関の施策を活用し、さらに民間ビジネスとの連携も強化していく必要があります。

　例えば国立国会図書館が収蔵する全資料の四分の三はまだデジタル化されていませんが、それを国立国会図書館の資金、技術だけで実施するのは困難なのです。特に、デジタル化資料の利活用に不可欠な本文内容の検索が可能なテキスト形式での電子書籍化は、国や民間の資源や研究開発成果を活用しなければ実現できないでしょう。この前提に立ったうえで、これから電子書籍の普及を目指すにあたり、出版社と国立国会図書館はいかに連携していくべきかを考えなければなりません。

　ここでポイントとなるのが、電子書籍ビジネスのプラットフォームの整備、電子書籍の著作権処理の集中管理や、電子書籍の仕様決め、すなわちデジタルフォーマットやメタデータ等の共通化などです。以下、順番に紹介していきたいと思います。

## 電子書籍の保存

　国立国会図書館では、電子書籍はデジタルアーカイブ、電子書庫によって永続的な保存を保障しています。分散ファイルシステムにより、段階的に拡張可能な形になっている電子書庫を、複数の拠点に分散して保存するわけです。この電子書庫は、出版社のデータベースのバックアップ機能、ディザスタリカバリ（disaster recovery ／災害で被害を受けたシステムを復旧・修復すること。または被害を最小限に押さえるための備えをすること）の役割を果たすものでもあると考えています。

## 資料や資源への一元的なアクセス

　国立国会図書館では、電子書籍を含めた国内の刊行物の目録を作成し、維持しています。目録では、紙資料・デジタル資料など形態の異なる書誌の情報も共通化し、統合的に検索できるだけでなく、その資料が保管されている場所までナビゲートできるシステムを準備しています。このシステムは「NDL search」?▼注4?という名前で公開されています。

## デジタル化したコンテンツの活用

　国立国会図書館でデジタル化した資料は、出版デジタル機構などを通じてその本の出版元に提供するモデルも想定されます。出版社は、それを自社で電子書籍化して販売することができます。今後は、出版社がデジタル化して出版したものを当館で収集し、当館の電子書庫の中に、ダークアーカイブとして保存し、電子書籍サイトで障害が発生した場合は、この電子書庫から提供していくということも考えられます。当館がデジタル化したコンテンツの利活用の仕組みに関しては、現在文化庁でも「著作物の流通と利用に関しての円滑化に関する実証実験」が行なわれており、その中でより有効な権利処理のガイドラインやビジネスモデルが検討されているところです。

私たちとしても、その実証実験の結果を当館のサービス構築に生かしていきたいと考えています

## 公共図書館へのデータ提供環境の共通化

　二〇一四年一月以降は、国立国会図書館がデジタル化した資料のうち、絶版本など市場で流通困難なものについては、当館から公共図書館にデータを配信できるようになる予定です。それに向けて現在、公共図書館における電子書籍の利用環境の整備を進めています。商用の電子書籍サイトや、国内の各種電子書籍販売サイトはもとより、アマゾンのキンドル、GoogleのGoogle Play ブックスなども含めて、公共図書館におけるデジタル資料の利用環境を共通化していきたいと考えています。

　電子書籍そのものの仕様についても、今後も関係機関と連携しながら共通化を進めていく予定です。

## 著作権情報の集中管理

　電子書籍の著作権情報管理については、国立国会図書館と出版業界が一緒に集中管理する仕組みを、両者共同で構築・運営できないかと考えています。両者が著作権情報を共有すれば、権利処理の効率が上がるだけでなく、双方が同じ資料を重複してデジタル化するなどの無駄も排除しやすくなるからです。著作権データベース構築にあたっては、当館がこれまで行なってきた著作権処理の経験を生かしながら、関係府省や民間と協力しあい、著作物の著作権ステイタスを一元的に把握できるシステムを作りたいと思っています。これも、現在文化庁では「著作権の集中管理の促進に関する実証実験」を行なっており、今後プラットフォームの構築や運用に関する提案が行なわれることになっています。

当館もその実験に連携することで、より有効な著作権情報の集中管理システムを実現したいと考えています。

　そもそも国立国会図書館では、紙かデジタルかを問わず、あらゆる資料の情報を登録できる書誌データベースを構築しています。ここには国立国会図書館の蔵書はも

とより、国内の公共図書館の蔵書、さらには商用サイトで確認できる資料の所蔵機関の蔵書情報も登録されており、現時点で七〇〇〇万件以上のメタデータ?▼注5?が管理されています。既にAPI?▼注6?を使って、その情報をインターネット上で公開できるようになっており、この書誌データベースと、新たに構築する著作権情報のデータベースを組み合わせれば、その本の書誌的事項と著作権の状況を常に把握できるようになります。そうすれば、書籍の権利処理を迅速に行なうことができ、デジタル化もより一層進めることができるでしょう。

## おわりに

　以上のことをふまえて、電子書籍ビジネスにおける国立国会図書館の役割を改めて考えるとすれば、それは「電子書籍の普及によって読者人口が増え、出版全体の市場が拡大して、出版ビジネスが加速されていく」という流れを支援することだと思います。その背景には、資料や情報などの文化的資産を長期保存し、その利用を将来にわたって保障する、つまり利用者にとってはラストリゾート（最後のよりどころ）、提供者にとってはダークアーカイブとなるという、国立国会図書館の役割があるわけです。私たちはそのいずれにおいても関係機関と連携しながら、今後さらに対応を進めていきたいと思います。

自己紹介

中山正樹●なかやま・まさき

国立国会図書館電子情報部長。

1955年生まれ。2002年国立国会図書館入館。電子図書館事業の企画立案、デジタルアーカイブシステム、図書館サービスシステムの構築・運用を管理。2011年より電子情報部長として、資料のデジタル化、オンライン資料の収集・提供等、関係機関と連携した電子情報サービスの企画・構築、東日本大震災アーカイブの構築事業を統括。また、CIO（情報化統括責任者）として、次世代図書館サービスの構築に向けた業務システム最適化計画を策定中。業務・システムを効率化し、より創造的なサービスの実現を目指す。

注4?NDL search……国立国会図書館サーチ。2012年1月6日よりスタートした、国立国会図書館の検索サービス。国立国会図書館をはじめ、全国の公共図書館、公文書館、美術館や学術研究機関等が持つ情報を検索できる。2014年1月末現在、90個のデータベースから収集した約9,400万件の文献情報等が検索可能。目次情報や資料の本文全文を対象とした検索、近日中に刊行される図書（ 近刊図書）の検索をすることができる。

http://iss.ndl.go.jp/

注5◆メタデータ……データそのものに付与されているデータのこと。

注6◆API（Application Programming Interface）……開発者が、外部のソフトウェアやウェブサービスの機能を利用するための仕組み。

# （2012年）資料組織化部会報告【将来構想会議】

平成24年11月19日

資料組織化部会

資料組織化部会では、「私たちの使命・目標2012-2016」で示された使命及び目標の内容を基本的な姿勢とし、紙、電子等の媒体の別又は当館における所蔵の有無に関わらず、利用者が多様な情報資源を一元的に入手及び活用できるために必要な組織化のあり方について検討した。その結果を報告する。

## 基本方針

#### 対象

1. 当館における組織化では、以下の情報資源を対象とする。

* 当館が収集、保存及び提供する多様な媒体の資料
* 書誌データ、典拠データ、個別資料に関する情報[[21]](#footnote-5)等の二次情報（以下「メタデータ」という。）

これら情報資源の収集（受入）、保存及び提供の各目的に応じた組織化の内容と方法を検討する。

1. 当館が収集・保存しない資料についても、利用者が確実にアクセスできるよう、他機関との交換及び共有を目的としたメタデータの調整を行う。
2. 特に、東日本大震災の記録等に関しては、収集範囲としている主題、形式のものを網羅的に組織化し一元的に検索できるように調整する。

#### 観点

　次の5つの観点から、今後、当館が進める組織化の方向性を検討する。

1. サービスの利便性向上
2. 媒体に関わらずメタデータを一元的に扱えるフレームワークの構築
3. メタデータの網羅性・信頼性の向上
4. メタデータ作成作業の効率化及び迅速化
5. 国立図書館としての貢献

## 具体的方向性

### サービスの利便性向上

#### 利用者の多様なニーズに応じたサービス実現に必要なメタデータの検討

　利用者が的確かつ容易に資料を発見、識別、選択及び入手できるサービスを実現するために必要なメタデータを検討する。また、当館が提供するメタデータの利活用による新たなサービスの創出を促進するため、インプロセス段階も含めたメタデータの多様な提供を実現する。

メタデータには、資料の検索及び識別に必要な書誌情報及び典拠情報に加え、提供、保存等に必要な管理情報及び権利情報等も含まれる。さらに、著作と体現形（電子媒体及び紙媒体）の関連性（例：同一著作のデジタル化した画像及び原資料）を表現するメタデータ、及び目次、要約、主題情報等を充実させる。また、刊行形態が電子媒体の雑誌記事索引採録誌についても継続的な記事情報の提供を目指す。

特に、電子媒体の資料（電子情報）について、資料本体のテキスト情報や付属するプロパティ情報等から必要な情報を自動的かつ正確に抽出し、メタデータとして補完できるような技術の適用可能性を調査するとともに、当館の組織化における適用範囲も検討する。

また、あいまい検索や連想検索、異なる媒体又は版の資料（著作）をまとめて表示するサービス等の要件を洗出し、組織化において実現可能な範囲を明確にし、その実現のための技術を調査する。

#### 電子情報及びメタデータの開放性向上

インターネット上で利用者が様々な方法で入手し活用できるような電子情報及びメタデータの提供を目指し、必要な取組や法的・技術的課題を検討する。また、「電子行政オープンデータ戦略」[[22]](#footnote-6)（2012年7月4日IT戦略本部）の基本原則に準じ、各省庁及び関連機関[[23]](#footnote-7)等の検討・取組状況を踏まえながら、他機関との連携も視野に入れ、当館が所有する電子情報及びメタデータの積極的な提供を目指す。必要に応じて、「国立国会図書館作成データベース提供方針」（平成20年国図総080717001号）等の当館内の方針も見直すこととする。

### 媒体に関わらずメタデータを一元的に扱えるフレームワークの構築

#### MARCに替わるフレームワークの構築

　従来扱ってきた資料だけでなく、電子情報も一元的に扱うためには、MARCフォーマットには限界があり、Linked Open Data[[24]](#footnote-8)に対応する等ウェブ環境に適したフレームワーク[[25]](#footnote-9)を構築する必要がある。既に米国議会図書館（LC）が進めているMARCフォーマットに替わる新しい書誌フレームワークの開発[[26]](#footnote-10)動向等に注意を払いつつ、日本の言語環境等に適したフレームワークを検討する。そのため、以下の取組を進める。

* メタデータのセマンティクス（意味）の共通化に留意し、各要素及び要素の記述方法の標準化を進める。
* データ間に同一著者や主題等の関連性を持たせることで、利用者が一つのデータから様々な資料にたどりつけるようにする。
* 書誌レコードの単位でなく、書誌レコードを構成するデータ間で対応付け（リンク）できるようにする。
* FRBRの概念モデルで表現されている「著作」「表現形」「体現形」及び「個別資料」等の実体間の関連性を記述できるようなフレームワークを構築する。
* ウェブ上の多様な情報資源と相互運用性が高いRDF（Resource Description Framework）で表現できるようにする。

#### メタデータ作成基準の作成

　今後の当館におけるメタデータ作成基準の方向性について検討する。具体的には、平成24年度末を目処に、日本目録規則（NCR）に準拠して作成している和図書等の書誌データについて、現在改訂中のNCR又は国際的な目録規則であるRDA（Resource Description and Access）のいずれに対応するか決定する。

　また、紙と電子の媒体の種別ではなく、インターネット資料、古典籍、地図資料等の資料の種別ごとの特性に応じ、書誌及び典拠コントロールの範囲や記述の精粗のレベル等を検討し、メタデータの作成基準を策定する。

### メタデータの網羅性・信頼性の向上

#### 統制語彙を用いた検索のための典拠データ等の拡充

信頼性の高い検索に資するよう、典拠データの作成対象の拡大、並びに主題情報及び各種コード類付与の拡充を行うため、以下の取組を進める。

* 典拠データの作成対象を日本語以外の外国刊行資料、博士論文、雑誌記事索引、電子情報等に拡大する。
* 典拠データの種類を統一タイトル、ジャンル形式等に拡充する。
* 現状においてNDLSHやNDCを付与していない資料群への付与や、運用していないコード類及び標準識別子等の運用を開始する。
* 検索語の推薦機能（サジェスト機能）の範囲と精度を向上させるため、統制語彙を充実させる。また、JST等の関連機関が管理するシソーラスや辞書サービスとの連携を検討する。
* 今後も全ての典拠データに一意の永続的な識別子を付与する。
* バーチャル国際典拠ファイル（VIAF）[[27]](#footnote-11)等を通じて、同一実体に関する別機関の典拠データ等と関連付ける。他言語で作成された典拠データを所有する関係機関との連携を促進し、海外からの日本語資料の入手可能性を向上させる。また、日本の利用者による海外の他言語資料の検索を容易にする。
* 複数の典拠データベース間の関連性についても公開、共有することで、その関連付けを利用者が活用できるようにする。

#### 永続的識別子の付与

　　利用者が情報資源を一意に識別し、確実にアクセスするため、永続的な識別子を付与する。識別子には、文字列ではなく、機械可読形式であるURI（Uniform Resource Identifier）を可能な限り用いることとする。URIを用いることで、情報資源を自動的に関連付けることができる。今後、当館における識別子付与の体系、対象及び単位を検討する。また、ジャパンリンクセンター（JaLC）で付与するデジタルオブジェクト識別子（DOI）や一般に普及している標準的な識別子との関連性に留意する。

#### メタデータの補完

メタデータの内容をより充実させ、利用者のアクセス可能性を向上させるため、メタデータ作成部署以外が付与する情報や、信頼性を担保した外部資源の活用可能性を調査・検討する。たとえば、当館が作成したレファレンス情報、解題等もメタデータに付加する。また、当館が提供するメタデータに対し、利用者がコメントや注釈を行ったり、既存の書評情報やWikipedia等の辞書と紐づけたりできるようにする。また、これらの付加された情報を活用し、利用者ニーズの把握や検索結果の表示順の改善等当館のサービス向上に資することも検討する。

### メタデータ作成作業の効率化及び迅速化

国内及び国際標準への準拠に留意するとともに、出版・流通業界、関係機関等と連携の上、様々な資源、知識、技術を活用し、効率化及び迅速化を図る。

#### メタデータの作成

　次の方法により、記述規則の簡素化・合理化を図る。

* メタデータの品質を確保するため、記述する要素は削減せず、記述の仕方を簡略化する。
* 紙と電子の媒体の種別ではなく、インターネット資料、古典籍、地図資料等の資料の種別ごとに定めた作成基準に基づき、メタデータの均質化を図る。
* 特に、全国書誌で提供するオンライン資料及び従来の図書館資料のメタデータについて、それぞれの粒度を検討する。
* 既にプロパティ情報等としてメタデータが付与されている電子情報を有効活用するための技術を調査する。
* タイトルや識別子等の検索や識別に必要な最小限のメタデータと、分類、件名等が付与された精緻なメタデータのそれぞれの内容を検討する。

#### メタデータの流通（外部作成書誌情報の活用と、外部への書誌情報の提供）

外部機関が作成する書誌情報（民間MARC・出版情報）の当館における活用方法を検討する。

民間MARCの調達が不安定なことに起因するデータ品質維持のコスト増大、また、民間MARC利用がメタデータのオープン化の妨げになる等の課題の解決方法を検討する必要がある。そのため、民間MARCや出版情報と当館が必要とするメタデータの内容や対象範囲について相違点を具体的に洗出し、利用可能性を調査する。必要に応じて、相互に利活用可能なデータとなるよう、調整することも視野に入れる。

　　新着書誌情報等の当館が作成したメタデータの外部への提供については、利用者（利用機関）のニーズに応じ、メタデータの開放性を高め、再利用、加工に関して限定を設けない等の提供方法を検討する。

### 国立図書館としての貢献

#### 全国書誌の提供範囲の拡大

国立国会図書館法第7条に規定する｢日本国内で刊行された出版物｣に該当する電子情報のメタデータについて、新たに全国書誌に収録し提供する。平成25年7月から開始されるオンライン資料の制度収集開始までに、IFLAガイドライン「デジタル時代の全国書誌」[[28]](#footnote-12)等を参考に、具体的な提供範囲や方式等を整理する。

これにより、従来の資料のみならず、電子情報のメタデータについても、国際的な流通に貢献することができる。

#### 国内で流通するメタデータの調整

　　当館が保有しているか否かを問わず、国内で流通・交換するメタデータの調整機能の確立を目指す。また、書誌情報だけでなく、関係機関における資料のデジタル化や著作権処理の状況についても調整・管理することで、重複作業を防ぎ、効率的に連携体制を構築することができる。そのほか、当館が日本センターとして管理しているISSNについて、電子雑誌への普及を促進する。

### 体制と資源配分の見直しに向けて

以上の観点から、当館における組織化業務の実施体制と方法を見直す必要がある。また、人的資源、外注予算等の資源配分についての見直しも必要である。

具体的な見直しについてはさらに検討を要するが、ここでは、見直しにあたっての基本的な考え方を示す。

#### 業務・作業体制の見直し

組織化業務体制を、大きく次の業務に分けて見直しを行う。

1. 書誌調整・標準化
2. 記述
3. 分類・標目付与

（ア）、（ウ）について、紙、電子等の媒体の別にかかわらず、組織・体制の統合による効率化について、検討を行う。

（イ）については、組織化の前段階である収集と連動する業務であり、資料形態・種別による業務内容等の差異が大きいため、収集業務に合わせた体制を想定する。

#### 業務・作業体制の見直し

組織化業務は、方向性として、システムによるノウハウの蓄積と自動組織化を目指すが、今後5年間において、人的資源なしで実施できる形にはなりえない。

職員の行う作業については、図書館員としての高度な専門知識やノウハウを要する作業を主に担う方向で見直しを行う。

* + - * 具体的には、著者名典拠の新規作成や主題分析及び主題アクセスポイント（件名、分類）の付与、典拠作成に主軸を置く方向で検討する。
      * 外注に関しては、これまでの作業に加え、メタデータ補正作業の一環として、メタデータと既存の著者名典拠とのリンク作業など、範囲を拡大する。
      * 記述に関しては、可能な限り、外部作成資源を活用する。また、既存のメタデータの記述内容の補正作業の負荷を減らすため、ある程度の正確性が担保された外部作成資源（民間MARC、出版情報等）の活用可能性を検討するとともに、メタデータの作成基準を見直し、表記のブレ等について可能な限り許容する方向で省力化を図る。それにより、紙、電子等の記述に関する資源配分を見直す。

# 

# （2013年）OpenGLAMセッションでの手持ち資料（図書館総合展）

平成25年10月30日

電子情報部　中山

## 私の理念

* 知識インフラの目的、文化と産業、権利と流通
* 2002年にNDL
  + 立ち位置は、図書館におけるイノベータ
* 知識インフラの構築を目指す
  + 様々な情報を保有している個別に保有している情報を、それぞれの業種業態で集約し、国全体で共有⇒分散アーカイブの構築とポータル。
* NDLは納本図書館であるが、国内刊行物の全てを所蔵していない
* デジタル化時代では、インターネット上の全ての情報を収集することは不可能。
* NDLは、全ての刊行物を保有していると思われているし期待されている
* 当館が所蔵しているものだけでサービスを行っていても、国民の期待には答えられない
* 関係機関との連携が必要。それが、分散収集・保存と検索サービスにより、利用者から見て一つの大きなデータベースとして見えるようにすること
  + 国全体で考えるとNDLはアーカイブ機関の１つ。
  + あらゆる分野でそれぞれの組織がそれぞれの使命によりアーカイブして提供しているものの集合体を目指す。
* これが知の共有化と、その知識の再利用による新たな知識の創造（再生産）を目指す「知識インフラ」に近づいていくと考える
* 知識インフラの構築を目指す
  + 様々な情報を保有している個別に保有している情報を、それぞれの業種業態で集約し、国全体で共有⇒分散アーカイブの構築とポータル。
    - 立法・行政・司法機関、自治体・公共図書館、学術・研究機関、マスコミ、民間、個人
  + あらゆる分野として、MLAは中核の１つ。
  + 大震災関連情報に特化したものが、東日本大震災アーカイブ（愛称ひなぎく）
* その理念の実現を目指して、デジタルアーカイブ事業を推進している
* NDLは、分散アーカイブの構築と集合体としてのポータル機能の構築のインキュベータの役割を果たす

## 連携が進まない理由

* そもそもデジタル化の目的が違う？
* 統合的に利用できるようになると、個別のサービスが使われなくなる？

## 本日の議論の方向性

* OpenGLAMの説明
* この後の発表の選択理由の説明
  + まち歩きアプリ：OpenGLAMによる開発・実装事例
  + Wikipedia Town：図書館がオープン化する事例
  + 総合資料館新システム：システム面でのオープン化事例
* 文化機関が拓く、文化機関を拓くオープンデータの世界
  + これらの文化機関（MALUI/MLA/MLAK/GLAM）を外と内の両面から拓いていくことを目指す。
  + 海外ではOpenGLAMという活動が、The Open Knowledge Foundationの活動の一環として展開されている
  + 【岡本】国内の図書館、博物館等のデータをオープン＆デジタルにし、それらをNDLのシステムと連携させる。
  + NDLとして
    - 趣旨に賛同
    - 当館が進めてきた、デジタル化、デジタルアーカイブ、ポータルの構築において、関係機関との連携を模索してきた活動は、まさに、OpenGLAMの方向性と一致していると思う。
    - 国内刊行物、ウェブサイト、オンライン資料（いわゆる電子書籍・電子雑誌に相当するもの）を収集して、アーカイブを構築している機関として、中核的な役割を果たすべきと考える

## OpenGLAMとNDL（概要）

### 知識インフラの構築に向けて

* 新たな知識の創造と還流のイメージ図

### 国としてのデジタルアーカイブの構築

* 2004年電子図書館中期計画
  + 当館としてのデジタルアーカイブの構築
  + 当館資料のデジタル化、ウェブアーカイブ、電子書籍・電子雑誌の収集により、デジタルアーカイブ構築
* NDLSearchのサービスの目指したところ
  + 従来の資料に加え、デジタル情報のさらなる活用
  + 国立国会図書館の多様な資料・情報、サービスの一元的な利用
  + 当館以外の機関等の情報・サービスに対する統合的なアクセス
* NDLSearchのコンセプト
  + 公共図書館、大学図書館、公文書館、美術館や学術研究機関等との連携や外国からの利用も目指す一元的な検索ポータル

### デジタル情報資源ラウンドテーブル

* 2008年準備会、2009～2011年、MLA関連機関で3回開催
  + 問題意識が共有されたものの、その後の具体的な活動に結び付いていない

### MLA連携に関連したNDLの活動

### 大震災とMLA

* 東日本大震災アーカイブの目的
  + 大震災の歴史的記録・記憶を、国全体で、散逸・消滅する前に収集し、文化遺産として後世においても利用できるように保存する
  + 蓄積した記録・記憶を教訓として、復旧・復興・減災の新たな対策の知見として活用する
* 関係機関との連携
* 刊行物に限らず、あらゆる記録を収集・保存することは、1つの機関で行うことは不可能
* 網羅的な収集・保存及び利活用のためには、様々な機関での分担と、連携が必要

### 国のオープンデータ戦略

* 「電子書籍の流通と利用の円滑化に関する検討会議報告」（平成23年12月21日文部科学省）
* 「知的財産推進計画2012」（平成24年5月29日知的財産戦略本部）
* 「新たなICT戦略に関する提言」（平成25年5月21日、自由民主党）
* 「二次利用促進のための府省のデータ公開に関する基本的な考え方（ガイドライン）（案）」（平成25年5月24日電子行政オープンデータ実務者会議）
* G8サミットにおけるオープンデータに関する合意事項の概要（平成25年6月18日）
* 「世界最先端IT国家創造」宣言（平成25年6月24日閣議決定）
* 日本のオープンデータ憲章アクションプラン（平成25年10月23日、CIO連絡会議）
* オープンデータの取組みに先行して取り組んでいる
  + 当館では、以前より、当館を含めて関係機関が保有する資料の書誌情報を国際標準に従った機械可読形式で公開し利活用されている。
  + また、当館保有の資料を可能な限りデジタル化し、公開している。さらにデジタル化資料の二次利用が、可能な限り自由に行えるように、また手続きが必要な場合でも簡素化に取り組んでいる。
  + さらに、当館が保有するデジタルアーカイブを含めた「ナショナルアーカイブ」の構築も目指している。
* Europeanaの動きとの比較
  + 組織の業種・業態の壁を越えて、デジタルアーカイブ、出版物の目録データベースを検索するためのポータルサイトと、各デジタルアーカイブが融合して、「ナショナルアーカイブ」を構築
  + 「日中韓電子図書館イニシアティブ会議」（CJKDLI）で合意された「各国のポータルの相互連携の実現」して、東アジアの「Europeana」と同様のサービスを構築

### 国としてのアーカイブの構築

* 国としての知識の貯蔵庫
  + 文化を後世に残すとともに、利活用を促進するエコシステムの実現形
  + 地域、業種業態毎の知的財産を、ハブとなるセンターに集合して、より網羅性の高い知的財産の貯蔵庫を構築
  + 現在の活動との関係
* 知識インフラと同義であり、NDLが進めているデジタルアーカイブ事業、大震災アーカイブの発展形
  + 関係機関は、デジタル情報資源ラウンドテーブルで想定している機関と同じ

### MLA連携で想定する方向性

* 知の交流により、知として共有
  + 点と点から線を、線と線から面に
  + 個々の組織・個人がハブとなる組織を中心に連合体となる
* そのハブ同士が繋がり、より大きな連合体となる
* 個々の図書館、博物館、美術館は、知の交流の拠点
  + 人々の交流や知的創造を促す拠点
* ライブラリアン、アーキビストの役割
  + 形式知を暗黙知で補う
* 階層型から対等な連携へ
  + 垂直統合から水平統合へ
* 意義の確認、認識の共有から、具体的な構築行動へ
  + 意義、連携の概念の議論でなく、具体的な連携のアクションへ

### NDLの役割

* 国の文化資源への結節点としてのハブ的な機能
* 将来に亘って利用を保証する中核的なアーカイブ機関の一つ

## 当館の姿勢

* 設置根拠、収集対象等の異なるMLAが、モノを媒介に連携することが困難な中、情報技術の進展とともに、館種を越えた連携の手段として、デジタル情報の共有による連携が注目されてきた。
* 当館では平成21年度から23年度にかけて、デジタル情報資源に関するMLAの連携の促進及び諸課題の解決を目指し、館種を越えた協議及び検討を行う場として、デジタル情報資源ラウンドテーブルを運営してきた。
* その場の議論で、MLAそれぞれが抱えるさまざま課題は抽出されたものの、それらの課題の解決はMLA各機関の自助努力によっているのが現状である。
* 今後はMLAの主要なデジタルアーカイブとNDLサーチとの連携を順次拡大する中で、実践的に諸課題の解決を目指すとともに、MLA連携の基盤となるデータベース整備の取組み（OpenGLAM等）に協力する。
* 東日本大震災関連の情報については、博物館・美術館、図書館、文書館、公民館の被災・救援情報を提供するSaveMLAKとひなぎくを連携させる。
* また、日中韓電子情報イニシアチブ等の既存の協力の枠組を活用し、アジアにおけるMLA連携の可能性を探る。

## 国としてのアーカイブの構築

### 方向性

* 知の交流により、知として共有
  + 点と点から線を、線と線から面に
  + 個々の組織・個人がハブとなる組織を中心に連合体となる
* そのハブ同士が繋がり、より大きな連合体となる
* 個々の図書館、博物館、美術館は、知の交流の拠点
  + 人々の交流や知的創造を促す拠点
* ライブラリアン、アーキビストの役割
  + 形式知を暗黙知で補う
* 階層型から対等な連携へ
  + 垂直統合から水平統合へ
* 意義の確認、認識の共有から、具体的な構築行動へ
  + 意義、連携の概念の議論でなく、具体的な連携のアクションへ

## ナショナルアーカイブとは

### 国としての知識の貯蔵庫

* 文化を後世に残すとともに、利活用を促進するエコシステムの実現形
* 地域、業種業態毎の知的財産を、ハブとなるセンターに集合して、より網羅性の高い知的財産の貯蔵庫を構築

### 現在の活動との関係

* 知識インフラと同義であり、NDLが進めているデジタルアーカイブ事業、大震災アーカイブの発展形
* 関係機関は、デジタル情報資源ラウンドテーブルで想定している機関と同じ
  + 特に出版物に中心に捉えたものが、中川勉強会で示されたイメージ図

### 関係機関が協力してWinWinの関係で構築

* 出版社
* 美術館、博物館、公文書館
* 大学
* ユーザ企業、個人
* 政府機関、公的機関

## デジタル文化資源ラウンドテーブルの設置と成果

### 背景

* 文書館、博物館、図書館においては、これまで大量に蓄積されてきた過去の文化資源のデジタル化は、未だ進んでいない。
* また、デジタル化形式の標準化や、各組織・機関の作成するデータベース間の連携についても、その利活用において、多くの課題を抱えている。

### デジタル情報資源ラウンドテーブルの設置

* 当館では平成21年度から23年度にかけて、デジタル情報資源に関するMLAの連携の促進及び諸課題の解決を目指し、館種を越えた協議及び検討を行う場として、デジタル情報資源ラウンドテーブルを運営してきた。

### 目的

* 利用者は、所在場所を意識せず必要な資料を探せるように。
* 検索結果から各機関のDBへ案内することにより、各機関のDBの利活用が進むように。さらに来館して実物を見る利用者が増えるように。

### 状況

* 平成20年3月準備会開催
  + 国立公文書館、東京国立博物館、国立情報学研究所、科学技術振興機構、慶応大学、奈良女子大、国立国会図書館
  + 横断的アーカイブズ論研究会
* 平成20年度前半に、ラウンドテーブルを発足させる
  + 日本国内の博物館、図書館、文書館が、館種を超えたデジタル化による仮想的なコレクションの構築を目指す。
  + デジタル化の進展に係る共通の課題について、その解決に向けた協議・検討を行う。
* 平成21年度にデジタル情報資源ラウンドテーブルを設置して以来、平成23年度まで定期的に会合を実施。
  + MLAそれぞれの現状や連携、メタデータとガイドライン、コンテンツの利活用と長期保存、人材の育成とその活躍の場等について議論がなされ、課題が提示された。
  + この会議では、主要関係機関代表者、関連分野の有識者により意見交換が行われ、それぞれの組織が抱えるさまざまな課題（別紙参照）が抽出され、問題意識が共有されたものの、その後の具体的な活動には結び付いていない。

### 提示された課題

* MLA関係者・研究者間での分野や立場を超えた課題や論点の共有
* 各種調査を通したデジタルアーカイブに関する基礎的な情報・事例の集積
* PORTA及びその後継となるNDLサーチにおける連携先の拡大等
* 近年のMLA連携に関する様々な議論の活性化への貢献

### まとめ

* デジタルアーカイブの枠を越えた組織の在り方的や政策提言的な議論が多くなり、課題の解決のためには、これまでとは異なる取組が必要となると考えられたことから、平成23年度の会合をもってとりあえず一区切りとした。

## デジタル文化資源ラウンドテーブルの今後

### 現状

* 連携の意義は浸透した？
* 第１ステップの統合検索のスキームはできた
  + 必要最低限のメタデータの共通化と、APIの実装を推進する
  + 現在、国立公文書館、国立美術館、人間文化研究機構の所蔵資料の目録の統合検索はできているが、東京国立博物館とは「e-国宝」として一部のコンテンツのみで、全体の統合検索できていない。博物館では、唯一、奈良国立博物館収蔵品データベースのみ。

### 次のステップは

* 所在場所を意識させない統合検索だけでなく、インターネット空間で仮想ミュージアムを。
  + 目録情報の連携から、コンテンツでの連携へ
* 今後の具体的な連携のあり方としては、
  + たとえば「対馬宗家文書」を所蔵する各機関（NDL、、独立行政法人国立博物館、東京国立博物館、東京大学史料編纂所、長崎県立対馬歴史民俗資料館等）が連携してデジタルアーカイブを構築するといった具体的コンテンツをめぐる連携なども検討して行くべきかもしれない。

## 次の協議の場の設置

### 設置の目的

* そこで今年度は、新たに協議会を設置し、各館が保有する文化情報資源のデジタル化の促進とその普及を目指し、平成26年度以降の活動方針、活動計画を策定するための協議を行う。

### 主な協議事項

* 所蔵品・所蔵資料のデジタル化と共有に関する協力領域について
* デジタル情報資源の利活用の促進に向けた協力領域について
* デジタル情報資源の保存に関する協力領域について
* デジタル化等の財源確保のための協力領域について
* デジタル情報時代におけるMLAの人材育成、人材交流について

## 今後のMLA連携の概要

### MLA連携の方向性

### ノウハウの共有

* 図書館の分類、目録等のメタデータ作成手法の一般化
* アーカイブズ手法（記録資料分析）の他分野適用
* 博物館の企画展示や国際交流に関わる手法の共有
* ?

### 活動の連携

* テーマ設定による共同展示
* アドボカシー（文化資源としての重要性アピール等）
* 民間企業等との連携
* 市民教育・生涯学習プログラム・地域振興
* ?

### 資源の共有化

* デジタル資源の統合的利用と長期保存分担
* 1次資料の保管と相互融通（図書館が文書館の役割を兼ねて行政文書を引き取る等）
* メタデータ標準化

### 人材の交流

* デジタル分野に関する研修の共有等
* より高度な共通資格の設定
* ?

## 今後のMLA連携の具体的項目

### 課題抽出

### メタデータ

* 簡易データの作成と公開
* 詳細データを含めた標準化ガイドライン作成
* 資料のタグ付け外部化（フォークソノミー）

### コンテンツ作成（デジタル化等）と利活用

* 各館におけるデジタル化促進と国としての優先順位付け
* 資料二次利用の枠組み整備
* クリエイティブコモンズライセンス等の活用
* ウェブサイト埋め込み機能の提供
* 資料情報APIの公開
* 官民の機能分担（官はデータを公開するにとどめ、利活用の形態は民に任せるなど課金範囲の）

### 法制度と予算

* 著作権処理の簡素化
* メタデータ作成軽減のためのテキスト化
* 民間企業との連携による資金調達
* クラウドサービスの共同利用
* デジタル化予算の獲得
* インセンティブ制度設計（寄付税制・補助金等の財政措置）

### 保存

* デジタルデータ規格標準化に向けた現状調査の共有
* 原資料の保存・修復に関わる活動の広報
* 「研究者資料」等の非従来型資料に関するテーマ別の共同保存体制構築

### 人材育成と雇用保障

* 雇用ニーズの把握
* 新しい専門職像の形成
* 研修・ワークショップの共同開催
* 継続的な事業（雇用）を通じた人材・スキル育成
* 大学との共同事業
* 新業態の拡大による利用拡大と雇用創出

## 今後のMLA連携の留意点

* プロジェクトの進展に応じて、大学、企業（いわゆるMALUI）へと協力・連携の範囲を拡大していくことを想定する。
* MLA諸機関の事情を鑑みて、具体的な成果が見込めるプロジェクト（文化遺産オンラインと国立国会図書館サーチとの連携、国文学研究資料館が中心となり推進する古典籍資料デジタル化プロジェクト等）から開始し、段階的に取組みを拡大していくことを想定する。
* 国立国会図書館と大学、研究機関等との研究開発による成果を、可能な限りMLA協力・連携によるプロジェクトに還元していくことに留意する。

## 知識インフラの構築

* 図4 新たな知識の創造と還流

## 国のオープンデータ戦略

### オープンデータの取組み

* 政府の直近の動きとして「世界最先端IT国家創造」（6月24日閣議決定）で、
  + 「公共データについては、オープン化を原則とする発想の転換を行い、ビジネスや官民協働のサービスでの利用がしやすいように、政府、独立行政法人、地方公共団体等が保有する多様で膨大なデータを、機械判読に適したデータ形式で、営利目的も含め自由な編集・加工等を認める利用ルールの下、インターネットを通じて公開する。」（p.6）
  + 「2013 年度から、公共データの自由な二次利用を認める利用ルールの見直しを行うとともに、機械判読に適した国際標準データ形式での公開の拡大に取り組む。」とされた。
* 当館では、以前より、当館を含めて関係機関が保有する資料の書誌情報を国際標準に従った機械可読形式で公開し利活用されている。
* また、当館保有の資料を可能な限りデジタル化し、公開している。さらにデジタル化資料の二次利用が、可能な限り自由に行えるように、また手続きが必要な場合でも簡素化に取り組んでいる。
* さらに、当館が保有するデジタルアーカイブを含めた「ナショナルアーカイブ」の構築も目指している。

### Europeanaの動きとの比較

* ヨーロピアナ（Europeana）は、絵画、書籍、映画、写真、地図、文献などのデジタル化された文化遺産を統合的に検索することができる電子図書館ポータルサイトである。 欧州連合の欧州委員会が公開しており、欧州連合加盟国(一部非加盟国含む)のデジタルアーカイブ群のアグリゲータを指向している　（Wikipediaより）
* 当館では、国立国会図書館サーチとして、図書館資料資源に留まらず、博物館、美術館、文書館等の文化機関、商用の電子書籍サイトの所蔵資料やデジタルコンテンツを、統合的に検索することができるポータルサイトを提供している。当館のデジタルアーカイブは、「国立国会図書館デジタル化資料」サイトへ、他機関の資料に関してもそれぞれのデジタルアーカイブサイトにナビゲートしている。組織の業種・業態の壁を越えて、デジタルアーカイブ、出版物の目録データベースを検索するためのポータルサイトと、各デジタルアーカイブが融合して、「ナショナルアーカイブ」が構築され、更に、「日中韓電子図書館イニシアティブ会議」（CJKDLI）で合意された「各国のポータルの相互連携の実現」して、東アジアの「Europeana」と同様のサービスの構築を目指す。

### 府省のデータ公開に関する基本的考え方（ガイドライン）

* 【基本原則】
  + 政府自ら積極的に公共データを公開すること
  + 機械判読可能な形式で公開すること
  + 営利目的、非営利目的を問わず活用を促進すること
  + 取組可能な公共データから速やかに公開等の具体的な取組に着手し、成果を確実に蓄積していくこと
* 具体的な取組内容
  + 二次利用を促進する利用ルールの在り方
  + 機械判読に適したデータ形式による公開の拡大の考え方
  + インターネットを通じて公開するデータの拡大についての考え方

### NDLが持つ文化資産

* ビッグデータ
  + 国内刊行物
  + 冊子体資料
  + デジタル化イメージ資料
  + 本文テキスト化資料
  + 電子書籍・電子雑誌
* 書誌・記事索引・メタデータ
* ウェブアーカイブ
* 大震災アーカイブ
  + 大震災に関連するあらゆる記録
* 著作者種別の利活用条件
  + 国が作成したもの
  + 民間が作成したもの
    - 著作権が切れたもの
    - 著作権が不明なもの
    - 著作権が切れていないもの
* 組織の枠を越えて、有償・無償問わず、所在場所へのナビゲート＆コンテンツの閲覧・提供

## NDLSearchの新たな挑戦　（１）目的

### 新しい検索サービスは、

* 2012年1月に、国立国会図書館は、当館の検索サービスの窓口
* 新たに「国立国会図書館サーチ」の提供を開始
* 利用者の情報の探索・発見を支援するためのもので
* これまでOPACが提供してきた蔵書検索サービスを大幅に拡張

### 第一に、当館資料の統合検索

* 電子ジャーナルやデジタル化資料等の様々なデジタルコンテンツを統合検索し、一元的に利用できるように
  + OPACがこれまで守備範囲としてきた紙資料に加えて、
* コンテンツの発見、コンテンツへのナビゲーションのためのツールとしての役割を果たす
  + 所蔵資料のデジタル化やインターネット上で提供、公開されている資料の制度収集等により大規模に蓄積が進んでいるデジタルコンテンツの利用を促すために、

### 第二に、関係機関の資料の統合検索

* 全国の公共図書館、大学図書館等の所蔵資料やデジタルコンテンツを統合検索できるようにする
  + 当館の蔵書やデジタルコンテンツだけではなく、
* 国立図書館として、日本の図書館資料資源を総覧できるようにする
  + 全国の図書館が相互に連携・協力して、国民の情報アクセス向上の基盤となる
* 図書館資料資源に止まらず、博物館、美術館、文書館等の文化機関の所蔵資料やデジタルコンテンツも統合検索の対象とする
  + 各文化機関が所蔵する資料を、空間的な制約を越えて共有し、相互に関連付けて利用できるようにすることで、文化情報資源の利活用のための基盤を整備し、新しい文化・知識の創造に貢献することを目的とする。

### 第三に、外部連携APIの提供

* 機械的連携機能（API）を提供することにより、「国立国会図書館サーチ」の機能やメタデータを活用
  + 国内外で様々な情報サービスが創造され、データが広く社会で活用されるようにすることを目指しています。
* 特に、当館が作成する書誌情報を、公共図書館、学校図書館等に迅速かつ無償で提供する
  + 日本の公共的な書誌情報基盤としての役割を担うことを目的とする

## NDLSearchの新たな挑戦　（2）現在有している機能

### まず、統合検索の対象は、

* 国内外のおよそ200のデータベースです。7,400万件余のメタデータを収集し、検索対象とする。検索件数は、月間170万件程度

### 国内刊行図書については、

* 出版界との連携により、刊行前の書誌情報（出版情報）から、刊行直後の新着情報（作成中書誌情報）、そして完成書誌までを一貫して提供
* これにより、出版から書誌情報提供までのタイムラグを解消。
* また、様々な機関から収集した多数のメタデータに対して、IFLAのFRBR（書誌レコードの機能要件）モデル（目録が果たす機能を利用者の観点から見直し、モデル化したもの）を指向した著作同定の仕組みを適用
* 同一資料や関連資料（形態を異にする同一著作）をグループ化して表示

### さらに、共通プラットフォームの提供

* 国立博物館や国立美術館、国立公文書館が所蔵する博物資料、美術品、文書等を図書館の文献とともに統合検索できるところも特長
  + 例えば、日本最古の歌集である『万葉集』を検索しますと、国立博物館が所蔵する古写本（国宝）の高精細デジタル画像と、全国の図書館や研究機関が所蔵する『万葉集』の研究書や解説書などを同時に探すことができる
  + このように、全国に分散して存在する文化情報資源を国民が有効に活用できるよう、博物館・文書館・図書館共通の検索プラットフォームとしての役割を果たす

### 一方、検索機能の充実

* 膨大な検索対象から必要な情報を的確に見出せるようにする
* ファセットによる検索対象の絞込みや検索結果の適合度順表示、関連キーワードによる検索対象の拡張等の検索支援機能を充実させました。
* 情報の所蔵機関・入手先に案内するという目録としての基本機能も意識して構築
* その他、検索語や検索結果を日中・日韓・日英間で機械翻訳する機能を設け、利用者の便を図っている。
* APIは、データ収集用APIと検索用APIの2種類を提供し、適用しているプロトコルはいずれも国際的な標準に準拠している

## NDLSearchの新たな挑戦　（3）今後の重点的な取組み

### 第一に、連携先の拡張による、統合検索対象の充実に取り組みます。

* 特に、昨年開催した日中韓電子図書館イニシアティブ会議（CJKDLI）で合意された「各国のポータルの相互連携の実現」に向けて、他機関の検索サービスとの相互連携を実現したい

### 第二に、検索機能の向上に取り組みます。

* 集合知を活用した検索支援機能や、データとデータとの関連付けの高度化、利用目的に応じた検索インターフェイスの提供、資料の全文を対象とした検索サービス等の実現を目指します。

### 第三に、メタデータの二次利用の促進に取り組みます。

* 文化情報資源への自由なアクセスが創造的な社会の基盤になるという確信に立って、データのオープン化に関係機関と歩調を合わせて取り組んでいきたいと考えています。

## 大震災とMLA

### 大震災とMLA

* 地域の歴史的記録・記憶の保存と共有
* 蓄積した記録・記憶を教訓とした復旧・復興・減災対策等の知識の創出
  + 研究成果、ファクトデータを研究者が活用

### NDL

* 各地域・機関が保有するの記録・記憶、復旧・復興に関する情報の集約と、連携の結節点（ハブ）としての役割
* 被災した資料の救済
* 被災地の図書館支援

### MALUI連携

* 最近は、MLAに加えて大学（U）と企業（I）を加えたMALUI連携という取組みも始まっており、当館は東日本大震災アーカイブの取組みを通じて、東北大学、東京大学、Google、Yahoo!、などとの連携を進めている。
* 大震災アーカイブ関係では、Lとの連携は進んだが、MとAは、むしろそれぞれの所蔵資料の救出、保存に重点を置いて活動するところが多かったので、具体的な連携には結びついていない。
* 課題あるいは展望（補足）
  + 今後の具体的な連携のあり方としては、たとえば「対馬宗家文書」を所蔵する各機関（NDL、、独立行政法人国立博物館、東京国立博物館、東京大学史料編纂所、長崎県立対馬歴史民俗資料館等）が連携してデジタルアーカイブを構築するといった具体的コンテンツをめぐる連携なども検討して行くべきかもしれない。
  + また、大震災の被災地域等におけるMLA連携も大きな課題となっており、NDLが何らかの関わりができないかについても検討すべきと考えられる。具体的には
* 2013年5月　仙台シンポジウム
  + MLAは、MALUI連携へ発展させる
    - 地域の歴史的・文化的資源の蓄積、記録・記憶を共有できることの重要性
    - 研究成果及びそれらの成果の前提となるデータ等を研究者が活用できることの重要性
    - 情報の共有と蓄積を推進し、新たな記録・記憶の共有を進める企業との連携の重要性
    - ⇒MALUI連携はこれらを結びつけ、文化的・知的活動を活性化し、復興の支えとする
  + NDLは連携の結節点となる⇒各アーカイブ機関への一元的なアクセスを可能とする
* 「大震災の被災地域等におけるMLA連携」とは具体的にどういうことか
  + 「被災地域を中心とした、各地域におけるMLA連携」の課題として、
    - 地域の文化的、歴史的資源の状態を把握し、災害時に対応する仕組みの弱体化
    - 目録の作成、公開のためのノウハウ、人材の不足
    - デジタル化や、地域内・全国への情報発信のための目利きの育成、確保の困難さ
* などがあり、財政的にも厳しい状況で、MLAが個別にこうした課題に取り組むことが難しくなっている。地域内のMLAが連携してこうした課題に対応し、地域の文化的・歴史的資源を、当館との連携によって全国的に共有する仕組みの構築につなげていくような支援策が、今後必要となってくるのではないか。
* これの実現だけでも相当な資源を必要とする

## ★知識インフラの構築に向けたNDLラボの活用

* 科審でのJST/NIIのトップからのプレゼン、第3期科学技術情報整備基本計画に記載されている「知識インフラの構築に向けた研究開発分野における協力・連携」に関して、期を逸せず、具体的なテーマの研究開発・実証実験の実現に向けて、準備を進めている。
* 当館からコンテンツを提供して、関係機関が研究開発して、その成果を利用して、NDLラボで実証実験するイメージ。地道に個別テーマで細々と進めてきた実証実験の範囲が加速されることを期待している。
* また、情報技術を活用した知識インフラの仕組み作りのために、関係機関に限らず、Code４Lib等とシステムの専門家（有識者）のWGを想定
* 当館は、人・物・金の資源のうち、物としてのコンテンツの提供と、場の提供で貢献することを想定。

# （2013年）要件定義の必要性と人材育成（IT研修内容）

平成25年12月17日

電子情報部　中山

## 要件定義の必要性

* 以前は、IPAでソフトウェア生産性向上に関連する事業、セキュリティセンターに従事
* 2002年に入館以来、当館のシステム開発に関して、直接的、間接的に関与させていただいている。
* その間、職員、業者の努力により、他の省庁と比べると、各段費用対効果の高いシステムを構築していると思う。
* しかしながら、予算が大幅に削減される状況において、システム開発の更なる効率化が必要と認識している。

## 知識情報基盤の構築

* 「科学技術基本政策策定の基本方針」として、平成22年6月総合科学技術会議基本政策専門調査会で、国の第4期科学技術基本計画（23年8月閣議決定）に知識インフラ関連の言及を入れてもらうために作られたイメージ図。（当時の科経の本吉課長）
* それを先行事業としての大震災に当てはめたもの。
* 外部プレゼンではよく使われるが、館内ではほとんど使われない。
* 当館の電子図書館中期計画2004、科審での知識インフラの構築、使命・目標で示される当館のデジタル関連の事業の目指すものは大きく変わっていない
* これの実現を目指すサービスの実現のために業務・システムを組み立てているところ
* この概念は、大きな枠組みの中では、当館は情報を集約して提供する組織のOneOfThem。しかしながら、大量の資料を保有しているが故に、分散保存、ポータルとして、中核的な役割を果たすことが求められている。

## 最適化計画の考え方

* 昨年度から策定してきた「業務・システムの最適化」の考え方（川島氏作成？）
* 目的と位置付け
* システムに求められる要件（仮定）
* 最適化の基本方針
* 目指すべきシステムの姿
* 取組み事項
* 実施方法

## 使命・目標達成に向けた考え方

* これも最適化計画策定に当たって示された図（川島氏作成？）
* 最適化計画は、サービスを決めるのではなく、サービスを可能にする計画
* 最適化を行うことで、費用対効果の高いサービスを構築し、それにより、新たなサービスを生む余地を作る

## 館の情報化推進体制

* 今年度の情報化推進委員会で提示した姿勢（25年4月館長作成）
* 情報化の推進を図るために、総合的な考慮が必要
* サービスを実現するために
  + 業務的観点（システム抜きには業務は成り立たない）
  + システム的観点（システム化は業務・サービスを向上させるためにある）
* 実施に当たっては、戦略的な経営資源の配分（人・物・金）
* CEOが行司役（システム化を通じた業務の変更、業務向上のためのシステム変更）
  + それにより、情報化推進委員会の委員長は総務部長とした。

