

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

国立国会図書館総務部情報システム課
中山正樹

東京外国語大学国際日本研究センター
比較日本文化部門・国際連携推進部門共催国際シンポジウム
e-Japanologyの構築に向けて
2010年12月11日

今日のお話し

● 背景

- 「知識はわれらを豊かにする」
- 当館の位置付けの再確認

● 電子図書館構想

- 電子図書館のあゆみ
- 当館資料のデジタル化
- デジタルコレクション
- 所蔵資料のデジタル化実施状況
- 学位論文のデジタル化

● 情報の利活用に関する連携協力

- デジタルアーカイブ関連での連携協力
- 出版物の利活用の推進に関する懇談会（三省懇）
- 「新ICT利活用サービス創出支援事業」
- デジタルアーカイブのポータル

● 学術デジタルコミュニケーションの現状

- 知識インフラの必要性
- 知識インフラの構築に向けて
- 知識インフラの課題と当館の役割
- 近い将来に取り組むべき事項
- 当館が取り組んできたこと

● 知識インフラのノードとしての社会的な機能の展開

- NDLSearchとは
 - NDL Searchが当面目指す方向性
 - NDLSearchの将来像
 - 24年1月のサービスイメージ
 - 22年8月公開の開発版の到達点
 - 22年度機能強化
 - NDL Searchでの連携協カイメージ
 - 知識の利活用の促進を目指して
- 終わりに

「知識はわれらを豊かにする」を デジタルアーカイブ・情報探索サービスの観点からみて

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)

(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します

収集・保存すること

(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します

(4) 利用者がどこにいても、来館者と同様のサービスが受けられるように努めます

提供可能にすること

情報資源

関係機関と協力して収集・提供

(1) 国会に対するサービスをより高度なものとし、立法補佐機能をさらに強化します

(5) 社会に多様で魅力的なサービスを提供し、国立国会図書館の認知度を高めます

(6) 公共図書館をはじめとする国内の各図書館とより密接な連携・協力を進めます

(7) 海外の図書館との密接な連携を行い、情報の共有・交換に努めます

冊子体資料



収集

情報

提供



関係機関と分担して、情報資源を知識として集積し提供

物の流通から情報の流通に軸足が移行する時代

電子図書館構想

● 電子図書館構想

- 電子図書館のあゆみ
- 当館資料のデジタル化
- デジタルコレクション
- 所蔵資料のデジタル化実施状況
- 学位論文のデジタル化

● 情報の利活用に関する連携協力

- デジタルアーカイブ関連での連携協力
- 出版物の利活用の推進に関する懇談会（三省懇）
- 「新ICT利活用サービス創出支援事業」
- デジタルアーカイブのポータル

電子図書館構築のあゆみ

第1ステージ【1994～1998】 - 揺籃期 -

電子図書館の実施に向けた計画策定と実証実験

- 1995年 パイロット電子図書館プロジェクトを実施
- 1997年 電子図書館推進会議報告書
- 1998年「電子図書館構想」の策定
- G7電子図書館プロジェクトへの協力

第2ステージ【1998～2002】 - 始動期 -

組織整備、コンテンツ作成開始、システム構築

- 1999年には電子図書館推進室を設置し、「電子図書館サービス実施基本計画」を策定。
- 2000年3月に、国立国会図書館ホームページにおいて「国会会議録」、「貴重書画像データベース」、「国立国会図書館蔵書目録」、電子展示会などの電子図書館のメニューを公開。
- 2000年5月に部分開館した国際子ども図書館でも、「絵本ギャラリー」など具体的なサービスを始める。

電子図書館構築のあゆみ

第3ステージ【2002～2009】 - サービス離陸期

● 専管的な部署の設置

- 関西館が2002年に開館し、本格的な電子図書館事業を開始。

● 各種サービスの提供

- 「近代デジタルライブラリー」（明治期以降の出版物のデジタル画像データベース）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
- レファレンス協同データベース
- 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開始
 - PORTA（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システム）に着手

第4ステージ【2009～】 - 新たな飛躍へ

● 国内外における各種連携・協力

- ⇒全国の公共図書館との連携の強化
- ⇒博物館や文書館などの文化機関との連携の強化
- ⇒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
- ⇒世界各国とのグローバルな協力の推進

● コンテンツの質・量・種類の拡充

- ⇒テキスト情報
- ⇒音楽・動画等のマルチメディア情報





- これらを実施するため、組織の改編とシステムの整備を予定

当館資料のデジタル化

- インターネットを通じていつでも、どこでも読める「電子図書館の蔵書」として構築。
 - そのため、著作権保護期間満了、著作権者の許諾を得る、文化庁長官裁定を受けて、公衆送信を可能としてきた。
- 2009年から保存目的でのデジタル化にも着手。
 - 2010年の1月の著作権法改正では、国立国会図書館が資料保存の目的でデジタル化することを、法律上に明確した。
 - 2009年に国の経済対策の一環として127億円（1.5億ドル）の補正予算が計上され、また2010年にも10億円（1200万ドル）が計上される。
- 出版者、著作権者等で構成する関係者協議会で提供範囲について、協議を重ねている。