## 情報化のプロセスと考慮点（サービス構築の流れに沿って）（流れ図参照）

電子情報部を設置する時に、各部局との役割分担を共有するために作成したもの。特にサービス要件定義の担当と、成果物の名称を示した。最適化計画本体にも記述されている。

#### 「私たちの使命・目標」、年度活動計画

* あるべき姿の検討、サービスの理念、活動の方向性、サービスの項目と概念、スケジュール感

#### 基本計画書、サービス実施計画書

* 使命・目標の達成のアプローチとして、サービスの提供の方向性、具体的なサービス、システムのイメージの想定
  + 何をしたいか？（サービス部門）
  + 何ができるか？（システム部門）
* カバレージとして、フロントエンドは利用者へのサービス、バックオフィスも職員へのサービス

#### サービス要件定義書

* サービスは、業務とシステムで構築
* 業務とシステムで連携・分担して、フロントエンドサービス、バックオフィス業務として何ができるようにするかを明確にする。
* 主たるサービス要件より、サービスの実現当たって例外的に行うべき要件の明示が重要（後述）

#### システム構築

* システム化要件定義書
  + 業務・機能要件定義、非機能要件定義
  + 開発プロジェクト標準、技術標準適用指針、技術標準適用ガイドラインに沿って
* 見積もり依頼書
  + 予算、実施内容を決定するに当たって重要な書類
  + 必要スキル、工数が算定できるレベルの仕様書の提示が必要
* 調達仕様書
  + 価格競争の場合は、実施すべき事項を明確に提示し、実施内容が業者によってブレないようにして、安い方を落札。
  + 仕様が明確に提示できない場合は、実施内容、方式を提案させる形での総合評価での競争にし、必須項目でも、優れた実施内容、将来への拡張性の配慮の内容には大きく加点する。
* 契約書
  + 総合評価での契約書に添付される仕様書は、提案書相当。
  + 提案書の内容をベースに検査仕様書を作成し、成果物の検収は、提案書の内容が実現しているかを検査する
* システム設計書
  + 開発担当者に向けた資料。
    - サービス要件に従ってシステムで実現することが明確になっている調達仕様書通り作られていれば、ユーザ部門は読む必要がない。
  + 調達仕様書、提案書通り、設計がなされないで、開発を行うと大きな手戻りが発生する
* システム運用手順書（マニュアル）
  + 開発時に、ほとんどゼロ知識でもできるレベルを求める

#### システム運用

以降、今回は時間がないので割愛

#### 業務構築

* 業務要件定義書
* 業務手順書（マニュアル）

#### 業務運用

#### 制度構築（制度に関しても同様）

* 制度設計

## 過去の問題事例

* 前の基盤システム（A社、H社）
  + *最低1ユニット500万での工数見積もり*
  + *同一業者で開発運用しているにも関わらず、影響度調査名目で大きな工数*
  + *ファンクションポイントで工数を算定したとき、同様の更新処理を、数分だけ積んだ。*
* 前のDAシステム（N社）
  + *開発の予算を認識せず、予算・内容ともに実施が不可能な設計*
  + *設計・開発の工程及び成果物のレベルの認識に齟齬（基本設計、概要設計、詳細設計、外部設計、内部設計）*
* NET4（H社）
  + *基本設計書が曖昧すぎて、詳細設計で手戻りが多くなりそうなので差し戻し。*
  + *運用に入ってから、ファイアウォール、端末の移設等、軽微な作業に対して、大きな工数を見積もられている*
* 今の業務基盤システム（E社、N社）
  + *総括メモ（F氏メモ）*
    - *前年度の成果物を要求要件としたが、設計書等が不十分だったため、後工程で大きな工数を要した。*
    - *当館、工程管理支援業者、設計開発業者での作業スコープ、役割、作業の責任範囲が一致していなかった。*
    - *一部担当者の暗黙知として蓄積されていた*
    - *リリースされたプログラムの品質が悪く手戻りが発生した*
  + リニューアル総括資料
    - リニューアル総括（25年6月）参照
    - （教訓）「企画段階で、サービス要件定義について、全館で合意を形成する」
    - システム化要件定義段階で、「サービス要件定義書と、業務・システムで実現可能な内容のFit&Gapについては、実現可能な範囲で合意を形成し、サービス要件定義書にフィードバックする
* NDLSearch（N社、H社）
  + 開発、機能強化、運用と、交互に落札しているようなシステム。ソースコードのリファクタリングが必要になっている。
* 大震災アーカイブ（N社）
  + NDLSearch、DAをベースに、発展形として、両システムの統合の足がかりになることを想定したが、違う方向を向いていった。

## 要件定義の重要性と考慮点

システム部門を担当して、部内、館内で示してきたことを、思いつくまま列挙したもの。

体系的な説明は、CIO補佐官の解説に委ねる。例外的なサービスを網羅的に（網羅性の確保）

#### 重要性は、ソフトウェア開発に限らない

* 組織間、部局課係間、担当者間での曖昧な内容での指示は、作業スキル、必要工数の算定で大きな齟齬が発生し、安全を見込むと大きな見積もりとなり、また、過小評価すると、後で大きな工数を要する。

#### サービス要件の定義は、基本計画書の行間を埋めるもの

* 使命・目標は概念であり、基本方針・基本計画書は、利害調整で玉虫色になっていることが多い。
* 合意形成のために、意識的に、サービスの実施内容が曖昧になっている
* 概念・方向性は認識できるが具体的なサービスの実施内容が判断できない
* 基本方針、基本計画の策定において検討したメモ、カットされた記述が重要。それがベースとなって、明確な要件となる。

#### 開発に大きな工数が掛かるのは、例外処理の規模と将来への拡張性への配慮

* 例外的なサービスを網羅的に（網羅性の確保）
  + 例外的なサービスは条件と実施内容を明確に提示
  + 例外的なサービスの内容が明確でないと、認識の齟齬により次工程でもめる

#### サービス要件定義段階での合意形成が最も重要

* 後工程でシステムの根幹に関わる変更は不可能

#### 業務、システムの構築は、論理的に明確なサービスの要件がなければ構築できない

（システムに限らず、法制度化等も同様）

* 曖昧なまま、業務構築、システム開発を行った場合、過大な見積もり、大きな手戻りが発生する
* 齟齬の顕在化が後工程になればなるほど、工数が大きくなる
* サービス部門とシステム部門で暗黙知でなく、形式知化した形で合意しておく必要がある⇒サービス要件定義書
* サービスの実施内容があいまいなまま、業務構築、システム開発を行った場合
  + システム化要件定義書が曖昧になる
    - 開発規模感が掴めない
  + 調達仕様書が曖昧になる
    - 開発者は安全を見込み、見積もり額が大幅に大きくなる
    - 開発工程で手戻りが発生する
    - 納品されて検収する際に、必ずもめる
    - 仕様書は担当者のスキルレベルに依存しないように明確に
  + やるべきことを明確に（曖昧性の排除）
    - 論理的な思考で、条件と内容を明確に。曖昧な文章表現ではなく、箇条書きで、判断要素は断定的な記述に

#### 業務とシステムでの分担は明確に

* システムでできないことの許容
  + 経営資源、適用可能な技術の面から、システムで実現可能な範囲が小さく、業務で行わなければならないことの負荷を許容できるかは大きな要件
* 業務とシステムの分担は図式化して共有
  + ユースケース図（人とシステムの役割）

#### 情報（データ）と機能の流れは明確に

* 機能情報関連図（DFD図：情報と機能の流れ（サブシステムレベル））、実体関連図（ER図）等で

#### サービス要件は、開発過程、検証中に、なんらかの変更があることを前提とする

* 事前に変更の可能性がある部分は明示が必要。
* 変更による後工程への影響を極力小さくするためにも、早期発見が必要。
* ⇒開発前に要件が明確にならない部分は、プロトタイピング手法が有効

#### サービス要件定義には、全体最適化の観点でシステムサイドからの助言が必要

* サービス部門が主体となってサービス要件定義書を策定するが、サービス要件にサービスの方法、手順が示されていても、それに捉われず、目的を確認して、全体最適化の観点で、より最適な方法がある場合は、システム構築の立場から、助言する

## 各工程での成果物の妥当性評価

#### 全ての工程での仕様書について

* サービス要件定義書、システム化要件定義書、調達仕様書、開発段階での仕様書、作業指示書等、全ての工程で、作業依頼書として、人から人へ伝えていくもの。成果物は、次のステップの作業依頼書となる。依頼書が曖昧であれば、成果物のレベル、内容もぶれる。
* 打合せ等で、意見でなく質問が出されるのは、ドキュメントが曖昧だから。全てが妥当な内容であるか評価が必要。

#### 曖昧性・不確実性の排除

* 「仕様素案」、「基本設計書」、「概要設計書」、「詳細設計書」、、、
  + プロセスと成果物の具体化度を明確にしないで、成果物名のみの提示は齟齬が生じる
  + 「柔軟に対応できること」「原則として～する」「～する場合もある。」はNG
  + 図表を駆使することも明確化に繋がる

#### 読むべき人が理解できるか？

* その仕様書をインプットとして、暗黙知なしで、難易度の認識、実施内容、妥当な工数見積もりできるかを評価

#### 将来のコストを削減するか、当面のコストを削減するか？ も重要な観点

* コストが高くても将来性、柔軟性のある方法
* コストが安いが、当面の課題は解決できるか。

## 参考見積書の評価

* 曖昧な仕様書では安全係数が大きくなり、高額な見積もりになる
* 業者の高額な参考見積もりを鵜呑みにして安易な要件緩和やスペックダウンはしない
* 提案もしくは指示した実施方法が明確な場合は、具体的な作業と要する工数を評価する
  + 評価に当たっては、ある程度の実地の経験は必要
* 競争入札になれば、適正な価格に近づく。随意契約の場合は、妥当と思われるまで調整する必要がある
* **「人件費単価が高いから見積もりが高くなる」という評価は正しくない。**
  + スキルの高い人が担当するので単価が高いなら、工数は少なくなるはず。

## ステークホルダーとの関係での留意点

* 館内での役割分担
  + 会議体
    - 意志決定、合意形成、実施調整・進捗管理
  + 企画部門
  + サービス部門
  + 開発部門
  + 運用部門
* 業者との役割分担
  + 調達支援業者
  + 工程管理業者
  + 開発業者
  + 保守業者
  + 統合運用業者

## まとめ

#### 方針

* これから、さらに財政事情は厳しくなる。
* システムの開発に限らず、費用対効果の高いサービスを構築していく必要がある
* 資源の無駄遣いの排除
  + 必要性の低いサービスの構築
  + 要件定義の不備による手戻り
  + 要件定義が曖昧なための高額見積もり
  + 考慮不足による機能改修の高額化
* 確保した人・物・金の資源の無駄遣いをなくして、よりサービスの向上を目指していただきたい。

#### ソフトウェア開発と取引の諸問題の再確認

* ステークホルダーと、密に連携を取りながら、それぞれの立場でシステム開発に参加、関与し、必要な役割と責任を果たさなければ、満足のいくシステムは開発できない
* 言葉
  + 開発工程と作業内容（成果物の詳細レベル）
  + 保守の範囲（瑕疵、軽微な改修、機能強化）
* 要件定義や役割分担
  + 企画から要件定義、開発、運用に至るライフサイクルを通して、タスク、アウトプットは、暗黙の了解でなく、必ず明文化し、関係者で共有
  + 誰が決定するかも明確にすることが重要
* ステークホルダーの合意
  + 誰がどの段階で何を行えばいいのか、ライフサイクル上で明らかにしておく
* 不明確な取引
  + 外部委託に関して、発注者側からお願いしたい範囲、内容を明確にして、この範囲でお願いするということを明確に伝える
  + 見積もり
    - 要求が曖昧なままでは、見積もりは高額になる
  + 仕様変更
    - 「仕様に含まれている、いない」、「ITのプロなら行間を読め」

## 標準類

#### 開発プロジェクト標準

* システム毎に異なる作業プロセス・成果物の共通化・標準化が必要
  + 受託業者毎の標準に依存しないもの
  + プロセスの区切りと成果物の具体化度を明確に
  + ステークホルダーとの関係、役割分担、対応姿勢を明確に
* 可能な限り、世の中の標準に準拠する

#### 情報システム調達のための技術参照モデル（TRM）平成24年度版

* 分離調達に関する留意点
* スキル
* 役務の分類

## 各種標準の概要

### 情報システム調達のための技術参照モデル（TRM）平成24年度版

#### 分離調達に関する留意点

* 全体アーキテクトの確保
* システムの要件はプロジェクトの初期段階では曖昧で、システム開発が進行するに従い徐々に明確になっていく
* 長期に亘るプロジェクトの場合、途中で要件が変化することも多い
* 全体アーキテクチャの策定を行った事業者がシステムのサービス開始までのアーキテクチャの維持・改善を実現することが望ましい

#### 役務の分類

* 新規開発、システム更改、ハードウェア更改、機能追加（サブシステムレベル）
* 改修（部分改修、機能追加、バグフィックス）
* 開発プロセスを伴わないAP保守単体での調達
* AP保守以外の保守
* ポイント
  + 瑕疵担保はどこでやるか？

#### スキル

* 情報化推進（企画・調整）
  + 経営資源の配分（人物金）
  + システム化を通じた業務の変更
  + 業務・サービスの向上のためのシステムの変更
* 全体管理アーキテクト
  + 技術面の整合性の確保だけではなく、業務面から見た整合性の確保も必要
  + 複数プロジェクト（業者）間の責任分界点を適切に定義できる能力
  + IT システムの設計に関する十分な経験に加え、システムがサポートする業務に対する知識
  + 長期に継続するシステム化対象の継続的な変更への対応が軸になるため、業務についての将来的な見通し
  + 企画・調整から、サービスの開始まで関与
  + アーキテクチャ策定と、個別事業者を統制するプロジェクトマネジメント能力

### 情報システム調達のための技術参照モデル（TRM）に基づく工程とアウトプット

#### サービスの構想

* 「使命・目標」
  + 戦略的目標、年度活動計画
  + サービス実施計画

#### サービスの実施計画⇒サービス要件定義

#### システム化構想立案⇒業務・システム最適化計画

* 現行システムの課題抽出（※現行システムが存在する場合）
* システム化要望の調査
* 技術動向の調査
* システム化構想書の作成 等

#### システム化計画立案⇒最適化実施計画（実施班作業）

* システムの概要検討
* 設計･開発･移行･運用スケジュールの策定
* 概算見積もりの実施
* システム化計画書の作成 等

#### 最適化計画の確認･評価･改善･効果算定（CIO補佐官）

* 最適化計画の確認、評価及び改善提案、最適化効果の算定

#### 要件定義（個別サービス毎）

* 調達対象となるシステム、業務に係る要件定義
* スケジュール定義
* 業務･機能要件定義⇒サービス要件定義から導かれるもの
* システム方式要件定義
* 情報･データ要件定義
* ユーザインタフェース要件定義
* 外部インタフェース要件定義
* ネットワーク要件定義
* ソフトウェア要件定義
* ハードウェア要件定義
* 情報セキュリティ要件定義
* 設計･開発要件定義
* テスト要件定義
* 移行要件定義
* 運用･保守要件定義
* 規模性能要件定義
* 信頼性等要件定義

#### システムコストの試算

* 要件定義に基づいたシステムの構築･運用コストの試算

#### 調達業務

* 調達計画書の作成、調達仕様書の作成、意見招請対応、事業者の評価等、調達関連業務

## 開発・運用の問題と解決に向けたアプローチ

#### 問題点

* 統合運用への引継ぎに手間がかかる
  + 職員担当者だけでなく、業者担当者も工程によって変わる
* 個別システム毎にドキュメントレベル、手順が違い過ぎる→分かりやすい文書とは

#### 再確認事項

* 新しいTRMでの方向性の再確認
* 統合運用の委託の在り方
  + 現在の契約の提案内容の再確認
    - SMO→全体SI→統合運用
* 体制
  + 統合運用業者と職員の役割分担
    - システム運用担当は、どの程度、システムの内容に関する知識、ノウハウが必要か
  + 開発とAP保守と統合運用の役割分担
    - 機能強化と軽微な改修
    - 個別システムの運用
  + 分担の再構成
* 実施準備
  + 開発／運用の標準化
  + 上流工程での運用設計

#### 見直すべき事項（案）

* 運用における業務システム運用係の負荷の低減
  + 個別システムの開発内容の詳細及びノウハウを要しない形にする
  + 障害発生時の個別システムの保守事業者対応は統合運用業者に任せる
  + 開発から統合運用への引継ぎは業者間で直接行う
* 統合運用業者の役割の明確化
  + 旧システムでのSMOの役割の実施
* 業者体制
  + 全体管理（全体SI、調達支援、工程管理支援）
  + 個別システム開発・構築事業者
  + 個別システム開発・構築成果保守事業者
  + 統合運用事業者
    - 個別保守事業者への作業指示、作業結果の取りまとめ、統合運用管理職員への報告
* 職員体制
  + 開発体制
    - 開発、機能強化、瑕疵対応、開発を伴う保守（軽微な改修）
  + 運用体制
    - 統合運用管理
    - 原課との対応

#### 見直しのスケジュール

## IT人材確保

### 国のIT戦略

#### 国の「新たなIT戦略」（IT総合戦略本部）

* 省エネ社会の実現、遠隔医療の実現、自宅で働ける環境の整備等、幅広い分野でIT技術が活用される世界最高水準のIT社会を実現する（「情報資源／データ立国」へ）
* 準備
  + 統廃合、政府共通プラットフォームへの移行によるクラウド化等、今後の方向性を検討する

#### ITの利活用は課題解決の横串ツール（情報資源／データの活用が鍵）

* データの収集（蓄積）／見える化／共有／連携／分析
* これらを可能とするシステムや仕組みづくりが必要

#### 世界最高水準のIT社会の実現に向けて

* 産業再興・経済活性化への貢献（イノベーティブな社会へ）
  + 国民の安全・安心への貢献（レジリエントな社会へ）
  + 行政機能や政策効果の向上を目指した「真の行政改革」への貢献（利用者視点に立った行政のデザインとガバナンスの強化へ）

### NDLにおけるIT活用の位置付け

#### ITの位置付け

* IT はあくまで手段。システムの導入にあたっては、既存業務の見直し、BPR を徹底的に行い、PDCA サイクルをしっかりと回していくべき。

#### ITの活用

* 戦略的経営判断のためのICT
* ERP、BPR、EPM
* 図書館におけるICTの位置付け
* 図書館サービスのためのICT
* 職員業務改善のためのICT
* 広報のためのICT

#### IT人材育成

* ライブラリアンからITスペシャリストへ
* Code4Libに参加できるレベルのスキル

### IT人材育成の必要性

#### IT融合を進めるIT技術者育成

* 作る技術者から価値を生む技術者へ

#### ビッグデータを利用したサービスの創出において

* これまでとは違った発想からアイデアを導き出す人材
* 技術からその産業にこれから求められるサービスまで横断的に分かる人材
* これらを社会に実装する人材など様々な人材が必要になる。

#### 短期的には

* 必要な人材は外から集めてくるというのは正しい

#### 長期的には

* 一定規模の母集団があればという前提であり、この母集団を膨らませるには中長期にわたる施策が必要である。
* また、その産業も魅力あるものでなければ、母集団を形作る若い人材も入ってこない。
* このような母集団の形成を一企業や業界団体でやるとなると、その業界特有の常識が、革新的なサービスを創出するネックになることもある。
* ビッグデータを利用する産業や情報処理についての知識は必要であるが、
* それだけに縛られない垣根を越えた考え方をするようなITによる産業の融合を進めるＩＴ技術者を育成する必要がある。

### 必要なITリテラシー

#### 図書館の組織の在り方

* 日本の図書館がいつまでも前時代的な図書館のままでいいはずがない。
* 経営学には「生き残れる組織」に対する考え方がある。ビジネスの世界で生き残れるのは，大きな組織でも強い組織でもない。それは「変化できる組織である」とよく言われる。時代のニーズ，社会のニーズをとらえ，それをきちんと提供できる。そのために変化することには果敢に挑める。そのような組織。
* 今の図書館を変えたくない，という方も少なからずいることは，耳にしている。ただ，変えないことは組織の老化現象となり，いずれ組織／機関そのものが老化し，いずれ死に至る病となる。
* ICTに明るいライブラリアン像のひとつは，「自らネットやデジタルの世界で遊び，自分で課題を見つけ，自分で検索するなどして解決し，試しては壊し，壊しては創りをできる人」である。

#### 必要なITリテラシー

* これからの図書館員に求められるICT素養（リテラシー）
  + レファレンスサービスとしてのICT素養，サーチャー（情報検索技術者）としてW e b資源，オンラインデータベース等から適切な情報を得て利用者に提供できるスキルが必要
* 利用者サービスのためのICT素養
  + TCP/IPによるネットワークの基礎知識，WiFi接続に対するノウハウ（特にWindowsマシンでは，本体内の無線LANのON/OFFや無線L A Nアダプタごとに異なる接続プログラムの操作方法など）である。また基本的な印刷やUSBメモリなどの取り扱いから始まり，We bサイトを作りたい，動画を編集したい，携帯音楽プレーヤーで音楽を聴きたいなど，実にさまざまなパソコンサポート
* リレーションシップとしてのICT素養
  + 本来のPublicRelations（社会との関係づくり）ととらえ，そこにデジタルやインターネットなどのスキルを使って情報を発信したり，利用者へのサービスを提供すること
* システムライブラリアンとして
  + 図書館としての基幹業務が正常に動作しているか，不具合が発生した場合に対処できるか図書館員自身が行なえるレベルと業者が責任を持って行なうレベルとを把握
* デジタル資料の収集／作成／編集／組織化
  + デジタルアーカイブからデジタル資源管理（デジタル・アセット・マネジメント）に対する考え方

#### 職員のITスキルと組織の現状

* 図書館員一人ひとりの素養は，諸外国の図書館員と比べてもそれほど大きな差は無いのではないか
* 意欲を阻害している要因があること
  + 一人ひとりの取り組みを活かせる仕組みがないのではないか
  + 全国すべての図書館にI C Tに明るい人材がいる，
* ICT素養を発揮し，図書館サービスを刷新している人たちをネットワークでつなぐこと

#### 図書館の意思決定者に対する人材活用のアドボカシー（政策提言）

* スキルを身につけたライブラリアンたちが，図書館の現場でその能力を遺憾なく発揮するための環境づくり

#### 前業務システム最適化、新利用者サービスの計画で目指したこと

* ICTに明るい図書館員がWebを活用した非来館型図書館サービスの充実を図り、各種APIを取り込んだサービスを構築しやすい土壌をつくるため、図書館情報システムに共通のAPIを持つことである。

# （2013年）我が国の知識インフラとしてのナショナル・アーカイブ構想（要約版）

平成26年2月4日

電子情報部

## ナショナル・アーカイブ構築の意義と目標

* 情報化の進展とインターネットの普及により、世界的にコンテンツのデジタル化が進展している。我が国で生産されるコンテンツは、質も高く、競争力があるが、それを体系的、統合的に集約し、活用するための枠組みは、必ずしも十分ではない。
* コンテンツの流出や散逸等を防ぎ、適正かつ継続的な利用と保存を図るためには、コンテンツの創出支援から収集・保存、配信・流通に至る全体の流れを総合的に集約し、国民的利用と世界に向けた発信を可能とするナショナル・アーカイブ（以下、本文中においては「NA」という。）の早急な構築が求められている。

## ナショナル・アーカイブの定義と社会的役割

* NAとは、「我が国で公表されたコンテンツのメタデータ（所蔵情報を含む）及びコンテンツの権利の所在情報、並びに、電子的に複製可能なものにあっては、コンテンツの複製データを統合的に集約するデータベースを構築・運用し、著作者等及びコンテンツ利用者等の間で、コンテンツの弾力的な利用と継続的な保存のための一定の手続を整備することにより、我が国のコンテンツの国民的利用と世界に向けた発信を可能とするための社会基盤」である。

## ナショナル・アーカイブを構成する主要機能

　NAは、次の5つの機能から構成される。

### コンテンツの生成機能

* 出版物のデジタル化、構造化、テキスト化、バリアフリー化等を支援し、EPUB、DAISY等をはじめとする電子書籍フォーマットの国際標準化を推進する。
* 出版物の二次利用が円滑に進むよう、裁定制度等の既存の手続の簡素化、オープンデータ化等に取り組む。
* 生成段階において、コンテンツの一意的な識別が可能となるよう、コンテンツのメタデータを、中央データベースに登録し、識別子の発行を行う。

### コンテンツの収集・一時保管機能

* コンテンツに識別子及び電子署名を付与し、メタデータ（書誌情報、管理情報及び権利情報）を中央データベースに登録する。
* コンテンツは、一定期間内に長期保存データベースにも複製、格納され、長期保存が図られる。

### 保存機能

* 保存対象は、保存用コンテンツ、一時保管コンテンツ、書誌・目次・索引情報、組織化情報、権利情報、保存のために必要な技術情報及び流通管理関係情報が考えられる。
* 保存機能を担うシステムとしては、OAIS参照モデルに準拠したシステムを視野に置く。
* ダークアーカイブは、とにかく収集し保存する上で有効な方式だと考えられる。どのような形態のダークアーカイブとするかは、当該コンテンツに対する利害関係者との合意に基づいて決定されることを想定する。
* 維持が困難になったデジタルアーカイブそのものあるいはアーカイブ内のコンテンツをNAが引き取ることも必要である。

### 権利情報・管理情報の収集・管理機能

* 著作物の識別子を管理するデータベースでは、各著作物の固有情報（名称、著作者等）のほかに、各著作物の著作権管理に関する情報（著作権者、著作権の管理を委託された著作権等管理事業者、使用料等）を参照できるようにする。
* 著作物の識別子を管理するシステムに著作権等管理事業者による使用料の徴収等を代行する機能を持たせることにより、著作権処理をより簡便に行えるようにすることも考えられる。

### 配信・流通機能

* ファイルフォーマットの異なる電子書籍やデジタルコンテンツを一元的に利用できる共通配信プラットフォームをNA上に構築する。
* 共通プラットフォーム上で、ユーザが様々な電子書籍やデジタルコンテンツを一元的に検索し、求めるコンテンツを、必要に応じて購入したり、図書館から借りたり、また、無償コンテンツについては、自由に閲覧することができるような統合検索サービス（ポータルサイト）を構築し提供する。

## 国立国会図書館の役割

　各セクター、関係機関との連携によってNAを構築することを前提とし、その中でNDLはおおむね次の役割を担う。

### コンテンツの生成機能

* 中央データベースの構築及び運用において、中核的な役割を果たす。具体的には、CIP（Cataloguing In Publication）、クラウドソーシング対応等を含め、電子書籍等のビジネスプラットフォームとして活用するための大幅なシステム機能拡張を行い、コンテンツの生成段階における情報の組織化を目指す。
* NDL資料（紙資料及びデジタル化資料）について、権利処理及び二次利用手続における簡素化と利便性向上を図り、NDL以外の主体によるNDL資料のデジタル化を推進する。
* 出版物のデジタル化、構造化、テキスト化、バリアフリー化等を支援し、EPUB、DAISY等をはじめとする電子書籍フォーマットの国際標準化を推進する。
* サイト、タイトル、巻号、記事、マイクロコンテンツ、ファイル等、多様な粒度で流通するデジタルコンテンツに対応した識別子体系を整備する。

### コンテンツの収集・一時保管機能

* コンテンツに識別子及び電子署名を付与し、中央データベースへの登録を行う。照会に応じ、コンテンツの受入等に係る証明を行う。
* 商用等コンテンツの複製データを収集する。具体的には、有償オンライン資料収集実証実験事業を、NAにおける商用等コンテンツの収集の先駆けとして位置付ける。
* 非商用等コンテンツについては、平成26年2月にリリース予定のオンライン資料の送信による収集機能を活用し、コンテンツの収集を図る。
* HTML等で構成される表層Webのコンテンツについて、我が国のインターネット文化の足跡を後世に残すため、.JPドメイン全体を広く体系的に収集・保存することも視野に入れる。

### 保存機能

* NDLが所蔵するコンテンツに関しては、OAIS参照モデルが想定するすべての機能を備えておくことがコンテンツ管理の観点から合理的である。
* ディザスタ・リカバリについては、他のNA構成機関との共同により実現することは考えられる。
* マイグレーションやエミュレーションについては、他の専門機関に委ねる、情報の所在、内容記述、権利情報、入手可能性などについての情報はNA全体のシステムが処理する、逆に当該情報についてはNDLが一元的に管理するといった分担方式も検討する。

### 権利情報・管理情報の収集・管理機能

* NDLは、中央データベースへの著作物の識別子の登録作業を、出版者等の協力を得て行う。
* 著作物の識別子等を管理するデータベースに対して、著作者情報（人名典拠）、コンテンツに表示されている情報（クリエイティブ・コモンズ・ライセンスの適用等）、NDLが申請した裁定結果の情報等を登録する役割を担う。

### 配信・流通機能

* NA配信プラットフォームを通じて、商業ベースで流通しないコンテンツの一部を配信する役割を担う。
* 配信プラットフォームを利用するに当たっては、当該プラットフォームの運営主体に使用料を支払うことで、一定の経費負担を行う。
* 配信プラットフォームを通じてデジタルコンテンツの配信を行う事業者・機関に対して、標準的な書誌データ（メタデータ）を提供する。

## 実現に向けた取り組みの方向性

### 今後の取組課題

#### 法規・制度面の整備

* 納本制度を参考にして、NAへの「納入」に関する規定を整備し、コンテンツの納入が事実上の著作権登録機能を担うようにする。
* NDLを含む公的機関所蔵情報のオープンデータ化は不可欠の要件であり、利用のための適正なルールを設定する必要がある。
* コンテンツの複製・公衆送信等について著作権の制限、裁定手続の簡素化等の措置を講じること、NAへの個人情報の提供について法令上の根拠を明確化することなどが考えられる。
* NAで保存するコンテンツについて、NDL等国の機関が行う場合は、その二次利用に関しては使用料を求めないことができるような法的規定を設ける必要がある。

#### 財政面の裏付け

* 出版物デジタル化、データの永久保存等基盤整備に係る分野については、公的資金の投入が不可欠と考えられるが、NA全体が発展的に機能するためには、民間流通部門のビジネスモデルを確立する必要がある。

#### 人的資源の整備

* NAのセンター的機能を果たすべき幾つかの機関では、NA全体のマネジメント、研究開発等を担う専門的人材を配置すると共に、参加各機関における担当要員の確保が不可欠である。そのための訓練・養成の仕組みも整備する必要がある。

#### 標準化等技術的対応

* 識別子の統一化・標準化の方向性を定めると共に、オープンデータへの識別子付与についても検討する。また、コンテンツの長期保存メディアのあり方についても、緊急に取り組むべき課題である。

#### 施設・設備の整備

* センター機能を果たすための組織・施設・設備の整備は不可欠である。また、参加各機関及びNAのサービス提供施設における利用環境整備を図る必要がある。

### 利用促進・啓発普及活動の必要性

* 関係者・関係機関の理解・参加を促すための啓発・利用促進事業、国民各層向け広報事業、利用促進事業に取り組む必要がある。

# （2013年）我が国の知識インフラとしてのナショナル・アーカイブ構想（本文）

平成26年2月4日

電子情報部

## ナショナル・アーカイブ構築の意義と目標

　情報化の進展とインターネットの普及により、世界的にコンテンツのデジタル化が進展している。我が国で生産されるコンテンツは、質も高く、競争力があるが、それを体系的、統合的に集約し、活用するための枠組みは、必ずしも十分ではない。

　例えば、コンテンツの所在が管理されていないために、その所在が不明である、若しくは、コンテンツ自体が消失する事例、コンテンツ自体は存在しても、その権利の所在が管理されていないためにコンテンツが死蔵される事例、又は、コンテンツの所在及びその権利の所在は明らかであっても、権利処理の手続やコスト等の問題により、コンテンツが活用できない事例等が多数生じており、競争力の高いコンテンツを、弾力的に流通させ、活用し、将来にわたり長期的に保存するための総合的枠組みは不十分である。

　IT産業、コンテンツ産業において、コンテンツのマス・デジタイゼーションが世界的に進行し、多国籍企業による寡占化が進む中で、我が国の出版物、音楽作品、美術作品等のコンテンツの流出や散逸等を防ぎ、適正かつ継続的な利用と保存を図るためには、コンテンツの生成から収集・保存、配信・流通に至る全体の流れを総合的に集約し、国民的利用と世界に向けた発信を可能とする「ナショナル・アーカイブ」（以下、本文中においては「NA」という。）の早急な構築が求められている。そのためには、産官学民の枠を超えた共通プラットフォーム基盤を作る必要がある。

　すでに出版コンテンツの分野では、国立国会図書館や国立情報学研究所による広範なデータベースの形成と横断的な検索、出版界における書誌・管理情報の統合化や図書館等への電子書籍流通モデル構築への取り組み、電子書籍流通基盤整備のための制度改正等を目的とする議員連盟の成立など、近年共通プラットフォーム形成に向かうと見られる動きが顕著であり、それらが整合性をもって、国として統合的に構築・運用されていくことが望ましい。

　NAは文字・静止画・動画・音像など各種のコンテンツを含む統合的な概念であるが、以下では、その核となるであろう出版コンテンツを中心に、今後のNA構築のあり方について構想したい。

## ナショナル・アーカイブの定義と社会的役割

　NAとは、「我が国で公表されたコンテンツのメタデータ（所蔵情報を含む）及びコンテンツの権利の所在情報、並びに、電子的に複製可能なものにあっては、コンテンツの複製データを統合的に集約するデータベースを構築・運用し、著作者等及びコンテンツ利用者等の間で、コンテンツの弾力的な利用と継続的な保存のための一定の手続を整備することにより、我が国のコンテンツの国民的利用と世界に向けた発信を可能とするための社会基盤」である。

　NAは、とりわけ出版分野においては、電子書籍のビジネスプラットフォーム整備、仕様の共通化、権利情報データベースの整備、図書館等公共施設での電子書籍利用促進、デジタル化資料の利活用等を通じ、我が国の出版文化の発展と国民の利便性を確保するためのエコシステムとして社会的に機能することをめざしている。

## ナショナル・アーカイブを構成する主要機能

　NAは、コンテンツの生成機能、収集・一時保管機能、保存機能、権利情報・管理情報の収集・管理機能、配信・流通機能の5つの機能から構成される（図1）。

　コンテンツの創出から収集・保存、配信・流通に至る全体の流れを一元的に集約するため、NA全体のメタデータを集約するデータベース（以下、「中央データベース」という。）においては、多様な主体が多様なデータを扱えるよう、柔軟かつ多層的なデータ構造を実現する。具体的には、著作物・著作者・出版者等の書誌情報、販売データ・販売者・所蔵機関等の所蔵情報、目次・索引・シソーラス等の情報探索情報等の多様なデータを一元的に管理可能な仕組みとする（図2）。

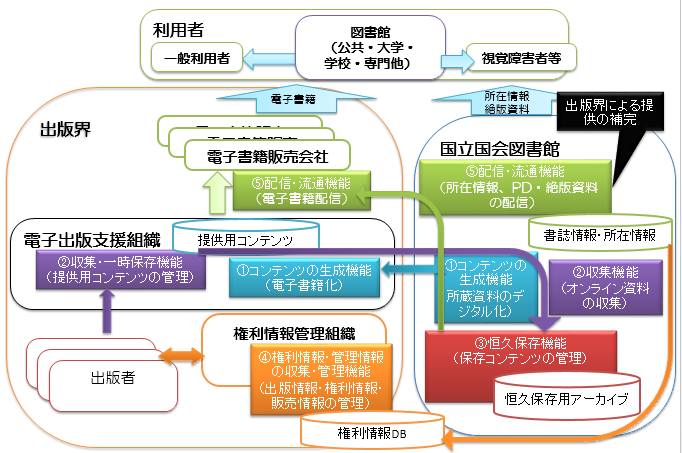


図1. NAの主な機能構成

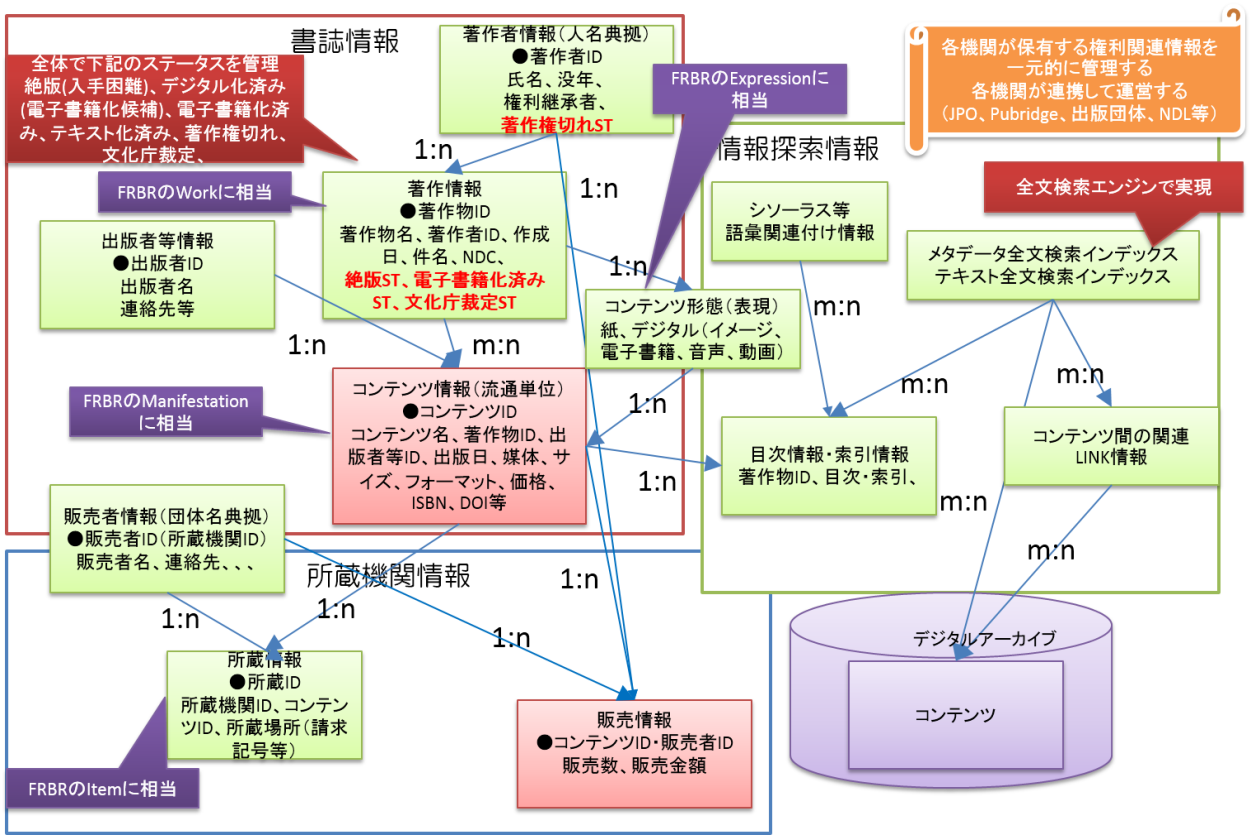


図2. NA権利情報データ関連図（ER図）

### コンテンツの生成機能（図3）

　我が国のコンテンツ産業の競争力と文化資源の充実によるソフトパワーを一層高めるため、コンテンツ創作に携わる著作者、出版者等が、余計なコストや手間をかけずに創作活動に専念できる支援機能をNAの中に組み込むと共に、既存の我が国の出版物、音楽作品、美術作品等のデジタル化を推進し、国全体のデジタルコレクションを、質、量ともに充実させることを目指す必要がある。そのためには、主要な資料・コンテンツ所蔵機関が積極的にNAに参加するための条件整備も求められる。

　とりわけ出版分野においては、既存の紙媒体の出版物のデジタル化、構造化、テキスト化、バリアフリー化等を支援し、EPUB、DAISY等をはじめとする電子書籍フォーマットの国際標準化を推進する。また、既存の出版物の二次利用が円滑に進むよう、裁定制度等の既存の手続の簡素化、オープンデータ化等に取り組む。

　また、コンテンツの制作主体、流通主体、流通経路等が多様化、複雑化する中で、海賊版の流布やデータの拡散を防止し、著作者から始まる権利情報の来歴等を着実に把握する必要性から、コンテンツ制作の上流工程である生成段階において、コンテンツの一意的な識別が可能となるよう、コンテンツのメタデータを、中央データベースに登録し、識別子の発行を行う。いわゆるCIP（Cataloguing In Publication）にとどまらず、コンテンツの生成段階で、LOD（Linked Open Data）、RDF（Resource Description Framework）等を活用する、コンテンツに権利情報を埋め込む、あるいは、クラウドソーシングによりメタデータを付与する等、インターネット等での流通や二次利用を意識した情報の組織化を図ることも視野に入れる。

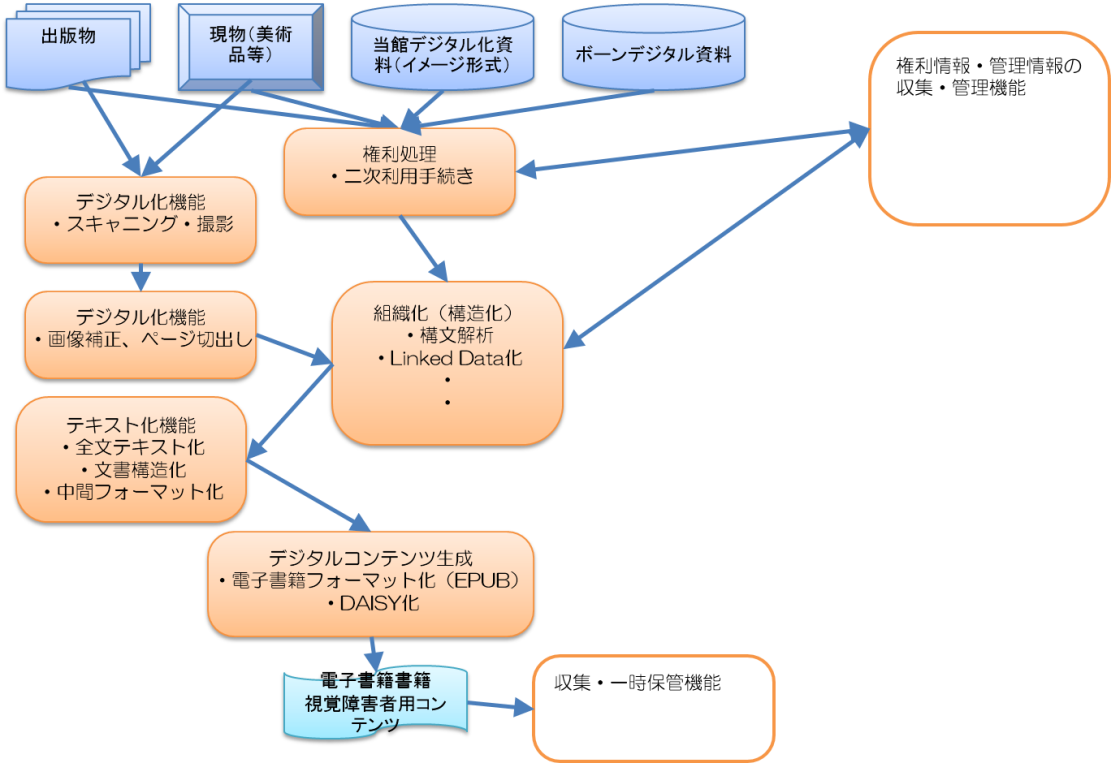


図3. コンテンツの生成機能

### コンテンツの収集・一時保管機能（図4）

　前項の機能により生成されたコンテンツが公表される直前において、コンテンツに識別子及び電子署名を付与し、メタデータ（書誌情報、管理情報及び権利情報）を中央データベースに登録する。

　コンテンツは、データを蓄積し、商用利用等の一次利用を担う、個別のデータベース（以下、「個別データベース」という。）に格納され、公表されるが、一定期間内に、コンテンツの複製データを公的に蓄積し、将来にわたる国民的利用を担う、長期保存用のデータベース（以下、「長期保存データベース」という。）にも複製、格納され、（３）の機能により、長期保存が図られる。

　また、個別データベースに格納されない非商用等コンテンツについては、長期保存データベースに直接収集、格納される。とりわけ、HTML等で構成される表層Webのコンテンツについては、我が国のインターネット文化の足跡を後世に残すため、.JPドメイン全体を広く体系的に収集・保存することも視野に入れる。

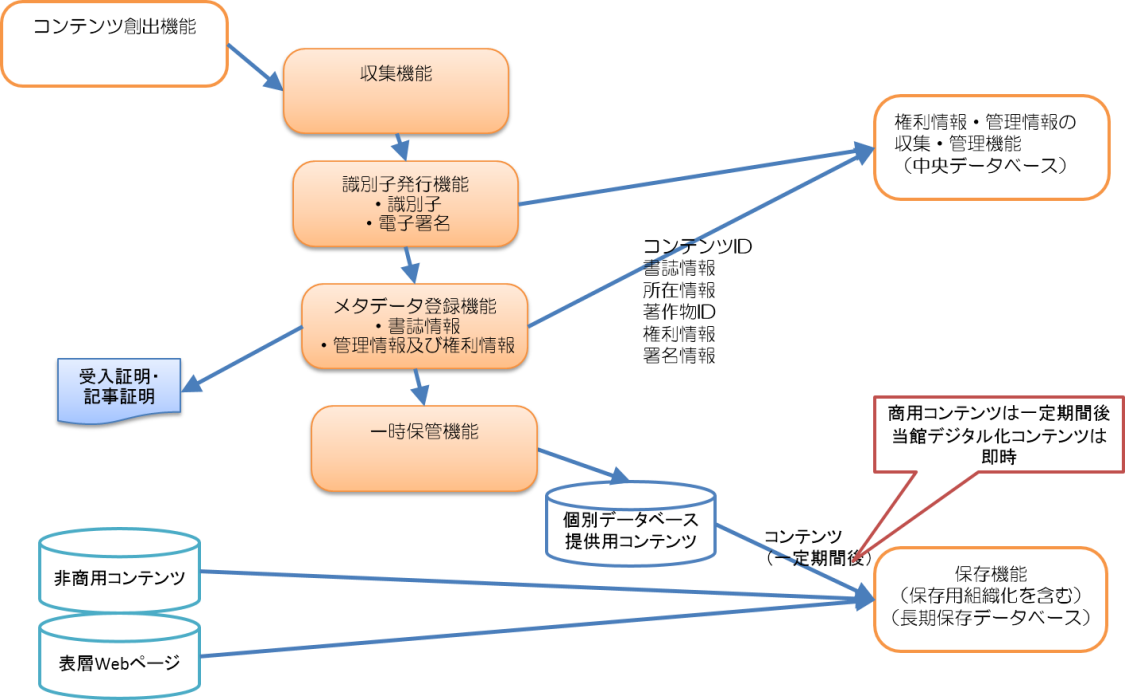


図4. コンテンツの収集・一時保管機能

### 保存機能（図5）

* 1. 保存対象

　NAの保存対象となるデータとしては、長期保存データベースに格納される「保存用コンテンツ」、商用利用等の一次利用を担う個別データベースに格納される「一時保管コンテンツ」、「書誌・目次・索引情報」、LODのように他の情報との有機的連携に有用な「組織化情報」、コンテンツに係る「権利情報」、保存のために必要な技術情報及び流通管理関係情報が考えられる。

② システム

　保存機能を担うシステムとしては、OAIS参照モデル（Open Archival Information System --reference model； ISO 14721:2003）に準拠したシステムを視野に置く。同モデルが備える機能としては、受入、データ管理、保管、アクセス、保存計画及び運用統括の各機能が想定される。

　デジタルコンテンツの複製を物理的に離れた場所に置き、災害に備える「ディザスタ・リカバリ」の仕組みは、「保管」機能の一つと考えられる。

　ただし、保存システムについては、多くの技術的課題、たとえば、適切な長期保存環境、長期保存に適した媒体、原本性を保証するマイグレーションやエミュレーションの技術といった課題が未解決のままである。

③ 運用

ａ．ダークアーカイブ

　ダークアーカイブとは、「出版社がコンテンツを提供している間はアーカイブの構築（収集・保存）のみを行い、出版社が何らかの理由でコンテンツを提供できなくなった際に初めて、保存したコンテンツを公開するという方式のアーカイブ」と説明されており、CLOCKSSやPorticoが知られている。世の中に登場したコンテンツが保存されないまま消滅してしまっては元も子もないので、とにかく収集し保存する上で有効な方式だと考えられる。

　本構想においては、ダークアーカイブとは言っても、当該コンテンツが市場に流通している限り決して公開しないということではなく、たとえば非公開期間や非公開の方法（施設内閲覧も含め全面非公開／施設内閲覧は可だが複写は禁止など）によっていくつかのバリエーションを設けることができると考える。どのような形態のダークアーカイブとするかは、当該コンテンツに対する利害関係者との合意に基づいて決定されることを想定する。その際、当該権利者等に対しては当該コンテンツの将来的な優先的利用を保証するなどの何らかの便宜供与の方法についても併せて検討する。

ｂ．維持が困難になったデジタルアーカイブの移管

　個別に構築・運用されている国内の各種デジタルアーカイブのうち、運用維持が困難になったものについて、アーカイブそのものあるいはアーカイブ内のコンテンツをNAが引き取ることも必要である。そのためには、個別アーカイブをNAが引き取ることを可能とする移管ルールを整備すると共に、移管を技術的に可能とするための標準仕様の整備が必要となる。

ｃ．商用サイトコンテンツを含めた長期保存と提供サービス

　NAの運用に際しては、民間の商業活動を阻害するのでなく振興に寄与するという視点を持つことが極めて重要である。その意味から、現用の商用サイトとの連携による提供サービス及び長期保存の実現に向けた取組みを行う。

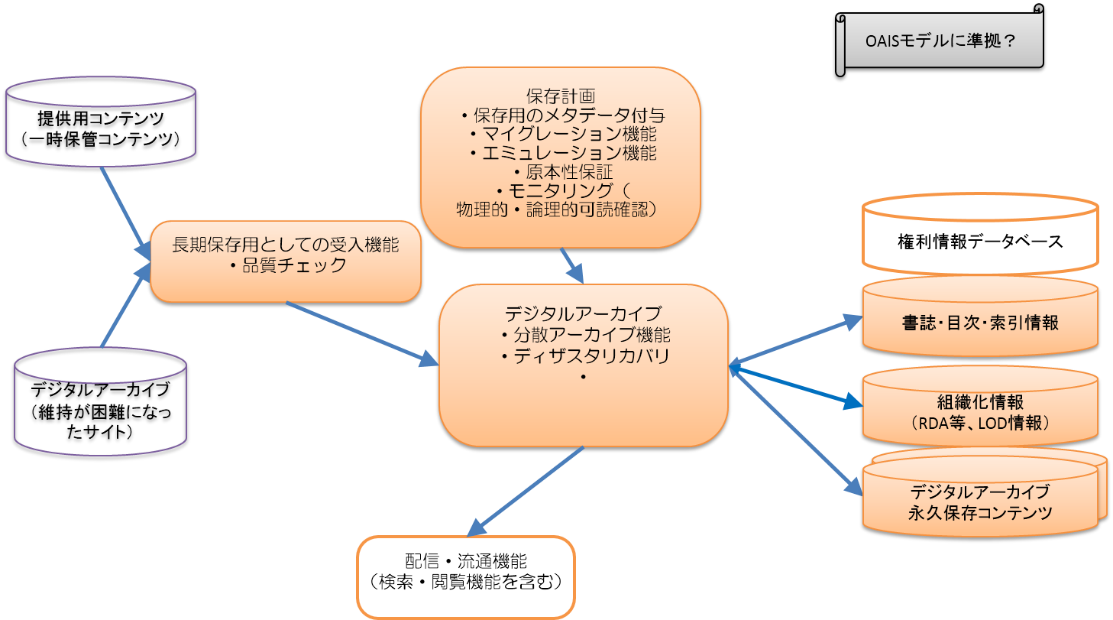


図5. 保存機能（保存計画を含む）

### 権利情報・管理情報の収集・管理機能（図6）

　コンテンツの円滑・適正な利用を促進するため、コンテンツの利用（二次利用）を希望する者が、そのコンテンツの利用に関わる著作権の存否や所在に関する情報を容易に得られるようにする仕組みを構築する。

　そのために、著作物に識別子を付与し、管理するシステムを構築・運営する。著作物の識別子を管理するデータベースでは、各著作物の固有情報（名称、著作者等）も管理・提供し、さらに著作権等管理事業者が管理する個別の著作権管理データベースと連携することで、各著作物の著作権管理に関する情報（著作権者、著作権の管理を委託された著作権等管理事業者、使用料等）を参照できるようにする。

　その上で、中央データベースに各コンテンツに関わる著作物の識別子を登録し、参照できるようにすることで、コンテンツの利用を希望する者に対して、必要な著作権関係の情報を提供することが可能になる。さらに、著作物の識別子を管理するシステムに著作権等管理事業者による使用料の徴収等を代行する機能を持たせることにより、著作権処理をより簡便に行えるようにすることも考えられる。

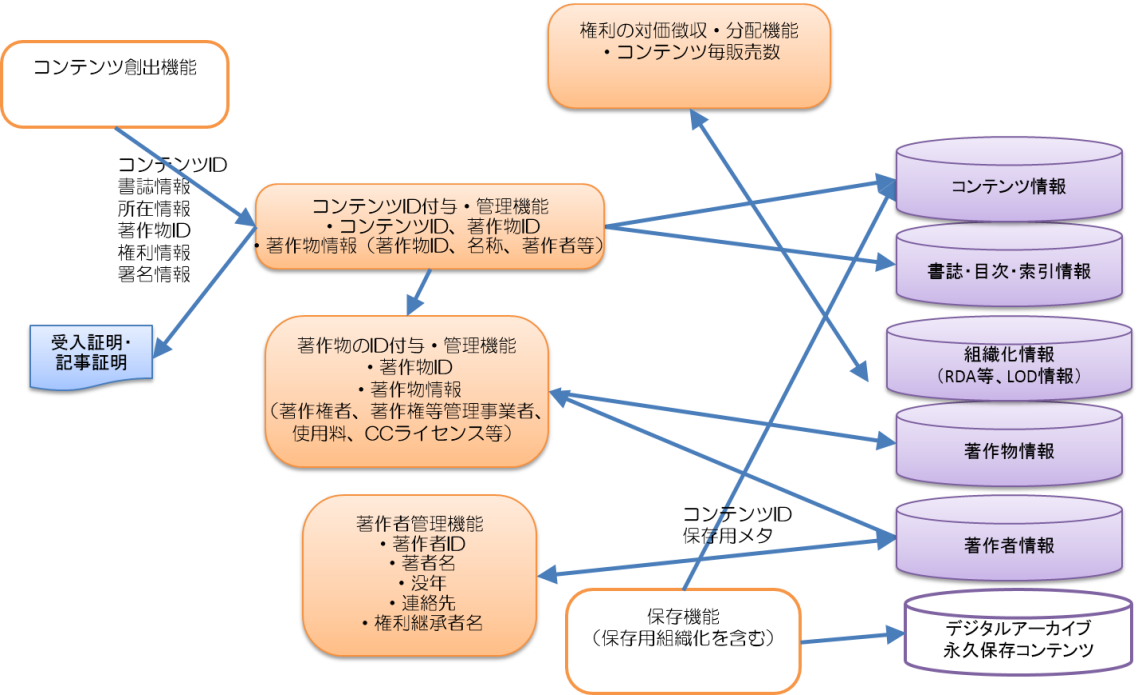


図6. 権利情報・管理情報の収集・管理機能

### 配信・流通機能（図7）

　ファイルフォーマットの異なる電子書籍やデジタルコンテンツを一元的に利用できる共通配信プラットフォームをNA上に構築することで、様々な出版社（又は配信サービス）等が提供する電子書籍・デジタルコンテンツを、国民が広く簡便に利用できる環境を整備することが必要である。

　この共通配信プラットフォームは、課金・決済機能を備えることにより、商用プラットフォームとしての役割を果たすものである。一方で、商業サービスだけでなく、図書館における電子書籍等の貸出サービスを実現するためのプラットフォームとして、また、各種図書館（国立国会図書館、公共図書館、大学図書館等）等のデジタル化資料を利用者に提供するためのプラットフォームとしても利用できる公共性を備えた配信プラットフォームである必要がある。さらに、読むことに困難を伴う国民の読書環境を充実させるために、読上げ機能や文字拡大機能などアクセシビリティに配慮するとともに、ユニバーサルデザインを志向したプラットフォームである必要がある。

　この共通プラットフォーム上で、ユーザが様々な電子書籍やデジタルコンテンツを一元的に検索し、求めるコンテンツを、必要に応じて購入したり、図書館から借りたり、また、無償コンテンツについては、自由に閲覧することができるような統合検索サービス（ポータルサイト）を構築し提供する。このサービスを使って、ユーザは、スマートフォンやタブレット端末など様々なデバイスに対応したユーザインターフェイスから、これまでにない新しい読書体験、コンテンツの利用体験を得ることができる。

　このNA共通配信プラットフォームは、商用プラットフォームとしての役割を果たすことから民間主導で構築されるのが望ましいと考えるが、上記に掲げた公共性を有するものでもあるため、出版文化の振興、図書館振興等の観点から一定の公的支援を検討する必要がある。なお、将来的にはファイルフォーマット、メタデータ等の標準化が図られることが、利用の観点及び電子書籍・デジタルコンテンツの保存・継承の観点から望まれる。

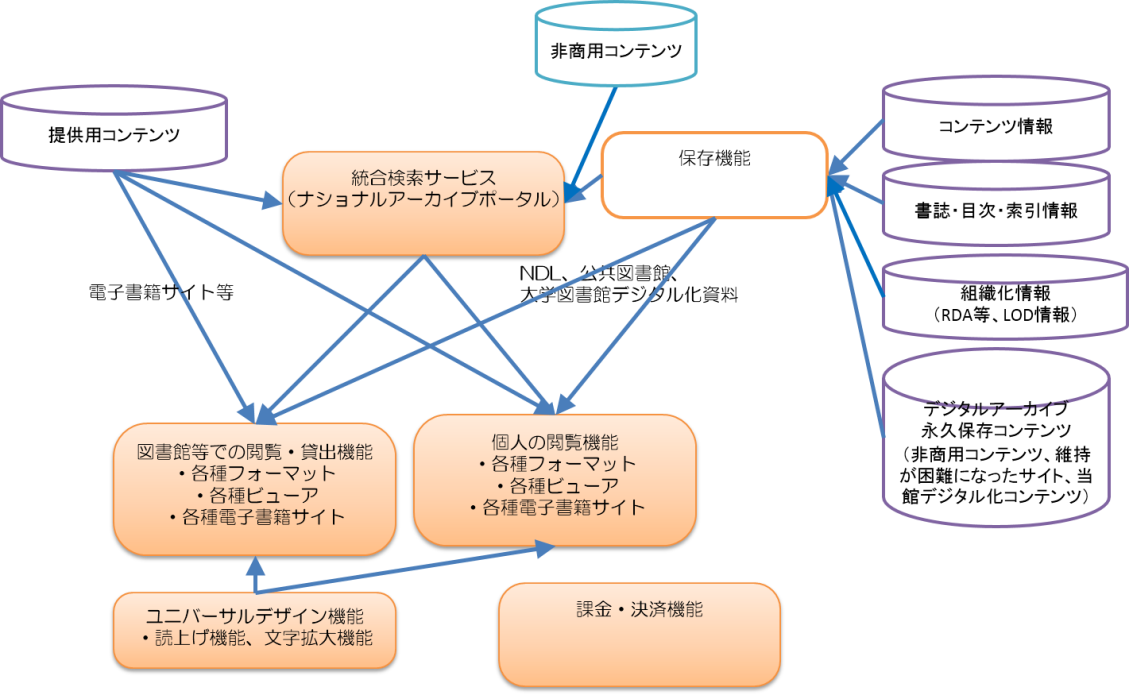


図7. 配信・流通機能（検索・閲覧機能を含む）

## 国立国会図書館及び関係機関の役割と分担

NAの構築は国家的課題として、産官学民を横断した広範な関係機関・関係者の連携・協力体制の下で行うべきものであるが、国立国会図書館もその中で主要な役割を果たす必要がある。

### 関係諸機関の役割

　NAにおいて、関係各機関はおおむね次の表に掲げるような役割を担うことが想定される。

表. NAにおける関係機関の役割

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 全体 | （１）コンテンツの生成機能 | （２）コンテンツの収集・一時保管機能 | （３）保存機能 | （４）権利情報・管理情報の収集・管理機能 | （５）配信・流通機能 |
| 出版者 | ・新規コンテンツを創出し、業界としてフォーマットの標準化を推進する。 | ・フォーマットの標準化推進 ・メタデータの登録 |  |  | ・出版に関わる関連情報の整備 | ・商用コンテンツの配信 |
| 著作権等管理事業者 | ・権利情報等を管理し、使用料の徴収や分配を行う。 |  |  |  | ・権利情報管理、使用料分配 |  |
| 取次 | ・コンテンツの創出を支援し、一時保管機能を担う。 | ・既存コンテンツのデジタル化等 | ・個別DBの維持管理 |  |  | ・商用コンテンツの配信 |
| 電子書籍配信事業者 | ・コンテンツの配信プラットフォームを提供し、円滑な流通を支える。 |  | ・個別DBの維持管理 |  |  | ・配信プラットフォームの提供 |
| 関連省庁 | ・制度の見直し等により、コンテンツの創出や円滑な利用を促進する。 | ・手続簡素化 |  |  | ・円滑・適正な利用の促進 |  |
| MLA | ・既存コンテンツのデジタル化を推進し、分担保存機能を担う。 | ・既存コンテンツのデジタル化等 |  | ・分担保存 |  |  |
| 研究開発機関 | ・構築段階におけるシステム開発支援を行う。 | ・中央DBの構築支援  ・デジタル化、テキスト化、構造化 |  | ・エミュレーション  ・マイグレーション | ・システム開発支援 | ・配信プラットフォーム構築支援 |
| 国立国会図書館 | ・既存コンテンツのデジタル化を推進し、コンテンツの長期保存を担う。 | ・既存コンテンツのデジタル化等 ・中央DBの構築推進 ・フォーマットの標準化推進 ・識別子体系の整備 | ・識別子の付与 ・非商用コンテンツの収集 | ・長期保存 | ・著作物ID等の登録 | ・一部非商用コンテンツの配信 |

### 国立国会図書館の役割

　NAの各機能において、国立国会図書館（以下、「NDL」という。）は次の役割を担う。① コンテンツの生成機能

ａ．NAの中央データベースの構築推進

　NA全体のメタデータの集約を担う中央データベースの構築及び運用において、中核的な役割を果たす。具体的には、CIP（Cataloguing In Publication）、クラウドソーシング対応等を含め、電子書籍等のビジネスプラットフォームとして活用するための大幅なシステム機能拡張を行い、コンテンツの生成段階における情報の組織化を目指す。

ｂ．NDL資料の二次利用の推進

　NDL資料について、紙の出版物に係る権利処理及び二次利用手続における簡素化と利便性向上を図り、NDL以外の主体によるNDL資料のデジタル化を推進するほか、既存のデジタル化資料についても、同様に、権利処理及び二次利用手続の簡素化及び利便性向上を図る。

ｃ．利用円滑化のための標準化の推進

　図書館界・出版界における、出版物のデジタル化、構造化、テキスト化、バリアフリー化等を支援し、EPUB、DAISY等をはじめとする電子書籍フォーマットの国際標準化を推進する。NDLの刊行物及びデジタル化資料についても、当該標準に即し、新しいフォーマットの採用やデータ変換等を実施する。

ｄ．識別子体系の整備

　サイト、タイトル、巻号、記事、マイクロコンテンツ、ファイル等、多様な粒度で流通するデジタルコンテンツに対応した識別子体系を整備する。NAの標準的な識別子体系として、どのようなものを採用するのが適切か、例えば、DOI採用の可能性も含めて検討する。なお、識別子体系の整備にあたっては、関係諸機関の協力が不可欠である。

② コンテンツの収集・一時保管機能

ａ．識別子及び電子署名の付与

　コンテンツに識別子及び電子署名を付与し、中央データベースへの登録を行う。照会に応じ、コンテンツの受入等に係る証明を行う。

ｂ．商用等コンテンツの収集

　個別データベースに格納され、公表されたコンテンツの複製データを収集する。具体的には、有償オンライン資料収集実証実験事業を、NAにおける商用等コンテンツの収集の先駆けとして位置付ける。

ｃ．非商用等コンテンツの収集

　個別データベースに格納されない非商用等コンテンツについては、平成26年2月にリリース予定のオンライン資料の送信による収集機能を活用し、コンテンツの収集を図る。

ｄ．.JPドメインの収集

　HTML等で構成される表層Webのコンテンツについて、我が国のインターネット文化の足跡を後世に残すため、.JPドメイン全体を広く体系的に収集・保存することも視野に入れる。

③ 保存機能

　NA全体の保存機能をNDLが単独で引き受けることは現実的ではないが、NDLが所蔵するコンテンツに関しては、OAIS参照モデルが想定するすべての機能を備えておくことがコンテンツ管理の観点から合理的であるが、ディザスタ・リカバリについては、他のNA構成機関との共同により実現することは考えられる。

また、マイグレーションやエミュレーションについては、他の専門機関に委ねる、情報の所在、内容記述、権利情報、入手可能性などについての情報はNA全体のシステムが処理する、逆に当該情報についてはNDLが一元的に管理するといった分担方式も検討する。

④ 権利情報・管理情報の収集・管理機能

　NDLは、中央データベースへの著作物の識別子の登録作業を、出版者等の協力を得て行う。

　また、著作物の識別子等を管理するデータベースに対して、著作者情報（人名典拠）、コンテンツに表示されている情報（クリエイティブ・コモンズ・ライセンスの適用等）、NDLが申請した裁定結果の情報等を登録する役割を担う。

⑤ 配信・流通機能

　NDLは、NA配信プラットフォームを通じて、商業ベースで流通しないコンテンツの一部を配信する役割を担う。具体的には、配信プラットフォームを通じて、著作権保護期間が満了した著作物のデジタルデータについては、誰もがインターネットからアクセスできる形で配信し、著作権保護期間内の著作物については、NDLの館内利用者端末に限定して配信する。

　なお、配信プラットフォームを利用するに当たっては、当該プラットフォームの運営主体に使用料を支払うことで、一定の経費負担を行う。

　また、配信プラットフォームを通じてデジタルコンテンツの配信を行う事業者・機関に対して、標準的な書誌データ（メタデータ）を提供する。

## 実現に向けた取り組みの方向性

### 今後の取組課題

構想するNAが本来の機能を果たすためには、制度面等における以下の条件整備が不可欠である。そのためには、関係者・関係機関に対して本構想を提示する場を設け、広く意見交換を行いながら、構想への合意形成を図る必要がある。

① 法規・制度面の整備

　NA構築にはNDLや出版社を始めとする官民に渡る多くの関係諸機関の協力・分担が不可欠であり、協力内容や権限に関する取り決め、関係機関間のデータ使用等手続きの簡素化、メタデータ等技術標準化、配信プラットフォームの共通化等についての協議と、こうした協力関係の裏付けとなる法規上の整備が必要である。

　長期保存データベースへの複製・格納については、納本制度を参考にして、NAへの「納入」に関する規定を整備し、コンテンツの納入が事実上の著作権登録機能を担うようにする。民間出版社等による出版関連情報の提供に合わせて、NDLを含む公的機関所蔵情報のオープンデータ化は不可欠の要件であり、利用のための適正なルールを設定する必要がある。

　また、コンテンツの収集・一時保管・長期保存や図書館送信等の弾力的な利用を推進するためには、コンテンツの複製・公衆送信等について著作権の制限、裁定手続の簡素化等の措置を講じること、NAへの個人情報の提供について法令上の根拠を明確化することなどが考えられる。さらに、NAで保存するコンテンツについて、NDL等国の機関が行う場合は、その二次利用に関しては使用料を求めないことができるような法的規定を設ける必要がある。

　複数の著作権管理事業者間でデータ処理上の齟齬が生じた場合の調整機能のありあり方について検討する。

② 財政面の裏付け

　NA全体の構築経費及びその後の運営経費について試算を行うと共に、その資金源における官民の役割分担の原則を明確にする必要がある。出版物デジタル化、データの永久保存等基盤整備に係る分野については、公的資金の投入が不可欠と考えられるが、NA全体が発展的に機能するためには、民間流通部門のビジネスモデルを確立する必要がある。

③ 人的資源の整備

　NAのセンター的機能を果たすべき幾つかの機関では、NA全体のマネジメント、研究開発等を担う専門的人材を配置すると共に、参加各機関における担当要員の確保が不可欠である。そのための訓練・養成の仕組みも整備する必要がある。

④ 標準化等技術的対応

　コンテンツのメタデータ識別子の統一化・標準化の方向性を定めると共に、オープンデータへの識別子付与についても検討する。また、コンテンツの長期保存メディアのあり方についても、緊急に取り組むべき課題である。

⑤ 施設・設備の整備

　NAのセンターとしては、必ずしも独立した施設は前提とせず、既存施設の活用も十分考えられるが、センター機能を果たすための組織・施設・設備の整備は不可欠である。また、参加各機関及びNAのサービス提供施設における利用環境整備を図る必要がある。

### 利用促進・啓発普及活動の必要性

　NAが本来の機能を果たすためには、関係者はもとより、最終的な利用者である国民各層の理解を得ることが大前提であり、そのための啓発・利用促進活動は、構想当初から取り組む必要がある。関係者が・関係機関がNAに自発的に参加・利用してもらうための啓発と、最終的に国民各層によってNAの諸機能が活用されるための普及事業の二つのフェイズが想定される。

① 関係者・関係機関の理解・参加を促すための啓発・利用促進事業

　NAの仕組みの理解・支援を目的とし、さらに出版関係者等に対しては実際の参加・利用を促進する必要がある。対象は以下のとおり。

* 国会議員、各府省、地方公共団体等
* 著作者・著作者団体・学術団体
* 出版社、編集者等出版関係者
* 出版流通・提供事業者
* 図書館、博物館、文書館等情報文化施設関係者
* 新聞等マスコミ関係者

② 国民各層向け広報事業

　国民全般、ネット世論、読書層・図書館利用者層、海外関係者等へ、それぞれの対象者特性にふさわしい広報手段を用いて、NAの存在と意義についての理解促進を図る。

③ 利用促進事業

　国民各層における利用の促進がNAの最終的な目標であり、NAコンテンツの学校教育教材への採用、図書館等公共施設における利用啓発等の利用促進事業に取り組む必要がある。

# （2014年）デジタル化資料送信サービスと今後の展開

―知の共有化と新たな知の創造を目指して―

神奈川県図書館協会　会報No.246

2014年4月1日発行

国立国会図書館（NDL）は、新たな図書館向けサービスとして、当館がデジタル化した資料のうち、絶版等の理由で入手が困難な資料を、公共図書館・大学図書館内で閲覧・複写ができる「図書館向けデジタル化資料送信サービス」を開始しました。当館による資料のデジタル化は、当館が収集した原資料を文化財として保存するとともに、デジタル化により利活用の機会を拡大することを目的としています。現時点において、当館が保有する資料のうち、和図書90万点、古典籍9万点、和雑誌107.5万点、博士論文14万点等、計230万点をデジタル化しました。このうち著作権が切れているもの48万点は、既にインターネットで公開しており、残りの182万点は従来、当館施設内でのみ閲覧可能でしたが、今回のサービス開始により131万点については、当館に来館しなくても、近くの公共図書館・大学図書館等の端末で閲覧できるようになりました。これにより著作権が切れていない絶版本の利活用の機会が拡大するとともに、利用統計を公開しますので、図書館等での閲覧頻度の高い資料は、出版社においてプリントオンデマンド（POD）、電子書籍等で復刻されることも期待されます。

このサービスを含め当館の電子図書館サービスでは、資料を文化財としてあらゆる人々が将来にわたって享受、活用できるようにするため、資料のデジタル化、インターネット情報の収集（WARP）及び電子書籍・電子雑誌等の収集（eデポ）に取り組み、当館の保有する資料を提供するとともに、国内の各機関が分散して保有する情報資源を活用するための統合検索サービス（国立国会図書館サーチ）を提供しています。今後は、これらの事業を発展させ、デジタル化された記録・記憶の本文内容で関連付け、知識として共有化を図り、新たな知識の創造活動に活用できる「知識インフラ」を構築しようと考えています。「NDL東日本大震災アーカイブ（ひなぎく）」は、知識インフラ構築の先行事業と位置付けています。

　デジタル化資料送信サービスを含めて情報の利活用の機会を拡大させるサービスと、新たな知識の創作活動を支援するサービスにより、著作者、出版社、将来にわたる利用者のそれぞれに利益が享受される形で共存共栄し、知識が国の文化財として蓄積していくことを目指しています。

# （2014年）第2回基本問題検討会議に関するメモ【本文】（私見）

平成26年6月12日

総務部　中山

## ナショナルアーカイブの考え方

* 1. 経緯
     1. 電書議連関係
        1. 電書議連、中川勉強会での「ナショナルアーカイブ構想」
        2. いつ具体的な議論に入っても、すぐに実現イメージを示せるように「構想の具体案」を作成
        3. 電書議連その他で、5つの柱があることを説明
        4. 電書議連で、5つの柱の出版界とNDLの役割分担を説明
     2. 文化財議連関係
        1. 当初、電子書籍も含めたアーカイブセンターをNDLの配下に設置する構想
        2. 5月の議連
           1. 恒久的保存基盤は、NDLDA＋WARP＋ひなぎくをベースに、文化財全体で分散アーカイブを構築する。利活用基盤は、その恒久的保存基盤を利用して、利用目的、利用者ニーズに沿って個別にポータル機能を構築
     3. 内閣官房知財本部の動き
     4. 文化庁の動き
     5. 当館の動き
  2. ナショナルアーカイブの目指す意義
     1. 当館の使命・目標、戦略的目標の達成
        1. 国内の出版物に相当する資料の網羅的な収集・保存と目録作成
        2. インターネット資料、オンライン資料の収集の網羅性確保
     2. 緊縮財政の中で、使命を実現するための戦術として
  3. 今後
  4. 基本問題として検討すべき事項
     1. 大局観の視野
        1. 当館が使命を果たすために与えられた権限を行使して、責任と義務を果たさなければ、その使命は他の機関へ移される（中山）
        2. 大局観はまだ絵に描いたものがない（館長）
           1. 情報資源の共同構築と役割分担で、まさに何が期待されているのか、何を優先して取り組むべきなのかの政策的な議論が必要（館長）

ネットワーク型活動を前提とした再構築（館長）

デジタル環境では、従来の文献提供、貸出、コピーよりも広いところで新たな事業が出てきている（館長）

* + - 1. 2020年、さらにその先10年は視野に入れる（中山）
      2. 文化財のナショナルアーカイブは、書籍のナショナルアーカイブを包含する（中山）
      3. ナショナルアーカイブ＝（当館が収集した資料＋当館が収集できない資料）≒国の蔵書＞当館の蔵書（中山）
      4. 伝統的な図書館業務の切り口で検討を始めるのは如何なものか？（中山）
         1. 紙のノウハウの延長でデジタルの特性を生かしたサービスを構築できるのか(中山)
         2. 利用者の発想や社会の動きを含めて、新しい図書館活動の方向を再構築する（館長）
         3. 抜本的なデジタル情報環境ということを常に大きく位置づけた上での再構築のための検討を念頭に。（館長）
         4. ならば、レガシーの体制で短期に検討を進めることは難しいだろう。だから電子情報部ができた（田中、中山）
      5. 組織のありよう
         1. 組織文化をどういう方向に向けていくのが今後不可欠なことかを浮かび上がらせることが必要（館長）
    1. 超短期
       1. 中長期を見据えた実現の目標とマイルストーンの社会的な合意形成
          1. 何が基本問題であるかを共有して優先順位付け（館長）
          2. 洗い出された問題事項を共通の柱立てでどのような解決の方向性を合意すべき（中山、館長）

使命・戦略的目標で、大きな方向性は出してある。解決の方向性くらいまで合意しておかないと、個別の議論でまたぶれる（中山）

* + - 1. 27年度予算確保策
         1. 長期的にどのような方向に向かうのかの合意なしで予算の獲得戦術が見いだせないのではないか（中山）
      2. 収書提示案件
    1. 短期（2年程度で実施すること）
    2. 中期（2020年頃に実現すること
    3. 長期（その先10年程度）
    4. 検討体制
       1. デジタル情報化時代に対応するためには、物の概念とは別に情報としての概念が必要。現状の物として扱う業務の延長線に据えて検討すると、既存の部の現状に引きずられた検討になる。

## 第2回検討会議の議論のポイント

* 1. 何を集めないのかということよりも、日本の情報資源として、どの範囲はＮＤＬの優先的カバーの範囲なのか。集めない部分は、社会的仕組みの中でどう満たされるかということを整理する必要がある。日本における情報ニーズというものの全体像があり、理想的にはすべて必要なものを満たすための資源投入があるべきだが、一方においては、優先化することで社会全体の仕組みが活性化するということがある。（館長）

## 第2回検討会議資料への意見

* 1. 「デジタル情報時代に。。。」
     1. これからデジタル情報時代になるのではない。世の中はすでに情報はデジタルが中心になってきている
  2. 基本問題検討の進め方
     1. まず、「基本問題論点整理」にある「検討内容」の方向性と、検討のマイルストーンを合意形成する
     2. その検討の方向性に沿って、個別検討部会でブレークダウンする
        1. 当館の今後の方向性が見えなければ、部会での検討が困難
     3. 運営管理も
        1. 運営管理は、当館資源の捻出が大きな目的であり、この前提がわからないと、個別の部会での将来イメージの検討が困難。
        2. 当館に資源がないからと、いきなりターゲットを絞った議論をするのではなく、資源が有効活用できていないところを洗い出して、見直しし、そこで浮いた資源を何に投入するかを検討する。それでも足りない資源は外部に求める。中の資源と外部の資源を使って、当館の使命を果たす。
  3. 紙とデジタル
     1. デジタル情報時代に事業のありようを考えるのであれば、紙の文化にデジタルを含めるのではなく、今後の流通形態に合わせてデジタルに紙を合わせていくのではないか。
     2. 現時点において、大分類の「収集・書誌・保存」を、収集書誌部を主管部局とした資料は合意形成できていない。紙とデジタルの収集・組織化・保存は、一緒に扱うことが困難であるとの結論から電子情報部で扱ってきたはず。現時点でデジタルのスキルを持たない部局が主管部局となることには無条件で賛同できない。中長期のビジョンに基づいた組織再編を検討すべき。
     3. 有形文化財は物。物として保存していく必要がある。物として保存するためにデジタル化する。利用はデジタル化したもので行う。
     4. 新しい本はデジタルで作られる。提供の体現形の１つとして、物としての本がある。当館が将来にわたって体現形の１つの物は残こしていくが、利用を保証するのは情報ではないのか。
     5. 組織化は、保存する物への書誌作成ではなく、利用する情報に対してのメタデータ付与。さらにメタデータが十分でなくても本文内容で検索できる形
     6. 基本問題事項では、今後5年、10年先を見据えて、大局観を持って、現在行うべきことを考えるのであれば、デジタル中心の業務へのソフトランディング。
  4. デジタル情報時代において、当館の使命は、デジタル情報を含めた国としての蔵書構築
     1. 蔵書構築は、「当館の蔵書構築」からスタートしない
     2. デジタル情報時代において、当館が全ての情報を集めることが不可能であることから、使命を果たすために、当館で収集できないものはどうするかを含めて、「国の蔵書構築」の観点で議論したい。
  5. 個人的な意見
     1. 物としての収集はまだまだ続くと思うが、提供はデジタルに移行しつつあり、組織化も物としてではなく、情報として行う時代である
     2. 収集、提供の主軸がデジタルコンテンツに移行しつつあるときに、紙の延長でデジタルの扱い方を検討することに疑問。
     3. ITリテラシーが高くなく、デジタルコンテンツの特性も十分に理解していない体制で、今後の主軸であるデジタルコンテンツを含めた収集・提供を考えることで、世の中の変化についていけるのか？
     4. 基本問題事項では、今後5年、10年先を見据えて、大局観を持って、現在行うべきことを考えるのであれば、デジタル中心の業務へのソフトランディング。
     5. 緊縮財政の折、当館資源で行う業務は否応なしに縮退しているので、人的資源に余裕があるはず。また、非効率な業務手続きの見直し、デジタル情報時代に過剰な質を維持している業務の質のレベルダウン、利用者ニーズの少ない業務の廃止等により、さらに人的資源を確保し、その資源によりデジタル化を進める。デジタル化が進めば、利用提供業務での資料の閲覧請求は少なくなり、外注費の削減につながる。
  6. 提案
     1. 「紙にデジタルを含めて検討する」ではなくて、「デジタルの検討に紙を含める形で検討する」のではないか。
     2. 紙のハンドリングの経費を削減して、デジタルのハンドリングの振り向ける
     3. 今後の組織体は、
        1. 今後5年を見据えるなら、デジタルの収集・書誌部、利用提供部と今後収集・提供が少なくなる紙資料を特化して扱う部
     4. という形

## 事業の目標

* 1. 「私たちの使命」、「知識インフラの構築」、「知識情報基盤」、「知の共有化」、「国としての蔵書構築」、「ナショナルアーカイブ」、「新たな知識の創造と還流」、、、の方向性は同じか？
  2. 様々な分野のあらゆる記録を情報として集約して、相互に関連付けて知識化し、将来にわたって利用を保証して、「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを構築する

## 検討の姿勢

* 1. 使命・目標の策定時点において、まだ方向性が見えなかったものも含めて、実現に向けた具体的な施策としてブレークダウンして、浮き出される課題を網羅的に洗い出す。（館議懇より）
  2. 限られた資源の中で、資源投入のための理論・実践を組み立てる。（館長）
  3. 業務見直しの方向性を問題提起する（館長）
  4. きちんと意識を共有する（館長）
  5. 全業務について理念を作成（関西館長）
  6. 必ずしも目標ごとである必要はない。また、現在の部局の仕切りで考えない。

## 問題検討にあたっての意識

* 1. **使命・目標の達成**
     1. 使命・目標に掲げた内容は、抽象的であり、2016年までに達成する目標レベルは明確ではないが、方向性は示されている
        1. 主な課題は、使命・目標の策定にあたっての将来構想会議の部会報告書で示されている。
        2. 戦略的目標は、特に推進すべき目標を提示しているが、内外のニーズの動向を踏まえて、既に見直しが必要な段階か？
  2. **電子図書館中期計画2004の達成**
     1. 電子図書館サービスに関しては、「電子図書館中期計画2004」が目指したことで達成できていないことを再認識する必要があるのではないか。
        1. 計画本文は、概念的な記述しかないが、参考資料その他で実現イメージ及びマイルストーンを記述している
        2. データプロバイダとしての位置づけ、連携のインキュベータとしての役割。
  3. **事業継続計画の観点**
     1. 「納本制度」、「インターネット資料収集」、「オンライン資料の収集」「保存のためのデジタル化」等、法律により当館に与えられた「権限」には、「実施の責任と義務」がある。
     2. その義務を果たしてきた実績に裏打ちされて、当館への新たな期待も寄せられている。その期待に応えなければ、国会・国民から「権限は持っているが、責任と義務を果たさない組織」として存立意義を疑われる事態になりかねない。
  4. **新たに検討すべき視点**
     1. 緊縮財政の中で、如何にして、利用可能な資源を確保し、事業を継続していくか
        1. 将来に向けた投資と、現在の経費の考え方を分けて整理する必要があるのではないか？
           1. 「要求なくして査定なし」。新しい事業を立ち上げなければ、予算は削減されるだけ。
           2. 中期的な予算確保は、外部資源の活用、国の施策の一環で進めることにより、他省庁の予算を活用
           3. システム開発を経費と考えるか、投資と考えるか

円滑で効率的な情報交換が行え、適切な意思決定を迅速に行える情報共有基盤などの戦略的なIT活用、業務・サービスの改善は投資

現状の業務・システム・サービスの維持は経費

* + - * 1. 人材育成費は将来のための投資
    1. 世界規模での文化的な情報資源、知識の共有化の方向性の中で当館の位置づけは？中核的な役割を果たすのか、OneOfThemとして一翼を担うのか？
       1. *「国民の文化志向の高度化と多様化に対応し、様々な文化財、美術品、地域文化、舞台芸術、重要な公文書等の歴史的資料等に関する情報が、地理的な制約を受けずにどこにおいても入手・利用できる環境」（e-japan重点計画2002）を実現する*
       2. 知識インフラ（科学技術基本計画）、ナショナルアーカイブに関するタスクフォース（知財）、文化財保存センター構想、ナショナルアーカイブ構想、大規模災害情報アーカイブス構想での方向性
       3. 知財本部
          1. 知的財産推進計画 2014（7月頃）

検証・評価・企画委員会

コンテンツを取り扱う会合

アーカイブに関するタスクフォース

* + - 1. IT戦略本部
         1. 世界最先端IT国家創造宣言
         2. 電子行政オープンデータ戦略

オープンデータ実務者会議

オープンデータ流通推進コンソーシアム

LODチャレンジ

* + - 1. 世界レベルでのアーカイブの構築を目指して分散保存と連携
         1. 日本のアーカイブ⇒CJKDLI⇒東アジア⇒Europeanaと連携により、世界規模での情報資源の共有を目指す。

CJKDLI

韓国によるポータルプロトタイプ構築への協力

タイトル等の読み仮名の項目

ポータル正式版の仕様調整

Europeana、IA等と連携して、世界レベルで、情報の保存と利用の保証を実現させる

* + - 1. 新たな産業や文化創造の基盤となる知的インフラを構築（タスクフォース報告書より）
         1. 文化資産及びこれらの関連資料などの情報のデジタルアーカイブ化と連携による知識インフラ化

現物保存と、利用のためのデジタル化

各アーカイブ間の連携を実現するための環境整備

* + - * 1. 海外発信を含めた利活用

二次利用の促進

文化力を更に強化し、世界へ発信

* + 1. サービス向上、業務・システムの効率化に資するIT関連の新しいソリューションをどのように取り込むか？（これらのソリューションを意識するかしないかで、将来像は大きく異なる可能性がある）
       1. クラウドソーシング：専門家グループから一般大衆による知識構築
       2. BYOD (Bring your own device)、テレワーク：外出先で自宅でもシームレスに業務遂行
       3. ビッグデータ：オープンデータから有益な情報を抽出し業務に活用
       4. セマンティックウェブ技術による情報の組織化も含む
       5. サービスのマッシュアップ：サービスのAPI連携
       6. ERP：企業内情報の一元的な管理が可能
  1. **デジタル情報時代を踏まえた基本的な考え方**
     1. 物の収集から情報（コンテンツ）の収集へ。
     2. 保存は物理媒体、利用は情報（コンテンツ）
     3. 図書館は本を集めて提供する組織から、コンテンツを集約して、コンテンツ同士を関連付けて知識として提供する機関へ
     4. NDLは、冊子体収集、デジタル化、ウェブアーカイブ、ポータルの実績を踏まえて、国レベルでの各種構想の中で中核的な役割を果たす
  2. **実施姿勢の再認識**
     1. 使命達成のためには様々な壁がある。既存の事業の継続、新しい事業の実施において、様々なリスクがあり、それを回避する努力をしてもリスクはゼロにはできない。残留リスクをどこまで許容？
        1. 法解釈や前例を越えて行った場合のリスクと、やらなくて非難を受けるリスクの評価
        2. 使命達成のためには様々な壁がある。既存の事業の継続、新しい事業の実施において、様々なリスクがあり、それを回避する努力をしてもリスクはゼロにはできない。残留リスクをどこまで許容するかの認識の共有できていたか？
     2. 今まで当館単独でやれることをやれる範囲でコツコツとやってきたのではないか？
     3. 創造力を駆使して、大局観で課題に取り組んできたか？
     4. 新たな事業展開にあたってのリスクの許容レベルは？

## 個別の課題提起

* 1. 連携協力
     1. 連携協力は、使命達成のための、横串の目標と各目標達成のための手段として位置付けられているが、連携協力のステークホルダーは、目標毎に存在するのではない。全体では共通の目標の達成に向けて、個々の対象のステークホルダー毎に、分野横断で連携協力を具体的に進める必要がある。
        1. 連携協力を各目標達成のための手段でなく、連携により**「知識インフラ」**としての**「アーカイブを構築」**と**「利活用を促進」**することを目的とする
           1. 個別の事業の実施においては、最終的な大きな枠組みの中で、整合性を持って効率的に組み込まれることを常に意識する。
     2. 国際連携
        1. 将来的に、世界レベルでの知識情報基盤の構築を目指して、Europeana等との連携協力を推進させる。
           1. インターナショナルアーカイブ（世界レベル）構築での連携協力の推進

Europeanaとの連携協力

* + - * 1. 東アジア版Europeanaとしての連携協力
        2. CJKDLIとしての連携協力の推進
    1. 国内での連携協力
       1. MLA連携
          1. 国としてのアーカイブ構築事業の枠組みの中で、MLA、MALUI、GLAM等との連携協力を推進する
       2. 出版界との連携の推進
          1. 読書機会の拡大による出版文化の発展を目指して、出版物、電子書籍等の販売拡大に直接繋がる形で出版界を支援。共存共栄の観点で出版界を補完したサービスを提供する

当館資料を利用した電子書籍のオンデマンド作成、電子書籍フォーマット業界標準策定、書誌の相互交換、出版サイトへのナビゲート、、、。

長期的な展望での出版文化の発展につながる施策と、短期的に出版界の利益になることでも選択的に行う。

例えば、JDLSのサービス構築・運用支援、助言、関係機関との調整の仲介。出版物、電子書籍等の販売拡大に直接繋がる支援の実施

* + - * 1. 出版情報と図書館書誌情報の相互交換（ONIX←→DC←→MARC）
        2. 電子書籍フォーマットの業界標準策定支援（EPUBサブセット仕様等）
        3. 当館デジタル化資料を活用したオンデマンドでの電子書籍作成、配信

イメージ画像→ページ切出し→テキスト化→構造化→標準EPUB→テキストDAISY、リフロー型EPUB、PDF（フィックス型EPUB）→配信

* + - * 1. 出版社、古書店、電子書籍サイトへのナビゲーション

統合検索もしくは検索結果からのナビゲーション

* + - 1. 公共図書館への支援及び連携の推進
         1. 知の共有化を目指して先進的なサービスの取組を行っている公共図書館もあり、当館が見習うべきところも多々ある。支援という観点だけでなく、公共図書館の協力の下で目標を達成させる施策を推進させる

ししょまろはんラボは、京都府立図書館、京都府立総合資料館の司書が、Webを通して実験的にいろんな情報を公開。京都が出てくる小説やマンガ・ライトノベル等の作品に出てくる京都の位置データに加え、おススメ度や内容紹介付き。

* + - * 1. 公共図書館で利用者に提供可能な蔵書構築支援

当館デジタル化資料図書館送信

商用電子書籍の導入支援

WARP内の他機関資料の活用

* + - * 1. 図書館での書誌作成を支援し、紙・デジタルを問わず、郷土資料を含めて、全ての国内刊行物を網羅した図書館蔵書の総合目録と、知識としての検索・閲覧ができる仕組みの構築
        2. 資料デジタル化及びアーカイブ構築支援
        3. 図書館間の情報共有（目録、レファレンス情報、相互貸借）
        4. 詳細

デジタル情報を含めた総合目録（所在情報DB）の作成（ゆにかねっと、全国書誌）

システム構築支援

ILSベンダーに標準通信規約（OAI-OMH）、メタデータ仕様（DC-NDL(RDF)）を実装するための情報提供、技術支援

各館OPAC→NDLSearch

選書および書誌作成支援

JPO近刊情報（まだ刊行されていない資料の書誌情報）の提供

当館納本資料の書誌情報をインプロセス段階から提供

インプロセスデータ取り込み機能を実装するための情報提供、技術支援

NDLSearch→各館OPAC

図書館における情報提供（閲覧サービス等）の拡充

デジタルアーカイブ構築支援

地域情報デジタルアーカイブ構築（ASPサービスも）

他機関が保有する資料の検索・ナビゲーション

NDLSearchの活用

県域総合目録の構築支援

デジタルを含めた総合目録

当館デジタル化資料の図書館送信サービスの利用

閲覧環境構築支援、制約条件の緩和

商用電子書籍の導入支援

図書館間との相互貸借

ゆにかねっとの活用

図書館間のレファレンス情報の共有

レファレンス協同データベースの構築・運用

* + - * 1. 他図書館と連携したサービスの提供

参考「公共図書館等への電子書籍配信に係る課題整理研究会」 検討結果まとめ

公共図書館等に出版社が提供する電子書籍の内容・種類・制作システムについて

出版社等による国立国会図書館大規模デジタル化データの商業利用の方向性

公共図書館等への出版コンテンツの有償配信・提供環境整備の在り方

公共図書館等で収集・制作した行政情報、地域情報等非商用コンテンツの活用方法

電子書籍へのアクセシビリティ向上

著作権処理の課題と権利の集中管理の在り方

* + - 1. 大学図書館、学術機関への支援及び連携の推進
         1. 「電子的学術情報資源を中心とする新たな基盤構築に向けた構想」（2012年3月付け）の示す方向での連携

①電子情報資源の確保、②メタデータのオープン化と相互接続性（相互運用性）の確保、③統合的発見環境とシステム基盤、その他、電子コレクション構築、システム整備

* + - * 1. 図書館電子書籍推進会議（仮称）への参画

大学図書館で利用する電子書籍ビジネスの確立と利活用の課題解決

* + - * 1. 大学学習資源コンソーシアムへの参画

電子的学習資源の製作、共有化を促進し、また学習・教育において著作物を最適に利用できる環境を整備するための検討を行い、具体化する

* + - * 1. ジャパンリンクセンター（JaLC）での書籍・雑誌記事、電子書籍・電子雑誌の永続的識別子管理
  1. 国会の活動の補佐：信頼性の高い専門的調査・分析と迅速かつ的確な情報提供
     1. 予測調査、レファレンス業務、外部のデータベース等の内容を統合的、横断的に検索・再利用できるナレッジベース化した業務システムを構築
     2. 「立法府のブレーン」「議員のための情報センター」としての役割の拡充強化
        1. 衆参、府省、シンクタンクの調査・分析と比較しての優位性は？
        2. 調査業務における資料・情報群を専門的に調査する業務は、効率的・効果的に行われているか。
           1. 予測調査、レファレンス業務、外部のデータベース等の内容を統合的、横断的に検索・再利用できるナレッジベース化した業務システムを必要としない？
     3. 「国会と国民とをつなぐ役割」の拡充強化
        1. 国会において生み出される情報への国民のアクセスを容易する情報提供サービスの内容は十分か？
           1. 何故、国会関係のシステムが、長期にわたり大幅な機能強化されずに利用されているか？
  2. 国内出版物の網羅的収集に努めるとともに、印刷出版物にとどまらず、電子的に流通する情報を含め、様々な資料・情報を文化的資産として収集し、保存⇒「当館としての蔵書構築」
     1. 収集
        1. 冊子体形式での国内出版物の収集の網羅性確保
           1. 納本義務により収集しているが、ボーンデジタルを含めて、全ての出版物を収集することは不可能。歴史的資料、郷土資料も含めて、出版者、公共図書館、大学図書館、博物館、公文書館、各種アーカイブ機関等で、分担収集し、網羅性を確保する

デジタルも含めた出版物の網羅的収集。紙もデジタルも現実的には分担収集・保存ではないか。

* + - 1. デジタル形式での電子書籍・電子雑誌の収集の網羅性確保
         1. 無償オンライン資料は、インターネット上から消失する可能性が高い。早期に収集範囲、フォーマットを拡大すべきではないか。また、当館単独でなく、ひなぎくのように、外部資源の活用による分担収集の道を探る必要はないか。
         2. 有償オンライン資料は、早期に、インターネット上から消失する可能性が低い。

⇒緊急度は高くない。まずは、電子書籍市場が拡大する形の出版界を支援する方策を取ることが先決ではないか。（図書館での電子書籍サービス等も含めて）

* + - 1. ウェブサイト上の様々な情報、記録の収集の網羅性確保
         1. ⇒国等の機関は制度的に収集可能であり、喪失する前に収集することが重要。当館だけでは不可能。外部資源の活用
      2. ウェブサイト上の情報からの著作物の抽出の網羅性確保
         1. ウェブサイト上の情報からの著作物の抽出の網羅性確保⇒当館職員による人手での切り出しには限度がある。中期的には各機関がデータベース化できるようにすることを目指す。また研究開発機関により、ビックデータ活用技術により、機械的に切り出させる仕組みの構築を目指す。それまでは、人海戦術であれば、クラウドソーシングによりブックマークし、それを起点としたロボット収集、電子書籍化も考えられる。
      3. 所蔵資料のデジタル化
         1. 情報量と質の確保（国と民間の役割分担、国の取組、支援）

情報の分類毎、出版年次毎の網羅性の確保⇒他機関のデジタル化と合わせて網羅性確保。他機関資料のデジタル化も視野に。

他機関でのデジタル化⇒Pubridge等での電子書籍化や、Google等が、非排他的な利用を認めるなら、協力を求めることも考えられる。

* + - * 1. ⇒当館所蔵資料のデジタル化は1/4。収集できていない資料も多く、それらも合わせて網羅的なデジタル化を目指すためには、国としての各種計画の中で、大きな施策の中で、所蔵機関も含めて、長期的に平年度化された予算、もしくは、イベント的な施策での補正予算等の確保を目指す。

古典籍資料（江戸期以前及び清代以前の和漢書）：21万冊、33億円、5～6年

雑誌（2000年までに受け入れた国内刊行雑誌）：50万冊、37億円、出版社がデジタル化してないもの

図書１（1969年以降、1980年代までに受け入れた国内刊行図書）：88万冊、84億円

1980年代以降は、出版社のデジタルコンテンツの保有状況による

市場での価値は高く、出版社でのデジタルデータの二次利用も期待できる

図書2（1969年以降、1980年までに受け入れた国内刊行図書）：48万冊、46億円

児童書（国内刊行の児童書）：10万冊、6億円＋α

* + 1. 保存
       1. 収集した冊子体資料等の物理媒体の長期保存
          1. ⇒利用はデジタル。原本の長期保存に関する新たな研究の成果は？

ビネガーシンドローム、中性紙保存、資料修復技術、、、

* + - 1. 収集した冊子体資料を、利活用のためにデジタル化
         1. 検索できないスキャン画像なら紙のほうが利便性はいい。デジタルだからこそできる利便性の訴求。
      2. デジタル情報の長期保存（将来的な利用の保証）
         1. 当館の資源だけで、ストレージの確保、保存技術の開発・適用していくことは不可能

⇒当館単独でなく、国全体での文化資産の保存の一環で進めるべき。外部資源の活用による保存環境の確保

* 1. 収集資料を簡便に利用し、また必要な情報に迅速かつ的確にアクセスできるように、新しい情報環境に対応して、資料のデジタル化、探索手段の向上など、誰もが利用しやすい環境・手段を整備
     1. 利用者へ情報が届きやすく（見つけやすく）すること、情報を用いた活動をしやすく（利用しやすく）すること（タスクフォース報告書）
     2. 利活用方法のサポート情報も含めた利用目的に応じたポータル機能の構築
        1. 各分野における様々なアーカイブを横断検索すること、さらに、分野を横断し横串で情報の連携を図る仕組み
           1. 利用目的毎に、所蔵場所、媒体形式、言語を問わず、また、書誌のみならず内容からの検索ができ、その資料自身が閲覧もしくはナビゲートできる機能
        2. 対象
           1. 文化財（写真・動画撮影したもの）、出版物、音楽・映像、放送
           2. 無形の記憶を記録したもの
           3. 観測データ、分析データ
           4. ポップカルチャー
        3. 目的別
           1. 調査研究