国立国会図書館のデジタルコレクション

(平成23年3月末想定)

資料種別	所蔵数 (H21年度末) (A)	デジタル化実施済*1 (B)	デジタル化未実施 (A-B)	実施割合 (B/A)
 古典籍	29万冊	6万冊	23万冊	1/5
 和図書	411万冊	89万冊	322万冊	1/5
 和雑誌	436万冊	121万冊	315万冊	1/4
 博士論文	39万冊*2	14万冊	25万冊	1/3
合計	915万冊	230万冊 (画像2億ファイル)	685万冊	1/4

*1 : 平成22年度第1次及び第2次調達 (平成23年3月納品予定) を含む。

デジタル化実施済刊行年代は次のとおり。

【古典籍】江戸期以前

【和図書】明治期～1968年刊行

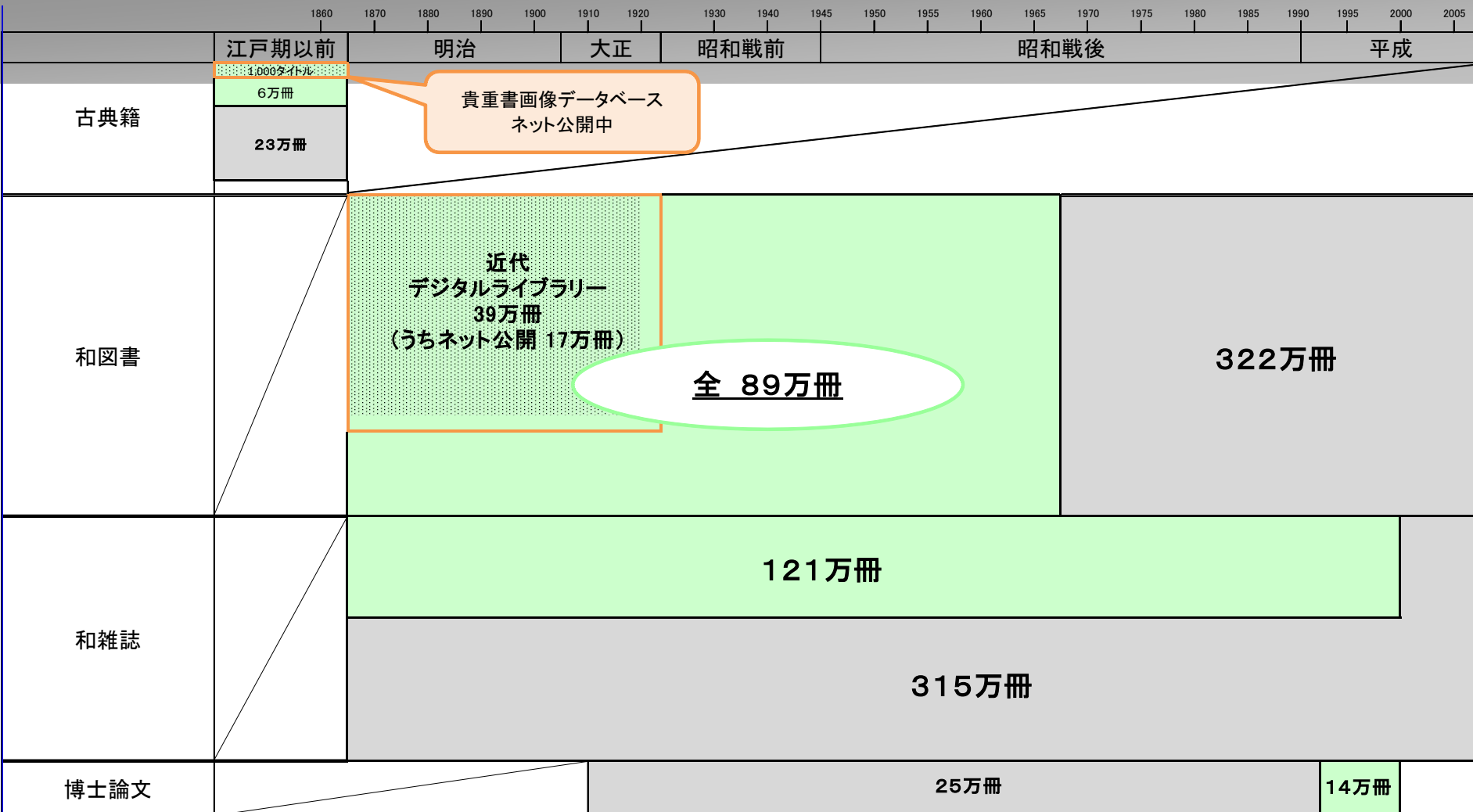
【和雑誌】明治期～2000年刊行 (商業出版との調整タイトル等を除く。)