目的毎の調査研究に資する情報の提供

* + - * 1. 教育利用

知識の習得（知）

情操教育（情）

課題解決能力（意）

問題分析力、発想力、プレゼン能力、創造力

* + - * 1. ビジネス利用

出版ビジネス

権利情報管理データベース

販売促進ツール

マーケット調査、新規事業開拓、、、、

* + - * 1. 日本の文化の発信

日本への関心の高まりへの対応、日本文化の魅力を海外に伝える

* + - * 1. ．．．
      1. 考慮すべきこと
         1. 情報提供サイトへのメリット享受

メタデータ提供、横断検索による情報提供サイトへのアクセス向上、利用拡大、販売促進等

* + - 1. 拡充されるべき機能
         1. メタデータのみならず、サービスとして連携するための標準的なWebAPI（デジタルコンテンツそのものを提供するAPI、DBを直接ハンドリングできるAPIを含む）
         2. 人工知能（AI）的な研究成果
      2. 付帯機能
         1. 利用者間のコミュニティの育成も含め、特定目的に応じたユーザーのニーズを捉えた情報発信（イベント、テーマ別、、、）
    1. サービスの利便性向上（資料組織化部会報告より）
       1. 情報の間を結び付けるための基盤としてのメタデータ整備及びWeb上のデータ間を機械的に関連付け可能なLinkedOpenDataに向けた取り組み
       2. サービス要件とその実現に必要なメタデータの検討
          1. 検索及び書誌の再利用の利便性を高めるメタデータの記述内容と記述レベルの設定

書誌ユーティリティとして。個別資料の権利情報管理用書誌として。

* + - * 1. 本文内容を活用した書誌・典拠情報、解題情報、資料の関連性、目次、要約、主題情報、雑誌記事索引等の機械付与の可能性等の検討
        2. ナショナルアーカイブでの権利処理データベースとの連携
        3. あいまい検索や連想検索等を実現できる技術の適用可能性の検討
      1. 電子情報及びメタデータの開放性向上
         1. 利用者が多様な方法で入手し活用できるデジタルコンテンツ、メタデータの提供API、フォーマットを検討する。
         2. 「電子行政オープンデータ戦略」の基本原則に基づいた、当館における取組内容を検討

「国立国会図書館作成データベース提供方針」（平成20年国図総080717001号）の更なる見直しの検討

* + 1. 媒体に関わらずメタデータを一元的に扱えるフレームワークの構築
       1. MARCに替わるフレームワークの具体的な検討
       2. RDAに対応したメタデータ作成基準の作成
    2. メタデータの網羅性・信頼性の向上
       1. 典拠データの作成対象拡大を目指した「典拠データ等拡充計画」の実施
    3. 永続的識別子の付与
       1. 当館が作成したレファレンス情報、解題等への永続的識別子付与方法の技術的調査及び業務フローの検討
       2. コメントや注釈、既存の書評情報や辞書サービス等外部の情報資源の付加方法の検討
    4. メタデータ作成作業の効率化及び迅速化

公共的書誌情報基盤

2010年（平成22年）から実施（システムリニューアルに先立って）

インプロセス書誌の早期提供

完成書誌の提供

書籍のデジタル化時代の書誌のあり方

MARCでカバーできない時代

電子書籍を含めた出版情報の活用

* + - 1. メタデータの作成工数の削減
         1. 外部作成情報資源を活用して、メタデータの安定した品質及び継続的な提供
         2. 外部作成情報資源の記述規則との整合（出版情報、電子書籍のメタデータ）
      2. 全国書誌の提供範囲の拡大
         1. 当館が収集したオンライン資料も全国書誌として提供できるようになったが、今後の全国書誌の利用想定は？誰がどんな目的で、どんな形での提供を望んでいるか？
    1. 書誌作成、メタデータ作成が効率的に行える体制への移行
       1. 電子情報のメタデータの品質向上と作成のために必要な資源の洗い出し
          1. メタデータの作成において活用する外部情報資源、職員の業務、外部委託内容を明確にして、必要な資源の量を算定する。
       2. メタデータの記述規則を調整する組織の業務内容と実施に必要な体制の洗い出し
    2. 当館所蔵資料の二次利用提供
       1. PDの自由利用
          1. ダウンロードを容易にする機能。パブリッククラウドから自由に利用できるように。
       2. 復刻・翻刻扱いの場合の原出版社への優遇策の可能性
    3. 国の刊行物の目録として、真の全国書誌（以前の書誌構築計画での記述）の提供
       1. 当館に納本されていない刊行物、電子書籍・電子雑誌を含む目録
  1. 未曽有の災害の記録・教訓を後世に確実に伝えるための東日本大震災アーカイブを構築
     1. 復興計画にあるあらゆる記録・記憶を収集して後世に残す
        1. ダークアーカイブ問題
           1. あらゆる記録・記憶を確実に後世に伝えていくこととされているにも関わらず、権利処理できていない記録は収集さえもできていない。
           2. 記録・記憶が消滅してしまう前に保存が必要。誰が何をすれば保存ができるのか？
        2. NDLクラウドとしての国有財産法の問題
           1. アーカイブの構築、保持が困難になった機関のアーカイブを、当館のアーカイブ環境において代替で保持すること。また、アーカイブ機関が利用に供している環境に障害が発生した場合、当館のアーカイブから代替で提供し、サービスを継続させる（ディザスタリカバリ―）
  2. 透明性が高く効率的な運営管理を行い、高度なサービス提供を担うことができる人材を育成し、また必要な施設を整備
     1. 組織運営のあり方の方向性の再認識
        1. 「社会環境変化への迅速な対応」および「業務プロセス改革」を実現する
           1. 経営戦略とIT戦略を密に連携・融合させ、最大限の効果を発揮させること
        2. 高付加価値化、差別化できる事業モデルの構築
           1. ITの活用を効率化から競争優位性の獲得へ
           2. 意思決定プロセスの迅速化

業務・サービスシステムにより収集した膨大な利用情報等を迅速に解析し、意思決定を迅速に行えるようにする（リアルタイムでの統計情報の出力と分析）

* + - 1. 課題解決の推進に必要な人材のスキル
         1. 組織の経営全般、事業全体に関する知識と変革に対する推進力、分析的・創造的思考力、様々な動きに対する感度の良いアンテナ、いつ、何のためITを導入するかを見極める力
         2. 感度の良いアンテナによって見つけ出した動きを深堀する専門性の高い知見と分析能力
         3. いつ、何のためにどういったITを導入するかを見極める力。

ITの活用を、業務コストの削減から、顧客満足度の高いサービスの向上へ

* + - 1. 今後の業務に必要な人材の育成
         1. 全般（韓国国立中央図書館へのResponse Sheetより）

ITリテラシーを持つこと。仕事の基礎となる法制度を正しく理解し、運用できること。社会的な水準で仕事全体をマネジメントできること。利用者やステークホルダー、また職員同士で、円滑かつ適切にコミュニケーションできること。

* + - * 1. システム部門

調達、プロジェクト管理、システム開発管理、システム運用管理のスキルを持った人材の育成が必要。委託業者と対等に議論できるレベル。

軽微な機能改修を自ら行えるレベルの職員の確保と保守契約の範囲の見直し。

* + - * 1. サービス提供部門

知識の分野毎に専門性の高い知見と分析能力を持つ人材

当館所蔵の文献だけでなく、インターネット上の多様な情報源から日々刻々と流れる膨大な情報の中から目的に合った情報を選別し、複数の情報を関連付けて編集し、新しい価値を持たせ「知識」として共有できるようにする（情報内容に関する知識と情報技術の両方の知識を持つキュレーター）

アーカイブとユーザーを繋ぐ人材

多様な図書館資料・情報を適切に扱うことができるような知識を深め、利用者に橋渡しできること

多様な情報リテラシーを持つ利用者に対する人的なサポート能力

* + - * 1. 保存部門

資料修復技術者

* 1. 緊縮財政の中で、業務・システムの効率化、サービスの向上を図るための資源の確保
     1. 人・物・金の資源の効率的な運用
        1. 業務・システムの効率化の候補の洗い出しと見直しの実施
           1. 事務手続き、業務手続きで簡素化できることはないか？
           2. ニーズが低いサービスで多くの資源を使っているものはないか？
        2. 業務・システムの効率化プロセスの効率化
           1. 業務・サービスの効率化は、改善のPDCAサイクルそのものも効率化させる必要がある
           2. 業務プロセス再構築の検討を効率化させるためのビジネスプロセス管理（BPM）ツール・手法の適用

従来型のOAツールを使った「業務・システム最適化計画の策定プロセス（AsIs、ToBe）、EA、、、」でなく、

ビジネスプロセス管理（BPM：Business Process Management）、BPMN（Business Process Model and Notation）の導入

*期待される効果*

*経営の高度化・最適化⇒価値の創出*

*業務部分の数値化、可視化により、経営の強化、業務の最適化、変化への迅速な対応が可能になる*

*モデリングの基盤*

*BPRやシステム構築を行うため、誰でも理解できる図表を段階的に作成可能で、構想初期段階から全体設計ができる*

*業務改革の基盤*

*作図による業務の可視化時間を短縮でき、また柔軟に修正できる*

*組織改革の基盤*

*BPM基盤上で業務を行うことで、日々の業務データを蓄積し、ボトルネック等の分析できるによるモデリングで、誰でも理解できる図表による業務・サービスの可視化*

* + - 1. アウトソーシングの見直し
         1. 将来構想会議では、アウトソーシングを推進する方向性を提示したが、緊縮財政の折、外注の見直しが必要

資料デジタル化、調達支援、進捗管理、システムの軽微な機能強化・改修等、、、。そのためには、人材の確保が必要。

* + - 1. IT関連資源の確保
         1. システム部門の人材育成

先進サービス動向、技術を把握してサービス要件、システム化要件を取りまとめ、構築・運用するために高いマネジメント能力を持った人材育成・確保を進める。基礎知識は研修でも習得できるが、実践的な技量はOJTが有効。しかしOJTは、スキルを持った教師役が必要。まず、教師役の人材を外部に求める。

* + - * 1. 外部人材（非常勤研究員、委嘱研究員、外部研究機関等）の活用による共同の実用化実証実験

当館が保有しているコンテンツ、システムを研究者に提供する。研究者は，それらの資源を活用して実用化システムを開発する。その成果を、NDLのシステムに実装して次世代のサービスを提供する。

これにより、社会のニーズに対応したサービスを、巨額の外部委託費をかけずに構築することができることを目指す。

情報の可視化技術、情報収集の効率化技術、情報の組織化技術、情報の集合知化技術、情報探索技術、閲覧表示技術

プロトタイプシステム構築

共通仕様策定の推進

メタデータ仕様、永続的識別子、電子書籍フォーマット、

* + 1. 新たな財源確保（国の施策での重点事業扱い、補正予算等）
       1. 国の新たな施策の中に当館が進めてきた事業が位置付けられ、その発展的な目標設定と予算措置がされることが重要

IT総合戦略本部（オープンデータ実務者会議）、知財計画2016（知的財産戦略本部（検証・評価・企画委員会アーカイブに関するタスクフォース））

ナショナルアーカイブ構想（電子書籍と出版文化の推進に関する議員連盟）、デジタル文化情報保存センター構想（デジタル文化資産推進議員連盟）、大規模災害情報アーカイブス構想（内閣府防災）等

* + - 1. 少なくとも27年度予算において、調査費程度でも得られることにより、今後数年間を展望することができる
         1. 当館は、ウェブアーカイブの構築、大規模デジタル化の実績が認められて、国のアーカイブ構築の中核的な役割を果たすこと期待されているのではないか。
         2. 当館の「使命・目標」が目指す方向と近いものと思っている。
         3. 当館の通常の予算獲得の方法で、更なるストレージの確保、長期保存の保証をすることは困難。
         4. このような、巨大なストレージを持った文化資産の保存センターが、当館で捻出可能な資源とは別に作られることは望ましいこと。
         5. 「要求なくして査定なし」の姿勢は、当館が予算要求可能な範囲で考えるよりも、オールジャパンで予算を確保することが効果的と思われる
      2. 国の施策の候補
         1. 2020年東京オリンピックにターゲットを当てた各種施策
         2. 名目

文化財保存のためのアーカイブ

日本の文化の発信のためのアーカイブ

社会のインフラとしてのアーカイブ

* + - * 1. 事業例
        2. ナショナルアーカイブ構想（電子書籍と出版文化の推進に関する議員連盟）

日本の文化資産全体を保存し、将来にわたって利用可能にするアーカイブを構築する

電子書籍・電子雑誌に相当する情報も、このアーカイブの中で保持する

出版界と

コンテンツの生成機能

デジタル化（イメージ画像化）

当館資料の大規模デジタル化

EPUB化

OCR→校正→構造化→EPUB化（DAISY化）

他機関のデジタル化支援

収集・一時保存機能

　提供用コンテンツの収集

保存機能

　デジタルアーカイブの構築

提供サイトが障害時、もしくは、提供を継続できなくなった時、代替で提供するバックアップサイトとなる

権利処理がされるまで、ダークアーカイブすることが許される必要がある

電子書籍の永久保存環境の構築

複数機関での重複投資削減、ディザスタリカバリ―

権利情報・管理情報の収集・管理機能

全ての著作物の権利、デジタル化状況が管理されるものを構築する

　当館書誌を提供する

出版者の販売実績集計及び権利者への著作権料の支払い処理が行えるものを想定する

著作権処理・デジタル化を円滑に行うための権利情報管理データベースの構築（権利処理の省力化）

著作権切れ、孤児作品の利用促進

著作権調査、デジタル化の重複投資削減

海賊版対策

配信・流通機能

当館デジタル化資料の二次利用の促進（復刻・翻刻、電子書籍の充実）

* + - * 1. デジタル文化情報保存センター構想（デジタル文化資産推進議員連盟）

多様なニーズに応える検索・閲覧環境の提供

複数拠点による分散アーカイブでの永久保存環境の構築

国全体でアーカイブセンターを構築することに賛同

インターネット上のWeb サイトの情報は、明らかに一般社会に対して公開する意図を持って出されているものであり、出版物の一種と見なしてよいものである。これらはほとんどが紙では出版されない各種の貴重な情報（文書、画像、音声、音楽、その他）である。

これらはわが国の様々な活動の記録であり、国内の出版物が納本義務によってすべて国立国会図書館に集められ、保存・利用されているように、これを残していくことは国としての義務である。一種の文化財とみなせる。

当館が求める資源

従来のスキームでは、NDLの資源だけで構築するには限度があり、とても国レベルのアーカイブ構築は見込めない

関係機関と分散アーカイブを構築・運用し、将来に亘って利用できる形での保存に要する費用

* + - * 1. 文化関係資料のアーカイブの構築（文化庁）

調査研究（文化関係資料のアーカイブに関する基準・手法の確立）（26年度）

* + - * 1. 大規模災害情報アーカイブス構想（内閣府防災）

ひなぎくの機能強化及び永続的な運用

* + - * 1. …
        2. アクション

各府省の計画に含まれるように、特にIT戦略本部、知財本部での検討会議で、積極的に提言する。

e-Japan戦略でのウェブアーカイブ構想、震災復興計画での東日本大震災アーカイブ等のように

* + 1. 事業の継続性の確保
       1. 不慮の事態の被害の拡大を防止と、サービスの早期復旧のための緊急時対応手順の再確認
       2. あらゆる不慮の事態に対して説明責任を果たした対応が行える事業継続計画の策定と運用

## ナショナルアーカイブに関しての理念

### 姿勢（留意点）

* + - 出版文化が隆盛する方向での社会的な合意づくり（特に出版界の納得）
      1. 出版界等、関係機関がNDLに期待する事業は、中核となって実施する
      2. 民間が主体で行える事業は、その事業を支援する形で一翼を担う
    - 産学官の各組織の非排他的な共同事業とする
    - 当館の既定の予算路線では財源（蔵書デジタル化、爆発的に増加するデータ保存の受け皿、データ長期保存策）確保が限界

### 文化情報保存センター構想

* + - 対象範囲
      1. 無形、有形文化財
    - 中核的な組織は？
      1. 文化財が中心のセンター構想では、NDLは一角（一翼）を担えるが、中核的な組織となることの総意にならない？

### ナショナルアーカイブ

* + - アーカイブの構築⇒情報資源を長期にわたって保存し、将来の利用を保証することは国の役割
      1. デジタル資産を豊富に
         1. 蔵書デジタル化←民間と分担して、重複なく、また網羅的にデジタル化
         2. デジタル出版活動←当館がデジタル化した資料の二次利用による電子書籍化
      2. デジタル資産への一元的なアクセス
         1. 日本版Europeanaを目指す←東アジア版←CJKDLI←GLAM（MALUI）←NDLSearch
      3. デジタル資産の収集・蓄積保存
         1. 保存の「集中」と「分散」←ノードとなる複数拠点と個々のアーカイブをP2Pネットワークで分散
      4. デジタル資産を扱う人材確保
         1. 必要な各種業務の人材確保・育成プログラム←オールマイティの人材でなく、個々のスペシャリストを確保
         2. 人的ネットワークの構築←物理的に一か所に集まる形でなくてもいい
      5. 施設のあり方
         1. 物理的な施設と論理的なネットワーク環境←全ての資源を拠点に集約する必要はない
    - 登録制度⇒民間主体で。
      1. デジタル資産の登録と権利管理←出版界、図書館界から書誌情報、典拠情報を提供
         1. 簡便で確実な登録制度

権利管理ビジネス

所要コスト循環モデル

* + - 利用・ビジネス⇒民間主体。民間ビジネスが成立しないところを国が補完
      1. デジタル資産の利活用・流通
         1. 館長：これから組み立てる課題でこれから考えていく
    - 国際化

### 著作権法の一部を改正する法律案の審議での説明・質疑

* + - 1. 中川議員
         1. 有体物、無体物でなく、コンテンツとして考えるべき
      2. 瀬尾氏
         1. 電子出版権の議論は、ナショナルアーカイブ構想の第1ステップ
         2. 後ろに大きな課題がある

ナショナルアーカイブ構築

海外のプラットフォーマとは対立でなく連携、融合

クールジャパンのインフラ構築

小さなクリエータ、地方のコンテンツもナショナルアーカイブに入れる

海外へのコンテンツ発信

世界規模で連携

ナショナルアーカイブで考えること

アーカイブの仕様の標準化

アーカイブをまとめたポータル

誰がどのように使って、どうペイするのか

ナショナルアーカイブ促進法で、旗印と方向性の提示

アジアをにらんで

各省庁が連携して

NDLを含めて、分散しているものをどうまとめていくか

権利情報の集中処理機構が必要

孤児作品の活用

ADR機構も

* + - 1. 相賀社長
         1. 中小出版社の電子書籍化、安心して預ける受け皿としてPubridgeの周辺で検討

# （2014年）ナショナル・アーカイブ構想の検討状況と今後の方向性（スライド説明用手持ち原稿）

平成26年6月26日

総務部　中山

## ナショナル・アーカイブ構想の実現を目指す意義

* 1. NDLの使命・目標、戦略的目標の達成
     1. 国内の出版物に相当する資料の網羅的な収集・保存と網羅的な目録・所在情報の提供
        1. デジタル情報時代において、当館が全ての情報を集めることが不可能であることから、NDLとして何を収集するかだけでなく、使命を果たすために、当館で収集できないものはどのように保存していくのか、利用を保証していくのか、議論したい。
     2. インターネット資料、オンライン資料の収集の網羅性確保
        1. 紙媒体でも困難であったが、デジタルではもはや当館だけで網羅性を確保することは不可能。
           1. 関係機関との分担収集、分担保存、保存場所を意識しない提供
  2. 緊縮財政の中で、使命を実現するための戦術として
     1. 関係府省庁の計画を取り込むことにより、NDL以外の資源（人物金）を活用して、業務システム最適化を実施する
        1. 実績：次世代室での外部研究員の成果、ひなぎくの構築等。。。
  3. 構想への具体的な対応の検討の中で、各基本問題事項の方向性を示唆する

## 問題意識と解決策の１つ

* 1. 何のために？
     1. 縮財政の折、「私たちの使命・目標2012-2016」及び「戦略的目標」の達成に必要な資源確保が困難な状況
        1. 業務システム最適化で示された機能強化
           1. 多様なニーズへの対応を容易にする検索UIサブシステムの構築費用
           2. デジタルアーカイブと東日本大震災アーカイブのストレージ及び保存システムの統合費用
           3. 館内サービス、図書館送信サービスでのコンテンツ提供環境の改善費用
           4. 東日本大震災アーカイブの運用費用
        2. 保存と利活用のためのデジタル化と組織化
           1. 当館資料のデジタル化費用
           2. 収集コンテンツのディザスタリカバリ対策費用
           3. メタデータの作成費用
           4. インターネット情報、オンライン資料の収集拡大費用
     2. 毎年度削られている当館の通常予算では、これら行うことはできない。使命・目標を果たすことができない。
  2. 何をする？
     1. 外部の資源を活用する
        1. 前提
           1. 国のアーカイブ構築関連の各分野の様々な施策では、当館のデジタルアーカイブ、ポータルサービス、デジタル化の実績が、高く評価されており、当館との連携を明示するものが多々ある。
        2. 各分野の施策を、当館の戦略的目標に掲げた方向性、業務システム最適化の実現形に近づけることにより、各施策の中で当館に期待する期待に応えることができる。
        3. それにより、当館が保有する資源では実施が困難な事業を、外部の資源を活用して実施できる可能性がある
  3. どのように？
     1. 各施策の動きを把握する
     2. 各分野の施策の目標をサービス要件として、国としてのアーカイブを構築する立場から業務要件、システム化要件を示し、積算が可能なレベルまで具体的なサービスの構築・運用の想定を落とし込む
        1. その際に、各種の分野毎の縦割り施策で、似て非なるシステムを複数作る羽目にならないように。
     3. 結果
        1. どれかの施策だけでも実現すれば、目標の達成に近づけられるように

## ナショナル・アーカイブ関連の施策

### ナショナル・アーカイブ関連の各施策等

* 1. ⑧世界最先端IT国家創造宣言（2014年6月14日閣議決定）
     1. <http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20130614/siryou1.pdf>
  2. ⑧電子行政オープンデータ推進のためのロードマップ(2013年6月14日高度情報通信ネットワーク社会推進戦略本部決定)
     1. <http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20130614/siryou3.pdf>
  3. ①「出版社の権利のあり方に関する提言（中山提言）」（2013年4月4日第7回「印刷文化・電子文化の基盤整備に関する勉強会（中川勉強会）別紙資料）
  4. *「我が国の知識インフラとしてのナショナル・アーカイブ構想（案）」（2014年2月4日電子情報部作成）*
     1. *二次利用検討の進め方の一環で、ナショナルアーカイブに関して当館の考え方を整理する（2013年7月12日館議懇）に基づき電子情報部内で検討し作成*
  5. ②日本の文化情報戦略基盤「国立デジタル文化資産振興センター（仮称）」設置構想提言（2014年5月23日デジタル文化資産推進議連資料）
     1. 要旨、本文
  6. ③「アーカイブに関するタスクフォース報告書」（2014年4月11日知的財産戦略本部検証・評価・企画委員会（第7回））
     1. <http://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kensho_hyoka_kikaku/dai7/siryou2-1.pdf>
     2. <http://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kensho_hyoka_kikaku/dai7/siryou2-2.pdf>
  7. ④知的財産戦略調査会の提言とりまとめ（2014年5月27日自由民主党知的財産戦略調査会）
     1. <https://www.jimin.jp/news/policy/pdf/pdf178_1.pdf>
  8. ⑤「知的財産推進計画２０１４」
     1. <http://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku2014.pdf>
  9. ⑥経済財政運営と改革の基本方針2014（2014年6月24日閣議決定）
     1. <http://www5.cao.go.jp/keizai-shimon/kaigi/cabinet/2014/2014_basicpolicies.pdf>
     2. <http://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2014/0613/shiryo_01_1.pdf>
  10. ⑦文化関係資料のアーカイブに関する有識者会議（2014年6～7月文化庁）
  11. ⑦大規模災害情報アーカイプス(仮称)（2014年6月内閣府防災）

### ①電子書籍・②文化資産の両議員連盟の動き

* 1. **電子書籍と出版文化の振興に関する議員連盟（電書議連）**
     1. 経緯
        1. 印刷文化・電子文化の基盤整備に関する勉強会（中川勉強会）での「ナショナルアーカイブ構想」を提示
        2. いつ具体的な議論に入っても、すぐに実現イメージを示せるように、電子情報部において「我が国の知識インフラとしてのナショナルアーカイブ構想（案）」を作成
        3. この資料をベースに、電書議連その他で、5つの柱があることを説明
        4. 電書議連で、5つの柱と、出版界とNDLの役割分担を説明
     2. 今後
        1. 権利処理関係は、出版界が中心に検討を行う
        2. 電子書籍の再販制度？、国際貢献？
  2. **デジタル文化資産推進議員連盟（文化財議連）（2012年6月～）**
     1. 文化資産議連（2013年11月12日）
        1. 電子書籍も含めたアーカイブセンターをNDLの配下に設置する構想
        2. 主管は、NDLと文化庁
     2. 文化資産議連（2013年3月24日）
        1. 検討委員会での検討状況の報告
     3. 文化資産議連（2013年5月23日）
        1. デジタル文化情報保存センター（仮称）設立構想についての検討委員会報告
           1. センター主管組織（設立推進体制）

産官学が合同で連携する推進体制（文化庁と国立国会図書館を含む）をスタートアップし「文化立国」のための国家戦略を立案

* + - * 1. ）恒久保存基盤整備

－多様な文化資源のデジタノレ化とデータ蓄積および各種デジタルアーカイブ相互連携基盤整備への早期着手が急務

・デジタルコンテンツを最も大量に取り扱う業務実績を有し、今後も書籍分野の文化資源の大規模デジタル化および段階的なウェブ情報の制度的収集を計画する国立国会図書館が、予算面・人員面の強化をはじめとする条件整備等の措置を前提として、この役割を担うことが望ましい

－恒久保存の技術的解決方策は継続検討（恒久保存メディア開発、マイグレーション・エミュレーション等）

* + - * 1. デジタル文化資産の「活用」

国内外に文化情報を発信する基盤（ポータル）構築・運用

集積したデジタル文化情報から二次的情報として新たなデジタル文化資産（コンテンツ）を創造

* + 1. 考察
       - 1. 構想恒久的保存基盤は、電子書籍も含めて、NDLDA＋WARP＋ひなぎくをベースに、文化財全体で分散アーカイブを構築する。利活用基盤は、その恒久的保存基盤を利用して、利用目的、利用者ニーズに沿って個別にポータル機能を構築
    2. 議連の各議員に、恒久的保存基盤関係の概算予算説明
       1. 2020年までの当館資料のデジタル化、文化財を含めた恒久的保存基盤の構築に必要な費用を提示。

### 「我が国の知識インフラとしてのナショナル・アーカイブ構想（案）」

* 1. ナショナル・アーカイブの定義と主要機能（参考：「電子情報部検討案」2014年2月4日）
     1. 「公表されたコンテンツとメタデータ（所蔵情報を含む）、権利の所在情報、コンテンツの複製データ等を統合的に集約するデータベースを構築・運用し、著作者等と利用者の間で一定の手続きを整備することで、コンテンツの利用と世界に向けた発信を可能とする社会基盤」
     2. ①コンテンツの生成機能、②収集・一時保管機能、③保存機能、④権利情報・管理情報の収集・管理機能、⑤配信・流通機能から構成される。
  2. 国立国会図書館の取組として期待される部分
     1. 所蔵コンテンツのデジタル化
     2. 商用・非商用コンテンツの収集
     3. メタデータの集約を担う中央DB構築
     4. 識別子の付与、著作物IDの登録
     5. 長期保存
     6. 一部商用コンテンツの配信
  3. 日本の文化資産全体を保存し、将来にわたって利用可能にするアーカイブを構築する
     1. 電子書籍・電子雑誌に相当する情報も、このアーカイブの中で保持する
     2. 出版界との共存共栄
  4. コンテンツの生成機能
     1. デジタル化（イメージ画像化）
        1. 当館資料の大規模デジタル化
     2. EPUB化
        1. OCR→校正→構造化→EPUB化（DAISY化）
     3. 他機関のデジタル化支援
  5. 収集・一時保存機能
     1. 提供用コンテンツの収集
  6. 保存機能
     1. デジタルアーカイブの構築
     2. 提供サイトが障害時、もしくは、提供を継続できなくなった時、代替で提供するバックアップサイトとなる
     3. 権利処理がされるまで、ダークアーカイブすることが許される必要がある
     4. 電子書籍の永久保存環境の構築
        1. 複数機関での重複投資削減、ディザスタリカバリ―
  7. 権利情報・管理情報の収集・管理機能
     1. 全ての著作物の権利、デジタル化状況が管理されるものを構築する
        1. 当館書誌を提供する
     2. 出版者の販売実績集計及び権利者への著作権料の支払い処理が行えるものを想定する
     3. 著作権処理・デジタル化を円滑に行うための権利情報管理データベースの構築（権利処理の省力化）⇒現在、JPOが中心に検討を行っている
        1. 著作権切れ、孤児作品の利用促進
        2. 著作権調査、デジタル化の重複投資削減
        3. 海賊版対策
  8. 配信・流通機能
     1. 情報資産の統合検索と所在場所へのナビゲーション
     2. 当館デジタル化資料の二次利用の促進（復刻・翻刻、電子書籍の充実）

### ナショナル・アーカイブと国立国会図書館

* 1. 電子書籍と出版文化の振興に関する議員連盟2014年第3回総会（平成26年6月4日）での報告資料
  2. ○国立国会図書館からは、著作権法の一部を改正する法律案に対する、国会の附帯決議におきまして、「図書を始めとする我が国の貴重な文化関係資料を次世代に継承し、その活用を図る上で重要な役割を果たす」ことが期待されております、ナショナル・アーカイブに関してご説明いたします。
  3. ○お配りした図をご覧ください。国立国会図書館としては、ナショナル・アーカイブに必要な機能は、次の５つだと考えています。
     1. ①コンテンツを生成する機能
     2. ②それらコンテンツを収集してとりまとめ、利用可能とする機能
     3. ③コンテンツを恒久的に保存する機能
     4. ④権利情報を管理し、権利処理を容易にする機能
     5. ⑤コンテンツや、コンテンツを探すための情報であるメタデータを提供する機能
  4. ○こうした機能全体を、官民が協力して、分担、協力しながら実現していくことが必要だと考えております。
  5. ○ 国立国会図書館としましては、公的機関として、民間における活動を補完し、支援する役割を果たすことができると考えております。図の右側をご覧ください。
     1. コンテンツの生成については、所蔵資料のデジタル化により、新たな出版の素材を提供することができます。
     2. 電子書籍の収集については、出版界と協議しながら、制度の整備を段階的に行っており、現在、無償かつ技術的保護手段が付されていないものについて収集を開始しています。
     3. こうして収集したものを長期にわたって保存することについては、技術的な困難はなおありますが、当館以外にこの役割を担える機関は他にないと自負しております。
     4. 権利情報の管理については、民間のしかるべき組織が担うのが適当と考えておりますが、我々が持つ著作権に関する情報を提供することで、その活動を支援することができると考えております。
     5. そして、著作権保護期間を満了し、パブリックドメインとして公共財となったコンテンツのインターネット公開や、商業的に流通していない絶版等書籍の図書館への送信など、日本の知的資産の流通において、商業出版を補いつつ、国際的な発信にも寄与することが可能です。
  6. ○ 一方で、こうした役割を十全に果たすためには、特に、図書・雑誌を合わせて少なくとも約140万冊のデジタル化がさらに必要だと考えておりますが、当館の現在の予算ではなお不足することは否めません。ご支援のほど、よろしくお願い申し上げます。

### ナショナル・アーカイブの機能情報関連図

* 1. 前の図では「出版活動の支援・補完」として概念的な連携イメージで、機能と分担に関しては賛同は得られるが、電子情報部では、出版界とは、より具体的な「日本の出版文化の発展に向けた共存共栄のための施策」が見えるように検討してきた。また、現在の事業との関係もわかるように。
  2. 中央の青の線：当館がデジタル化した画像データを二次利用のために提供する（当館資料の二次利用の促進）
  3. その画像データを利用して、出版社が電子書籍化する
  4. 中央の左から右への紫の線：出版社等が提供する電子書籍を、オンライン資料として収集し、恒久保存アーカイブで保存する（有償オンライン資料の制度収集）
  5. NDLがデジタル化もしくは収集したコンテンツは、恒久保存アーカイブで管理する。（デジタルコンテンツの仕様の標準化（Jpeg、EPUB等））
  6. 中央から上に向かう青の線：図書館向け電子書籍配信サービス、保護期間満了、絶版資料は当館から提供する。図書館利用者は共通の環境で閲覧できるようにすることを目指す。
  7. また、NDLは全ての情報の検索と所在場所へのナビゲートを行う
  8. 右のオレンジ色の線：権利情報DBへ、当館が保有している書誌メタデータ、著作権データを提供することを想定（データベース提供方針）（メタデータの標準化）

### ③アーカイブに関するタスクフォース報告書 内閣官房知的財産戦略本部（知財本部）

* 1. アーカイブに関するタスクフォース報告書（知的財産戦略本部　検証・評価・企画委員会）
     1. 基本的な考え方
        1. 公的アーカイブは情報社会における社会インフラとしての役割を期待。
        2. アーカイブの利活用促進のため、情報が届きやすく（見つけやすく）、利活用がしやすくすることが必要である。
        3. 利活用を図るべき分野を特定するなど戦略的な利活用の活性化を図るべきである。
        4. アーカイブの構築・運営に係る各機関の体制の強化が望まれる。
        5. メディア芸術のアーカイブについての充実に向けた取組が必要である。
     2. 取り組むべき施策
        1. 連携強化のための横断的取組
           1. 分野横断的検索システム
           2. 目的別利用促進の取組
           3. 分野別ポータルサイト
           4. 利用者とアーカイブをつなぐ人材の育成
           5. 権利処理の円滑化（一元化等）
        2. 各アーカイブ機関に求められる取組
           1. 連携を目指したメタデータの採用
           2. 外部ソフトウエア向けインターフェイスの公開や共通化（APIの開放など）
           3. 二次的な利用の手続の円滑化
           4. 利活用のビジネスモデルの開発
        3. アーカイブ促進のための基盤の整備
           1. アーカイブ機関における体制の充実
           2. アーカイブを担う人材の育成
           3. 長期保存等の基盤技術の開発
           4. 孤児著作物のデジタル化ルールの整備などアーカイブに係る著作権制度改善

## ④知的財産戦略調査会の提言とりまとめ（2014年5月27日自由民主党知的財産戦略調査会）

* 1. アーカイブの利活用促進に向けた整備の加速化（p.8）
     1. 目標：コンテンツのアーカイブ化を進め、文化、地域情報の海外発信を積極的に展開する。
     2. 現状：書籍、映画、放送番組、音楽、アニメ、漫画、ゲーム、写真など9の文化資産及びこれらの関連資料などのアーカイブ化が遅れている。これらのコンテンツのアーカイブ化を促進することは、新たな産業や文化創造の基盤となる知的インフラ構築のために必要不可欠である。
     3. 施策：アーカイブの利活用を促進するため、必要な資金や人材の確保、アーカイブシステムを支える基盤技術の開発・関連法制度の整備等の措置を積極的に講じる。

### ⑤知的財産推進計画２０１４（2014年6月20日知財本部決定）

* 1. 知財計画2014
     1. ①職務発明制度の見直し、②営業秘密保護の強化、③中小・ベンチャー企業等における知的財産活動の支援強化を、コンテンツ分野に関して、④アーカイブに関する取組の強化、⑤音楽産業をモデルケースとしたコンテンツの海外展開 を重点施策とする
     2. 各タスクフォースの報告書がベースとなって取りまとめられた。
  2. 第３. デジタル・ネットワーク社会に対応した環境整備
     1. １．デジタル・ネットワークの発達に対応した法制度等の基盤整備（p.39）
        1. （新しい産業の創出環境の形成に向けた制度等の構築・整備） （文部科学省、経済産業省）
        2. （コンテンツ提供のプラットフォーム構築） （経済産業省、総務省、文部科学省）
        3. （電子書籍の本格的な普及促進） （文部科学省・経済産業省）
        4. （公共データの二次利用の促進・ビッグデータビジネスの振興等） （内閣官房、総務省、文部科学省、経済産業省）
     2. ２．アーカイブの利活用促進に向けた整備の加速化（p.43）
        1. （アーカイブの戦略的利活用の推進） （文部科学省、経済産業省）
           1. アーカイブを通じた日本の文化情報の海外への発信の強化の取組
           2. 多言語化やユーザーコミュニティと連動したポータルサイトの整備などを促進
        2. （各種アーカイブの充実）
           1. 文化創造や新たな産業の基盤となる知的インフラを構築するため、映画、音楽、アニメ、マンガ、ゲーム、デザイン、写真、書籍、文化財等の文化資産及びこれらの関連資料等のデジタルアーカイブ化等を、国立国会図書館等の関係機関と連携しつつ促進する。（短期・中期）（文部科学省）
           2. 我が国のアーカイブ関連施策の推進において、出版物等の分野で国立国会図書館が果たしてきた中核的な役割に鑑み、国立国会図書館による各種資料の収集の充実、デジタルアーカイブ化及び適切な利活用の促進等が、関係府省と連携しつつ、引き続き行なわれることを期待する。
        3. （アーカイブの利活用促進のための環境整備等）（総務省、文部科学省、経済産業省）
        4. （アーカイブの利活用促進のための著作権制度の見直し）
        5. （メディア芸術分野等における取組の加速化） （文部科学省）
        6. （アーカイブに関する基盤技術の開発等）（総務省、文部科学省、経済産業省）
        7. （アーカイブ関連人財の育成等） （文部科学省、総務省）

### ⑥経済財政運営と改革の基本方針2014（2014年6月24日閣議決定）

* + - 1. 経済財政運営と改革の基本方針2014（仮称）（素案）（2014年6月13日内閣府）
         1. 「文化芸術立国を目指した取組を推進する。」（P.12）
      2. （2014年6月24日閣議決定）
         1. （スポーツ・文化芸術の振興） （p.8）

また、文化芸術立国を目指し、地方公共団体や民間団体等、文化芸術の振興に取り組む様々な主体との適切な連携の下、観光等他の分野との協働や産業振興等の視点も踏まえつつ、「日本遺産（Japan Heritage）」など魅力ある日本文化の発信、子どもの文化芸術体験機会の確保、国立文化施設の機能強化、文化芸術の担い手の育成、文化財の保存・活用・継承等に取り組む。

### ⑧内閣官房高度情報通信ネットワーク社会推進戦略本部（IT戦略本部）の動き

* + 1. 内閣官房高度情報通信ネットワーク社会推進戦略本部（IT戦略本部）の動き
       1. 世界最先端IT国家創造宣言（2014年6月14日閣議決定）
          1. オープンデータ・ビッグデータの活用の推進(p.6)

公共データの民間開放（オープンデータ）の推進

ビッグデータ利活用による新事業・新サービス創出の促進⇒アーカイブの構築

* + - * 1. 国・地方を通じた行政情報システムの改革（p.18）

モバイル端末の利活用等を通じて、情報のデジタル化（ペーパーレス化）の推進と生産性向上

ワーク・ライフ・バランスや災害時等の業務継続性に配慮

2021 年度を目途に原則全ての政府情報システムをクラウド化し、拠点分散を図りつつ、災害や情報セキュリティに強い行政基盤を構築

* + - * 1. 国民全体の IT リテラシーの向上（p.21）

情報モラルや情報セキュリティに関する知識を含め、国民全体の IT リテラシーの向上

* + - * 1. 国際的にも通用・リードする実践的な高度な IT 人材の育成（p.21）

高度なIT人材の育成が必要。実践の中で技術を習得。

IT人材のスキルを共通尺度で明確化するスキル標準を、ITの技術変化等を踏まえて適切に整備・活用

* + - * 1. 研究開発の推進・研究開発成果との連携（p.23）

研究成果を、迅速かつ的確に IT 戦略と連携させる

* + - 1. 電子行政オープンデータ推進のためのロードマップ(2013年6月14日高度情報通信ネットワーク社会推進戦略本部決定)
         1. 二次利用を促進する利用ルールの整備（p.2）
         2. 機械判読に適したデータ形式での公開の拡大(p.3)
         3. データカタログ（ポータルサイト）の整備(p.４)
         4. 公開データの拡大(p.5)
         5. 普及・啓発、評価(p.6)

### ⑦文化庁・内閣府防災の動き

* 1. 文化庁の動き
     1. 文化関係資料のアーカイブに関する有識者会議
        1. 我が国の文化関係資料のアーカイブに関する現状、課題等を踏まえ、中長期的な視点も含めた文化関係資料のアーカイブの取組の総合的な推進方策を検討
        2. 文化遺産オンライン、メディア芸術アーカイブ事業（マンガ・アニメ・ゲーム）の今後の展開に関して方策を検討。
           1. 「MANGAナショナルアーカイブ」構想を提示
        3. NDLからは、P.20のPPTを提示
  2. 内閣防災の動き
     1. 大規模災害情報アーカイプス構想
        1. 大規模災害に関する情報（記録）を後世に残すとともに、調査、研究、対策検討に活用できるようにする
        2. 「ひなぎく」のシステムをベースとして、対象とする災害を拡大し、記録に残すシステムを恒久化する

## 国立デジタル文化情報保存センターのカバレージは？

* + 1. 昨年1１月に電子書籍と文化遺産のアーカイブとNDLの現状のアーカイブとの関係を図示したもの。

## 各種アーカイブ構築施策の一元化

* + 1. 最近の国のアーカイブ構築の動きとして、左にあるようなものがある。収集対象、見せ方はそれぞれの目的毎に違うが、効率的・効果的に構築することを目指して、一体のものとして構築されるべき

## 日本の文化情報基盤の強化のための費用試算

* 1. 国のアーカイブ構築に必要な要素
     1. 大きく３つ。
     2. ３の恒久的保存基盤は、利用目的を問わず一元的に保存できる環境
     3. ２の活用基盤は、利用目的毎に付加価値を付けて発信していく
     4. １のデジタル化は、産学官で分担してデジタル化
  2. 分散アーカイブの構成
     1. 恒久的保存基盤は、１つのサイトにすべてを集約するのではなく、各機関のアーカイブを集約する分野毎のノードがあり、それを集約拠点に集め、これらが相互にコンテンツを持ち合って、障害、災害に備えて将来にわたって利用を保証する仕組みを想定。（いわゆるP2Pネットワークのイメージ）

## 国立デジタル文化資産振興センター 基盤整備の全体像

* + 1. 下の恒久的保存基盤は、分野を問わず共通のプラットフォームを提供し、そのプラットフォームの上に、文化財及びメディア芸術としてのポータルを構築する。文化財に特化したプラットフォーム機能は、そのポータルの中で実装することを想定したもの。

## 文化資産として関連付けて保存すべきインスタンス【源氏物語を例に】

* + 1. これは、左上の源氏物語を例に、1つの古典作品がベースとなって、オレンジの部分にあるように、書籍、電子書籍・静止画・音声、映像・動画のように、様々な形態のコンテンツが生成されている。さらに、この作品に関する参考情報、有形のグッズ、無形の文化財も存在する。これら全てが、関連付けられて利用できるように、また、文化財として恒久的に保存される必要がある。
    2. これらのコンテンツは下の緑の部分にあるように、様々な機関がコンテンツホルダーとなって、所有している。利用に当たっては、様々な形態のコンテンツの中から利用者が選択して、それらのコンテンツホルダーへナビゲートして閲覧できるような仕組みが必要となる。
    3. このように、探し出す仕組み、保存する仕組みとして、アーカイブの分野に関わらず、共通的なものの構築を目指すべきと考える

## ナショナル・アーカイブの検討に当たっての考察

* + 1. 様々な分野のあらゆる記録を情報として集約、相互に関連付けて知識化し、将来にわたって利用を保証。「新たな知識の創造と還流」の仕組みを構築する
       1. 知識情報基盤の構築
          1. 2011年の科学技術情報整備基本計画で掲げた知識インフラのイメージ
    2. 各項目
       1. 国としてのアーカイブ構築
       2. 利用目的毎のポータルの提供
       3. サービスの高度化
       4. 法的課題の解決

## 基本問題事項

### 基本問題事項としての論点

* 1. 資料デジタル化と利活用
     1. 大規模デジタル化の推進（予算確保）
     2. 商業出版との相補的・互恵的な関係構築
     3. 著作権の登録制度等と連動した制度検討
  2. オンライン出版物のアーカイブ
     1. 制度的収集の実現
     2. 利活用の基盤となる制度設計
  3. 国レベルの「デジタル情報資源」政策との関係
     1. 機関連携による文化情報資源全体のアーカイブ基盤の形成
     2. 安定的な長期保存の実現
  4. 蔵書構築とは？
     1. ナショナル・アーカイブ≒国の蔵書＝（NDLの蔵書＋NDLが収集できない情報）
     2. 全国書誌の位置づけ（真の全国書誌の意義と、当館所蔵資料の書誌を全国書誌ということのギャップ）
  5. デジタル情報時代の業務のありよう
     1. 今後数年間で、新規の著作物の主たる体現形はデジタル形態になり、ニーズに応じてPOD等により、紙を発行する形へ移行することが予想される。
        1. デジタル情報時代において、当館の使命は、デジタル情報を含めた国としての蔵書構築を関係機関と分担して
     2. 媒体の収集から、情報としての収集へ
        1. 紙の文化にデジタルを含めるのではなく、今後の流通形態に合わせてデジタルに紙を合わせていくのではないか。
     3. 媒体の組織化と情報の組織化の違い
     4. 媒体の保存と情報の保存の違い
     5. 媒体は提供せず、情報を提供

### 基本問題事項の検討の進め方

* 1. 基本問題事項の方向性
     1. 中長期を見据えた実現の目標を設定
     2. 洗い出された問題事項を共通の柱立てでどのような方向性で解決するかを合意する
     3. その方向性に沿って個別検討
     4. マイルストーンの設定
        1. 超短期（数ヶ月）
        2. 短期（2年程度）
        3. 中期（2020年頃）
        4. 長期（その先10年程度）
  2. その他
     1. 外部の資源の活用で他に期待できることは？
        1. 研究開発、技術開発成果の適用（総務省、文部科学省、経済産業省）
        2. 選考採用、非常勤研究員、委嘱研究員
        3. 外部資源によるデジタル化
     2. 併せて、内部の資源の活用は？
        1. 事務の簡素化
           1. 運営管理は、当館資源の捻出が大きな目的であり、この前提がわからないと、個別の部会での将来イメージの検討が困難。
           2. 当館に資源がないからと、いきなりターゲットを絞った議論をするのではなく、資源が有効活用できていないところを洗い出して、見直しし、そこで浮いた資源を何に投入するかを検討する。それでも足りない資源は外部に求める。中の資源と外部の資源を使って、当館の使命を果たす。
        2. 人材育成（文部科学省、総務省）
        3. 職員によるデジタル化
           1. 緊縮財政の折、当館資源で行う業務は否応なしに縮退しているので、人的資源に余裕があるはず。また、非効率な業務手続きの見直し、デジタル情報時代に過剰な質を維持している業務の質のレベルダウン、利用者ニーズの少ない業務の廃止等により、さらに人的資源を確保し、その資源によりデジタル化を進める。デジタル化が進めば、利用提供業務での資料の閲覧請求は少なくなり、外注費の削減につながる。

# （2014年）デジタル時代の国立国会図書館の活動

　第33回日中業務交流

基調報告

2014年11月19日

デジタル時代の国立国会図書館の活動  
―電子図書館事業20年を迎えた新たな方向性の模索―

　国立国会図書館

中山正樹

## はじめに

*昨年度は、「利用者サービスの新たな展開」として、2012年以降の利用者サービスの動向、新時代のレファレンスサービス、国立国会図書館サーチの新しい挑戦をご報告しました。*

*今年度は、当館での電子図書館サービスの構築に着手して20年を経過したところであり、今までの歩み（目指してきたこと、実現できたこと）、そして、今後、デジタル化、IT技術の進展に対応したサービスの構築に向けて、どのような方向性で進めるべきかについてについてご報告します。*

2014年は、当館での電子図書館サービスの構築に着手して20年を経過したところであり、今までの歩みと今後の活動の方向性について４つのステージに分けてご報告します。

## この20年間の電子図書館の構築の歩み（1994年～2012年）

*当館では、1990年代に始まるインターネットの急速な発展を背景に、1994年に我が国で最初の電子図書館の実証実験プロジェクトを開始し、2002年から本格的なサービスとして離陸し発展させてきました。*

### 第1ステージ【1994～2002】 揺籃期・始動期

1994年に我が国で最初の電子図書館の実証実験プロジェクトを開始し、資料のデジタル化と公共図書館の所蔵資料の総合目録の構築を進めました。

この実験では、図書館における情報の蓄積と提供の可能性を検証し、将来の電子図書館を想定したプロトタイプ環境**を**構築することを目的に、1000万ページに及ぶ資料をデジタル化しました。*その上で利用を予測して大規模ストレージに格納し、大容量のマルチメディア通信回線を用いて試行提供しました。*この実験での*技術課題の*検証を踏まえ、1998年「国立国会図書館電子図書館構想」を策定し、それ以降の電子図書館構築の骨格を明らかにしました。

### 第2ステージ【2002～2012】 サービス離陸期・発展期

電子図書館サービスを開始し、発展させてきた時期で、2002年に開館した関西館を拠点として、「近代デジタルライブラリー」の公開、インターネット資源の選択的収集事業（WARP）、各種の電子展示会等を公開・提供しました。

2003年には、政府においてe-Japan重点計画2003が策定され、「国のデジタルアーカイブ構想」、「ジャパン・ウェブ・アーカイブ構想」が掲げられ、それにあわせた形で、2004年2月に、「国立国会図書館電子図書館中期計画2004」が策定されました。

この中期計画では、デジタルコンテンツを広く提供するために、当館が国のデジタルアーカイブの重要な拠点となること、日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築し、利用者がワンストップで利用できるようにすることを目指しました。【スライド2】

この計画に基づき、2004年10月から、様々なデジタルアーカイブ内の情報を統合検索する仕組みの実用性を検証するために、デジタルアーカイブポータルプロトタイプ（ndldap）を開発し、2005年7月に試験公開しました。その後、実用システムとして「PORTA」を構築し、2007年10月に正式公開しました。以降、全国の図書館や、博物館、美術館、公文書館等の文化機関との連携先の拡大を進めると同時に、「PORTA」の後継として、国立国会図書館サーチ（NDL Search）の開発を進め、2010年8月に試験公開、2012年1月のシステムリニューアル時に、新NDL-OPAC等と併せて、正式運用を開始しました。

資料のデジタル化に関しては、2009年5月から大規模なデジタル化が開始され、2011年までに、冊子体としては230万冊、約2億枚の画像をデジタル化しました。また、2010年4月には 国等のウェブサイトの制度的な収集が開始されました。

### 知識インフラの構築を目指して

また、2010年 に、我が国の第4期科学技術基本計画の策定に向けて決定された「科学技術基本政策策定の基本方針」（2010年6月総合科学技術会議基本政策専門調査会決定 ）で、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」という方向性が提示されました。これを踏まえて、当館において、2011年に「第三期科学技術情報整備基本計画」を策定しました。「知識インフラ」とは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を知識として集約し、新たな知識の創造を促進し、知識の集積・流通・活用と創造するサイクルの構築を目指すものです。

*新たな知識の創造のためには、分野を越えた知識の関連付けが必要であり、日本中に散在するコンテンツの所在を集中管理し、そこに検索をかければ、関連する全ての必要なコンテンツが得られるようにするものです。そこでは、単に情報を集めたものではなく、関連するものが有機的に結合され、ネットワーク的に統合化されたものであり、日本中にある芸術を含んだあらゆる学問・研究のコンテンツ、研究ツール、社会状況データ等が知識の形に組織化され、これらの知識・情報が公開され、全ての人が共有できることを目指すこととされました。*

## 2012年以降進めてきたこと

### 第3ステージ【2012～2014】 総括と再始動期、見直し期

　これまで進めてきたサービスの総括と新たな方向性を示す時期です。新たな事業・サービスとして、民間オンライン資料制度収集（2013年7月）、図書館向けデジタル化資料送信サービス（2014年1月）を開始しました。

### 2012年1月のシステムリニューアルの総括と、次期業務システム最適化計画（2013～2017）の策定

2012年のシステムリニューアルにおいて、パッケージ製品を核として構築したシステムに、基幹業務・サービスを適合させました。これにより、システムの最適化は進み、またデジタルコンテンツによる利用者サービスも拡充しました。*必要なサービス要件を満たすパッケージの選定、パッケージに合わせた業務フローの見直しが十分とは言えず、結果として、多くの機能を外付け開発し、また、新しいシステムに対する利用者からの苦情も多く寄せられましたが、チューニングを行うことにより、現在は安定的に稼働しています。*2012年後半にシステムリニューアルの総括を行い、更なる最適化とサービスの向上を目指した2017年度までの業務システム最適化計画を策定しました。

* *システムリニューアル*
  + *2008年度から2012年度までのシステムリニューアルの実施では、パッケージ製品を核として構築したシステムに、基幹業務・サービスを適合させ、同時に類似のシステム群を統合する取組を行った。*
  + *システムリニューアルの総括として、企画（サービス要件定義）、業務要件・システム化要件定義、業務・システム構築、移行・研修の各段階での合意形成、各工程での実施内容の妥当性を評価できるスキル向上が教訓として洗い出された。*
  + *システムのリニューアルにより、当館の業務・システムの最適化は進んだが、緊縮予算の中にあっても発展していく当館のサービスを業務・システムが支えるためには、更なる最適化が不可欠である。次期の業務システム最適化計画を策定した。*
* *業務システム最適化計画*
  + *当館システムに求められる要件*
    - *①関係機関との分散収集・保存を可能にすること*
    - *②効率よく電子資料を保管・閲覧できる電子書庫*
    - *③利用者のニーズに合ったユーザインタフェース*
    - *④当館の情報を活かした高度な検索*
    - *⑤外部システムからのスムーズなデータアクセス*
  + *方針と目標*
    - *①特性を考慮したシステムの統合*
    - *②共通するシステム要素の集約と汎用化*
    - *③継続的な業務・システムの最適化*
    - *④業務を効率化する小さなツールの推進*
    - *⑤外部委託範囲の見直し*
    - *⑥システムハードウェアの最適化*
      * *ⅰ）サーバの仮想化*
      * *ⅱ）クラウドサービスの利用*
      * *ⅲ）信頼性要件の見直し*
    - *⑦システム運用の効率化*
    - *⑧分散ファイルシステムの活用*

### 「私たちの使命・目標2012-2016」の策定

当館として今後果たすべき使命とおおむね5年間にわたって取り組むと目標して、「私たちの使命・目標2012-2016」を策定し、デジタル時代に適合した図書館を目指しています。

*「私たちの使命・目標2012-2016」では、「国立国会図書館は、出版物を中心に国内外の資料・情報を広く収集し、保存して、知識・文化の基盤となり、国会の活動を補佐するとともに、行政・司法及び国民に図書館サービスを提供することを通じ、国民の創造的な活動に貢献し、民主主義の発展に寄与します。」としています。*

* 基本問題の検討

*デジタル情報時代における課題を洗い出し、緊縮財政の中で、如何にして資源を確保して当館の事業を進めていくかの検討を始めており、まず、当館としての蔵書構築、利用者者サービスのあり方から着手しています。*

### 東日本大震災アーカイブ

2011年3月には、東日本大震災が発災し、甚大な被害をもたらしました。この被災・復興の記録を後世に残すとともに、今後の防災・減災に役立てるように、知識インフラの構築の一環で、分野を特定した実現形の先行事例として、2013年3月には、大震災に関するあらゆる記録、記憶を保存し、一元的に検索できるようにする「東日本大震災アーカイブ（ひなぎく）」を構築しました。構築に当たっては、図書館サービスを効率的、効果的に進められるように、次世代技術の実用化実証実験に取り組み、成果の積極的な活用を図りました。

*アーカイブでは、クローラによる収集のみならず各機関が収集した映像・観測データ等を受け取ることを想定して、理能力、ストレージ容量を必要に応じて増強できる仕組み、として、分散処理サーバ、分散ファイルシステムを導入しました。また、大震災アーカイブ自身が、災害で消失してしまわないように、ディザスタリカバリも考慮しました。また、画像・映像なども的確に検索ができるように、明確なメタデータが付与されていない情報にも可能な限りメタデータを自動付与する仕組み、本文も含めたテキストの全文インデキシング等の実装を考慮しました。*

## 今後10年で目指すところ（2015年～2024年）

### 第4ステージ【2015～2024】 本格的なデジタル情報の普及期、サービスの変革期

次のシステムリニューアルが予定される2020年から展開されるサービスの構築に当たっては、今後の社会の要請、情報技術の実用化動向を想定した図書館サービスの姿と、その実現に向けて実施すべき事項を明確にする必要があります。

知識インフラの実現形として、様々な分野のあらゆる記録を情報として集約し、相互に関連付けて知識化し、将来にわたって利用を保証して、「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを構築することを想定しています。

### 2020年から数年の予測と、国全体での対応

我が国では、「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部）等により、今後10年を見据えた政府の取組が示されました。従来の事業モデルの「改善」だけでなく、事業モデルそのものを創造・転換する「イノベーション」を創出すること、あらゆる分野の知的財産を対象としたビッグデータも含めたクラウド化、*非営利目的での利用のみならず産業利用も含めたコンテンツの多岐にわたる利活用の仕組み（*オープン化*）*、*ユーザが共同でコンテンツを制作する*クラウドソーシング、*産業競争力強化のための*知財人財の育成・確保、*新しい産業の創出環境の形成に向けた制度整備、*コンテンツの権利処理の円滑化、電子書籍の普及促進、*コンテンツがプラットフォームをリードする*エコシステムの実現、*データの収集・蓄積・分析による*多様な付加価値の創造の研究開発、日本の伝統や文化*に根ざした魅力あるコンテンツ・製品など*の発掘・創造等が示されています。

*このビジョンが想定する10年後の社会の姿の実現に向けて、国全体の方針として、「経済財政運営と改革の基本方針2014」、「学術研究の大型プロジェクトの推進に関する基本構想ロードマップ」等、具体的な実施に繋がる計画が検討されています。*

*学術情報ネットワーク（SINET）の整備に関する提言、ロードマップは、大学を中心とした機関での学術分野にフォーカスしたものではあるが、文化資産を含めたデジタルコンテンツの保存・活用基盤での課題と整備すべき内容と、かなり重なっているもので。国全体のアーカイブ構築の施策は学術分野の成果を活用することで、効率的・効果的に構築できると思われます。*

* *今後10年を予測*
  + *出版物*
    - *今後10年は、出版形態の主体が紙からデジタルへ移行する時期*
      * *出版物としての流通から情報としての流通へ。⇒収集の主体がデジタルへ移行*
      * *収集の主体がデジタル、情報としての組織化、物と情報の保存、提供はデジタルが主体*
  + *コンテンツ生成*
    - *当館所蔵資料のデジタルコンテンツ化率はかなり高くなっているであろう*
      * *（関係者協議内容が見直されて）全文テキスト化も可能になって、ハイブリッド型の電子書籍化も行えているであろう*
    - *出版社による電子書籍販売サービスのビジネスモデルは成熟しているであろう*
      * *少量出版物は、PODに移行しているであろう*
    - *孤児作品や絶版作品、フェアユースのための法制度が導入され、容易に電子書籍化ができるようになっているであろう*
    - *公的資金で作成された出版物は、オープンアクセスになっているであろう？*
  + *収集*
    - *現在のような閲覧を不可にするDRMでなく、電子透かし的なDRMに代わっているだろう*
    - *収集の主体はデジタル化コンテンツになっているであろう*
      * *当館によるオンライン資料（電子書籍・電子雑誌等）の収集は軌道に乗っているであろう*
  + *組織化*
    - *物としての組織化から、マイクロコンテンツとしての情報の組織化へ移行？*
  + *保存*
    - *紙とデジタルでなく、媒体としての長期保存、情報としての利用、情報の長期保存*
    - *原本性を保証する役割を果たすことが必要である*
      * *一つの識別子で、バージョンを管理することは現時点においても可能となっている*
  + *提供*
    - *提供の主体はデジタルコンテンツになっているであろう*
      * *来館者の閲覧・複写申込みは、デジタル化コンテンツが主流となっているであろう*
    - *当館、公共図書館では、商用の電子書籍販売サービスを活用しているであろう。*
      * *閲覧提供は、有償オンラインとして収集した情報か、電子書籍サイトかは、予想できない*
    - *当館のデジタル化資料送信サービスもほとんどの図書館で利用されているであろう。（海外の図書館も、出版社への支援活動が功を奏していれば、自宅への配信も可能になっているかも知れない）*
    - *登録利用者への特別サービスが考えられる*
      * *有償で、付加サービスも*
      * *デジタル複写サービス*
      * *認証も、マイナンバー、Open－ID*
    - *図書館資料を有償で利用することの結論は10年では出ないだろう*
  + *ファクトデータ*
    - *研究データは、学術情報基盤でビッグデータとして利用可能になっているであろう*
  + *ウェブアーカイブ*
    - *収集するインターネット情報は、公共機関、NPO、民間に広げている。*
    - *セマンティックウェブ技術や、新しい技術開発が進み、人工知能（AI）も実用レベルになっているであろう*
    - *ウェブアーカイブとデジタルアーカイブの情報が関連付けられる技術も相当進展している*
    - *その意味でも、現時点では、可能な限り収集を拡大していくべき*

### アーカイブに関連した国の活動の方向性

*これからの当館の電子図書館事業の展開を考える上で、国全体の議論として押さえておきたい。*「知的財産政策ビジョン」に基づく国の施策の中で、当館に直接関連する計画として、電子書籍化と利活用の促進に関する構想【スライド3】、デジタル文化資産の保存・活用に関する構想、学術情報の公開と共有の拡充に関する計画、大規模災害の記録と記憶の保存などのアーカイブの構築等が検討されています。

*その中で取り上げられている文化的資産は、無形・有形を問わず、また媒体を問わず、多岐にわたっており、様々な機関により分散してアーカイブされています。*

これらは、目的は異なっていても、対象とする文化的資産は相互に関連するものであり、知識インフラを目指した仕組みとして、国全体で*文化関係資料の価値を高め、新たな文化や情報を生み出す*社会基盤として、「恒久保存・継承・公開・活用」が可能な、ナショナルアーカイブとなることが望まれます。【スライド4】

### 必要な機能の想定

*このようなアーカイブは、我が国の多様な文化を知識として保存・継承する役割、様々な分野の専門家が参加し、新たな文化を知識として創造していくための社会的な基盤としての役割、それらの知識を利用目的に応じて発信する役割、そして、これらの仕組みを統括し運用していく役割が考えられ、文化的資産を館種毎に集約している拠点と、新たな知識を創造し発信するしている拠点等が分担して構築・運用していくことが必要です。*これを実現するために、以下の役割を持つ基盤が想定されています。

#### 恒久的保存基盤

様々な分野・機関*での個別目的のアーカイブ*に共通のプラットフォームを提供し、恒久的な分担保存を行うとともに、*それぞれのアーカイブの中にあるコンテンツの中から*必要な情報を*分野横断で的確に絞り込んで検索し*取り出せるようにしたバックヤード的な役割です。*当館が収集したデジタルコンテンツのアーカイブである*国立国会図書館デジタルコレクション、インターネット情報収集蓄積事業（WARP）、東日本大震災アーカイブ（ひなぎく）、全国のアーカイブをネットワーク化した国立国会図書館サーチ（NDLサーチ）の*アプリケーション・プログラミング・インターフェース(API)機能*をベースとして、情報間を*機械的に*連携させるセマンティックウェブ技術*、リンクデータ化*等による実現が想定されます。

* *恒久保存と利活用のための共通プラットフォームを作る*
* *分散アーカイブの構成*
  + *恒久的保存基盤は、１つのサイトにすべてを集約するのではなく、各機関のアーカイブを集約する分野毎のノードがあり、それを集約拠点に集め、これらが相互にコンテンツを持ち合って、障害、災害に備えて将来にわたって利用を保証する仕組みを想定。（いわゆるP2Pネットワークのイメージ）*

#### コンテンツ創造基盤

*現在は、個々の分野のアーカイブは、閉じた世界での価値として創造され蓄積されているので、それの原作の書籍等、他の分野の情報と関連付けられていません。*

異種のアーカイブと結び付いて、*お互いの*コンテンツが人を媒介にして、*それにより*新しい価値を生み出すものとして、*それぞれの分野の専門家である*キュレーター、ライブラリアンの能力の発揮の場で、保有する情報に付加価値を付けたり、他の分野のコンテンツと関連付けて、二次的著作物を創造したり、元になったコンテンツへナビゲートできるようにして*、ここで生成された情報が、*新たな知識として恒久的保存基盤に蓄積することが想定されます。

* *現在のヨーロピアーナとの違い*
  + *現在のヨーロピアーナは、どこにどういうものがありますというのが検索できるもの。一部ギャラリ的なものも提供（NDLの電子展示会レベル）、あとは考えてごらんなさいというもの。⇒現在のNDLサーチのレベル*
    - *玉も石も一緒に出てくるので、基本的に件数は多いけれども、そこから何かを学ぼうとはなかなかいかないレベル*
  + *日本がやるからには、そのレベルを超えて、玉石混交ではなくて、ちゃんとキュレーションのレイヤーが入る。全国の博物館、美術館の人たちの本当の仕事はここを作ること。*
    - *データを磨いて、自分たちの持っているものの価値をいろいろなところに発信してつないでいくことが、これからのキュレーター、ライブラリアンの仕事*

#### 情報発信基盤

*「見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」として、*広く国民による新たな知識の創造、新産業の創出、科学技術イノベーション、教育活用、地域活性化、国際文化交流、防災情報等、様々な利用者毎の目的に応じて、恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする基盤です。

国民が文化芸術を鑑賞し、また、*日本遺産（Japan Heritage）としての魅力ある*日本文化を発信する分野では、文化庁が構築している「文化遺産オンライン」をベースとして、単なる分野横断的な検索だけでなく、新たに創造された知識と併せて発信する日本らしいEuropeanaの形を目指します。

また、大規模災害防災・減災情報に関しては、「*東日本大量災アーカイブ（*ひなぎく*）*」をベースに、大規模自然災害の地震、津波、降雨・水位等の*ストックデータ、リアルタイム*データを発信するポータルの新設、*学術分野の高度な研究成果の発信は、国立情報学研究所が運営している「学術情報ネットワーク（SINET）」をベースに、また、*文献情報に関しては、国立国会図書館サーチの*国立図書館デジタルコレクションのユーザインタフェース（GUI）*をベースとすること想定します。

#### 運用基盤

*データ保存課題、人材不足問題、資金不足問題、権利処理コスト問題など、デジタルアーカイブの活性化を阻害する課題を制度的に克服する方策(人材育成策・予算措置・権利処理に関する法改正など)を導き出し、解決を図っていく中核的な役割を担う推進母体(司令塔)としての組織・体制も必要です。この基盤では、*ナショナルアーカイブ全体の戦略企画、デジタル情報の保存や利活用のための調査研究、研究開発、デジタル化支援、アーカイブに所蔵された資料に関する知識、読解力とIT技術の知識等も備えた高度な専門的人材の育成、孤児著作物の権利処理や、*絶版作品のデジタルアーカイブ化における所有権、肖像権問題も含めた*権利情報DBの構築を促進する等の役割を持つことを想定します。

* *デジタルアーカイブとして必要な機能*
  + *デジタル化機能*
    - *デジタル化（スキャンニング）、OCR化、構造化（文書構造、LOD化）、テキスト化（EPUB、DAISY等）*
    - *デジタル化ガイドライン*
  + *コンテンツの収集・一時保管機能*
    - *識別子付与、メタデータDB登録、提供用コンテンツ*
  + *保存機能*
    - *デジタルデータの恒久保存環境整備*
      * *デジタル化、収集・組織化・保存・提供のすべての機能を研究開発成果の活用。（産学官での共同による実用化実証実験）*
    - *分散デジタルアーカイブ、ディザスタリカバリ環境*
      * *アーカイブは民間で継続的に行うことは困難。また、１つの機関で維持することも困難。*
    - *物理保存（消えてしまわないように）*
      * *物理媒体の長寿命化*
      * *寿命前の他の物理媒体への移行対策*
    - *論理保存（使えるように）*
      * *フォーマット変換*
      * *（マイグレーションとエミュレーション）*
        + *OAISフレームに準拠した保存管理機能（保存用のメタデータの管理、アーカイブ情報パッケージ（AIP）*
        + *標準仕様の策定を推進した機関が、新たな標準仕様へのマイグレーションを保証する責務を持たせる*
    - *恒久的なデータ保存（ダークアーカイブ（100年1000年メディア：原則触らない））*
      * *民生品としてマーケットが見えないため、民間での技術開発が進まない*
    - *保存と利活用のためのデータ保存*
      * *短期的な対処として、物理媒体のマイグレーションを自動化した分散ファイルシステムの適用*
      * *長期的な対処として、拠点分散を自動化したネットワークシステム（P2Pネットワーク等）*
  + *権利情報・管理情報の管理機能*
    - *識別子を付与し、コンテンツの利用に関わる著作権の存否、所在に関する情報、デジタル化状況を管理*
    - *関係機関が保有する情報を関連付けて、出版界、図書館界で活用する*
  + *利活用機能*
    - *一元的に関連情報を含めて検索し、閲覧及び二次利用できる仕組み*
      * *分野横断的検索システム（NDLSearchをベースに）*
      * *目的別利用促進の取り組み（文化遺産オンラインをベースに）*
  + *研究開発、実用化実証実験（産学官の研究開発成果を活用）*
    - *デジタル化、収集・組織化・保存・提供のすべての機能の構築にあたっては産学官の研究開発成果の実用化実証実験を行い、有用性が認識された技術を適用*

### 今後のNDLの活動の方向性

*国全体のアーカイブ構築の中で、当館は何をしていくべきかを考えます。*

当館は、唯一の国立図書館として、納本制度、インターネット資料**の**制度収集、オンライン資料の制度収集、保存のためのデジタル化等、法律により「権限」が与えられており、確実な収集・保存・提供の実施の「責任と義務」を負っております。しかし、紙媒体の情報については国内の出版物に限られており、しかも、100パーセント完全に収集できているわけではありません。更に、デジタル化された出版物は、全てを１つの組織で収集・保存すること自体が不可能です。当館が主体的にアーカイブできる部分は「選択的」にならざるを得ませんが、可能な範囲ではなく、他の機関に分担の「協力」を求め、国全体で、*文化資産のアーカイブの*網羅性を確保することにより、当館の使命・目標の達成を目指すことが重要です。

### 選択と協力

当館が、あらゆる情報のうち「選択」して収集、保存する情報の種別は、文献情報**が中心**であり、国内で刊行される冊子体の出版物、デジタル化されて流通する出版物です。各府省や関係機関等には、個々のアーカイブ構築において、標準的な仕様（通信規約、メタデータ、永続的識別子等）の適用によるリンクデータ化、オープンデータ化を推進することの「協力」を求めたいと考えています。

国としてのアーカイブの基盤においては、当館は、「恒久的保存基盤」を主管し、必要な情報を的確に取り出せるようにすること、「コンテンツ創造基盤」の分担として、文献情報等に基づいてライブラリアンが作成するレファレンス情報、調査報告書及び、電子展示会の創作を行うこと、「情報発信基盤」の分担として、国会関連情報、文献情報、ウェブアーカイブ情報等の発信サービスを行うことを想定します。その概念は、この20年、当館が進めてきたデジタルアーカイブ構築の発展形であり、また、「ひなぎく」で進めてきたあらゆる記録・記憶を保存する役割と同義のものとして位置づけられます。

## 図書館における情報システムの役割

*当館の使命・目標の達成を目指して、社会的な存立意義を維持向上できるようにするため、業務・システムの効率化に留まらず、市場ニーズ、利用動向の変化に対して、実績・予測情報の確認、分析、対応の指示を速やかに行えるように、業務・組織を再構築するための計画を策定することが重要です。*

*情報システムは、今までは、人海戦術的な業務を情報システムに置き換えることによって、図書館サービスの高度化、業務・サービスの運営の効率化・コスト削減のための手段として活用してきました。政府における「電子政府構築計画」も同様です。今後は、電子情報の特性（共有・連携のしやすさ、検索の高度化、情報移動の容易さ等々）を生かしたデジタル情報時代の電子図書館サービスの構築を目指すとともに、民間企業で積極的に導入されている組織の持つ様々な資源(人材、資金、設備、資材、情報など)を統合的に管理・配分し、業務の効率化や経営の全体最適を目指す手法を支援するツール（ERP）、ビジネスプロセスを可視化しプロセスの分析、業務改善検討、モニタリングを行うことを通じて、継続的にビジネスプロセスの改革を推進するマネジメントツール（BPM：Business Process Management）の導入など、業務実施状況の迅速な数値化、可視化により、経営の強化、業務の最適化、変化への対応を速やかに行えるようにすることが、情報システムに求められていると考えます。1994年に国内機関の中で先行する形で始めて20年を経た当館の電子図書館サービスの進展が加速されないのは、人員・予算・著作権などの問題がありつつも、当館の使命・目標の達成を目指して、社会的な存立意義を維持向上できるようにするため、業務・システムの効率化に留まらず、市場ニーズ、利用動向の変化に対して、実績・予測情報の確認、分析、対応の指示を速やかに行えるように、業務・組織を再構築するための計画を策定する経営判断のための情報を速やかに提示できなかった情報システムにも一因があったとも考えられます。今後は、社会環境変化への迅速な対応」および「業務プロセス改革」を実現するために、経営戦略とIT戦略を密に連携・融合させ、最大限の効果を発揮させることが重要です。*

## おわりに

デジタル情報時代を踏まえ、物理媒体としての原資料は文化財として保存するために、デジタル化していく必要があり、一方では、収集すべき出版物は、物理媒体からデジタル化コンテンツへ移行しつつあります。このようにデジタル化が進む状況において、文化的資産をあらゆる人々が将来にわたり享受、活用できるようにし、人々の創造的な活用に貢献するためには、社会全体でデジタル情報資源の充実に取り組む必要があり、デジタルアーカイブは重要な役割を果たすことになります。

産学官のそれぞれの組織は、それぞれの所掌範囲で「デジタルアーカイブの構築と利活用」の施策を計画していますが、これらの施策が同一の方向性を持って、相互に資源を補完し合っていく必要があります。当館は、ナショナルアーカイブを構築し、さらに、日中韓電子図書館イニシアティブ（CJKDLI）でのポータル、世界レベルでの「インターナショナルアーカイブ」の構築へと発展することを目指し、その一翼を担っていきたいと考えています。

# （2015年）国立国会図書館のサービスシステムの歩みと今後の方向性 ―電子図書館事業20年を迎えた新たな方向性の模索―

月報用原稿素材

2015年1月19日

国立国会図書館

専門調査員　中山正樹

## はじめに

2014年は、当館での電子図書館サービスの構築に着手して20年を経過したところであり、今までの歩みを振り返るとともに、デジタル情報化時代が本格化する中で、図書館サービスシステムの今後の方向性についてについて考察します。

## この20年間の電子図書館構築の歩み

1980年代から、海外の多くの先進的な図書館同様、日本でも電子図書館事業に取り組んできました。1988年、NDLは関西地区に新たな大規模図書館を設置する構想の検討を開始し、1992年、21世紀初頭に関西学術文化研究都市の一角に関西館を設置することとし、具体的な構想を取りまとめました。関西館の予定する機能が、電子図書館的な機能であったこと、また国の産業構造審議会情報産業部会が公共部門の情報化を積極的に進めるべきとの提案を行ったことで、1994年に我が国で最初の大規模な電子図書館の実証実験プロジェクトを実施することとなりました。その後、2002年から本格的なサービスとして離陸し発展させて、現在に至っています。

### 第1ステージ【1994～2002】 揺籃期・始動期

NDLは、1994年1月、通商産業省（現：経済産業省）の高度情報化プロジェクト事業の一環で、情報処理振興事業協会（現情報処理推進機構（IPA）と国立国会図書館（NDL）が協力して、パイロット電子図書館プロジェクトを開始しました。パイロット電子図書館プロジェクトの目的は、21世紀の高度情報社会において、地球規模の知的財産を誰でも容易に利用できるように、地球上に広く分散して個々に収集・蓄積されている知的資源を、空間的・時間的制約を越えてアクセス可能とする環境を提供するための実証実験です。国立国会図書館を含む58館の公共図書館が総合目録ネットワーク実験に参加し、約1,400万件の書誌データが蓄積され、また、パイロット電子図書館実証実験では、当館所蔵の貴重書、明治期刊行図書・雑誌、出版社から提供を受けた資料等をデジタル化した約1,000万ページ（CD-ROM約2000枚）のコンテンツを利用して、効率的な検索・利用方法、ユーザインタフェース、電子化データの高度利用、効果的な電子図書館の構築支援の手法等について実証実験を行いました。その成果であるシステム及びデジタル化コンテンツは、その後の電子図書館サービスに引き継がれています。

また、1995年には、先進7カ国首脳会議（G7）の共同研究テーマの一つに電子図書館が取り上げられ、日本がG7電子図書館プロジェクトの共同幹事国をフランスと共管しました。

### 第2ステージ【2002～2012】 サービス離陸期・発展期

実証実験の成果を踏まえて、2002年10月に開館された関西館を拠点として、近代デジタルライブラリー、インターネット資源の選択的収集事業（WARP）、各種の電子展示会を公開・提供しました。

2003年には、e-Japan重点計画2003、e-JAPAN戦略Ⅱ加速化パッケージ、（内閣官房IT戦略本部）において、「国のデジタルアーカイブ構想」、「ジャパン・ウェブ・アーカイブ構想」の実現を、また2004年には、e-Japan重点計画2004において「国立デジタル・アーカイブ・ポータル構想」を一層推進することとされました。このような動きを踏まえて、NDLでは、2004年2月に「電子図書館中期計画2004」を策定しました。この中期計画において、デジタルコンテンツを広汎な利用者に提供するために、当館が国のデジタルアーカイブの重要な拠点となるということ、また国内外の多様な利用者層の需要に応じ、日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築し、利用者がワンストップで利用できるようにすることを目指すとされました。

この計画に基づき、2004年10月から、様々なデジタルアーカイブ内の情報を統合検索する仕組みの実用性を検証するために、デジタルアーカイブポータルプロトタイプ（ndldap）を開発・試験公開し、その後、実用システムとして「PORTA」を構築し、2007年10月に正式公開しました。システムは、*Web2.0時代の****サービス指向アーキテクチャ（*SOA）を適用し**、仮想OS上でOSSを組み合わせ、国際標準のメタデータ記述要素、記述規則、メタデータ交換の共通APIを利用して、商用を含めた外部サービスと**マッシュアップによるサービス連携**、コンテンツマネジメントシステム（CMS）による機能追加の容易性を確保しました。検索機能として連想検索検索エンジン（GETA）等を利用した**あいまい検索**も実装しました。統合検索対象は、近代デジタルライブラリ、青空文庫、国立公文書館アジア歴史資料センター、NDL-OPAC等の媒体形式か異なるもの、分散所蔵しているもの、**地域情報の情報をまとめているアグリゲータ**としてデジタル岡山大百科等を**ワンストップ**で検索し、コンテンツに辿り着けるようにしました。以降、科学技術振興機構（JST）、国立情報学研究所（NII）をはじめ、全国の個々の図書館や、博物館、美術館、公文書館等の文化機関との連携の拡大を進め、「PORTA」の後継として、国立国会図書館サーチ（NDL サーチ）の開発を進めました。NDLサーチでは、人間文化研究機構、**商用電子書籍ポータル**のhon.jpとの連携等のPORTAが担ってきたデジタル情報のポータルとしての役割を継続するとともに、ゆにかねっとが担ってきた各地の図書館蔵書の総合目録としての役割を引き継いだうえ、NDL-OPAC、インターネット資料収集保存事業（WARP）、国立国会図書館デジタルコレクション、国会会議録検索システム、リサーチ・ナビ（調べ方案内）、レファレンス協同データベース等の当館の主要データベースを統合検索することを可能としています。システムの構築に当たっては、外部専門家の参画、OSSの更なる適用によるコストダウン、次世代技術の試行、共通APIの実装の働きかけによる連携先拡大の加速化、著作単位でのグルーピング表示、キーワードサジェスト、障害者向け機能、日中韓英翻訳機能、パーソナライズ機能、スマートフォン対応を実装しました。また、GUIでの提供だけでなく、収集したメタデータをオープンデータとして、APIで外部システムに提供する「**情報ハブ**」の役割も提供しています。

2010年8月に試験公開、2012年1月のシステムリニューアル時に、新NDL-OPAC等と併せて、正式運用を開始しました。冊子体刊行物の収集・組織化業務及び、蔵書検索・申込システム「NDL-OPAC」を、パッケージに切り替えることによる開発・運用コストの大幅な削減するとともに、デジタル化資料の来館者サービスシステム等、デジタル時代の図書館利用者サービスの基本機能の充実を図ってきました。

また、デジタルアーカイブでは、資料のデジタル化に関しては、2009年5月から大規模なデジタル化が開始され、2011年までに、冊子体としては230万冊、約2億枚の画像をデジタル化しており、また、2010年4月には 国等のウェブサイトの制度的な収集が開始されました。

### 知識インフラの構築を目指して

2010年 に、我が国の第4期科学技術基本計画の策定に向けて決定された「科学技術基本政策策定の基本方針」*（2010年6月総合科学技術会議基本政策専門調査会決定 ）*で、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」という方向性が提示されました。これを踏まえて、当館において、2011年に「第三期科学技術情報整備基本計画」を策定し、**国の知識インフラの構築の一翼を担う**こととしました。「知識インフラ」とは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を知識として集約し、新たな知識の創造を促進し、知識の集積・流通・活用と創造するサイクルの構築を目指すものです。

### 東日本大震災アーカイブ

2011年3月には、東日本大震災が発災し、甚大な被害をもたらしました。この被災・復興の記録を後世に残すとともに、今後の防災・減災に役立てるように、**知識インフラの構築の実現形の先行事例**として、2013年3月には、大震災に関するあらゆる記録、記憶を保存し、一元的に検索できるようにする「東日本大震災アーカイブ（ひなぎく）」を構築しました。

### 第3ステージ【2012～2014】 総括と再始動期、見直し期

　これまで進めてきたサービスの総括と新たな方向性を示す時期です。

また、新たな事業・サービスとして、民間オンライン資料制度収集（2013年7月）、図書館向けデジタル化資料送信サービス（2014年1月）を開始しました。

## 今後10年で目指すところ（2015年～2024年）

### 第4ステージ【2015～2024】 本格的なデジタル情報の普及期、サービスの変革期

次のシステムリニューアルが予定される2020年から展開されるサービスの構築に当たっては、今後10年を踏まえて、社会の要請、情報技術の実用化動向を想定した図書館サービスの姿と、その実現に向けて実施すべき事項を明確にする必要があります。

分野を越えた知識インフラの実現形として、あらゆる記録を情報として集約し、相互に関連付けて知識化し、将来にわたって利用を保証するとともに、「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを構築することを想定しています。

### 2020年から数年の予測と、国全体での対応

我が国では、「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部）等により、文化資産のデジタルアーカイブ化を推進する政策も含めて、今後10年を見据えた知的財産関連の政府の取組としての目標が掲げられました。政策としては、従来の事業モデルの「改善」だけでなく、事業モデルそのものを創造・転換する「イノベーション」を創出すること、あらゆる分野の知的財産を対象としたビッグデータも含めたクラウド化、オープン化、ユーザが共同でコンテンツを制作するクラウドソーシング、知財人財の育成・確保、コンテンツの権利処理の円滑化、電子書籍の普及促進、データの収集・蓄積・分析による多様な付加価値の創造の研究開発、日本の伝統や文化に根ざした魅力あるコンテンツ・製品などの発掘・創造等の取組が示されています。

## 国全体の活動の方向性

### アーカイブに関連した国の活動の方向性

「知的財産政策ビジョン」に基づく国の施策の中で、当館に直接関連する計画として、電子書籍化と利活用の促進に関する構想、デジタル文化資産の保存・活用の基盤の整備に関する構想、学術情報の公開と共有の拡充に関する計画、大規模災害の記録と記憶の保存などのアーカイブの構築等が検討され、個々の施策の目的は異なっていても、対象とする文化的資産は相互に関連するものであり、知識インフラを目指した仕組みとして、国全体で*文化関係資料の価値を高め、新たな文化や情報を生み出す*社会基盤として、「恒久保存・継承・公開・活用」が可能な、ナショナルアーカイブとなることが望まれます。「資料・情報を文化資産として収集・保存する」ということは、従来からの出版物に相当する情報の範囲ではなく、美術館、博物館、文書館等が保有する無形・有形の文化財をデジタル化した情報を含め、インターネット上で流通している著作物全てを文化資産としてアーカイブすることです。

文化財のナショナルアーカイブの構築と運用に当たっては、様々な機関において、必要性の認識の共有やアーカイブ立国宣言の提案、制度的な課題解決のために、アーカイブ基本法の法制化、推進体制作りなどが議論されています。また、具体的なサービスシステムの仕組みとして、技術的課題のための研究開発が進んでいます。そのような状況を踏まえて、実現形として国のデジタル文化財のアーカイブとして、実現形のサービスシステムの基盤を想定します。

### 恒久的保存基盤

恒久的保存基盤は、恒久保存と利活用のための共通プラットフォームとして、１つの機関にすべてを集約するのではなく、各分野のアーカイブを集約する拠点が中核となって分散アーカイブを構築し、各機関の情報を相互に持ち合って、障害、災害に備えるとともに、情報のフォーマットのマイグレーションを行うことにより、将来にわたって利用を保証する仕組みです。その分散アーカイブを集合して、あたかも１つのアーカイブとして見えるようにして網羅性、完全性を確保し、個々の情報同士を意味的に関連付けて、情報間のネットワークを構築することを想定します。

このようにネットワーク化された情報に対して、分野を越えて網羅性を保証した検索インターフェースとして、本文の全文検索、あいまい検索、シソーラス検索などを組み合わせた検索で情報を取り出すだけでなく、取り出された情報から芋づる式に関連する情報を取り出せるようにします。

### 知識創造基盤

知識創造基盤は、それぞれの分野の専門家のみならず、広く国民も含めて、様々な分野の網羅的な知識を活用して、新たな著作物を創造する場です。創造活動を支援する基盤として、情報全体の基本情報としてのメタデータを付与する活動、記事、章節項、文節等の単位で組織化・構造化する活動、情報間を意味的に関連付けるための基本情報として、用語辞書、典拠、シソーラス辞書等を作成する活動を想定します。新たな知識を創造する活動は、まず、恒久的保存基盤に格納された網羅的な情報を活用して新たな知識を創作する活動があります。関連付けて利用できる情報の幅が広がるため、より高度な創造性が期待できます。

また、歴史的な文化財や現代文化を映像化、画像化、テキスト化する活動、構造化された情報に解題情報等を付与する活動、情報間を意味的に関連付ける活動、テーマを設定してデジタルギャラリを構築する活動等が含まれます。

*ここで生成された情報は、*新たな知識として恒久的保存基盤に蓄積されていきます。

### 情報発信基盤

*「見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」として、*広く国民による新たな知識の創造、新産業の創出、地域活性化、防災・減災、教育活用、教養・娯楽、観光、国際文化交流等、様々な利用者毎の目的に応じて、恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする基盤です。網羅的な情報から、利用目的に応じてあらかじめ適切に絞り込み、利用者の属性、スキル、利用場所に応じて、様々な画面インターフェースを用意して、利用者が必要とする情報、参考となる関連する情報を容易に得られるようにするもので、レファレンスサービスによる情報探索支援、オンラインレファレンスなども含まれます。

### 運用基盤

*データ保存課題、人材不足問題、資金不足問題、権利処理コスト問題など、デジタルアーカイブの活性化を阻害する課題を制度的に克服する方策(人材育成策・予算措置・権利処理に関する法改正など)を導き出し、解決を図っていく中核的な役割を担う推進母体(司令塔)としての組織・体制も必要です。この基盤では、ナショナルアーカイブ全体の戦略企画、デジタル情報の保存や利活用のための調査研究、研究開発、デジタル化支援、アーカイブに所蔵された資料に関する知識、読解力とIT技術の知識等も備えた高度な専門的人材の育成、孤児著作物の権利処理や、絶版作品のデジタルアーカイブ化における所有権、肖像権問題も含めた権利情報DBの構築を促進する等の役割を持つことを想定します。*

## 今後のNDLの活動の方向性

NDLは、「私たちの使命・目標2012-2016」の中で、「印刷出版物にとどまらず、電子的に流通する情報を含め、様々な資料・情報を文化的資産として収集し、保存します。」としています。

NDLは、唯一の国立図書館として、国内刊行物の納本制度、公的機関のインターネット資料（ウェブサイト情報）の制度収集、民間のオンライン資料（電子書籍・電子雑誌に相当する情報）の制度収集、保存のためのデジタル化等、法律により「権限」が与えられ、確実な収集・保存・提供の実施の「責任と義務」を負っており、その責任と義務は、可能な範囲で行えばいいということではありません。しかし、物としての紙媒体の出版物については、公共図書館が保有する郷土資料、美術館、博物館、文書館が保有する典籍資料をはじめとして、全てを収集できているわけではなく、また、電子書籍・電子雑誌に相当する情報は、セルフパブリッシングも含めて指数的に増加しており、ウェブサイト全体を丸ごとアーカイブするインターネット資料収集事業（WARP）において、もはや全てを１つの組織で収集・保存すること自体が不可能です。

当館が主体的にアーカイブするとともに他の機関の情報を併せて利活用できるようにしてきた範囲は、人・物・金の資源の制約により「選択的」にならざるを得ませんでしたが、国全体で、文化資産のアーカイブの網羅性を確保できるようにするためには、アーカイブの共通基盤の仕様等を提示し、他の機関に対して、その適用と分担の「協力」を求めることが重要です。

その概念は、パイロット電子図書館実証実験に始まって、電子図書館中期計画2004に基づいて進めてきたアーカイブ構築の考え方そのものです。「ひなぎく」で進めてきたあらゆる記録・記憶を保存する役割の発展系として、ナショナルアーカイブを推進することが、国民の期待に応える活動であり、NDLの使命・目標を達成することとなると思われます。

## おわりに

デジタル情報時代において、出版物は、冊子体から動画・音声等を含むマルチメディア化されたコンテンツへ移行しつつあり、また、冊子体の原資料は文化財として保存するために、デジタル化していくことが求められています。他の文化財の保有機関のデジタル化が進む状況において、文化的資産をあらゆる人々が将来にわたり享受、活用できるようにし、人々の創造的な活用に貢献するためには、社会全体でデジタル情報資源の「見える化」はもとより、より効率的なアクセスの保証に取り組む必要があり、組織を越えたナショナルアーカイブは重要な役割を果たすことになります。産学官のそれぞれの組織は、これらの施策が同一の方向性を持って、相互に資源を補完し合っていく必要があります。NDLは、ナショナルアーカイブの構築、さらに、世界レベルでの「インターナショナルアーカイブ」の構築へと発展することを目指し、その中核的な役割を担っていくべきと考えています。

今後10年のデジタル情報化の進展を見据えて、このようなナショナルアーカイブを利用して知識創造のための情報が入手できる状況において、知識創造を支援する図書館の役割と必要な機能の検討を加速させる必要があると考えます。

# （2015年）【カット分戻し】国立国会図書館のサービスシステムの歩みと新たな方向性の模索

## はじめに

　私は、情報処理振興事業協会（現 情報処理推進機構（IPA））に在籍中の1995年からパイロット電子図書館プロジェクトに関わり、2002年から国立国会図書館（以下、NDL）において電子図書館事業、情報システム関連事業に従事してきました。

　2014年は、NDLでの電子図書館サービスの構築に着手して20年を経過したところであり、事業を担当してきた立場から今までの歩みを振り返るとともに、デジタル情報化時代が本格化する中で、これまでの経験を踏まえた私見として、図書館サービスシステムの今後の方向性について考察します。

[ここに出典を記載します。]

*この20年間、国の施策、社会での表現、事業の括りは、微妙に変化しているが、大きな施策としての方向性は変わっておらず、確実に進んできたと認識している。この10年間の実績により、NDLのアーカイブの提供を含めて、MLA等の各アーカイブ機関と合わせて、国としてのアーカイブを構築し、保有しているコンテンツへのナビゲーションサービスの道筋ができた。この間の活動成果により、NDLは、国の文化的資産の保存と利活用を促進するための中核的な役割が期待されるようになった。*

*今後は、出版領域を含む文化資産のナショナルアーカイブとして、各機関が保有しているコンテンツを国全体のアーカイブとして保存し、かつ見つけやすくするとともに、アーカイブ内の知識を利活用した新たな知識が創造される環境の整備と、目的に応じた知識が自由に利用できるようにすることにより、創造された知識がアーカイブされ、より豊かな知識として循環し共有化が進むようにする。*

*そのような施策の１つとして、「デジタル文化資産の「保存・活用」の基盤の整備」がある。*

*電子書籍関連の利活用の促進に関する議論と、文化財、ポップカルチャーに関する保存と利活用に関する機論が進み、国としての文化的資産の保存と活用基盤の概念と構築のフレームワークが見えてきた。その概念は、NDLの電子図書館中期計画に基づいて進めてきたアーカイブ構築の発展形として考えられる。2020年東京オリンピックの開催の時点での実現を目指すアーカイブのフレームワークとサービスを想定する。*

## この20年間の電子図書館構築の歩み

　1980年代から、海外の多くの先進的な図書館同様、日本でも電子図書館事業に取り組んできました。NDLは1992年、21世紀初頭に関西学術文化研究都市の一画に関西館を設置するために、具体的な構想を取りまとめました。関西館の予定する機能が、電子図書館的な機能であったこと、また国の産業構造審議会情報産業部会が公共部門の情報化を積極的に進めるべきとの提案を行ったことで、1994年に我が国で最初の大規模な電子図書館の実証実験プロジェクトを実施することとなりました。その後、2002年から本格的なサービスとして離陸し発展させて、現在に至っています。

[ここに出典を記載します。]

### 第1ステージ【1994～2002】

揺籃期・始動期

* パイロット電子図書館プロジェクト

　NDLは、1994年1月、通商産業省（現 経済産業省）の高度情報化プロジェクト事業の一環で、情報処理振興事業協会と協力して、パイロット電子図書館プロジェクトを開始しました。このプロジェクトの目的は、21世紀の高度情報社会において、地球規模の知的財産を誰でも容易に利用できるように、広く分散して個々に収集・蓄積されている知的資源を、空間的・時間的制約を越えてアクセス可能とする環境を提供するというものです。

①総合目録ネットワークプロジェクト

　NDLを含む都道府県立・政令指定都市立の58館の公共図書館がこのプロジェクトに参加し、約1,400万件の書誌データを蓄積しました。

②電子図書館実証実験プロジェクト

　NDL所蔵の貴重書、明治期刊行図書、国内刊行雑誌、出版社から提供を受けた資料等約1,000万ページ（CD-ROM約2000枚相当）をデジタル化しました。そのコンテンツを利用して、効率的な検索・利用方法、ユーザインタフェース、電子化データの高度利用、効果的な電子図書館の構築支援の手法等について実証実験を行いました。

　これらの成果であるシステムおよびデジタル化コンテンツは、その後の電子図書館サービスに引き継がれています。

[ここに出典を記載します。]

* 全国公共図書館総合目録ネットワーク事業

NDL、都道府県立・政令指定都市を含む**58館の公共図書館**が総合目録ネットワーク実験に参加し、約1,400万件の書誌データを蓄積しました。全国的な公共図書館間の資料相互貸借に貢献するところとなっています。

* パイロット電子図書館実証実験プロジェクト

NDL所蔵の貴重書、明治期刊行図書・雑誌、出版社から提供を受けた資料等をデジタル化した**約1,000万ページ（CD-ROM約2000枚）**のコンテンツを利用して、効率的な検索・利用方法、ユーザインタフェース、電子化データの高度利用、効果的な電子図書館の構築支援の手法等について実証実験を行いました。その成果であるシステム及びデジタル化コンテンツは、その後の電子図書館サービスに引き継がれています。

また、1995年には、先進7カ国首脳会議（G7）の共同研究テーマの一つに電子図書館が取り上げられ、日本がG7電子図書館プロジェクトの共同幹事国をフランスと共管しました。

### 第2ステージ【2002～2008】

サービス離陸期

　実証実験の成果を踏まえて、2002年10月に開館した関西館を拠点として、近代デジタルライブラリー、インターネット資源選択的蓄積実験事業（WARP）、各種の電子展示会を公開・提供しました。

　国の取り組みとしては、e-Japan重点計画―2003、e-Japan戦略Ⅱ加速化パッケージ（内閣官房IT戦略本部）策定の議論の中で、「国のデジタルアーカイブ構想」、「ジャパン・ウェブ・アーカイブ構想」の実現を、またe-Japan重点計画―2004策定時には「国立デジタル・アーカイブ・ポータル構想」を一層推進することとされました。

[ここに出典を記載します。]

* 「国立国会図書館電子図書館中期計画2004」

　このような動きを踏まえて、NDLでは、2004年2月に「電子図書館中期計画2004」を策定しました。この中期計画において、デジタルコンテンツを広汎な利用者に提供するために、NDLが国のデジタルアーカイブの重要な拠点となるということ、また国内外の多様な利用者層の需要に応じ、日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築し、利用者がワンストップで利用できるようにすることを目指すとしました。

　この計画に基づき、様々なデジタルアーカイブ内の情報を統合検索する仕組みの実用性を検証するために、デジタルアーカイブポータルプロトタイプ（ndldap）を開発・試験公開し、その後、実用システムとしてPORTAを構築、2007年10月に正式公開しました。この開発にあたっては、可能な限り先進技術の適用を意識したのも特徴です。

[ここに出典を記載します。]

* PORTA

　PORTAの統合検索対象は、近代デジタルライブラリー、青空文庫、国立公文書館アジア歴史資料センター、NDL-OPAC、デジタル岡山大百科等です。媒体形式が異なるもの、分散所蔵しているもの、地域情報をまとめているアグリゲータ等を統合検索し、コンテンツにたどりつけるようにしました。PORTAの公開以降も、科学技術振興機構（JST）、国立情報学研究所（NII）をはじめ、全国の図書館や、博物館、美術館、公文書館等の文化機関との連携の拡大を進め、PORTAの後継として、国立国会図書館サーチの開発を進めました。

[ここに出典を記載します。]

システムは、*Web2.0時代の****サービス指向アーキテクチャ（*SOA）を適用し**、仮想OS上でOSSを組み合わせ、国際標準のメタデータ記述要素（ダブリンコア（DC）をベース）、記述規則、メタデータ交換の共通API（OAI-PMH、SRU、OpenSearch、リンクリゾルバとしてのOpenURL等）を利用して、商用を含めた外部サービスと**マッシュアップによるサービス連携**、 RSSによる情報のアグリゲーション、コンテンツマネジメントシステム（CMS）による機能追加の容易性を確保しました。検索機能として連想検索検索エンジン（GETA）等を利用した**あいまい検索**も実装しました。統合検索対象は、近代デジタルライブラリ、青空文庫、国立公文書館アジア歴史資料センター、NDL-OPAC等の媒体形式か異なるもの、分散所蔵しているもの、**地域情報の情報をまとめているアグリゲータ**としてデジタル岡山大百科等を**ワンストップ**で検索し、コンテンツに辿り着けるようにしました。以降、科学技術振興機構（JST）、国立情報学研究所（NII）をはじめ、全国の個々の図書館や、博物館、美術館、公文書館等の文化機関との連携の拡大を進め、「PORTA」の後継として、国立国会図書館サーチ（NDL サーチ）の開発を進めました。

### 第3ステージ【2008～2012】

発展期

* 国立国会図書館サーチ

　PORTAが担ってきたデジタル情報のポータルとしての役割を継続するとともに、総合目録ネットワークシステム（ゆにかねっと）が担ってきた各地の図書館蔵書の総合目録としての役割を引き継いだうえ、NDLの主要データベース（NDL-OPAC、インターネット資料収集保存事業（WARP）、国立国会図書館デジタルコレクション、国会会議録検索システム、リサーチ・ナビ（調べ方案内）、レファレンス協同データベース等）を統合検索することを可能としています。また、収集したメタデータをオープンデータとして、APIで外部システムに提供する「情報ハブ」の役割も提供しています。

　2010年8月に試験公開、2012年1月のシステムリニューアル時に、新NDL-OPAC等と併せて、正式運用を開始しました。このシステムリニューアルでは、冊子体刊行物の収集・組織化業務およびNDL-OPACをパッケージに切り替えることにより、開発・運用コストを大幅に削減するとともに、デジタル化資料を利用しやすい来館者サービスシステム等、デジタル時代の図書館利用者サービスの基本機能の充実を図りました。

[ここに出典を記載します。]

PORTAが担ってきたデジタル情報のポータルとしての役割を継続して、人間文化研究機構、**商用電子書籍ポータル**のhon.jpとの連携を進めるとともに、ゆにかねっとが担ってきた各地の図書館蔵書の総合目録としての役割を引き継いだうえ、NDL-OPAC、インターネット資料収集保存事業（WARP）、国立国会図書館デジタルコレクション、国会会議録検索システム、リサーチ・ナビ（調べ方案内）、レファレンス協同データベース等の当館の主要データベースを統合検索することを可能としています。システムの構築に当たっては、外部専門家の参画、OSSの更なる適用によるコストダウン、次世代技術の試行、共通APIの実装の働きかけによる連携先拡大の加速化、著作単位でのグルーピング表示、キーワードサジェスト、障害者向け機能、日中韓英翻訳機能、パーソナライズ機能、スマートフォン対応を実装しました。また、GUIでの提供だけでなく、収集したメタデータをオープンデータとして、APIで外部システムに提供する「**情報ハブ**」の役割も提供しています。

2010年8月に試験公開、2012年1月のシステムリニューアル時に、新NDL-OPAC等と併せて、正式運用を開始しました。冊子体刊行物の収集・組織化業務及び、蔵書検索・申込システム「NDL-OPAC」を、パッケージに切り替えることによる開発・運用コストの大幅な削減するとともに、デジタル化資料の来館者サービスシステム等、デジタル時代の図書館利用者サービスの基本機能の充実を図ってきました。

また、デジタルアーカイブでは、資料のデジタル化に関しては、2009年5月から大規模なデジタル化が開始され、2011年までに、冊子体としては230万冊、約2億枚の画像をデジタル化しており、また、2010年4月には 国等のウェブサイトの制度的な収集が開始されました。

* デジタルアーカイブ

　資料のデジタル化に関しては、2009年5月から大規模なデジタル化を開始し、2015年1月までに、冊子体としては約246万冊、約2億枚の画像をデジタル化しました。これらは現在、国立国会図書館デジタルコレクションで提供しています。

　また、2010年4月には 国等の公的機関のウェブサイトの制度収集を開始しました。

[ここに出典を記載します。]

* 知識インフラの構築を目指して

「第三期科学技術情報整備基本計画」

　2010年に、我が国の第4期科学技術基本計画の策定に向けて決定された「科学技術基本政策策定の基本方針」（2010年6月総合科学技術会議基本政策専門調査会決定）で、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」という方向性が提示されました。これを踏まえて、NDLにおいて、2011年に「第三期科学技術情報整備基本計画」を策定し、国の知識インフラの構築の一翼を担うこととしました。「知識インフラ」とは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を意味的に関連づけて知識として集約し、新たな知識の創造を促進し、知識の集積・流通・活用と創造のサイクル構築を目指すものです。

　知識インフラ構築の実現形の先行事例として、2013年3月には、「東日本大震災アーカイブ（ひなぎく）」を構築しました。大震災に関連する災害現象そのもの、災害前・災害直後・復興の過程、災害時の対応、他地域・次世代への教訓等のあらゆる記録を後世に残すとともに今後の防災に生かすため、関係府省、各種震災関連情報の保有機関と協力して分担収集・保存し、一元的に検索・閲覧できるようにしたものです。知識インフラ構築に必要なフレームワークと次世代技術を積極的に適用しています。

[ここに出典を記載します。]

2010年 に、我が国の第4期科学技術基本計画の策定に向けて決定された「科学技術基本政策策定の基本方針」*（2010年6月総合科学技術会議基本政策専門調査会決定 ）*で、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」という方向性が提示されました。これを踏まえて、当館において、2011年に「第三期科学技術情報整備基本計画」を策定し、**国の知識インフラの構築の一翼を担う**こととしました。「知識インフラ」とは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を知識として集約し、新たな知識の創造を促進し、知識の集積・流通・活用と創造するサイクルの構築を目指すものです。

*新たな知識の創造のためには、分野を越えた知識の関連付けが必要であり、日本中に散在するコンテンツの所在を集中管理し、そこに検索をかければ、関連する全ての必要なコンテンツが得られるようにする。知識は関連するものが有機的に結合され、ネットワーク的に統合化されたものであり、単に情報を集めたものではない。日本中にある芸術を含んだあらゆる学問・研究のコンテンツ、研究ツール、社会状況データ等が知識の形に組織化され、これらの知識・情報が公開され、全ての人が共有できることが大切である。*

* 東日本大震災アーカイブ（ひなぎく）

**知識インフの構築の実現形の先行事例**として、2013年3月には、「東日本大震災アーカイブ（ひなぎく）」を構築しました。大震災に関連する災害現象そのもの、災害前・災害直後・復興の過程、災害時の対応、他地域・次世代への教訓等のあらゆる記録を後世に残すとともに今後の防災に生かすため、関係府省、各種震災関連情報の保有機関と協力して分担収集・保存し、一元的に検索・閲覧できるようにするものです。知識インフラ構築にに必要なフレームワークと次世代技術を積極的に適用しています。

### 第4ステージ【2012～2014】

総括と再始動期、見直し期

　2012年1月にリニューアルしたサービスの総括を行い、更なる最適化とサービスの向上を目指した2017年度までの「業務・システム最適化計画」を策定しました。新たな事業・サービスとして、2013年7月から、民間のオンライン資料のうち、無償でかつDRMがかかっていない電子書籍・電子雑誌を制度的に収集・保存することが可能になりました。また、2014年1月には、NDLデジタル化資料のうち、著作権保護期間内のためNDLに来館しなければ閲覧できなかった資料を、入手困難な資料に限って送信する「図書館向けデジタル化資料送信サービス」を開始しました。これにより、公共図書館等において、NDLデジタル化資料も各図書館の蔵書と同様に閲覧・複写サービスを提供できるようになりました。

[ここに出典を記載します。]

2012年1月にリニューアル公開したサービスの総括を行い、更なる最適化とサービスの向上を目指した2017年度までの業務システム最適化計画を策定しました。

新たな事業・サービスとして、「民間オンライン資料制度収集」（2013年7月）により、電子書籍、電子雑誌として流通する資料を、無償でかつDRMが掛かっていない資料に限って、制度的に収集・保存することが可能になりました。また、「図書館向けデジタル化資料送信サービス」（2014年1月）により、NDLデジタル化資料のうち、著作権保護期間内のためNDLに来館しなければ閲覧できなかった資料を、入手困難な資料に限って、送信するサービスを開始しました。これにより、公共図書館において、NDLデジタル化資料も各図書館の蔵書と同様に閲覧・複写サービスを提供できるようになりました。　これまで進めてきたサービスの総括と新たな方向性を示す時期です。

また、新たな事業・サービスとして、民間オンライン資料制度収集（2013年7月）、図書館向けデジタル化資料送信サービス（2014年1月）を開始しました。

## 今後10年で目指すところ（2015年～2024年）

### 第5ステージ【2015～2024】

本格的なデジタル情報の普及期、サービスの変革期

　次のシステムリニューアルが予定される2020年から展開されるサービスの構築にあたっては、今後10年の社会の要請を踏まえて、情報技術の実用化動向を想定した図書館サービスの姿と、その実現に向けて実施すべき事項を明確にする必要があります。

　分野を越えた知識インフラの実現形として、あらゆる記録を情報として集約し、相互に関連づけて知識化し、将来にわたって利用を保障するとともに、「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを構築することを想定しています。

[ここに出典を記載します。]

*国立国会図書館中期計画2004がが策定されてから、10年以上経過した。当時の想定や、現在に至るまでの活動を振り返ってみると、各機関が想定通り事業を展開できたこと、社会情勢の変化などで事業ができなかったことなど、いろいろ見えてくる。総括すると、今後のナショナルアーカイブ関連の議論をする際に、参考になるのではないか。*

*また技術的な観点から、実現に当たって、2003年頃は、データの相互運用性を確保して検索性能の向上や利便性を高める次世代Web技術として、「セマンティックウェブ」関連の技術要素である、XML、RDF、OWL、URI等の適用を掲げてきたところである。昨今は、セマンティックウェブの実現への新しい潮流のなかで、Linked Open Data(LOD)として利用できる形での公開と識別子での相互関連付けを広めようという表現になっている。2003年頃に目指したアーカイブ構築の実現方法は、技術的な仕組みとしては微妙に変わってきているが、より一般にもわかりやすい「相互にデータを連携できる形での公開・利用」をLODという表現が定着することにより、普及が加速される時期に来ているように思える。*

## 国全体の活動の方向性

* 考察

*知の共有化を目指した知識情報の保存基盤と、新たな知識の創造を目指した利活用基盤の構築を目指して、今後は、国としてのアーカイブを更に利用しやすくするとともに、アーカイブ内のコンテンツを利活用した新たな知識の創造と、目的に応じた知識の発信を充実させることにより知の循環が更に進むようにする施策の実施の一翼を担えるようにしていくことが重要と考える。*

*そのような施策の１つとして、「デジタル文化資産の「保存・活用」の基盤の整備」がある。*

*電子書籍関連の利活用の促進に関する議論と、文化財、ポップカルチャーに関する保存と利活用に関する機論が進み、国としての文化的資産の保存と活用基盤の概念と構築のフレームワークが見えてきた。その概念は、NDLの中期計画に基づいて進めてきたアーカイブ構築の発展形として考えられる。*

*このビジョンが想定する10年後の社会の姿の実現に向けて、国全体の方針として、「経済財政運営と改革の基本方針2014」、「学術研究の大型プロジェクトの推進に関する基本構想ロードマップ」等、具体的な実施に繋がる計画が検討されています。*

*学術情報ネットワーク（SINET）の整備に関する提言、ロードマップは、大学を中心とした機関での学術分野にフォーカスしたものではあるが、文化資産を含めたデジタルコンテンツの保存・活用基盤での課題と整備すべき内容と、かなり重なっているもので。国全体のアーカイブ構築の施策は学術分野の成果を活用することで、効率的・効果的に構築できると思われます。*

### 今後10年の国全体の取り組み

　我が国では、「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部）等により、文化資産のデジタルアーカイブ化を推進する政策も含めて、今後10年を見据えた知的財産関連の政府の取り組みとしての目標が掲げられました。政策としては、従来の事業モデルの「改善」だけでなく、事業モデルそのものを創造・転換するイノベーションの創出、日本の伝統や文化に根ざした魅力あるコンテンツ・製品などの発掘・創造を目指すとともに、実現に向けて、ビッグデータの収集・蓄積・分析による多様な付加価値の創造に資する研究開発、オープンデータ化、権利処理の円滑化、人材の育成・確保等の取り組みが示されています。

[ここに出典を記載します。]

* *（予測）2020年から数年を見据えて、情報の提供と利用はどうなっているか？*

*NDLの今までの活動は、あらゆる情報、記録を対象とすることを目指してきましたが、制度面、予算面、人材面等の課題により、網羅性、完全性の確保としては不十分でした。しかし、今後10年を見据えると、現在の制度的課題、技術的課題はかなり解決されていると思われます。その状況において、情報の保存と利活用はどうなっているか想定します。*

### アーカイブに関連した国の活動の方向性

　「知的財産政策ビジョン」に基づく国の施策の中で、NDLに直接関連する計画として、デジタル画像の電子書籍化と利活用の促進に関する構想、デジタル文化資産の保存・活用の基盤の整備に関する構想、学術情報の公開と共有の拡充に関する計画、大規模災害の記録と記憶の保存などのアーカイブの構築等が検討されました。

　一方、個々の施策の目的は異なっていても、対象とする文化資産は相互に関連するものであり、知識インフラを目指した仕組みとして、国全体で文化資産の価値を高め、新たな文化や情報を生み出す社会基盤として「恒久保存・継承・公開・活用」が可能な、出版領域を含む文化資産のナショナルアーカイブとなることが望まれ

ます。「資料・情報を文化資産として収集・保存する」ということは、従来からの出版物に相当する情報の範囲だけではなく、美術館、博物館、文書館等が保有する無形・有形の文化財をデジタル化した情報を含め、インターネット上で流通している著作物全てを文化資産としてアーカイブすることです。

　文化資産のナショナルアーカイブの構築と運用にあたっては、制度的な課題解決のために、アーカイブ基本法の法制化、推進体制作りなどが議論されています。また、具体的なサービスシステムの仕組みとして、技術的課題のための研究開発が進んでいます。

　そのような状況を踏まえて、私見として、国のデジタル文化資産のアーカイブに必要なサービスシステムのフレームワークと主な役割を想定すると、次のとおりです。

[ここに出典を記載します。]

* *必要な機能の想定*

*このようなアーカイブは、我が国の多様な文化を知識として保存・継承する役割、様々な分野の専門家が参加し、新たな文化を知識として創造していくための社会的な基盤としての役割、それらの知識を利用目的に応じて発信する役割、そして、これらの仕組みを統括し運用していく役割が考えられ、文化的資産を館種毎に集約している拠点と、新たな知識を創造し発信するしている拠点等が分担して構築・運用していくことが必要です。これを実現するために、以下の役割を持つ基盤が想定されます。*

* ＜恒久的保存基盤＞

――恒久保存と利活用のための共通プラットフォーム

　１つの機関にすべてを集約するのではなく、各分野のアーカイブを集約する拠点が中核となって分散アーカイブを構築し、各機関の情報を相互に持ち合って、将来にわたって利用を保障する仕組みです。

　その分散アーカイブを集合して、あたかも１つのアーカイブとして見えるようにし、個々の情報同士を意味的に関連づけて、情報間のネットワークを構築することを想定します。

　このようにネットワーク化された情報に対して、本文の全文検索、あいまい検索、シソーラス検索などを組み合わせた検索で情報を取り出すだけでなく、取り出された情報から芋づる式に関連する情報を取り出せるようにします。

[ここに出典を記載します。]

* ＜知識創造基盤＞

――それぞれの分野の専門家のみならず、広く国民も含めて、様々な分野の網羅的な知識を活用して、新たな著作物を創造する場

　創造活動を支援する基盤です。情報全体の基本情報としてのメタデータを付与する活動、記事、章節項、文節等の単位で組織化・構造化する活動、情報間を意味的に関連づけるための基本情報としての用語辞書、典拠、シソーラス辞書等を作成する活動を想定します。

　関連づけて利用できる情報の幅が広がるため、より高度な創造性が期待できます。

　また、歴史的な文化財や現代文化を映像化、画像化、テキスト化する活動、構造化された情報に解題情報等を付与する活動、情報間を意味的に関連づける活動、テーマを設定してデジタルギャラリを構築する活動等が含まれます。

　ここで生成された情報は、新たな知識として恒久的保存基盤に蓄積されていきます。

[ここに出典を記載します。]

知識創造基盤は、それぞれの分野の専門家のみならず、広く国民も含めて、様々な分野の網羅的な知識を活用して、新たな著作物を創造する場です。創造活動を支援する基盤として、情報全体の基本情報としてのメタデータを付与する活動、記事、章節項、文節等の単位で組織化・構造化する活動、情報間を意味的に関連付けるための基本情報として、用語辞書、典拠、シソーラス辞書等を作成する活動を想定します。新たな知識を創造する活動は、まず、恒久的保存基盤に格納された網羅的な情報を活用して新たな知識を創作する活動があります。関連付けて利用できる情報の幅が広がるため、より高度な創造性が期待できます。

また、歴史的な文化財や現代文化を映像化、画像化、テキスト化する活動、構造化された情報に解題情報等を付与する活動、情報間を意味的に関連付ける活動、テーマを設定してデジタルギャラリを構築する活動等が含まれます。

*ここで生成された情報は、新たな知識として恒久的保存基盤に蓄積されていきます。*

*あたかも1つのアーカイブとして利用できる異種のアーカイブと結び付いて、お互いのコンテンツが人を媒介にして、それにより新しい価値を生み出すものとして、それぞれの分野の専門家であるキュレーター、ライブラリアンの能力の発揮の場で、保有する情報に付加価値を付けたり、他の分野のコンテンツと関連付けて、二次的著作物を創造したり、元になったコンテンツへナビゲートできるようにして、ここで生成された情報が、新たな知識として恒久的保存基盤に蓄積することが想定されます。*

* ＜情報発信基盤＞

――「見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」へ。

　様々な利用者ごとの目的に応じて、恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする基盤です。

　網羅的な情報から、利用目的に応じてあらかじめ適切に絞り込み、利用者の属性、スキル、利用場所に応じて、様々な画面インターフェースを用意して、利用者が必要とする情報、参考となる関連する情報を容易に得られるようにするものです。レファレンスサービスによる情報探索支援、オンラインレファレンスなども含まれます。

[ここに出典を記載します。]

*「見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」として、*広く国民による新たな知識の創造、新産業の創出、地域活性化、防災・減災、教育活用、教養・娯楽、観光、国際文化交流等、様々な利用者毎の目的に応じて、恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする基盤です。網羅的な情報から、利用目的に応じてあらかじめ適切に絞り込み、利用者の属性、スキル、利用場所に応じて、様々な画面インターフェースを用意して、利用者が必要とする情報、参考となる関連する情報を容易に得られるようにするもので、レファレンスサービスによる情報探索支援、オンラインレファレンスなども含まれます。

*国民が文化芸術を鑑賞し、また、日本遺産（Japan Heritage）としての魅力ある日本文化を発信する分野では、単なる分野横断的な検索だけでなく、新たに創造された知識と併せて発信する日本らしいEuropeanaの形を目指します。*

* ＜推進体制の構築＞

　ナショナルアーカイブ事業の推進のために、関係府省の協力のもと、中核的な役割を担う組織を設置して、全体の戦略企画、デジタル情報の保存や利活用のための研究開発、アーカイブに所蔵された資料に関する知識とIT技術の知識等も備えた高度な専門的人材の育成、孤児著作物の権利処理、絶版作品も含めた権利情報DBの構築促進等のための方策を導き出し、解決を図っていく必要があります。

[ここに出典を記載します。]

ナショナルアーカイブ事業の推進のために、関係府省の協力のもと、中核的な役割を担う組織を設置して、全体の戦略企画、デジタル情報の保存や利活用のための研究開発、アーカイブに所蔵された資料に関する知識とIT技術の知識等も備えた高度な専門的人材の育成、孤児著作物の権利処理、絶版作品も含めた権利情報DBの構築を促進等のための方策を導き出し、解決を図っていく必要があります。

## 今後のNDLの活動の方向性

　NDLは、2012年に策定した「私たちの使命・目標2012-2016」の中で、「印刷出版物にとどまらず、電子的に流通する情報を含め、様々な資料・情報を文化的資産として収集し、保存します。」としています。

　NDLは、唯一の国立図書館として、国内出版物の納本制度、公的機関のインターネット資料（ウェブサイト情報）の制度収集、民間のオンライン資料（電子書籍・電子雑誌に相当する情報）の制度収集、保存のためのデジタル化等、法律により「権限」が与えられ、確実な収集・保存・提供の実施の「責任と義務」を負っており、その責任と義務は、可能な範囲で行えばいいということではありません。しかし、物としての紙媒体の出版物については、公共図書館が保有する地域資料、美術館、博物館、文書館が保有する典籍資料をはじめとして、全てを収集できているわけではなく、また、電子書籍・電子雑誌に相当する情報は、セルフパブリッシングも含めて指数関数的に増加しており、ウェブサイト全体を丸ごとアーカイブするインターネット資料収集事業（WARP）においても、もはや全てを１つの組織で収集・保存すること自体が不可能です。

　NDLが主体的にアーカイブするとともに他の機関の情報を併せて利活用できるようにしてきた範囲は、ヒト・モノ・カネの資源の制約により「選択的」にならざるを得ませんでしたが、国全体で、文化資産のアーカイブの網羅性を確保できるようにするためには、アーカイブの共通基盤の仕様等を提示し、他の機関に対して、その適用と分担の「協力」を求めることが重要です。

　その概念は、パイロット電子図書館プロジェクトに始まって、「電子図書館中期計画2004」に基づいて進めてきたアーカイブ構築の考え方そのものです。「ひなぎく」で進めてきたあらゆる記録・記憶を保存する役割の発展系として、ナショナルアーカイブを推進することが、国民の期待に応える活動であり、NDLの使命・目標を達成することとなると思われます。

[ここに出典を記載します。]

NDLは、「私たちの使命・目標2012-2016」の中で、「印刷出版物にとどまらず、電子的に流通する情報を含め、様々な資料・情報を文化的資産として収集し、保存します。」としています。

NDLは、唯一の国立図書館として、国内出版物の納本制度、公的機関のインターネット資料（ウェブサイト情報）の制度収集、民間のオンライン資料（電子書籍・電子雑誌に相当する情報）の制度収集、保存のためのデジタル化等、法律により「権限」が与えられ、確実な収集・保存・提供の実施の「責任と義務」を負っており、その責任と義務は、可能な範囲で行えばいいということではありません。しかし、物としての紙媒体の出版物については、公共図書館が保有する地域資料、美術館、博物館、文書館が保有する典籍資料をはじめとして、全てを収集できているわけではなく、また、電子書籍・電子雑誌に相当する情報は、セルフパブリッシングも含めて指数的に増加しており、ウェブサイト全体を丸ごとアーカイブするインターネット資料収集事業（WARP）において、もはや全てを１つの組織で収集・保存すること自体が不可能です。

当館が主体的にアーカイブするとともに他の機関の情報を併せて利活用できるようにしてきた範囲は、ヒト・モノ・カネの資源の制約により「選択的」にならざるを得ませんでしたが、国全体で、文化資産のアーカイブの網羅性を確保できるようにするためには、アーカイブの共通基盤の仕様等を提示し、他の機関に対して、その適用と分担の「協力」を求めることが重要です。

その概念は、パイロット電子図書館実証実験に始まって、電子図書館中期計画2004に基づいて進めてきたアーカイブ構築の考え方そのものです。「ひなぎく」で進めてきたあらゆる記録・記憶を保存する役割の発展系として、ナショナルアーカイブを推進することが、国民の期待に応える活動であり、NDLの使命・目標を達成することとなると思われます。

* 図書館における情報システムの役割

図書館における情報システムは、業務・システムの効率化に留まらず、今後は、社会環境変化への対応および業務プロセス改革を迅速に行い、社会的な存立意義を維持向上できるようにするために、経営戦略とIT戦略を密に連携・融合させ、最大限の効果を発揮させる業務・サービスを構築することが重要です。

## おわりに

　デジタル情報時代において、出版物は、冊子体から動画・音声等を含むマルチメディア化されたコンテンツへ移行しつつあります。また、冊子体の原資料は文化財として保存するために、デジタル化していくことが求められています。他の文化財も保有機関においてデジタル化が進む状況において、文化的資産をあらゆる人々が将来にわたり享受、活用できるようにし、人々の創造的な活用に貢献するためには、社会全体でデジタル情報資源の「見える化」はもとより、より効率的なアクセスの保障に取り組む必要があり、組織を越えたナショナルアーカイブは重要な役割を果たすことになります。

　産学官のそれぞれの組織は、これらの施策が同一の方向性を持って、相互に資源を補完し合っていく必要があります。NDLは、ナショナルアーカイブの構築、さらに、世界レベルでの「インターナショナルアーカイブ」の構築へと発展することを目指し、その中核的な役割を担っていくべきと考えています。

　同時に、今後10年のデジタル情報化の進展を見据えつつ、このようなナショナルアーカイブを利用して知識創造のための情報が入手できる状況になったときに、知識創造を支援する図書館の役割は何か、また図書館に必要な機能の検討を加速させる必要があると考えます。

[ここに出典を記載します。]

今後10年を見据えると、「社会・経済的な価値の創出」を目指した「新たな知識の創造と還流」の仕組み作りは、今後10年の目標と言える。人・物・金の資源、そして制度的課題は1つずつ解決していかなければならない課題であるが、その課題は、様々な人たちの努力により解決され、「新たな知識を創造する国民」のニーズは、より高度化していくと思われます。そのようなニーズを満足させられる図書館等の役割は大きく変革していくと思われます。

知識インフラの1つの実現形としての「文化財のナショナルアーカイブ構築の夢」は、様々な人たちの努力により、実現している可能性が高いと思われます。

そのような時代においては、文献を中心に保存し提供してきた図書館は、どんな役割を果たすべきか、知識を保存し提供する機関の一翼を担う社会全体において「新たな知識の創造と還流」が行える情報化社会が相当進展していると思われます。

図書館等の類縁機関は、文化資産をあらゆる人々が将来にわたり享受・活用できるようにし、人々の創造的な活動に貢献することを目指しています。

昨今、政策化する活動、ベンチャー等の若手技術者を含むIT技術者、研究者も参加して、具体的なサービスの創造、具体化を地道にかつ確実に進めるアイデアソン、ハッカソン、オープンソース、オープンデータ化の大きな２つの動きがある。出版界でも、今後10年を見据えて、今後の方向性、新しい技術を適用して、電子書籍に取り組む動きも地道に進められている。

これらの活動が同一の方向性を持って、補完しあいながら進められることが大切。

このような国全体での文化財全般にわたる大きな変革期において、当初より使命に掲げてきたNDLをはじめ、民間も含めて多くの機関が分担して担ってきたアーカイブがあたかも一つの多きなアーカイブとして、利用できるように、具体的な活動を進めてほしい。その中でそれぞれの機関が久貝的な役割を果たしていってほしい。

また図書館界においては、主たるサービスを、物としての書籍から、情報としての電子書籍を含むデジタルコンテンツへ、デジタルコンテンツを活用したコミュニケーションの場へと変革していく必要がある。それを実現するためのサービスシステムを実装し目に見える形にし、ユーザの声を聴いて試行錯誤を重ねて、実現をしていってほしい。

この20年の活動の成果により、*デジタル文化財のシステムイメージが見えてきており、*NDLは今後のデジタル文化財のアーカイブの中核となることが期待されている。

*図書館があらゆる情報を集約して利用できるようにし、コミュニケーションのハブ的役割を果たせば、信頼性の高い情報を発信する出版界をはじめとしたコンテンツクリエータ、コンテンツプロバイダーの市場が拡大して発展が期待できるのではないか。*

## 年表

### 第0ステージ

* 1988 21世紀型図書館としての関西館の設立構想の開始
* 1992 国立国会図書館関西館を設置する構想のとりまとめ
  + 関西館の基本機能を情報発信拠点とし、中でも電子文献提供サービスをその中心軸に据えた。
* 1993 産業構造審議会情報産業部会が公共部門の情報化を積極的に進めるべきとの提案を行った
  + 関西館の予定する機能が、電子図書館的な機能だったため、この提案の中に位置づけられた。

[ここに出典を記載します。]

### 第1ステージ【1994～2002】揺籃期・始動期

* 1994 パイロット電子図書館プロジェクト開始
* 1995.10 電子図書館実証実験プロジェクト、総合目録ネットワークプロジェクト開始
  + 将来の電子図書館を想定したプロトタイプ環境を構築し、1000万ページに及ぶ資料をデジタル化。大規模ストレージに格納し、大容量のマルチメディア通信回線を用いて試行提供した。
* 1995 G7電子図書館プロジェクト
  + 先進7カ国首脳会議（G7）の共同研究テーマの一つに電子図書館が取り上げられ、日本がG7電子図書館プロジェクトの共同幹事国をフランスと共管した。
* 1995 新世代通信網実験協議会（BBCC）との協力実験
  + 成果として、「ディジタル貴重書展」（1998）等を公開。
* 1996 次世代電子図書館研究開発プロジェクト(JIPDEC)
* 1998 国立国会図書館電子図書館構想
  + 以降の電子図書館構築の骨格を示した。
* 2000 国立国会図書館蔵書目録（Web-OPAC）、国会会議録検索システム、貴重書画像データベースを公開

[ここに出典を記載します。]

### 第2ステージ【2002～2008】サービス離陸期

* 2002.10 国立国会図書館関西館開館、近代デジタルライブラリー、インターネット資源選択的蓄積実験事業（WARP）、データベース・ナビゲーション・サービス（Dnavi）を公開
* 2003 e-Japan重点計画－2003（IT戦略本部）
* 2004.2 国立国会図書館電子図書館中期計画2004策定
* 2004.10 デジタルアーカイブポータルプロトタイプの開発に着手
  + サービス指向アーキテクチャ（SOA）を適用。オープンソースソフトウェア（Linux, Apache, Xoops, MySQL, PHP, Dspace, chasen等）の適用、標準プロトコル（OAI-PMH, RSS, SRU, SRW等）の実装。
* 2005.4 NDLデジタルアーカイブシステムの開発に着手
* 2005.7 デジタルアーカイブポータルプロトタイプ試験公開
* 2007.10 PORTA正式公開
  + 大量アクセス、大量データ、大量ユーザに対応。拡張容易性、障害時運用継続性、環境変更容易性、直感的操作性の確保（ベンダーに依存しないパッケージ、OSSの適用）。
  + 館種を問わない全国の図書館との連携の強化と、博物館や文書館などの文化機関との連携の強化。国際標準のメタデータ記述要素（DCNDL）、記述規則、メタデータ交換の共通APIを利用した、商用を含めた外部サービスとマッシュアップによるサービス連携。連想検索エンジン（GETA）等を利用したあいまい検索も実装。

[ここに出典を記載します。]

### 第3ステージ【2008～2012】 発展期

* 2008.6 知的財産推進計画2008（知的財産戦略本部）
* 2009. 5 大規模デジタル化の実施開始（補正予算）
* 2010.2 「電子書籍の標準化の調査」をJEPAに委託
  + EPUBの日本語固有の縦書き、ルビ付与の仕様に関しての調査。
* 2010.4 国等の公的機関のインターネット資料（ウェブサイト）の制度収集開始
* 2010.8 日中韓電子図書館イニシアチブ（CJKDLI）協定締結
  + 海外に対しては東アジアの日中韓3カ国を初めとするアジア諸国との連携の強化や、世界各国とのグローバルな協力の推進。
* 2010.8 国立国会図書館サーチ試験公開
* 2011.12 「電子書籍の流通と利用の円滑化に関する検討会議」の報告
  + 「（NDLの）デジタル化資料を活用した新たなビジネスモデルの開発が必要」であり、「事業化に意欲のある関係者による有償配信サービスの限定的、実験的な事業の実施なども検討することが必要」。
* 2012.1 国立国会図書館サーチ、新NDL-OPAC、来館者管理システム等、全面リニューアル公開
  + 国立国会図書館サーチの構築にあたっては、外部専門家の参画、OSSの更なる適用によるコストダウン、次世代技術の試行、共通APIの実装の働きかけによる連携先拡大の加速化を図った。また、著作単位でのグルーピング表示、キーワードサジェスト、障害者向け機能、日中韓英翻訳機能、パーソナライズ機能、スマートフォン対応といった機能を実装した。

[ここに出典を記載します。]

### 第4ステージ【2012～2014】　総括と再始動期、見直し期

* 2012.4 東日本大震災アーカイブ開発開始
* 2012.7 「私たちの使命・目標2012-2016」を策定
* 2013.1 「電子書籍フォーマット適用調査」をJEPAに委託
* 2013.3 「電子書籍の流通と利用の円滑化に関する実証実験報告書」（文化庁eBooksプロジェクト（2012.10～2013.3））
* 2013.3 東日本大震災アーカイブ（ひなぎく）公開
* 2013.5 「戦略的目標」を策定
  + 「私たちの使命・目標2012-2016」を実現する中期的目標として、6つの目標の下にそれぞれ「戦略的目標」を策定。
* 2013.5 リニューアル総括および次期業務・システム最適化計画策定
* 2013.7 民間のオンライン資料の制度収集開始
* 2014.1 図書館向けデジタル化資料送信サービス開始

[ここに出典を記載します。]

# （2015年）文化的資産を将来に継承し、新たな知識が創造される社会を目指して【枠組みのみ】

## 骨子

## 中期計画で目指したこと

## 実施してきたこと

## 今後のナショナル・アーカイブの効率的・効果的な構築

## 文化的資産のアーカイブ

## ICTの活用によるイノベーションのためのアーカイブ

## デジタルネットワーク社会の進展に対応したサービスの構築・運用

# （2016年）電子図書館サービスからナショナルアーカイブの構築へ－LOD化によるデジタル文化財の利活用を目指して－

平成28年8月31日

元国立国会図書館　専門調査員・電子情報部長　中山正樹

1. **標題（和文および英文）**

電子図書館サービスからナショナルアーカイブの構築へ

－LOD化によるデジタル文化財の利活用を目指して－

From digital library services to the construction of National Archives

－Aiming at utilization of digital cultural heritage by making LOD－

1. **著者名（日本語およびローマ字）**

中山　正樹

NAKAYAMA Masaki

1. **著者の所属機関名および部署名（和文および英文）**

元国立国会図書館　専門調査員・電子情報部長

（同志社大学大学院総合政策科学研究科　嘱託講師）

Former Director General of Digital Information Department of the National Diet Library

1. **著者の所属機関の住所（和文および英文）、E-mail アドレス**

nakayama@slis.doshisha.ac.jp

1. **著者抄録（和文および英文）**

筆者は、国立国会図書館に在職中、電子図書館サービスシステムの設計・構築と運用に携わり、今後の方向性としてのナショナルアーカイブ構想の立案に関わってきた。

国立国会図書館は、電子図書館構想を策定し、「電子図書館サービス」を進展させてきたが、今後は、「電子図書館サービス」から、図書館の枠を越えてあらゆる分野の情報を関連付けて利活用する基盤として「知識インフラ」の構築へ発展させ、その一翼を担うこととした。その実現形が、「ナショナルアーカイブ」の構築である。

「電子図書館サービス」の設計・構築において、筆者が意識的に適用してきたセマンティックウェブに関連する技術を振り返る。また今後、ナショナルアーカイブの構築において、関係機関が連携して何を目指し、LOD化技術を含めたセマンティックウェブの技術を、どのように体系的に適用していくべきかを展望する。

The author、 during his tenure at the National Diet Library、 had been involved in the design、 construction and management of digital library service system、 and in making plans of National Archives initiative.

The National Diet Library、 formulating “Electronic Library Concept”in 1998、 has developed digital library services、 and is now seeking to build "knowledge infrastructure" as the basis for associating the information in all fields beyond the framework of library and utilizing them.

This paper reflects on the technology related to the Semantic Web、 which the author had consciously applied in the design and construction of digital library service system、 and also puts in perspective what goals the NDL should aim at、in cooperation with related institutions for the construction of the National Archives、 and how to systematically apply the technology related to the Semantic Web including LOD.

1. **著者付与キーワード（和文および英文）**

ナショナルアーカイブ、 デジタルアーカイブ、 知識インフラ、 セマンティックウェブ、 Linked Open Data、 クラウドコンピューティング、 WebAPI電子図書館、 国立国会図書館、 国立国会図書館サーチ、

National archive、 Digital archive、 knowledgeinfrastructure、 Semantic web、 Linked Open Data、 Cloud computing、 Web API、 Digital library、 Natonal Diet Library(NDL)、 NDL Search

## はじめに

筆者は、民間企業、政府関係の独立行政法人において、情報システム化により、業務サービス及びソフトウェア開発の効率化に取り組んできた。2002年に国立国会図書館（NDL）に入館して以降は、電子図書館事業の企画立案、図書館サービスシステムの構築・運用に携わり、2011年より電子情報部長および情報化統括責任者（CIO）として、電子図書館事業、情報システム全般を統括し、次世代図書館サービスの構築に向けて検討を進めてきた。筆者の信条は、一貫して、労働集約的な業務を情報システムにより効率化して、人がより創造的な業務へ移行できるようにすることである。

図書館に限らず、様々な業種・業態の組織において、大量の情報の中から効率的に適切な情報を取り出す情報提供システムが構築・運用されている。情報システムの観点では、図書館システムは、そのような情報システムの中での１アプリケーション分野であり、情報提供サービスとしても、利用者にとってOneOfThemのサービスであると考えている。

筆者は、そのような基本的な考え方で、NDL以外の情報資源、技術ノウハウを活用して、利用者の情報取得、知識創造を効率化するサービスを、如何にして効率的に構築・運用できるようにするかを念頭に進めてきた。次世代図書館サービスシステムは、従来型の図書館の延長線で考えるのではなく、世の中の多くの情報提供サービスの中で信頼性の高い情報を提供する機関としての責任と義務を果たせるサービスの提供を目指すべきとしてきた構想を論述する。

昨今、クラウドコンピューティングの普及に伴って、ウェブ上の情報を意味的に理解できる形で公開・共有を目指すセマンティックウェブ技術の中で、構造化された情報同士をリンクさせるための技術や方法論の総称としてのLinked Open Data(LOD)が注目されてきた。この概念は、NDLが1990年代に構想した電子図書館構想の「地球規模の知的財産を誰でも容易に利用できるようにする」という理念の実現のために、電子図書館サービスの構築の仕組みの考え方や、NDLが個別に適用してきた要素技術の集合としてみることができる。

筆者が、電子図書館サービスの構築において、セマンティックウェブ、LODの考え方を、どのように適用してきたか、今までの歩みを顧みるとともに、今後、電子図書館サービスの発展形としての国の文化財全体の利活用を目指したNDLを中核とした情報保有機関によるデジタル文化財のアーカイブ構築において、LODの考え方がどのように適用されていくべきかの方向性について、技術的な観点にフォーカスして考察する。

## 電子図書館20年の歩み

1980年代から、海外の多くの先進的な図書館同様、日本でも電子図書館事業に取り組み、1994年に我が国で最初の大規模な電子図書館の実証実験プロジェクトがNDLにより実施された。その後、NDLは、2002年から本格的なサービスとして離陸し発展させて、現在に至っている。

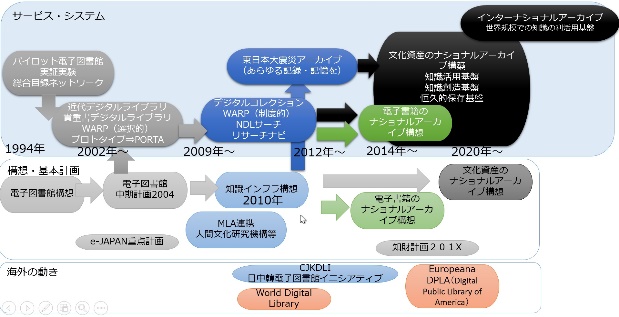


図1電子図書館に関連する構想とサービスの変遷

### 電子図書館構想と実証実験

1994年1月、通商産業省（現：経済産業省）の高度情報化プロジェクト事業の一環で、情報処理振興事業協会（現情報処理推進機構（IPA）は、慶應義塾大学の湘南藤沢キャンパス(SFC)内に、情報基盤センター（CII）を設置した。1）当時としては高速のネットワーク（6Mbps）、大容量サーバ（4TB）を配備して、各省庁の白書のSGML化や、各種統計情報を串刺しにして連関を分析する「新産業創造データベース事業」、学校教育の高度化を図るための教育ツールとしてのオーサリングシステム、エグゼキュータシステムを提供する「100校プロジェクト」とともに、NDLと連携協力して、パイロット電子図書館プロジェクトを開始した。筆者は、本事業には、IPA側の担当として参画した。

パイロット電子図書館プロジェクトの目的は、21世紀の高度情報社会において、地球規模の知的財産を誰でも容易に利用できるように、地球上に広く分散して個々に収集・蓄積されている知的資源を、空間的・時間的制約を越えてアクセス可能とする環境を提供するための実証実験であり、NDLは、全国公共図書館総合目録ネットワークプロジェクトとパイロット電子図書館実証実験プロジェクトを実施した。

総合目録ネットワーク実験は、出版物の総合目録と図書館間の相互貸借のためのシステムとして、NDLを含む58館の政令指定都市および都道府県立図書館が参加して実施された。共通のメタデータ記述規則としてMARCベースの総合目録共通フォーマット、通信プロトコルとしてファイル転送規約を設定し、各図書館からの大量の初期データ、週次の差分データを受け取ることにより、1、400万件の書誌データがRDBMSの１つであるOracleデータベースに蓄積された。ISBN等の識別子がないメタデータも多く、タイトル、著者等で同一の書籍のメタデータであることを認識する機械的な同定作業を行った。所定の時間内に処理するためには、レスポンスを低下させないデータベースのテーブル設計、SQL言語によるプログラミングでは様々な工夫が実証実験された。

パイロット電子図書館実証実験では、NDL所蔵の貴重書、明治期刊行図書・雑誌、出版社から提供を受けた資料等を、マイクロフィルムからのスキャニング、デジタルカメラでの撮影により、当時としては高精細画像（4、000×5、000ピクセル）で、tiffおよびJpegフォーマットで約1、000万ページ（CD-ROM約2000枚分）のコンテンツをデジタル化し、画像データベースに登録。個々の画像データにメタデータを付与してメタデータデータベースを構築して、効率的な検索・利用方法、ユーザインタフェース（UI）、電子化データの高度利用、効果的な電子図書館の構築支援の手法等について実証実験を行った。その成果であるシステム及びデジタル化コンテンツは、その以降の電子図書館サービスのベースとして引き継がれている。

### 国立国会図書館中期計画2004

NDLは、パイロット電子図書館プロジェクトの成果を踏まえて、2002年10月に開館された関西館を拠点として、近代デジタルライブラリー、インターネット資源の選択的収集事業（WARP）、各種の電子展示会を公開・提供した。

2003年には、e-Japan重点計画2003、e-JAPAN戦略Ⅱ加速化パッケージ、（内閣官房IT戦略本部）において、「国のデジタルアーカイブ構想」、「ジャパン・ウェブ・アーカイブ構想」の実現を、また2004年には、e-Japan重点計画2004において「国立デジタル・アーカイブ・ポータル構想」を一層推進することが明記された。

　このような動きを踏まえて、NDLは、2004年2月に「電子図書館中期計画2004」（中期計画）を策定した。この中期計画において、デジタルコンテンツを広汎な利用者に提供するために、NDLが国のデジタルアーカイブの重要な拠点となるということ、また国内外の多様な利用者層の需要に応じ、日本のデジタル情報全体へのナビゲーションを行う総合サイトを構築し、利用者がワンストップで利用できるようにすることを目指すこととした。2）

### アクションプランとしての3つの柱

１つ目は、 「デジタル･アーカイブの構築」 である。これには、書籍のデジタル化の推進と、インターネット情報の収集と恒久的保存の２つがある。２つ目は、「情報資源に関する情報の充実」で、レファレンス情報などナレッジの蓄積を蓄積し提供提供すること、3つ目は、「デジタル･アーカイブのポータル機能」である。

### 電子図書館中期計画2004の実現イメージ



図2 NDL電子図書館サービスの全体像

図2は、2004年にNDLが作成した電子図書館サービスの実現のイメージである。

右下の「デジタルアーカイブ」は、デジタル化したリソースやインターネット情報などを蓄積し、長期保存する。

その左は、既存システムのNDL-OPACなどの「蔵書目録」は、これらが書誌情報やデータベース内のメタデータを提供する機能を備えることにより、デジタルアーカイブポータルから統合的に利用できるようにする。

上の「ポータル」は、これらNDLの資源だけでなく、左下にある、他機関のデジタルアーカイブや目録などの有用情報も合わせて案内できるようにするもの。

左上の「ポータルのインタフェース」は、GUIにより利用者へのニーズに合わせた検索手段の提供と、他システムがポータルの機能を利用するためのAPI機能を持つ。

また、右の、「主題情報、目次情報、解題情報、レファレンス情報など検索を支援するための情報、知識としての「ナレッジデータベース」である。

2004年に描いたイメージの実現形として、ポータルを発展させたものが、PORTAであり、また、現在の「国立国会図書館サーチ（NDLサーチ）」である。

### デジタルアーカイブの構築

資料のデジタル化に関しては、NDLは、2009年5月から大規模なデジタル化を開始し、2011年までに、冊子体としては230万冊、約2億枚の画像をデジタル化した。また、インターネット資料の収集に関しては、NDLは、2010年4月に国等のウェブサイトの制度的な収集を開始した。さらに、いわゆる電子書籍・電子雑誌に類するオンライン資料は、2013年1月から、無償でDRMによるアクセス制限が掛かっていない資料についてのみの収集を開始した。デジタル化および収集した資料の提供に関しては、著作権保護期間の満了もしくは著作権者から許諾を得た資料は、インターネット上に公開している。デジタル化した資料のうち、市場で入手困難な資料に関しては、公共図書館等で閲覧が可能になっている。ここでは詳細は割愛するので、他の文献を参照していただきたい。本稿では、保存システムの機能として、メタデータに関連した施策として、「長期保存システム」のみ紹介する。

### デジタル情報の長期保存（OAIS、 情報パッケージ(SIP、 AIP、 DIP)）

電子情報を長期的に保存するためのシステムの構築に当たっては、NDLは、この長期保存システムとして必要な機能やプロセスを抽象的に定めたモデルを参照しつつ構築している。

その一つとして、OAIS（Open Archival Information System）レファレンスモデル(ISO 14721:2003)モデルがある。

このレファレンスモデルに沿ったパッケージとして、米Hewlett-Packard 社のDspace、IBM社のDIASがあった。受入機能として適切なメタデータの付与、アーカイブ（保存）機能として定期的なエラーチェックや保存媒体の更新、データ管理機能としてメタデータの管理、保存管理・運営機能として長期保存計画の立案、提供者との交渉、マイグレーションの実施、さらに保存計画機能として旧式化の検知、対応策の提示、アクセス機能として再生環境の保存、エミュレーションが想定されている。

このうち、保存計画を実現する要素技術であるフォーマットレジストリやフォーマット識別ツールの開発が海外で行われている。

情報パッケージとしては、MODSをベースにしたメタデータをMetsで括ったSIP（受入れメタデータ）、AIP（保存メタデータ）、DIP（提供メタデータ）を想定した。

### デジタルアーカイブポータルのプロトタイプ

NDLは、中期計画での3つ目の柱であるポータル機能について、2004年10月から開発を開始した。様々なデジタルアーカイブ内の情報を統合検索する仕組みの実用性を検証するために、デジタルアーカイブポータルのプロトタイプシステムとして、2005年7月に試験公開し、機能の検証を開始した。これが、現在の分野を越えたナショナルアーカイブ構想の原点であったと考える。

### 開発当初のサービス要件とシステム化

#### サービス要件定義において

* 利用イメージとして、利用者の情報探索行動の全般をサポートすることを目指すこととした。  
  情報探索行動として、問題の定義、情報ニーズの識別、情報探索戦略の策定、情報の獲得（情報源の咀嚼、情報の抽出）、情報の統合（情報の組織化、提示）、評価（成果の評価、プロセスの評価）等のステップを想定した。このようにして目的の情報の存在を知るだけでなく、情報そのものの閲覧および入手先までナビゲートを目指すこととした。
* それを実現するために、各デジタルアーカイブの構築の在り方にも言及した。交換用の標準的なメタデータ記述要素、記述規則の適用、交換用APIの実装を求めることとした。また、巨大な知識ベースとして「意味ある情報資源」として利用できるように、本文内容を組織化、検索できるように研究開発、技術開発が進むことも求めた。

#### システム化において

* 先進的かつ将来標準的な仕様となることが見込まれる技術を適用すること。
* オープンソースソフトウェア（OSS）・パッケージソフトを活用することとして、特に、適用事例が多いOSSを活用し、可能な限り新規開発はせず適用したソフトウェアのカスタマイズは必要最低限とすることとした。
* Webサービスとして他サービスから利用しやすいものとすることとして、各々の機能は独立したWebサービス機能として提供し、図書館界のみならず、デジタルアーカイブの世界で標準となり得る仕様の採用を目指した。

### PORTAの構築

NDLは、プロトタイプでの検証を踏まえて、実用システムとして「PORTA」を構築し、2007年10月に正式公開した。3)

### システムアーキテクチャ

NDLは、システム基盤として、Web2.0時代のサービス指向アーキテクチャ（SOA）を適用し、仮想OS上でOSSを組み合わせ、Dublin Core（DC）に準拠した国際標準のメタデータ記述要素、記述規則、メタデータ交換の共通API（OAI-PMH、SRW（SRU/SOAP）、RSS、Z39.50等）を利用して、商用を含めた外部サービスとマッシュアップによるサービス連携と、コンテンツマネジメントシステム（CMS）による機能追加の容易性の確保を目指した。先進的な検索機能として連想検索検索エンジン（GETA）等を利用したあいまい検索も実装した。

### 統合検索対象の拡大

統合検索対象は、画像イメージ形式で提供している近代デジタルライブラリーとテキスト形式で提供している青空文庫の統合検索から始め、同時代の資料が分散保存されている国立公文書館アジア歴史資料センターとNDL-OPAC、地域情報をまとめているアグリゲータサイトとしてデジタル岡山大百科等を、ワンストップで検索しコンテンツに辿り着けるようにした。

### 統合検索のために苦労した点

* データ提供機関に対して、統合検索への賛同を得ること。大きな意義は、各機関のコンテンツの利活用が促進されることであり、また、利用者の利便性が向上することであったが、サービスが横取りされると誤解されることが多々あった、また、サイトへのアクセスが増えることを不安視され、なかなか賛同が得られなかった。
* また、連携には、提供機関のサイトで標準プロトコルの実装が必要であったが、ほとんどのサイトで、外部提供インタフェースを持っておらず、各サイト側に実装が必要であったが、開発ベンダーが法外な開発費を提示し、その費用が捻出できない状況であった。
* さらに、メタデータのマッピングの調整においては、それぞれの機関において、 Dublin Core（DC）をベースにした記述要素を適用できても、記述規則が異なり、統合しても内容として認識できない状況であった。特に、横断検索においては、メタデータの表記のゆれをカバーした検索はできず、調整には、膨大な時間を費やした。汎用的なメタデータのマッピング、記述規則等のルールの普及が必要であった。

### インキュベータとしてのPORTAの役割

筆者は、PORTAは唯一の統合検索サービスではなく、PORTAがインキュベータとなって、各利用者ニーズにあった様々な統合検索サービスが出現することを期待した。

* まず、継続して、デジタルコンテンツの可視化を目指すこととし、インキュベータの役割として、各データプロバイダに検索やサービス連携のためのAPIが実装されることの普及啓発活動を継続するとともに、外部機関においてPORTAの外部APIを利用した様々なサービスが出現することを期待した。
* また、利用者の情報探索行動において、統括可能情報の網羅性を高めるために、様々な機関から提供されるデータベースを分散デジタルアーカイブと位置付け、コンテンツの体系的な構築と、網羅性の高い統合検索サービスの実現を目指すこととした。
* さらに、単なるキーワードでの検索サービスではなく、意味的に関連付けるセマンティックウェブ・サービスでの連携を目指すこととした。

### Web3.0の時代を見据えて適用すべき技術と進め方

筆者は、今後の機能強化のため、セマンティックウェブ技術を中核として、適用すべき技術として3つに分類した。

#### 収集を容易にする技術

収集効率を高める技術（差分収集と再現技術を含む）、収集品質を高める技術、収集したサイトイメージから著作物を切り出す技術の適用を想定した。

#### 管理を容易にする技術

* メタデータの付与を自動化・省力化する機能

書誌的事項は、情報の作成元が作成することを基本としつつ、管理、検索に十分なメタデータが付与されていない場合は、多少精度が低くてもメタデータは可能な限り自動付与する機能、自動付与されたメタデータの確認作業を省力化する技術の適用を想定した。

* セマンティックウェブ技術の活用

当時、集合知を利用するWeb2.0から、意味的情報を相互利用するWeb3.0が叫ばれた時期であり、Web2.0の要素として、Blog、Forksonomy、Ajax、Mashup等の技術、更に

Web3.0の要素として、RDF、SKOS、オントロジー、GRDDL、SPARQLの技術の適用を想定した。

* 蔵書管理指向からサービス指向へ

将来的には、目録を利用者側の視点から見直すモデル（FRBR）に基づくメタデータの再構築も想定した。

* 検索を容易にする技術

従来型のキーワードマッチングによる検索だけでなく、キーワードサジェストやクラスタリング検索等により、検索候補を推定する技術を活用した検索機能も想定した。

#### PORTA構築の進め方

筆者は、進め方として、図書館の枠を超えた国際標準、業界標準を積極的に適用するとともに、政府機関、民間を問わず、国内外の研究機関、研究者等と連携して調査研究を進め、関係機関との共同構築、オープンソースの利用、単独でのトータルシステムを目指すのではなく、各機関の研究成果を組み合わせてマッシュアップによるサービス構築を想定した。

まずは、海外ではIIPC、国内では様々な技術の研究開発性を保有しているJST、NII、NICT、AIST、IPA等の政府機関、大学の研究室との連携、有用な検索サービスを実施している商用ポータル機関（Google、Yahoo等）との連携によるウェブ協調型のサービスの提供を目指した。

### トータルな次期図書館システムを目指して

2009年当時の冊子体資料を扱う従来型の図書館システムの最大の問題は、運用・保守作業、機能拡張コストが肥大化して、機能拡張により新たな利用者ニーズに応えるための機能拡張ができなくなったことである。そのために、NDLは、下記のような方針により、全面リニューアルすることとした。

* 紙・デジタルを問わずトータルな次期図書館システムの構築と提供の実現を目指す。
* NDLサーチを、図書館ポータル、次世代OPAC、PORTA、ゆにかねっとを統合した入口機能とする。
* 新規開発を極力なくす。
* 実現すべき機能とレベルを明確にしたシステム化要件定義書により、見積もり内容を的確に判断する。
* 特定ベンダーしか提案できないRFPにならないようにして、競争原理が働く調達を行う。
* 従来型業務・サービスは、パッケージで実現できない業務は、業務の内容や流れを見直し、可能な限りパッケージの仕様に合わせる。

　　この方針に基づきシステム開発を進めた。

冊子体刊行物の収集・組織化業務及び、蔵書検索・申込システム「NDL-OPAC」を、パッケージに切り替えることによる開発・運用コストを大幅に削減し、デジタル化資料の来館者サービスシステム、国立国会図書館サーチ等、デジタル時代の次期図書館利用者サービスの基本機能を充実させた。

### 国立国会図書館サーチ（NDLサーチ）

NDLは、PORTAの後継の情報探索サービスシステムとして、2012年から正式に提供を開始した。のちに提起される「知識インフラ」のフレームワーク作りのベースとしたサービスであり、筆者が振り返るとLinked Data、Linked Open Dataと同一の方向性を指向したものとなっていた。

NDLは、システムの構築に当たっては、様々な視点を持つ専門家の参画のもとで、サービス要件を確定させ、サービス要件を満たす機能を実装するためのシステム化要件を定義した。段階的な実装を想定し、OSSを含めて、その時点での最適な適用技術、パッケージを示す技術標準適用ガイドラインを踏まえて、開発を進めた。

### 機能の概要

NDLは、FRBRのモデルに沿って同一著作、原作から様々な媒体や形式に派生して発行されている著作物を書誌事項で体系的に把握できるように著作単位でグルーピング表示する機能、検索キーワードが明確でなくても想定されるキーワードサジェスト機能、障害者向け機能、検索ワード、外部サービスを活用して内容表示を日中韓英に翻訳して表示する機能、個人ごとの検索要件、検索範囲をあらかじめ特定できるパーソナライズ機能、スマートフォン対応機能等を実装した。

NDLは、交換用メタデータとして、冊子体出版物以外の情報への適用を想定して従来の仕様をよりシンプルにしたDC-NDL(Simple)と、セマンティックウェブに適合した記述モデルであるRDFをデータの記述に使用し、FRBRのWork、Manifestation、Itemを体系的に定義したDC-NDL(RDF)を交換用メタデータ記述要素・規則として定義した。

NDLは、メタデータハーベスト、横断検索のためのWebAPIを設定し、そのAPIの実装を働きかけることにより、連携先の一層の拡大を図るとともに、外部システムでメタデータを利活用のためのWebAPIとして、検索結果や新着書誌情報をFeedly等のニュースリーダで確認できるように、RSS配信する機能、DC-NDL(RDF版)、JSON形式でダウンロードする機能等の実装を強化し、「情報ハブ」の役割を提供した。4)

また、著者名典拠、件名典拠情報に関しては、国立国会図書館典拠データ検索・提供サービス（WebNDLA）として、キーワード検索、NDC分類記号検索により検索できるほか、NDLサーチと連携し、関連キーワードの提示、典拠コントロールされた各標目による書誌データの再検索など多様な検索機能を提供した。また、全典拠データの個別ダウンロード、国立国会図書館件名標目表（NDLSH）の収録範囲となる典拠データを、RDF/XML形式、 RDF/Turtle形式、 JSON形式で、ダウンロードすることを可能とした。

### サービスの状況

NDLサーチでは、人間文化研究機構の6構成機関の研究資源共有化システム（nihuINT）、国立美術館の4つの美術館が所蔵している作品の総合目録、商用の電子書籍ポータル（hon.jp）との連携等のPORTAが担ってきたデジタル情報のポータルとしての役割を継続するとともに、ゆにかねっとが担ってきた各地の公共図書館蔵書の総合目録としての役割を引き継いだうえ、NDL-OPAC、インターネット資料収集保存事業（WARP）、国立国会図書館デジタルコレクション、国会会議録検索システム、リサーチ・ナビ（調べ方案内）、レファレンス協同データベース等のNDLの主要データベースを統合検索し、図書館、美術館、博物館、文書館、政府機関、新刊書籍の出版予定情報としてJPOの近刊図書情報、Japan Knowledge等の商用データベースなど、200DB、約1億件の書誌・メタデータを横断的に検索できるようにした。検索対象が必ずしもデジタル化されてインターネットで閲覧可能なコンテンツでなく、所在情報により入手手段へのアプローチに留まるものも多いが、今後各機関によるデジタル化が進めば、インターネットで閲覧できるコンテンツへのナビゲーションも充実すると思われる。

また、統合検索のほかに、検索キーワードを使った外部サービス検索として、WorldCat、WebCat Plus、Wikipediaでの再検索、更に入手手段として、各種オンライン書籍・電子書籍販売サイト、古本屋ポータルサイトへのナビゲートも行った。

最近の進展としては、公共図書館からのデータ提供方式のOAI-PMH / DC-NDL(RDF) への切替、ジャパンリンクセンター（JaLC）との連携によりDOIでの固定URLでの検索、WebNDLAからVIAFへのナビゲートも可能になっている。

### 連携方針・連携モデル

統合検索先の拡大においては、連携可能なところから進め、網羅性の確保のため、分野ごとの中核的なデータベースに標準的な連携機能の実装を働きかけてきたが、標準仕様が浸透してきた状況において、NDLは、改めて連携方針・連携モデルを提示し明確にした。

#### 連携実施計画策定の課題・背景

* 日本におけるメタデータ提供のプラットフォームとしての認知度が高まりつつある今、連携可能な範囲での拡張ではなく、目標と計画を設定して内外に示す必要がある。
* ナショナルアーカイブ（構想）におけるコンテンツ検索・提供機能を担うことを見据える必要がある。
* 図書館等の情報機関が保持するメタデータの集約及びAPIを通じた一般への提供により、政府が推進する「公共データの民間開放（オープンデータ）」の一翼を担う必要がある。

#### 連携方針・連携モデル

次の5項目を連携方針の柱とする。

* 日本の刊行物及び刊行物と同等の内容を有するコンテンツの網羅を志向。
* 公的機関、学術研究機関、図書館・文書館・博物館・美術館等の文化機関が作成し、インターネット上で提供している一次情報（コンテンツ）、二次情報（メタデータ）及び参考情報等を対象とする。
* 一般利用者にとって有用性が高いコンテンツを持つシステムを優先。その際、一次情報の入手までの障壁が低いシステムを重視。
* APIを実装したシステムを優先。メタデータ授受に用いるフォーマットとして、DC-NDL(RDF)を推奨。
* 効率的に検索対象を拡大するために、個別のシステム（=data provider）との連携よりも、それらを集約した統合検索サービス（=aggregator）との連携を優先。

つまり、APIは普及期にありインキュベータとしての役割を終了し、既にAPIが実装されているシステムとの連携を拡大し、人間文化研究機構、JST、NII等のように、各所管において、統合データベースを提供しているサービスとの連携により、効率的に網羅性を確保するということである。

### 公共的書誌情報基盤

情報内容として整理・要約されている出版物に関する情報の共有化と利活用は最も重要視される。

公共的書誌情報基盤は、出版界が作成する出版情報（販売促進情報）を活用した図書館での書誌作成の省力化を目指すもので、出版界と図書館界が持つ著者、著作の書誌・書評等の情報の共通識別子によるLinked Data化により、出版情報の網羅的なデータベースの構築と共有が期待できる。

### 公共的書誌情報基盤の整備（2010年）

NDLが、出版関係機関と協力し、我が国を代表する標準的な書誌情報を作成・提供する公共的基盤を整備するものである。出版文化の基礎となる質の高い出版・書誌情報が、無償もしくは廉価にて、迅速またタイムリーに読者、利用者に届けられることを目的とするものであった。5)

また、この事業によりNDLの納本事務がより網羅的かつ円滑に行われることが期待された。

当時の想定は、次のようなものであった。

* 近刊情報は、出版物のISBNとNDL書誌IDが関連付けられるように、出版社からISBNが付与された書誌情報が、JPOを経由してNDLに送られ、JPNO、 書誌IDを付与して、近刊情報として公開する。その近刊情報をJPOを経由して取次に送ることにより、共通の識別子で近刊情報が管理される。
* 新刊情報は、取次において、JPO近刊情報と現物を突合し、出版情報追記して、新刊情報とする。その情報をNDLにおいて、近刊情報を新刊情報として置換え、最低限の書誌事項を追記して新刊情報として公開すること。

しかしながら、NDLでの実装においては、納本された新着情報の提供は、出版情報をインプロセス情報として提供することにより早まったものの、出版界と図書館界での情報の共有には繋がっていない。

### 全国書誌情報の利活用に向けた超党派勉強会での検討（活字文化議員連盟　2015年9月）

図書館などで広く活用されているMARCのもととなる書誌データを国民が無料で利用できるように政策や予算なども含め総合的な施策を検討するもの。

背景としては、NDLが作成する書誌データは、民間の書誌データを活用することで迅速な提供に向けて取り組んでいるが、民間の提供スピードには及ばず、公共図書館などでの利活用が進まない現状があり、2016年春をめどに課題のたたき台をとりまとめるとされている。

筆者が期待する結論としては、近刊情報、新刊情報が図書館で活用され、また、図書館での書誌情報、典拠情報が出版界での出版物に関するデータベースで利活用されることにより、出版情報を活用した書誌作成が効率化されるとともに、出版物の検索の網羅性が確保されることにより利用者が出版物を見つけやすくサービスの提供が容易になることである。

### 知識インフラの構築を目指して

新たな知識の創造のためには、分野を越えた知識の関連付けが必要であり、日本中に散在するコンテンツの所在を集中管理し、そこに検索をかければ、関連する全ての必要なコンテンツが得られるようにするものである。そこでは、単に情報を集めたものではなく、関連するものが有機的に結合され、ネットワーク的に統合化されたものであり、日本中にある芸術を含んだあらゆる学問・研究のコンテンツ、研究ツール、社会状況データ等が知識の形に組織化され、これらの知識・情報が公開され、全ての人が共有できる。

### 知識インフラとは

「知識インフラ」とは、情報資源を統合して検索、抽出することが可能な基盤で、国内の各機関が保有する情報を知識として集約し、新たな知識の創造を促進し、知識の集積・流通・活用と創造するサイクルの構築を目指すものである。

### 知識インフラの必要性

第４期科学技術基本計画「科学技術に関する基本政策について」（内閣府　総合科学技術会議2010年12月24日）が答申され、文献等研究情報のデジタル化、オープンアクセスの推進等とともに、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」とされた。6)これを踏まえて、NDLは、2011年に、第三期科学技術情報整備基本計画－知識インフラ構築に向けて」を策定した。7)

### 知識インフラの構築の目的

知識インフラは、科学技術研究活動の実践を根本で支え、科学、技術、学術、文化活動によって生み出される多様なデータ、情報、文献、関連する情報が組織化され、それらへの迅速で適切なアクセスを可能にすることで、次の研究、開発、教育、その他の社会的・文化的実践へとつなげる動的サイクルを形成することを目的としている。つまり、情報の生産→流通→アクセス→再生産という知識の循環を促進するネットワーク、プラットフォームとなることを目指すものであり、また、組織や個別学術分野を越えた知識の融合を可能とし、学際的な新しい知識やイノベーションの創造を容易にするものである。

### 知識インフラの機能

知識インフラにおいては、文字データだけでなく多様な形式で表現されるデータや情報を対象とし、収集、保存、識別、組織化、検索、表示、公開といった機能を実現させる必要がある。利用者は、大量のデータに対して特定条件に適合したデータだけを抽出したり、多様な分野の情報を一括して検索したり、自分の関心に合わせて実体のリンクやネットワークを形成したりといったことが自由にできることが求められる。

また、単語等による検索だけでなく、自動分類や収集された全体を見通した上での体系化や秩序化がなされることが期待される。8)

### 知識情報基盤の構築モデル

NDLは、NDLサーチの前身のPORTAと、NDLデジタルデポジットシステムが進めてきた概念をベースとして、知識情報基盤の構築モデルを描き、後の「東日本大震災アーカイブ（ひなぎく）」により、大震災の分野に絞った「知識インフラ」の実現を目指した。

「ひなぎく」の実現に当たっては、情報の可視化技術、情報の収集の効率化技術、情報の組織化技術、情報の集合知化技術、情報検索技術、閲覧・表示技術等の次世代技術の研究開発成果の活用なしには実現できないと考える。9)

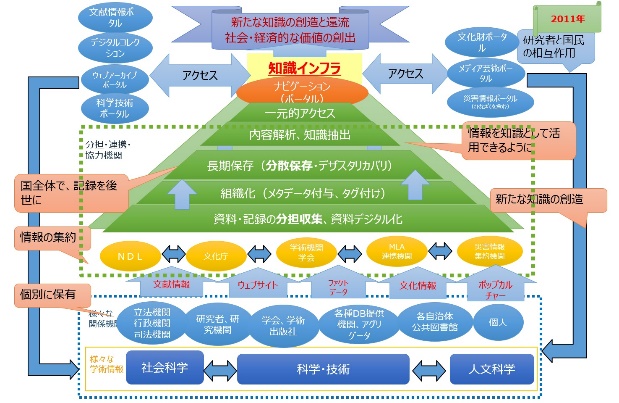


図3 知識情報基盤の構築モデル

### 東日本大震災アーカイブ

2011年3月には、東日本大震災が発災し、甚大な被害がもたらされた。国は、この大震災に関連する災害現象そのもの、災害前・災害直後・復興の過程、災害時の対応、他地域・次世代への教訓等のあらゆる記録を後世に残すこととした。10)NDLは、それを実現するために、大震災関連にフォーカスした知識インフラの構築の実現形の先行事例として、2013年3月には、大震災に関するあらゆる記録、記憶を保存し、一元的に検索できるようにする「東日本大震災アーカイブ（ひなぎく）」を構築した。11)

大震災アーカイブの構築は、従来の図書館の枠を越え「あらゆる情報、記録を収集、永久保存して、提供する」というNDLの使命を達成するために、今後数年間で取り組むべき事項の多くが含まれていると考える。

### 収集の基本的な考え方

大震災に関連する情報のすべてを、NDLだけで収集し保存することは不可能であり、国全体で分担して収集し、分担して保存する形を目指すべきである。

#### 収集範囲

国全体で．様々な機関が保有している大震災前の記録、大震災後の事象・被害、状況の記録、今後の復旧・復興の記録等、過去から、現在、未来に亘って、可能な限り収集・保存する。

#### 収集方法

* NDLによる直接収集

NDLは、制度的に収集可能な記録のほか、他の機関が保存の対象としていない記録等については、積極的に受け入れる。

* 他機関による記録等の保存の推進

記録の保有機関、アーカイブ機関、NDLのいずれかで記録等を保存し、NDLは、その所在情報を把握できるようにする。

* メタデータの収集又は検索の機械的連携

NDLは、現在のNDLサーチと同様に、可能な限り、メタデータの収集、もしくは横断検索による機械的連携を行う。

* 働き掛け・協議

NDLは、国等の各機関で、国全体としての記録等の保存の必要性の認識を共有し、関係機関に協力を要請する。

### 東日本大震災アーカイブのシステムイメージ

NDLは、これまでの成果であるNDLサーチ、ウェブアーカイブシステム、デジタルデポジットシステム等の既存のシステムをベースに、機能拡張する形でシステムを構築した。

構築に当たっては、効率的、効果的に進められるように、また、利便性の高いサービスにするために、次世代技術の実用化実証実験に取り組み、成果の積極的な活用が図れるように調査研究を進めた。



図4 東日本大震災アーカイブのシステムイメージ

* 映像・観測データ等を受け取ることを想定して、処理能力、ストレージ容量を必要に応じて増強できる仕組みとして、分散処理サーバ(Hadoop)、分散ファイルシステム(GlusterFS)を導入した。
* 大震災アーカイブ自身が、災害で消失してしまわないように、ディザスタリカバリも考慮した。
* 画像・映像なども的確に検索ができるように、明確なメタデータが付与されていない情報にも可能な限りメタデータを自動付与する仕組みの導入を検討した。
* 本文も含めたテキストの全文インデキシング等を試行した。
* 次世代の図書館システムのモデルとなることを指向した。

### 「ひなぎく」での課題

構築・運用においては様々な課題が顕在化した。

#### 技術面での課題として、

* 大規模分散ファイルシステム、分散処理システムを適用したが、実際には、想定ほどコンテンツが集まっておらず、将来的な実用化レベルでの検証ができていない。
* 計測機等から得られた生データ、分析の中間データなどのファクトデータの永久保存も目指しているが、情報に関するメタデータまでしか収集できていないため、ビッグデータへの対応の検証ができていない。
* ビッグデータとしてのウェブアーカイブから大震災関連の情報を自動的に切出し、コンテンツ単位で精緻な検索をする仕組みが構築できていない。
* 各アーカイブにおいて、検索およびハーベスト用のWebAPIが未実装なものが依然として多い。
* 収集前データに永続的識別子がない。
* 各アーカイブのメタデータとの記述規則の差異、付与されたメタデータに記述要素の不足が多い。
* 不完全なメタデータへの自動付与機能はない。
* 利活用のための検索・閲覧機能として、本文テキスト、イメージ認識技術等、内容を関連付けた検索技術が未成熟である。

#### 人材面での課題としては、

* 専門分野に関する知見、文化資産の収集・保存・修復・公開の技能の不足。
* 文化資産を取り扱うための知識・技能の不足（プリザベーションエンジニア、コーディネータ、エンベデッドライブラリアン）。
* デジタル技術を活用したアーカイブ化のための知見の不足（アーキビスト、法規担当）。
* システム開発・運用管理の一般的な知識・技能の不足（システムライブラリアン、ITエンジニア）。

#### 協力体制の課題として

* 縦割り行政のため、立法府であるNDLが進めることに警戒感。
* 「放送アーカイブ」とする構想に対する報道圧力・事後検閲の可能性に対する警戒感。

#### 制度面・運用面での課題として

* 情報公開法では、各文書に対して最低保存期間が設定されるが、運用では行政文書管理簿に掲載されていない軽微な資料の保存期限は1年未満とされ、1年を経過すると廃棄される。
* 公文書管理法では、軽微な資料を含めて、歴史的公文書と指定されなければ公文書館に移管されない。
* 肖像権、プライバシー権、人格権等の権利処理がされなければ提供できないが、NDLは、提供を前提としない、いわゆるダークアーカイブとしての収集は行わないため、継続して維持できなくなったアーカイブのコンテンツを預かることもできない。
* 国有財産法、財政法の解釈では、アーカイブシステム機器等の運用維持が困難になったアーカイブ機関のために、Webサーバやストレージを無償で貸出すことができない。
* 現在のインターネット情報の制度的収集は国等のサイトに限定されているため、民間および個人のサイトは個別許諾により収集しなければならないが、個別に許諾手続きが必要なため、悉皆的な収集が困難。民間サイト、個人サイトは、ハーバード大学と連携し、インターネットアーカイブ社に収集を依頼している。

このように課題は多いが、協力体制、制度面・運用面の課題は、大震災に限らない共通の課題であり、今後、国全体のナショナルアーカイブ構想の議論の中で、具体的な解決策の検討が進められることを期待している。

## 今後10年で目指すところ（2015年～2024年）

NDLは、このような歩みで電子図書館事業を進展させてきたが、筆者は、今後10年を踏まえて、社会の要請、情報技術の実用化動向を想定した図書館サービスの姿と、その実現に向けて実施すべき事項を明確にし、速やかに対応していく必要があると考える。

東日本大震災アーカイブを踏まえて、分野を越えた知識インフラの実現形として、あらゆる記録を情報として集約し、相互に関連付けて知識化し、将来にわたって利用を保証するとともに、「社会・経済的な価値を創出」できる「新たな知識の創造と還流」の仕組みを現実的に構築していく時期と考える。

### 「知的財産政策ビジョン」での今後10年の目標

我が国では、「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部）等において、文化資産のデジタルアーカイブ化を推進する政策も含めて、今後10年を見据えた知的財産関連の政府の取組としての目標を掲げた。12)政策としては、従来の事業モデルの延長線上での「改善」だけでなく、事業モデルそのものを創造・転換する「イノベーション」を創出すること、あらゆる分野の知的財産を対象としたビッグデータも含めたクラウド化、オープン化、更にユーザが共同でコンテンツを制作するクラウドソーシングの普及が見込まれている。そのための知財人財の育成・確保、コンテンツの権利処理の円滑化、電子書籍の普及促進、データの収集・蓄積・分析による多様な付加価値の創造の研究開発、日本の伝統や文化に根ざした魅力あるコンテンツ・製品などの発掘・創造等の取組も示された。

### 電子書籍ナショナルアーカイブ

有形の出版物を中心とした物としてのアーカイブ機能の延長に、無形の情報としての電子書籍のナショナルアーカイブがある。

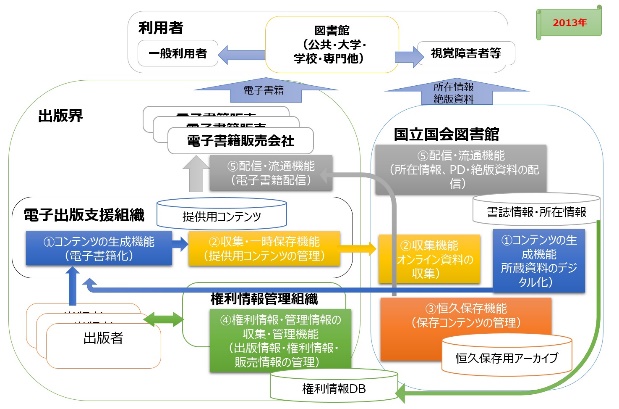


図5 電子書籍分野のアーカイブの機能モデル

### 電子書籍分野のアーカイブの役割

筆者は、電子書籍分野のナショナルアーカイブでの必要な機能を5つに分解し、それぞれの出版界と図書館界の役割分担を例示した。

長尾前館長により、2008年に提案され、改めて2015年3月に示された電子出版物流通に関する「長尾モデル」13）として提案されたモデルと類似したプラットフォームである。

提供用の商用コンテンツは、電子書籍サイトが保持し提供する。NDLは、DRMによるアクセス制限が掛かっていないコンテンツを収集し、ダークアーカイブとして恒久的保存の役割を担い、電子書籍サイトがコンテンツを保持できない、もしくはサービスを継続できない場合、NDLの恒久保存用アーカイブを利用して電子書籍サイトから提供する。また、NDLが原資料保存のために画像イメージでデジタル化したコンテンツを、出版界が二次利用して電子書籍化・プリントオンデマンド（POD）化し、電子書籍サイトから提供することにより、いわゆる絶版本の電子書籍化を促進する、さらに出版物・著作権者の情報は、出版界・図書館界の共通のデータベースとして維持・更新し、相互に利活用することも想定したモデルである。

### 電子書籍分野のアーカイブの機能モデル

電子書籍に関するナショナルアーカイブは、コンテンツの生成機能、収集・一時保管機能、保存機能、権利情報・管理情報の収集・管理機能、配信・流通機能の5つの機能を想定する

　コンテンツの創出から収集・保存、配信・流通に至る全体の流れを一元的に集約するため、恒久保存のアーカイブと共に、ナショナルアーカイブ全体のメタデータを集約するデータベースにおいては、多様な主体が多様なデータを扱えるよう、柔軟かつ多層的なデータ構造を実現する。具体的には、著作物・著作者・出版者等の書誌情報、販売データ・販売者・所蔵機関等の所蔵情報、目次・索引・シソーラス等の情報探索情報等の多様なデータを一元的に管理可能な仕組みとすることを想定した。

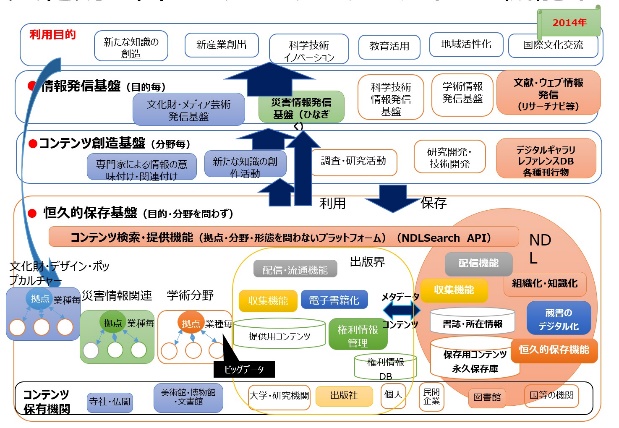
### 国の文化資産のアーカイブ

以下は、筆者がNDL在職中に、知識インフラの実現形として国のデジタル文化財のアーカイブシステムの基盤として想定したものであり、現時点において、NDLを含めて関係機関のコンセンサスが得られたものではない。

### アーカイブに関連した国の活動の方向性

国は、「知的財産政策ビジョン」（2013年6月7日知的財産戦略本部決定）に基づく国の施策の中で、図書館に直接関連する計画として、電子書籍化と利活用の促進に関する構想（電子書籍と出版文化の振興に関する議員連盟）、デジタル文化資産の保存・活用の基盤の整備に関する構想（デジタル文化資産推進議員連盟）14)、2020年までの6年間の文化芸術分野での振興に関する基本的な方針（文化庁）15)、学術情報の公開と共有の拡充に関する計画（文部科学省）16)、大規模災害の記録と記憶の保存などのアーカイブの構築等を検討している。個々の施策の目的は異なっていても、対象とする文化的資産は相互に関連するものであり、知識インフラを目指した仕組みとして、国全体で文化関係資料の価値を高め、新たな文化や情報を生み出す社会基盤として、「恒久保存・継承・公開・活用」が可能なナショナルアーカイブとなることが望まれる。「資料・情報を文化資産として収集・保存する」ということは、従来からの出版物に相当する情報の範囲ではなく、美術館、博物館、文書館等が保有する無形・有形の文化財をデジタル化した情報を含め、インターネット上で流通している著作物全てを文化資産としてアーカイブすることであると筆者は考える。

文化財のナショナルアーカイブの構築と運用に関しては、様々な機関において、必要性の認識の共有やアーカイブ立国宣言の提案17)、制度的な課題解決のために、アーカイブ基本法の法制化、推進体制作りなどが議論されている。また、具体的なサービスシステムの仕組みとして、技術的課題解決のための研究開発も進んでいる。18)

図6 文化財を含めたナショナルアーカイブの機能イメージ

恒久的保存基盤の部分は、文献も含めて、すべての文化財、災害情報が相互に関連付けられて蓄積されている。恒久的保存基盤を活用して新たな知識創造活動が行われ、知識創造活動の成果を、様々な目的に応じて利活用されていくことを想定する。

### 恒久的保存基盤

恒久的保存基盤は、恒久保存と利活用のための共通プラットフォームとして、１つの機関にすべてを集約するのではなく、各分野のアーカイブを集約する拠点が中核となって分散アーカイブを構築し、各機関の情報を相互に持ち合って、障害、災害に備えるとともに、情報のマイグレーションを行うことにより、将来にわたって利用を保証する仕組みである。その分散アーカイブを集合して、あたかも１つのアーカイブとして見えるようにして網羅性、完全性を確保し、個々の情報同士を意味的に関連付けて、情報間のネットワークを構築することを想定する。

このようにネットワーク化された情報に対して、分野を越えて網羅性を保証した検索GUIとして、本文の全文検索、あいまい検索、シソーラス検索などを組み合わせた検索で情報を取り出すだけでなく、取り出された情報から芋づる式に関連する情報を取り出せるようにすることである。

### 知識創造基盤

知識創造基盤は、それぞれの分野の専門家のみならず、広く国民も含めて、様々な分野の網羅的な知識を活用して、新たな著作物を創造する場である。創造活動を支援する基盤として、情報全体の基本情報としてのメタデータを付与する活動、記事、章節項、文節等の単位で組織化・構造化する活動、情報間を意味的に関連付けるための基本情報として、用語辞書、典拠、シソーラス辞書等を作成する活動を想定する。

新たな知識を創造する活動は、まず、恒久的保存基盤に格納された網羅的な情報を活用して新たな知識を創作する活動がある。関連付けて利用できる情報の幅が広がるため、より高度な創造性が期待できる。

また、歴史的な文化財や現代文化を映像化、画像化、テキスト化する活動、構造化された情報に解題情報等を付与する活動、情報間を意味的に関連付ける活動、テーマを設定してデジタルギャラリを構築する活動等が含まれる。

ここで生成された情報は、新たな知識として恒久的保存基盤に蓄積されていくことを想定する。

### 知識利活用基盤

「見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」として、広く国民による新たな知識の創造、新産業の創出、地域活性化、防災・減災、教育活用、教養・娯楽、観光、国際文化交流等、様々な利用者毎の目的に応じて、恒久的保存基盤に格納された一次情報、知識創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする基盤である。網羅的な情報から、利用目的に応じてあらかじめ適切に絞り込み、利用者の属性、スキル、利用場所に応じて、様々なGUIを用意して、利用者が必要とする情報、参考となる関連する情報を容易に得られるようにするもので、レファレンスサービスによる情報探索支援、オンラインレファレンスなども含まれる。

### 運用基盤

データ保存課題、人材不足問題、資金不足問題、権利処理コスト問題など、デジタルアーカイブの活性化を阻害する課題を制度的に克服する方策(人材育成策・予算措置・権利処理に関する法改正など)を導き出し、解決を図っていく中核的な役割を担う推進母体(司令塔)としての組織・体制も必要である。この基盤では、ナショナルアーカイブ全体の戦略企画、デジタル情報の保存や利活用のための調査研究、研究開発、デジタル化支援、アーカイブに所蔵された資料に関する知識、読解力とIT技術の知識等も備えた高度な専門的人材の育成、孤児著作物の権利処理や、絶版作品のデジタルアーカイブ化における所有権、肖像権問題も含めた権利情報DBの構築を促進する等の役割を持つことを想定する。

## 文化財でのLOD化の目指すところ

従来は、NDLは、同分野の情報の網羅性を意識して、各機関の情報に付けられたメタデータによる網羅的な情報探索、情報検索を効率的に実践できることに注力してきた。今後の国の文化財全体のナショナルアーカイブでは、異分野の様々な情報も含めて情報を関連付けることにより、思わぬところで自分の興味を喚起する情報に出会え、新たな発想の知識の創出が可能になる環境の実現が重要であると筆者は考える。

Linked Data化は、様々な情報を関連付けて知識として再利用するための具体的な実現方法を示したものであり、LOD化が進むことにより、あらゆる情報が関連付けられて利活用が可能になる。筆者は、LOD化の動きが、分野を越えて進展することにより、文化財全体のアーカイブがナショナルアーカイブとして実現されていくものと考える。

### クラウドコンピューティングの世界でのサービスの連携（2007年想定）

図7は、筆者が2007年に想定した、クラウドコンピューティングの普及を見越した将来像である。

　様々な機関から提供されているWebサービスが、連携することにより、より効果的なサービスとなっていく。サービスが連携できる情報提供環境として、当時の大きなパラダイムシフトの一つとしてのクラウドコンピューティングは、今や個人の情報管理のインフラとして普及している。

クラウドコンピューティングは、ネットワークコンピューティング、グリッドコンピューティング、SaaS（Software as a Service）、Webサービス等の技術を融合させた基盤である。

個別の組織が、単独でコンピュータ資源（ハードウェア、ソフトウェア）や情報を抱えてサービスを提供し、利用者が、個別のサービスを利用するような形ではない。クラウドコンピューティングの普及により、様々な機関の情報の共有だけでなく、個別のサービスを組み合わせて利用することが容易に実現できる環境が整った。

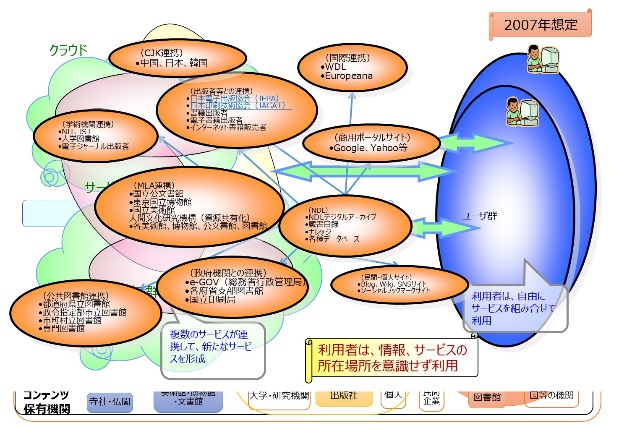
各情報保有機関のサービスも、このパラダイムシフトに対応していかなければ、孤立したサービスになっていく。クラウドコンピューティングが普及期に入り、また、デジタルトランスフォーメーションの考え方が一般化してきた社会において、各機関が保有するデジタル情報の管理の仕方、顧客サービスのあり方を再考する必要がある。今後ますます、情報と情報が情報群、サービスとサービスがサービス群、人と人がユーザ群として有機的に関連付けられていく。

いわゆるWeb2.0の時代では、パーソナライズ、人と人の関係を活用して情報を選択してきた。また、サービスも先進的な機関が提供するWebサービスを個別に利用してきた。Web3.0の時代以降は、それに加えて、情報と情報、サービスとサービスの関係を組み合わせて、ユーザ群と情報群が多対多の関係で、より的確な情報の活用が可能になってくる。

従来のように個別のデータベースを検索し、コンテンツを閲覧するような形から、クラウドコンピューティングの世界での複数のリソースから得た情報に、分析情報などの付加価値を付けて、利用者へ助言するような内容を知識として提供するサービスが実現されつつある。例えば、近年のカーナビの情報は、単に目的に行くルートを情報として提供しているのではなく、様々な手段により動的に変化する状況を集約し、その結果を踏まえてのリルート情報や、ルート上の様々な位置でのスポット情報などを合わせて提供している。ナショナルアーカイブの枠組みの中で、Linked Open Data化が進むと、更に利便性の高いサービスとして発展する。

このようなインターネットの世界での情報提供の状況の変革が、NDLで進めているウェブアーカイブ、情報探索サービスのあり方にも大きく影響してくる。従来のウェブアーカイブは各機関のウェブ情報を収集保存してきているが、このようなサービスの形をそのまま保存することは困難であり、また、関連付けられたメタデータ等、どのようにしてサービスの中にある情報を収集していくのかを考える必要がある。

クラウドコンピューティングの普及により、想定したサービスが現実になってきたが、今後、NDLは、情報を収集し将来にわたって保存する使命をどのように果たしていくか、考えていく時代が来ているように思える。

図7 クラウドの世界でのサービスの連携

### 文化財のアーカイブ構築のために何に留意するべきか

最後に、ナショナルアーカイブの構築を目指す中で、司令塔的な役割を果たす組織および文化財を保有する各機関は何に留意すべきか、筆者が想定する事項を列挙する。

* 今後10年を見据えて、個別の組織の論理ではなく、利用者視点でサービスを構築する。
* アーカイブの網羅性の確保のために、業種業態毎のアグリゲータ機関を設置してメタデータを集約する。更に、Linked Data化するために、情報を一意にする識別子を付与する。
* メタデータはもとより、コンテンツをOpen Data化する。
* 分野を越えたアーカイブ構築のために、メタデータの記述要素、記述規則および情報交換の通信規約（API）は、共通仕様を実装する。
* メタデータおよび情報内容の用語表現の違いを吸収する辞書を構築する。
* アーカイブの構築・運用に必要な人材を育成・確保する。

これを実現するするために、国の役割は次の点にある。

* 知的財産政策ビジョン、知財計画に掲げた施策を確実に実施する。
* 文化資源のアーカイブの利活用に必要な制度を整備する。
* アーカイブの構築・運用に必要なスキル・知識を習得するための体系的なカリキュラムと教材を作成するとともに、教育体制を確立する。

## まとめ

NDLは、今までは「電子図書館構想」（1998年）の実現を目指し、具体的な実施の柱立てとして「電子図書館中期計画」（2004年）を策定し、デジタルアーカイブ構築、情報に関する情報の充実、デジタルアーカイブポータルの構築を行ってきた。

筆者は、電子図書館としての知識・情報・文化の新しい基盤を含めて、図書館の枠を越えた基盤作りが、「知識インフラ」であり、知識インフラの実現形の１つが「ひなぎく」、知識インフラ構築の全体の実現形が、「文化資源のナショナルアーカイブ」と位置付けている。1990年代から目指す方向性は明確であり、構築に当たっては、セマンティックウェブ技術の適用を常に意識して進めてきており、具体的な実現手段として浸透しつつあるLinked Data化は、ナショナルアーカイブの構築においても、実現手段として中核的な技術の一つとなると考える。

## おわりに

個々の文化財の保有機関でのデジタル化が進む状況において、文化的資産をあらゆる人々が将来にわたり享受、活用できるようにし、人々の創造的な活用に貢献するためには、個々に情報資源の「見える化」をするだけでなく、全体のそれぞれの施策が同一の方向性を持って、相互に資源を補完し合って、社会全体でより効率的な利活用の保証に取り組む必要があり、ナショナルアーカイブは重要な基盤の役割を果たすことになる。ナショナルアーカイブの構築の先には、世界レベルでの「インターナショナルアーカイブ」があり、その構築へと発展することを目指していくべきと考えている。

今後10年のデジタル情報化の進展を見据えると、現行の図書館業務は根幹から変革が求められると思われる。知識創造のための情報の多くがインターネットにより、入手できる状況において、知識創造を支援する図書館の役割と必要な機能を再検討し、速やかに対応していかなければ、情報提供機関としての図書館は存立意義を失う可能性がある。

～～～

■参考文献

1. 田屋裕之.「パイロット電子図書館システム」の現状と課題.出版ニュース.1996、 96 1/上・中、p.14-17
2. 中山正樹.国立国会図書館電子図書館中期計画2004の実施に向けて.情報の科学と技術. 2004、Vol.54、No.9、p.453-460.
3. 久古聡美・吉田曉・中山正樹. Web2.0世代のデジタルアーカイブポータルの提供を目指して.情報管理. 2006、Vol.49、No.6、p.313-323.
4. 国立国会図書館. “機能概要≪国立国会図書館サーチについて”. 国立国会図書館サーチ、 　<http://iss.ndl.go.jp/information/function/>、　(accessed 2016-1-15).
5. 国立国会図書館. “日本全国書誌の在り方に関する検討会議について”. 国立国会図書館インターネット資料収集事業(WARP). <http://warp.da.ndl.go.jp/info:ndljp/pid/8703998/www.ndl.go.jp/jp/library/data/jnbconf_20100303.html>、 (accessed 2016-1-15).
6. 内閣府科学技術政策・イノベーション担当. “第4期科学技術基本計画（平成23年8月19日　閣議決定）”. 内閣府. <http://www8.cao.go.jp/cstp/kihonkeikaku/4honbun.pdf>、(accessed 2016-1-15).
7. 国立国会図書館. “第三期科学技術情報整備基本計画”. 国立国会図書館インターネット資料収集事業(WARP). <http://dl.ndl.go.jp/view/download/digidepo_8262067_po_basic_plan03.pdf?contentNo=1>、 (accessed 2016-1-11).
8. 長尾真.未来の図書館を作るとは.長尾真、 2012、 27p
9. 中山正樹. 国立国会図書館におけるデジタルアーカイブ構築－知の共有化を目指して－. 情報管理. 2012、Vol.54、No.11、p.715-724.
10. 内閣官房東日本大震災復興構想会議. 復興への提言～悲惨のなかの希望～. 2011-06-25、 第12回東日本大震災復興構想会議. 57p. <http://www.cas.go.jp/jp/fukkou/pdf/kousou12/teigen.pdf>、 (accessed 2016-1-15).
11. 国立国会図書館. “東日本大震災アーカイブ構築プロジェクト”. 国立国会図書館.<http://kn.ndl.go.jp/static/about/project>、(accessed 2016-1-15).
12. 内閣官房知的財産戦略本部. “知的財産政策ビジョン”.首相官邸政策会議.<https://www.kantei.go.jp/jp/singi/titeki2/kettei/vision2013.pdf>、(accessed 2016-1-15)
13. 長尾真. “「電子図書館と電子出版の今後」JEPAセミナー”. 電子出版協会.<http://www.slideshare.net/JEPAslide/20150303-drnagao-45401397>、 (accessed 2016-1-15)
14. デジタル文化財創出機構.“「国立デジタル文化資産振興センター(仮称)」設立構想提言”.文化庁文化関係資料のアーカイブに関する有識者会議.<http://www.bunka.go.jp/seisaku/bunkashingikai/kondankaito/bunka_archive/01/pdf/shiryo_5-2.pdf>、(accessed 2016-1-15)
15. 文化庁.「文化芸術の振興に関する基本的な方針（第4次基本方針）」.文化庁文化審議会.<http://www.bunka.go.jp/koho_hodo_oshirase/hodohappyo/pdf/2015052201.pdf>、(accessed 2016-1-15)
16. 日本学術会議情報学委員会. 我が国の学術情報基盤の在り方について －SINET の持続的整備に向けて－.日本学術会議.<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-t192-2.pdf>、(accessed 2016-1-15).
17. 「アーカイブ立国宣言」編集委員会 編. アーカイブ立国宣言. ポット出版、 2014、272p.
18. 中山正樹.国立国会図書館のサービスシステムの歩みと新たな方向性の模索－電子図書館事業20年の歩みー. 国立国会図書館月報. 2015、Vol.648、p.18-24.

**■図表、写真のキャプション**

図1 電子図書館に関連する構想とサービスの変遷

図2 NDL電子図書館サービスの全体像

図3 知識情報基盤の構築モデル

図4 東日本大震災アーカイブのシステムイメージ

図5 電子書籍分野のアーカイブの機能モデル

図6 文化財を含めたナショナルアーカイブの機能イメージ

図7 クラウドの世界でのサービスの連携

≪討議報告≫

LOD化によるデジタル文化財の利活用を目指して

**（発表：中山正樹）**

**和中　幹雄（大阪学院大学）**

フォーラム当日の中山氏による発表題目は、「LOD化によるデジタル文化財の利活用を目指して」であった。本誌掲載の論文ではそれは副題目となり、発表内容をより端的に示す題目を用いている。この発表の趣旨は、「国立国会図書館（NDL）の電子図書館事業の20年間を情報技術の観点から振り返り、今後のナショナルアーカイブに向けた展望と課題を概観したもの」ということができる。

フォーラム当日のプレゼン資料には、「NDL電子図書館サービスの全体像」「長期保存システムの概念モデル」「国立国会図書館サーチのコンセプト」「文化財を含めたナショナルアーカイブの機能イメージ」「ナショナルアーカイブにおけるLOD化」などなど、数多くのモデル図やイメージ図で溢れ、個々のスライド一枚を取ってみても、優に一つの発表内容となるものが多く見られた。そのため、予定された時間で終了できず、発表の趣旨の理解も不十分だったように感じられた。しかしこの点については、本誌掲載の論文を読めば氷解するであろう。

彼の発表が自分の職歴の紹介から始まるのも故なしとしない。NDLの電子図書館事業は、1994年に情報処理振興事業協会（IPA）との協力で実施したパイロット電子図書館プロジェクトによって始まる。彼はこの事業にIPA側の担当者として参画する。その後、関西館が開館し電子図書館サービスが開始される2002年にNDLに入館し、2004年に、デジタル・アーカイブ構築とデジタル・アーカイブ・ポータルの考え方を明示した「電子図書館中期計画2004」策定の中心メンバーとなり、その後、PORTAからNDLサーチの開発に携わり、2015年3月末に、電子図書館事業や情報システム全般を統括する立場にある者として定年退職を迎える。このように、NDL電子図書館事業の20年と中山氏の図書館システムに関わるキャリアは全く重なるからである。

1994年から始まる現在までの20年間のシステム構築に当たっては、セマンティックウェブ技術の適用を常に意識して進めてきたと彼は述べている。つまり、LODも含めたセマンティックウェブの要素技術をどのような考え方で個別に適用してきたかを振り返ることにより、今後の方向性を示したのが本発表であったと言える。

情報技術の観点からの発表ではあるが、特に今後のナショナルアーカイブ構築に向けた課題整理において、内閣に設置された知的財産戦略本部による「知的財産政策ビジョン」などの国の政策を背景とした提言が多く見られる。彼は、発表を終わるにあたって、次のように述べている。「LOD化されたナショナルアーカイブ構築に当たっては、それらの意義・理念の確認、認識の共有、草の根的な試作など、やれるところからやるという段階から、大所高所の方針・計画に基づいた、具体的な実用サービス構築を行うという次の段階に移行する必要がある。そのためには、司令塔が必要である」と。

我が国にどのような司令塔が可能であろうか。

彼が示したモデル図やイメージ図は、これからの図書館を考える上で示唆に富む情報が数多く含まれている宝の山と言える。

司会者の仕切りが悪かったこともあり、討議時間はほとんど取れず、質疑もNDLサーチの個別問題に終始したことは残念であった。当日のプレゼン資料にあるさまざまなモデル図やイメージ図を基にまとめられた本誌掲載の論文をお読みいただき、今後、活発な議論がなされることを期待したい。

# （2016年）AIを活用した「知の共有化」システムの方向性―「未来の図書館を作るとは」の実現に向けて－

2017年10月3日

元国立国会図書館　専門調査員・電子情報部長

中山正樹

## はじめに

1980年代後半に検討が開始された電子図書館事業は、1990年前半に策定された電子図書館関連の構想において「地球規模の知的財産を誰でも容易に利用できるようにする」という目標を掲げると同時に、1990年代中旬に我が国で最初の実用化実証システムとして、「ネットワーク環境における情報と文献の利用のための高度検索システム（Ariadne）」（京都大学）1）、パイロット電子図書館実証実験プロジェクト（国立国会図書館（NDL））での実用化実証実験2）が行われた。国の動きでは、2003年には、e-Japan重点計画2003、e-JAPAN戦略Ⅱ加速化パッケージ、（内閣官房IT戦略本部）において、「国のデジタルアーカイブ構想」、「ジャパンウェブアーカイブ構想」の実現を、また2004年には、e-Japan重点計画2004において「国立デジタルアーカイブポータル構想」を一層推進することが明記された。このような動きと同期して、国立国会図書館では、国立国会電子図書館中期計画20043）を策定し、①「デジタル･アーカイブの構築」として、国立国会図書館デジタルコレクション、国立国会図書館インターネット資料収集保存事業（WARP）」、 ②「情報資源に関する情報の充実」として「リサーチナビ、レファレンス協同データベース」③「デジタル･アーカイブのポータル機能」として「国立国会図書館サーチ」を構築し提供した。2005年7月に試験公開したPORTAのプロトタイプは、ナショナルアーカイブポータルの原点と言え、インキュベータの役割を果たした。4）さらに「東日本大震災アーカイブ（ひなぎく）」により大震災関連の記録・記憶に限定はされているが、従来の図書館の枠を越えたデジタル知識基盤の構築を進めてきた。

現在は、分野を特定せず、図書館を含めて文化情報資源、知的情報資源を保有する機関（以降、「アーカイブ機関」という。）が連携して、日本全体でのデジタル知識基盤を構築し、様々分野の情報を知識として保存し活用できるようにする機運が高まり、活動が活性化してきた。

デジタル知識基盤を構築するに当たっては、ビジネス、制度、組織、技術等の観点から相互に関連し合う課題が多い。情報システムの構築とサービス展開、知的情報のデジタル化においては、いわゆる「第4次産業革命」、「デジタルトランスフォーメーション（デジタル革命）」と言われる技術革新とビジネス変革の時期に来ている。そのような現状と今後の展開を考慮すると、長尾元国立国会図書館長が2012年3月にNDLを退官される際に職員に配布された「未来の図書館を作るとは」5）の中で示された「未来には実現できるだろう」とされたことが、「今、この時代」の図書館で実現できるレベルにあり、これから構築する「デジタル知識基盤」は、第4次産業革命、デジタル革命の方向性に沿うことが必要と考える。

　各アーカイブ機関が、デジタル知識基盤に対応したデジタルアーカイブシステムの構築を、適正な内容と費用で行うためには、有効性が確認されている標準的な調達手順（プロセス）で行うことが重要である。「政府情報システムの整備及び管理に関する標準ガイドライン」（各府省情報化統括責任者（CIO）会議）6）は、政府機関全てでのシステム構築のプロセスと成果物を規定しているガイドラインであり、これを参考にすることにより過不足のない調達要件を受託者に提示して、適正な手順で、適切な技術、パッケージを活用したシステム開発を効率的・効果的に行うことができる。

さらに、「政府情報システムの整備及び管理に関する標準ガイドライン」に沿った開発プロセス（タスク）を遂行できる情報システム関連の担当者の人材育成は、開発工程の個々のタスクに必要なスキルと知識の項目が示されている「iコンピテンシ・ディクショナリ(iCD)」（2016年情報処理振興機構（IPA））7）が参考になる。iCDで網羅的に示されたタスク毎のスキル、知識の中から、担当するタスクに応じて選択的に身に付けることにより、実践的な人材を効率的に育成することができる。

## 「知の共有化」とは

図書館が保有する冊子体資料目録のデジタル化から始まり、冊子体資料のデジタル化、ボーンデジタルの著作物を扱う「電子図書館」、「デジタルライブラリ」、学術情報、研究データも含めあらゆる情報を知識として活用できるようにする「知識インフラ」、美術館、博物館等を含めて文化情報資源全体の「文化資産ナショナルアーカイブ」等、更に同義語的に、インターネット資料、オンライン資料、電子出版物、電子書籍、電子雑誌、知的情報基盤、文化情報資源、文化情報資産、デジタルアーカイブ、ナショナルアーカイブ、インターナショナルアーカイブ等、時の流れの中で、目指すところは同様でありながら全体集合、部分集合の違いにより、様々な呼び方をされてきたが、知識全般を扱う適切な言葉が共有されていないのが現状である。

本稿では、アーカイブ機関に限らず、あらゆる組織、個人がデジタル化した情報を共通的に保存し利活用できるようにする基盤を「デジタル知識基盤」といい、「デジタル知識基盤」により、新たな知識を創造し還流することを「知の共有化」として記述する。

## デジタルトランスフォーメーション(デジタル変革)

今は、IoT、ビッグデータ、ロボット、AI等の技術革新による、いわゆる「第4次産業革命」の入り口にいる。あらゆるものがインターネットに接続するIoTの広がり、あらゆる情報がビッグデータとして活用され、AI技術により、様々な分野で定型的な業務はもとより、人海戦術では不可能だった業務まで、AI技術を適用したサービス、ロボットの適用が始まっている。既存のビジネスや業務に新技術を取り入れるだけでなく、ビジネスモデルを変え、経済活用のみならず、個人の生活や社会構造まで影響が及ぶ。8）

デジタルフォーメーション（デジタル変革）とは、あらゆる情報がデジタル化され、IT技術によって、社会や産業、企業、人のあり方や働き方が変わっていくこと。第4次産業革命が進むにつれて、発展するビジネスと縮小するビジネスが明確になっていく。時代環境が大きく変わる時、それにそぐわないビジネスは淘汰されていく。匠の技的な高度な伝統的技能を要する作業や、旧来の延長線で仕組みの高度化、洗練により、生き残れるビジネスもあるが、現状維持のビジネスの多くは、相対的に意義を失う可能性が高い。

しかし、IoT、ビッグデータ、ロボット、AI等の技術を、クラウドコンピューティングやモバイル環境で容易に活用できるようになったことは、少ない投資で事業や先進サービスを立ち上げることが可能になった。各アーカイブ機関においても同様であり、利用者へのサービスを向上して、機関の存在意義を高める大きなチャンスでもある。

## 「未来の図書館を作るとは」【長尾先生】を読み返して

「未来の図書館」の概念

* 「人間の持っている知識は頭脳の中にあり、種々の知識が何らかの関係性によってつながれていて、連想的に関係する知識が取りだされている」
* 「図書館においてもぼう大な書物の中に存在する知識が関連性をもって書物という単位を超えてつなげられ、それが取り出されることが大切であろう。」
* 「本のある部分に存在する単語や概念を集め、それらに近い単語や概念が存在する部分を他の本について網羅的に調べる」
* 関連する知識を人間頭脳の中のネットワークのようにつないで、利用者の要求に応じて提示できるような形の電子図書館の内容の組織化が望まれているのである。
* 電子図書館における図書・資料は部品に解体され、それぞれが種々の観点からリンク付けされた巨大なネットワーク構造が作られるようにする。これは1つの社会で共有する中立的な知識構造、知識システムである。
* 個人によって違った知識の構造の部分については、その人の力によって種々の検索方式を試み、自分の必要とする情報をとり出して中立的な知識の構造に付加してゆくことが出来ねばならないし、またそれによって自分に合った知識の構造を作りあげてゆくことができるだろう。
* 現実世界の本や情報の大切さ以上にヴァーチュアルな世界における情報処理と表現力の可能性にもっと大きな関心を持つべき時代に来ていると言えるのではないだろうか。

「未来の図書館を作るとは」（長尾真）より抜粋

「電子図書館」（1994年長尾真著）では、「既存の図書や資料をデジタル化すればそれで電子図書館が実現するかといえばそうではない。あるべき姿はデジタル化された情報を縦横に使いこなし、まったく新しい知的空間を創造するための図書館である。」9）とされており、「Ariadne」はその理念に基づいた実用化実証実験システムである。

「未来の図書館を作るには」5）が発行された2012年初めは、まだ第3次人工知能（AI）ブームの前で、AIはまだブレークスルーしていなかった。しかし、2012年以降のAIにおける機械学習はディープラーニング手法等により飛躍的に進展し、また、アーカイブ機関での資料のデジタル化、デジタルコンテンツのオープンデータ化、LOD化の加速化により、AIが扱える質の高いビッグデータが揃いつつある状況で、「未来の図書館を作るには」の中で「未来」と示唆されていた相当な範囲の仕組みが、今後5年程度で実用化を見通せるようになった。そこで、2017年時点での第4次産業革命等の動向を踏まえて、「未来の図書館を作るには」で示唆された提言の実現性について考察したい。

## 知的情報資源のビッグデータとしての活用の可能性

図書館等のアーカイブ機関は、従来から専門員の知識・ノウハウにより統制され形式知化してきた情報を保有している。二次情報としてのメタデータ、書誌情報、情報を見つけ出すための情報としてのレファレンス情報（参考情報）、Q&A、調べ方案内情報等は、既にビッグデータとして活用できる状況であるが、一次情報としてイメージデータ化されたデジタルコンテンツの内容本文は、検索にはほとんど利用されていない。

今後、ビッグデータとして活用が期待される情報として、アーカイブ機関のデジタルアーカイブ内で保有している一次情報があり、オープンデータ化されていない情報も含めて、全文テキストを活用した検索インデックスの作成をはじめ、AIの学習データとして活用するサービスが認められる方向である。

## 「知の共有化」システムの方向性

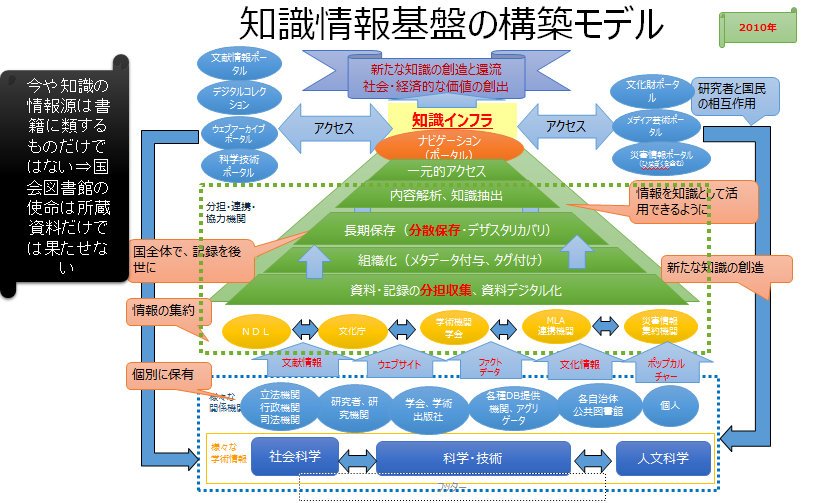
Web2.0時代と言われた頃、図書館界ではOpac2.0と称するサービスで、いわゆるAIとまでは言えないが、AIを指向した組織化、検索サービス等の実現を目指してきた。

既にインターネット上に広範な情報が公開されているが、ここ数年で、文化機関、公的機関が保有している情報もオープンデータ化が進みつつあり、ビッグデータとして利活用できる方向に向かっている。このような時代に、従来の業務やサービスは、AIが組み込まれたシステムやロボットに支援されてサービスが省力化され、さらにビッグデータとして網羅性が確保されることにより、「知の共有化」が可能な新たなサービスが生まれようとしている。

### 知識インフラの概念

* 知識インフラ構築における最も大切な概念は、情報を集め、これを知識化し活用することによって新しい情報・知識を創出し、知識インフラに加えるという形で循環的にこのシステムを強化・拡大してゆくことによって社会・経済に貢献することである。
* 知識が種々の観点から組織化、構造化されて利用しやすい形で蓄積されて利用に供される

「未来の図書館を作るとは」（長尾真）より抜粋



既に、第４期科学技術基本計画「科学技術に関する基本政策について」（内閣府　総合科学技術会議2010年）では、文献等研究情報のデジタル化、オープンアクセスの推進等とともに、「文献から研究データまでの学術情報全体を統合して検索・抽出が可能なシステム（「知識インフラ」）の展開を図る」とされていた。10)

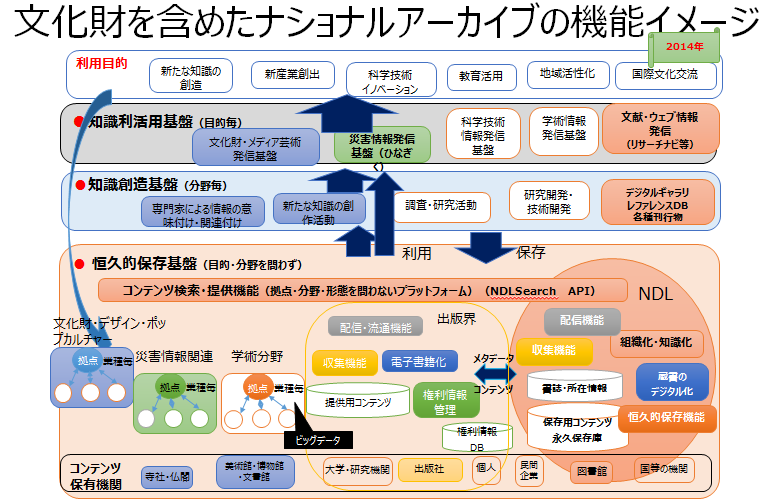
図1 知識情報基盤の構築モデル

図書館の視点で見れば、専門家によって作成されてきた、書誌、辞書、典拠（シソーラス）類に加えて、著作物のイメージ画像、全文フルテキスト、目次、索引、あらすじ、書評等、出版界や図書館界で従来から形式知化してきた情報を、信頼性の高いビッグデータとして利用可能にし、AIの教師用データとして投入・学習することにより、長尾先生が1994年に提唱した「電子図書館」が、信頼性の高い知識を豊富に持った「AI化されたデジタル知識基盤」として実現することになる。

### 文化情報資源のナショナルアーカイブ

* 【長尾先生】図書館は、書物等の情報資源を収集・保存し提供する場であるとともに、それらを使って関心のある人が集まって議論し新しい知識を創造する場
* 【長尾先生】日本としてのナショナルアーカイブは、日本中に存在する知識情報が有機的に結合され、日本中の人が自由に使える日本の「知識インフラ」

「未来の図書館を作るとは」（長尾真）より抜粋

　「文化財を含めたナショナルアーカイブの機能イメージ」は、2012～2014年に描いた図である。11）12）

我が国の多様な文化を知識として保存・継承する役割、様々な分野の専門家が参加し、新たな文化を知識として創造していくための社会的な基盤としての役割、それらの知識を利用目的に応じて発信する役割、そして、これらの仕組みを統括し運用していく役割が考えられ、文化的資産を館種毎に集約している拠点と、新たな知識を創造し発信している拠点等が分担して構築・運用していくことを想定した。13）

図2 文化財を含めたナショナルアーカイブの機能イメージ

### 恒久的保存基盤（あらゆる情報の恒久的保存）

恒久保存と利活用のための共通プラットフォームで、１つの機関にすべてを集約するのではなく、各分野のアーカイブを集約する拠点が中核となって分散アーカイブを構築する。あたかも１つのアーカイブとして見えるように、個々の情報同士を意味的に関連付け、検索で情報を取り出すだけでなく、取り出された情報から芋づる式に関連する情報を取り出せるようにする。分野での情報の網羅性・完全性が重要であり、また異分野の情報との関連付けにより、AIとしての知識として活用できる。

### 知識創造基盤（新たな知識創造活動の場）

キュレーター、ライブラリアン等の支援のもとで、それぞれの分野の専門家のみならず、広く国民も含めて、情報に付加価値を付けたり、他の分野の情報と関連付けて、二次的著作物を創造する場として想定した。

### 知識創造支援機能

情報全体の基本情報としてのメタデータを付与する活動、記事、章節項、文節等の単位で組織化・構造化する活動、情報間を意味的に関連付けるための基本情報として、用語辞書、典拠、シソーラス辞書等を作成する活動である。

* 辞書類作成

辞書類の作成において、**コーパス**は、AIの支援により省力化が進むと思われる。

**件名典拠、シソーラス**は、本文フルテキストもAIへの入力データとして特徴量を抽出することにより、書誌データのみならず、本文テキストから標目を抽出することが可能になり、また、コーパスを利用することにより、同義語、上位語、下位語、関連語、分類記号の関係づけをすることも可能になると思われる。**固有名典拠**は、既に刊行されている書籍の著者に関して、AIを利用して、著作名、著者、出版年、NDC分類等の特徴量を抽出・分析することにより、著者の自動同定の精度は相当向上すると思われる。AIにより、省力化、自動化ができることにより、書籍だけでなく雑誌記事等も含めて、固有名典拠の自動作成が可能になる。

**NDC分類**は、件名典拠をベースにした分類の番号体系は、人間の判断によるところが大きいので、容易にAI化することは困難と思われるが、分類を番号体系でなく意味的なシソーラス体系として維持し、改訂していく作業は、AIにより体系を整理することは可能になる。また番号体系の版毎の対比表の維持も人手を介さずに行えるようになると思われる。

ただ、配架のための分類という考え方、資料検索のための分類体系は、全文検索を含めて様々なレベルでの検索が容易になってくる状況において、役割を問い直す必要があるのではないか。

* **情報を関連付けるための情報**

情報を関連付けるための情報として、まず共通語彙がある。分野を超えた情報交換を行うには、個々の単語について、表記・意味・データ構造を統一し、互いに意味が通じるようにすることが必要であり、AIの知識として蓄積する場合、共通語彙になっていることが学習の効率を高められる。

　情報と情報を結びつけるための情報として、コーパス**、**オントロジー**、**分野別シソーラス、各種用語辞書、DBPedia等があり、それらが、永続的識別子となって、Linked Data化されていると、より精度の高い関連付けが可能になる。データに用いる文字や用語を共通化し、情報の共有や活用を円滑に行うための基盤であるIMI(Infrastructure for Multilayer Interoperability:情報共有基盤)の構築への参画、連携が重要である。

* **情報に関する情報の組織化**

情報の組織化は、二次情報としての書誌、メタデータの作成があるが、新しい著作物に対するNDC分類、件名などは、既存の書誌、典拠を教師データとして蓄積することにより、AIによりかなりの精度の自動付与が可能になる。専門家は、AIによる精度に応じてチェック、補正することが役割になるのではないか。また、雑誌記事に関しても、記事を著作単位とした書誌作成がAIにより省力化して付与が可能になる。

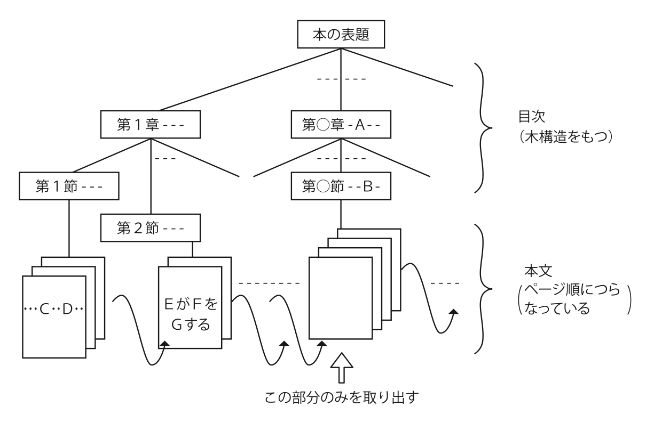
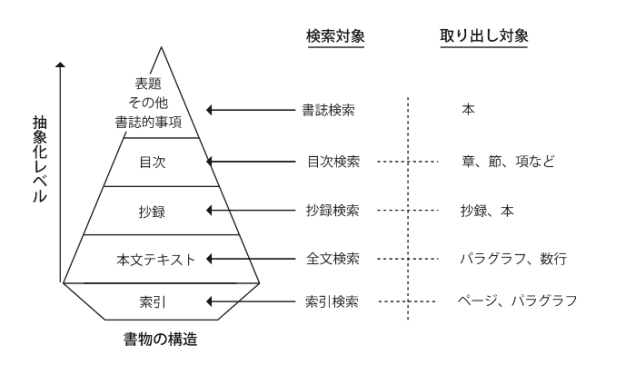


図4 書籍の構造化（「未来の図書館を作るとは」より）

図3 種々の検索システム（「未来の図書館を作るとは」より）

* 従来の書誌情報の考え方をマルチメディア情報に対応できるように拡張かつ詳細化するとともに、対象資料に関係する様々な種類の情報にリンクをはり、それらの情報をたどってゆけるようにする

「未来の図書館を作るとは」（長尾真）より抜粋

* 図書・資料は部品に解体され、それぞれが種々の観点からリンク付けされた巨大なネットワーク構造が作られるようにする
* 目次情報を付けたり、本の表紙の画像や数行の簡単な要旨を付ける
* 動画・静止画・音声は、画像認識機能により自動タグ付け

「未来の図書館を作るとは」（長尾真）より抜粋

### 知識創造活動

新たな知識を創造する活動は、図書館等のアーカイブ機関の主たる目的の一つであり、デジタル知識基盤の中で、恒久的保存基盤に格納された網羅的な情報を活用して新たな知識を創作する活動である。歴史的な文化財や現代文化を映像化、画像化、テキスト化する活動、構造化された情報に解題情報等を付与する活動、情報間を意味的に関連付ける活動、テーマを設定してデジタルギャラリを構築する活動等も創造活動の一つと位置づけられる。創造活動が、効率的に行えることにより、質、量ともに多くの知識が再生産され、創造された知識は、恒久的保存基盤でアーカイブされることを想定した。

* 既存の知識から自分の必要とする部分を切り出して、新たな発想の著作を創造し、新たな知識とする

「未来の図書館を作るとは」（長尾真）より抜粋

### 新たな知識創造の場としての図書館

図書館等の機関は、自館が保有している資料に加えて、「デジタル知識基盤」で共有されている知識を組合わせて提供し、知識創造の場として、施設としての実空間と、インターネットを利用した仮想空間が、相互にシームレスに、「場」として利用できるものを想定する。

情報を媒介して専門家と専門家を繋ぐこと、様々な組織が持つ情報が関連付けられて、それぞれの組織を中心に活動していた人が繋がっていくこと、専門家を媒介して辞書と辞書を繋ぐことを可能にして、様々な分野の人たちが保有する知識が有機的に関連付けられ、それを活用して、更に利用者同士がつながっていく「場」としての機能を期待する。

* 司書の世話にならなくてもある程度のレファレンスサービスが受けられるように
* 自然言語による質問要求を受け付けて、取り出したものがその要求に対応するものであるかどうかを自然言語処理技術によって調べ、できるだけ質問要求に近いものだけを選択する

「未来の図書館を作るとは」（長尾真）より抜粋

* 議論の場の提供という図書館機能
* 解決したいという人達と研究者、そして図書館司書のグループが種々の角度から資料をもとに議論
* 考え方の違う人達が知識を共有し、その違いを議論を通じて明らかにすると共に、新しい知識・思想を作り出してゆく場
* 出版社、著者と読者をつなぐ場としての図書館
* 図書館が司書による相談サービス、あるいは自動的な案内サービスをする場合でも、自分の電子書棚を作りたいという人の多くは図書館から借りるのではなく出版社のデータベースの方に行き、書物を購入することになるわけで、図書館は出版社と読者を結合する接続業者のようになってゆく

「未来の図書館を作るとは」（長尾真）より抜粋

* レファレンスサービス

利用者からの問い合わせに対して、AIを利用して、必要とする内容に最もマッチングする資料名、所在場所へナビゲーションする。資料がオープンデータであれば、参考となる内容の記載部分を列挙し、さらに答えとしての事実まで提示が可能になる。

自然言語処理機能を利用することにより、音声での自然言語の問いに対して、「デジタル知識基盤」を知識としたAIが、解の候補を抽出し、それを音声での自然言語で回答することが可能になる。

* 事実に基づいた報告書

依頼または予測調査に関しては、AIを利用した文献の意味的内容検索により、最もマッチングする資料名、資材場所を探し出し、その資料がオープンデータ化されていた場合は、記述された内容を洗い出し、内容を要約して報告書の原案を作成することも数年で実用可能と思われる。

### 知識利活用基盤（情報発信）

恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報が、デジタル知識基盤の中でAIの機能を活用して有機的にネットワーク化され、あらゆる分野で利活用できるようにする。

見るだけのアーカイブ」から「使い、創り、繋がり、伝えるアーカイブ」として、広く国民による新たな知識の創造、新産業の創出、地域活性化、防災・減災、教育活用、教養・娯楽、観光、国際文化交流等、様々な利用者毎の目的に応じて、恒久的保存基盤に格納された一次情報、コンテンツ創造基盤で創出された二次的情報を有機的に組み合わせて、利用できるようにする。

* 電子図書館になって取り出す単位が書籍の単位ではなく、書籍の中の章や節、パラグラフ、あるいはこんな内容が書かれている部分のみ
* 自動的な形で適切な知識の所在にまでナビゲートしてゆくシステム
* その本のどこに書かれているかを探すというのではなく、自分の欲しい情報そのものが出てくることになる。
* 書誌検索のような単純、単一の検索でなく、種々の検索のモードを提供する
* 種々のあいまいさを許すあいまい検索の工夫

「未来の図書館を作るとは」（長尾真）より抜粋

知識検索機能として、網羅的な情報から、利用目的に応じてあらかじめ適切に絞り込み、利用者の属性、スキル、利用場所に応じて、様々な画面インターフェースを用意して、利用者が必要とする情報、参考となる関連する情報を容易に得られるようにする。

### 自分の知識の外部記憶装置

* 個人によって違った知識の構造の部分については、その人の力によって種々の検索方式を試み、自分の必要とする情報をとり出して中立的な知識の構造に付加してゆくことが出来ねばならないし、またそれによって自分に合った知識の構造を作りあげてゆく
* 自分の書棚を電子的に作れば、自分の連想に基づいて自分の本や本の部分部分をリンク付けして自分の知識の構造に合った自分図書館を作って楽しむことができる

「未来の図書館を作るとは」（長尾真）より抜粋

　現在のAIの技術で、デジタル知識基盤にある共通的な知識から個人のスキル・嗜好に合わせてパーソナライズするとともに、個人が保有する知識を組み合わせて、自分専用のインテリジェンスは外部記憶装置を持つことが可能になる。

スマートフォン等からこの外部記憶装置にアクセスして、①可能な限り「事実情報」を提供する、出来ない場合は、②参考になりそうな文献候補を提示する、③その文献候補の所蔵先へ案内する。④デジタル化された情報があれば、閲覧サービスまでナビゲート、⑤デジタル化された情報がなければ、入手可能なサービスへナビゲートという仕組みは、もう実現可能な段階になっている。

## 「知の共有化」システムの構築業務と、人材の資質とスキル

AIと人間の能力と役割の一般論として、AIが実用段階に達した今、今まで人が担ってきた部分の作業も、精密化するとAIを活用したほうが効率的なことが多々ある。それはアーカイブ機関でも同様であり、マニュアル化、ルーチン化されている業務は、一気にAIに置き換わらなくても、専門職員から、外部委託、非常勤職員、等へシフトされていくものも多い。その中でアーカイブ機関の職員としての仕事を見極める必要がある。

アーカイブ機関において、システムの調達・構築・運用のスキル・知識を持った人材が不足しているのが現状であり、適正な調達を行うために、政府機関での調達の標準ガイドラインに沿って開発タスクとドキュメントをひな形に進めることが効果的である。また、開発タスクを遂行する職員のスキル・知識の習得は、iコンピテンシ・ディクショナリ等を活用することが効率的である。特に、デジタルアーカイブ構築は、ビッグデータや人工知能の活用が必須であり、データサイエンス領域のスキル標準を活用すること有効と考える。

### 政府標準ガイドラインに沿ったシステム構築と運用

「政府情報システムの整備及び管理に関する標準ガイドライン」（政府標準ガイドライン）6）は、業務の効率化及び高度化、情報セキュリティを含む情報システムの運用リスクへの適切な対応等、具体的な取組を政府横断的に進めるため、情報システムの標準的な整備及び管理について、その手続・手順に関する基本的な方針及び事項並びに各組織の役割等を定める体系的な共通のルールとして策定し、政府機関全体でこのルールに沿って運用されている。

これは政府機関における調達の共通ルールであるが、知の共有化を進める公的機関でも適正な調達を行う際の手引きとして参考にし、調達手続き、調達仕様の内容を必要に応じて取捨選択して、システム構築。運用を進めることにより、重要事項の考慮漏れをなくすことができ、発注者と受注者の認識の齟齬を減らすことができる。

### 政府標準ガイドラインに沿った開発タスクとドキュメント

この図は、政府標準ガイドラインに沿って、組織としての事業計画に基づいた、業務・サービスの企画段階から、運用・保守、その後のシステム監査まで様々な業務（タスク）があるが、その各工程でのドキュメントを抜き出したもの。工程ごとに様々な種類の仕様書類があるが、テンプレート的にデフォルトとなる記述内容が示されており、計画書・企画書をベースに、必要な手続き、仕様内容を加筆訂正していく形でドキュメント化していくことが可能である。

アーカイブ機関のデジタルアーカイブ構築においては、特に、要求要件を明確にして、個々の機能要件部分を精緻化していくことに注力する必要がある。

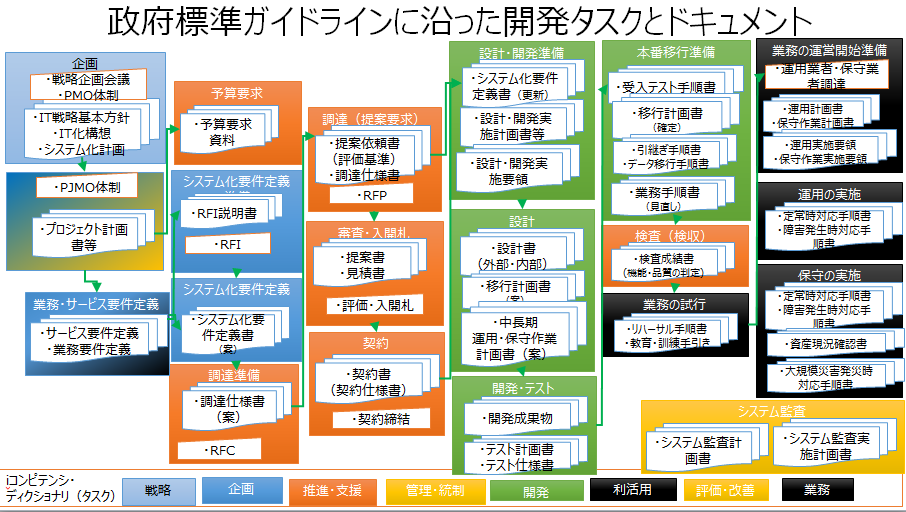


図5 政府標準ガイドラインに沿った開発タスクとドキュメント

### システムで何ができるようにするのか（個別機能要件）

機能に関する事項では、①アーカイブ化機能として、メタデータ、画像データの登録・変更・公開機能、AIを活用したレファレンスに必要な情報の蓄積機能、②サービス提供機能として、利用者に効率的で快適な体験を提供するカスタマーエクスペリエンス（CX）デザイン思考8）でのサービスを実現するユーザインタフェース、内部処理機能、外部サービスとの連携などの要求要件を明確にする必要がある。

　特に、AIを活用できる外部サービスを利用することとして、音声の自然言語による入出力、知識として学習の効率を高めるための教師データの選択、教師データを最適に学習できる機械学習APIを持つAI処理モジュールの選択が重要である。それにより、従来からのレファレンスの回答として、参考になる資料の選択と、その資料の所在情報の提示だけでなく、答えそのものの検索（事実検索）が可能になる。現在もスマートフォン等で利用可能になっているGoogle Assistant、 Apple Siri、Microsoft Cortana、Amazon Alexa、IBM Watson等の知識にデジタルアーカイブ機関が組織化して保有した情報が知識として蓄積されていれば、より的確の事実検索が可能になり、この外部サービスのAPIを利用することにより、アーカイブ機関でのAIを活用したサービスの構築が効率化する。

### どのようなデジタルコンテンツを用意するか（コンテンツ構築要件）

アーカイブ機関がデジタルコンテンツとして構築するコンテンツの要件も明確にして、①二次情報としての蓄積情報・データに関する事項である、永続的識別子、メタデータ、目次・索引データ、関連データ、②一次情報としてのコンテンツそのものである、動画、音声、画像データ、全文テキストデータ、更にIoT等からの収集データ等が他のアーカイブ機関のコンテンツと合わせて、ビッグデータとして利活用しやすいようにするために、より一層の共通的な仕様を適用することにより、AIでの活用を加速させられる。

　著作物の電子書籍化に関しては、原資料からのデジタル化（イメージ化、テキスト化）と、文献の作成段階からデジタル化され電子書籍と印刷書籍が同時並行で進むものがあるが、作成過程及び最終成果物の仕様を標準化することにより、シングルソース・マルチユースが容易になり、利用者の読書環境に応じて様々な形態で提供されることにより、著作物の利活用が促進される。

　現時点でのシングルソースとしてのマスター原稿段階では、章節項、目次、索引、引用等の構造を明示するXSLが適用されたXMLテキストとし、電子書籍化する場合は、様々な閲覧環境に最適な書式とできるように、html5+CSS3をベースとしたEPUB3.1の仕様を適用する。AIシステムの教師データ付き学習データとしては、マスター原稿段階のXMLテキストを利用することにより、構造化された知識として活用が容易になる。例えば、平成28年度情報通信白書（総務省編）は、マスター原稿から、ページレイアウト固定版（PDF版）、スマートフォン・PC・タブレット向けのリフロー版電子書籍（EPUB版）、Webブラウザ版（html版）が用意されている。

### iコンピテンシ・ディクショナリ（iCD）

i コンピテンシディクショナリ（iCD）は、情報処理振興機構（IPA）が作成した、システム開発におけるタスクとスキル・知識を体系的に洗い出したものである。それぞれ、タスクの種類を列挙したものが、「タスクディクショナリ」、スキルの種類を列挙したものが、「スキルディクショナリ」、知識の種類を列挙したものが、「知識ディクショナリ」として辞書化したものである。個々のタスクに必要なスキル、スキルに必要な知識が紐付けられている。また、事業を実施するために様々な業務があり、その業務も複数のタスクにより遂行される。そのタスクの固まりが「タスクプロフィール」として例示的に示されている

各アーカイブ機関が人材育成について検討する際、事業の内容に合わせて、「タスクディクショナリ」からタスクを選択することにより、そのタスクを遂行するために必要なスキル、知識が提示される。タスクを担当する人材は、そのスキル・知識を絞り込んで習得するようにすることにより、短期間に効率的に人材育成ができる。

### 業務遂行のタスクとスキル・知識の蓄積のスキーム

全てのスキルと知識を身に付いていれば、どんなタスクもこなせるが、そんな人材の確保は困難である。

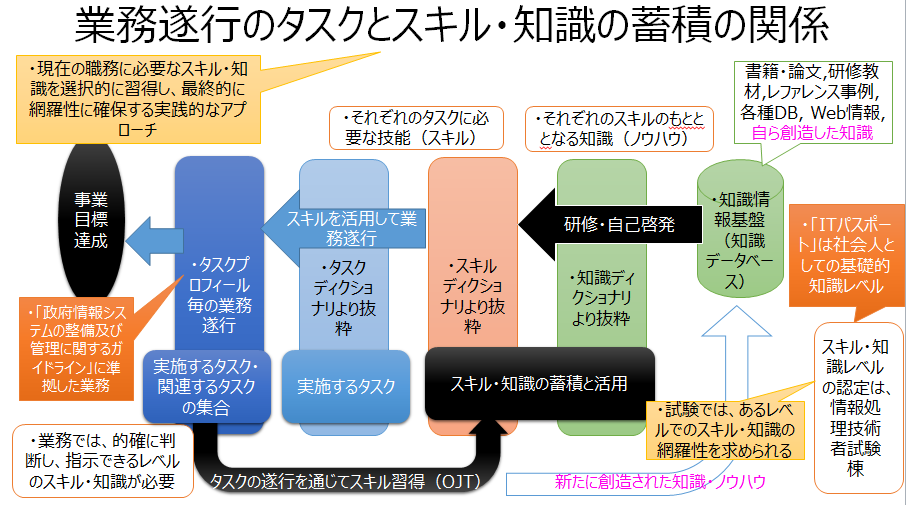


図6 業務遂行のタスクとスキル・知識の蓄積の関係

１つのタスクを実施するためには、様々なスキルが必要であり、１つのスキルは、経験と様々な知識により身に付く。現在の職務に必要なスキル・知識を選択的に習得し、最終的に網羅性に確保する実践的なアプローチが、iコンピテンシ・ディクショナリの考え方である。

業務を遂行するためには、的確に判断し、指示できるレベルのスキル・知識が必要であり、業務に必要なスキル・知識を、事前に選択的に習得して、業務を遂行することが望ましいが、業務を遂行する前に持っていないスキル・知識は、タスクの遂行を通じて習得（OJT）することで、業務を円滑に進められるようになる。

### iコンピテンシ・ディクショナリ（iCD）を活用した業務の遂行とスキル・知識の選択的習得

図7は、政府の標準ガイドラインで示された開発プロセスでの業務をタスクとして、それぞれのタスクの遂行に必要なスキル・知識の選択の概念を示したものである。

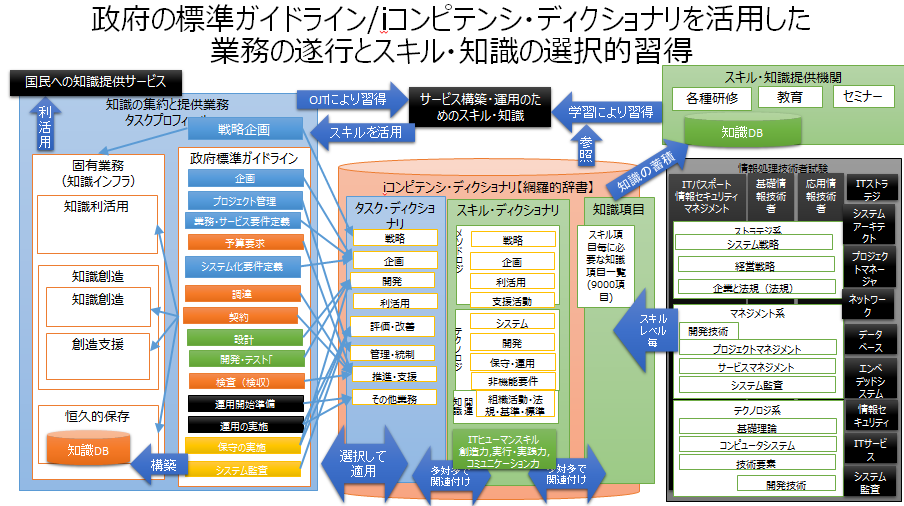


図7 政府標準ガイドライン／iコンピテンシ・ディクショナリを活用した業務の遂行とスキル・知識の選択的習得

更に一定レベルの網羅的な知識の習得状況を評価する情報処理技術者試験との関係も示している。

iCDでは、網羅的なスキルの領域として、①ビジネス活動の様々な手法、方法のスキルとして「メソドロジ」、②IT関連技法などのスキルとして「テクノロジ」、③対象となる業務を進めるための関連知識や業務固有のスキルとして「関連業務知識」、④ IT に関するタスクを実行する際に必要となる実行力・実践力、創造力、コミュニケーション力等のスキルとして「ITヒューマンスキル」とに分類されている。

「④ITヒューマンスキル」は、AIが実用化され、社会の構造が大きく変革しそうな時代において、「価値の創造・問題解決」を着実に遂行する力としての「創造力」は特に重要なスキルである。

また、第4次産業革命に対応した新スキル標準（ITSS+）14）として、「セキュリティ領域」とともに、「データサイエンス領域」に関して、大量データを分析し、その分析結果を活用するための一連のタスクとそのために習得しておくべきスキルカテゴリ、タスク構造が網羅的示されている。業務部門のタスクとして、業務設計、データの作成と収集、構造化データ加工、解析用データ準備、データの準備、データ解析、データ可視化、非構造化データ処理、評価、業務への組み込みと評価の工程が定義されている。



図8 データサイエンス領域のタスク構造

これからのデジタルアーカイブの構築は、ビッグデータやAIを活用が必須であり、各アーカイブ機関の業務部門が中心となって「データサイエンス領域」のタスク工程に沿って確実に進め、業務への適用を評価していくことが重要であり、業務部門での人材育成、人材確保が課題となる。

## まとめ

個々のアーカイブ機関でのデジタル化が進む状況において、文化的資産をあらゆる人々が将来にわたり享受、活用できるようにし、人々の創造的な活用に貢献するためには、個々の知識情報の「見える化」をするだけでなく、全体のそれぞれの施策が同一の方向性を持って、相互に資源を補完し合って、社会全体でより効率的な利活用の保証に取り組む必要があり、デジタル知識基盤は重要な役割を果たすことになる。文化情報資源を中核としたデジタル知識基盤を利活用したサービスの方向性を見極めて、個々のアーカイブ機関は、構成する一機関としてデジタルアーカイブの構築と提供を考えていくことが重要と考える。

今後5～10年のデジタル情報化、AIの実用化レベルの進展を見据えると、アーカイブ機関の業務は根幹から変革が求められると思われる。社会の変革の中でのアーカイブ機関は、従来の延長線上での業務を行っていては、有益なサービスを提供する機関としての存立が危ぶまれる。

AIは人の仕事が奪うよりも、人間がより人間らしい仕事にシフトしていくきっかけでもあり、仕事が奪われるとしたら、むしろ、AIを活用した省力化と新しいサービスに取り組まなかったために事業を継続できなくなる機関・組織であろう。知識創造のための情報の多くがインターネットから入手できる状況において、知識創造の支援の一翼を担う機関としての役割と必要な機能を再検討し、速やかに対応して、「知識インフラ」で目標とした「新たな知識の創造と還流により社会・経済的な価値の創出する仕組みが確立する」が加速され、「未来の図書館」が作られることを期待する。

# 参考文献

1. 兼松芳之，BBCC における電子図書館の研究：けいはんな学研都市フェスティバルより，カレントアウェアネス，CA979，http://current.ndl.go.jp/ca979（accessed 2017-9-1）
2. 田屋裕之．「パイロット電子図書館システム」の現状と課題．出版ニュース．1996，96 1/ 上・中，p.14-17
3. 中山正樹．国立国会図書館電子図書館中期計画2004の実施に向けて，情報の科学と技術，2004，Vol.54，No.9，p.453-460．
4. 久古聡美・吉田曉・中山正樹．Web2.0世代のデジタルアーカイブポータルの提供を目指して．情報管理．2006，Vol.49，No.6，p.313-323．
5. 長尾真，未来の図書館を作るとは．長尾真，2012，27p
6. 政府情報システムの整備及び管理に関する標準ガイドライン．各府省情報化統括責任者（CIO）連絡会議，2014，<http://www.soumu.go.jp/main_sosiki/gyoukan/kanri/infosystem-guide.html>
7. ｉコンピテンシ・ディクショナリ，情報処理振興機構，2015，https://www.ipa.go.jp/jinzai/hrd/i\_competency\_dictionary/icd.html（accessed 2017-9-1）
8. ベイカレント・コンサルティング．デジタルトランスフォーメーション．日経BP 社．2016．
9. 長尾真．「電子図書館」．岩波書店．1994年
10. 内閣府科学技術政策・イノベーション担当．“第４期科学技術基本計画（平成23年８月19日閣議決定）”．内閣府．<http://www8.cao.go.jp/cstp/kihonkeikaku/4honbun.pdf>　(accessed 2017-9-1）．
11. 中山正樹．国立国会図書館におけるデジタルアーカイブ構築―知の共有化を目指して―．情報管理．2012，Vol.54，No.11，p.715-724．
12. 中山正樹．国立国会図書館のサービスシステムの歩みと新たな方向性の模索―電子図書館事業20年の歩み―．国立国会図書館月報．2015，Vol.648，p.18-24
13. 中山正樹，電子図書館サービスからナショナルアーカイブの構築へ― LOD 化によるデジタル文化財の利活用を目指して―，TP&D フォーラム実行委員会，2016．25巻．p.13-33
14. データサイエンス領域のスキル標準「ITSS＋」，情報処理振興機構，2017，https://www.ipa.go.jp/jinzai/itss/itssplus.html（accessed 2017-9-1）

～～～～～～～～～～～～～～～～～～

1. http://www.ndl.go.jp/jp/aboutus/elib\_plan2004.html [↑](#endnote-ref-2)
2. RSSとは、Rich Site SummaryもしくはRDF Site Summaryの略。Webサイトの見出しや要約などのメタデータを構造化して記述するXMLベースのフォーマットで主にサイトの更新情報を公開するのに使われている。 [↑](#endnote-ref-3)
3. OAI-PMHとは、Open Archives Initiative Protocol for Metadata Harvestingの略。OAI（Open Archives Initiative）が策定した、ウェブ上のメタデータを収集（メタデータ・ハーベスティング）するためのプロトコル [↑](#endnote-ref-4)
4. SRWとは、Search/Retrieve Web Serviceの略。データの検索およびデータの取得に使用することを想定したWebサービスのインタフェース仕様 [↑](#endnote-ref-5)
5. CMSとは、コンテンツマネージメントシステム。Webコンテンツの作成、更新及び管理を行うシステムで、トップページも最新情報を含めて動的なページとして容易に配信することができる。 [↑](#endnote-ref-6)
6. XOOPSとは、コンテンツマネジメントシステムの１つで、オープンソースライセンスに基づいたソフト。コミュニティサイトを容易に構築できる。 [↑](#endnote-ref-7)
7. GETA(Generic Engine for Transposable Association)とは、連想検索のエンジンで、入力した文章から「連想」されるような、文章と関連性の高いと思われるコンテンツを検索するもの [↑](#endnote-ref-8)
8. OAIS（Open Archival Information System）とは、情報を長期保存するために必要とされる機能要素と、そこで扱う情報の構成の規定がなされた技術標準 [↑](#endnote-ref-9)
9. DSpaceとは、マサチューセッツ工科大学図書館とヒューレット・パッカード研究所が共同開発した,OAIS参照モデルに準拠したデジタル情報保存システム [↑](#endnote-ref-10)
10. DIAS(Digital Information Archiving System)とは、オランダ国立図書館とオランダIBMが共同開発した、OAIS参照モデルに準拠したデジタル情報保存システム [↑](#endnote-ref-11)
11. JPEG2000は、画像圧縮方式の一つ。国際標準規格であるJPEGを発展させた形で、高圧縮かつ高品質なデータ圧縮が可能 [↑](#endnote-ref-12)
12. [↑](#endnote-ref-13)
13. Next-L Enjuとは、大学等の研究者が開発したオープンソースの図書館システム。 [↑](#footnote-ref-1)
14. GETAssocとは、国立情報学研究所（NII）連想情報学研究開発センターが開発した連想検索エンジン。 [↑](#footnote-ref-2)
15. 参考文献

    1. 中山正樹：国立国会図書館におけるデジタルアーカイブ構築、情報管理Vol.54,No.11, pp.715-724(2012).

    [↑](#endnote-ref-14)
16. 1. 知的財産戦略本部：知的財産推進計画2012．知的財産戦略本部会合議事次第、http://www.kantei.go.jp/jp/singi/titeki2/120529/gijisidai.html (2012)

    [↑](#endnote-ref-15)
17. 1. 高度情報通信ネットワーク社会推進戦略本部：電子行政オープンデータ戦略、http://www.kantei.go.jp/jp/singi/it2/pdf/120704\_siryou2.pdf(2012)．

    [↑](#endnote-ref-16)
18. 1. 総合科学技術会議：第５回科学技術イノベーション政策推進専門調査会議事次第、http://www8.cao.go.jp/cstp/tyousakai/innovation/5kai/index.html(2012)．

    [↑](#endnote-ref-17)
19. Digital Repository Audit Method Based on Risk Assessment , http://www.repositoryaudit.eu/

    （参考）http://current.ndl.go.jp/ca1681 [↑](#footnote-ref-3)
20. Life Cycle Information for E-Literature, http://www.life.ac.uk/,

    （参考）http://current.ndl.go.jp/files/ca/ca1696.pdf [↑](#footnote-ref-4)
21. 書誌レコードの機能要件（Functional Requirements for Bibliographic Records: FRBR）概念モデルの"item"に相当する情報。物理的な資料一点ごとに関する所蔵情報等。 [↑](#footnote-ref-5)
22. http://www.kantei.go.jp/jp/singi/it2/pdf/120704\_siryou2.pdf

    基本原則は、次の4つである。①政府自ら積極的に公共データを公開すること②機械判読可能な形式で公開すること③営利目的、非営利目的を問わず活用を促進すること④取組可能な公共データから速やかに公開等の具体的な取組に着手し、成果を確実に蓄積していくこと [↑](#footnote-ref-6)
23. 主な取組機関として、総務省等をはじめとする官民連携で設立された「オープンデータ流通推進コンソーシアム」等があげられる。 [↑](#footnote-ref-7)
24. 個々のデータや概念に対し一意の識別子であるURIを与え、コンピュータ処理に適した形式で公開し、相互に関連づけることで、ウェブ上でのデータの共有・利用を促進するための技術の総称。これらの技術で作成され、法的な制約がなく、いかなる目的においても自由に利用可能なデータの集合を呼称することもある。 [↑](#footnote-ref-8)
25. ここでいう「フレームワーク」とは、メタデータの記録、流通及び交換等のための「容れ物」である。 [↑](#footnote-ref-9)
26. http://www.loc.gov/marc/transition/ [↑](#footnote-ref-10)
27. 平成24年10月1日にVIAFへの参加についての協定（平成24年国図収1209181号）を締結した。 [↑](#footnote-ref-11)
28. National bibliographies in the digital age : guidance and new directions / IFLA Working Group on Guidelines for National Bibliographies ; edited by Maja Žumer. -- München : K.G. Saur, 2009. -- (IFLA series on bibliographic control ; v. 39) -- ISBN 978-3-598-24287-8.

    日本語訳は右記参照。<http://www.ndl.go.jp/jp/library/data/nbda\_jp.pdf> [↑](#footnote-ref-12)