【博士論文】平成3 (1991) 年度～平成12 (2000) 年度受入れ

*2 : 平成21年度までの所蔵数から平成13 (2001) 年度～平成21 (2009) 年度整理数を除いた数。

所蔵資料のデジタル化の実施状況

(平成23年3月末想定)



貴重書画像データベース
ネット公開中

近代
デジタルライブラリー
39万冊
(うちネット公開 17万冊)
全 89万冊

■ デジタル化実施済み(平成22年度第1次及び第2次調達分を含む。)の範囲

■ デジタル化未実施の範囲

2010/12/11

■ インターネット公開済分

■ 館内公開済分

* 冊数は平成21年度末所蔵数より。

学位論文のデジタル化

● 背景

- 国立国会図書館と大学図書館との連絡会議（平成18年2月）
 - 共通する課題について、政策的及び実務的な面から問題を協議
- 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（平成20年3月）
 - 学位論文の保存・蓄積及び利用・提供に係る考え方、役割分担の枠組み、メタデータの標準化と相互運用、制度面における課題の整理及び取組みの方策について検討
 - 過去分（学位授与日が基準日以前であるもの）は、国立国会図書館が電子化及び保存を行う
 - 学位論文の電子的利用に必要な著作権許諾を得るために大学及び国立国会図書館が許諾書の統一的書式等のガイドラインを作成する
- 著作権法の改正
 - 資料の保存を目的とする国立国会図書館所蔵資料のデジタル化

学位論文のデジタル化及び著作権処理

● 基本方針

- 国公立大学図書館協力委員会との確認事項

● デジタル化の対象範囲・実施方法

- 1991～2000年度に国立国会図書館が受け入れた学位論文
- 画像データ

● 著作権処理

- 学位授与大学の学長（または図書館長）及び国立国会図書館長が共同して、学位論文の著作者に当該著作者の学位論文のデジタル化（複製）、デジタル化した学位論文の譲渡及び公衆送信の許諾を依頼
- 当該著者が単一の許諾書によって、許諾すること（共通許諾）を基本とする

● デジタル化・著作権処理は、国立国会図書館が実施する。

● 共通許諾書の内容

- 改正著作権法に基づいてデジタル化した学位論文を広く利用（全文複写提供、公衆送信）に供すること
- 国立国会図書館がデジタル化した学位論文を複製して学位授与大学に譲渡すること
- その学位論文を学位授与大学が利用（全文複写提供、公衆送信）に供すること
- 今回の共通許諾の枠組みの適用範囲は、平成22年度デジタル化

デジタル・ネットワーク社会における 出版物の利活用の推進に関する懇談会

1 背景・目的

デジタル・ネットワーク社会に対応した知の拡大再生産を実現し、我が国の豊かな出版文化を次代へ着実に継承しつつ、広く国民が出版物にアクセスできる環境を整備することが重要な課題となっている。

そのため、関係者が広く集まり、デジタル・ネットワーク社会における出版物の利活用の推進に向けた検討を行う懇談会（総務省、文部科学省、経済産業省の副大臣・大臣政務官の共同懇談会）を開催する。

2 検討内容

- 1 デジタル・ネットワーク社会における出版物の収集・保存の在り方
- 2 デジタル・ネットワーク社会における出版物の円滑な利活用の在り方
- 3 国民の誰もが出版物にアクセスできる環境の整備 等

3 運用

懇談会の下に、技術に関するワーキングチーム、利活用の在り方に関するワーキングチームを設置し、検討。

4 開催期間

平成22年3月17日に第1回会合を開催。6月28日に一定の取りまとめとして、懇談会報告を公表。

「新ICT利活用サービス創出支援事業」 (総務省)

(電子出版の環境整備)

日本語基本フォーマットの確立

- ①国内ファイルフォーマット(中間(交換)フォーマット)の共通化に向けた環境整備
(「電子出版日本語フォーマット統一規格会議(仮称)」の設置・運営を含む。)

検索技術の確立

- ②書誌情報(MARC等)フォーマットの確立に向けた環境整備
(「電子出版書誌データフォーマット標準化会議(仮称)」の設置・運営を含む。)
- ③メタデータの相互運用性の確保に向けた環境整備
- ④記事、目次等の単位で細分化されたコンテンツ配信等の実現に向けた環境整備

障がい者・高齢者等の利用促進

- ⑤電子出版のアクセシビリティの確保

知のインフラへのアクセス環境の整備

- ⑥書店を通じた電子出版と紙の出版物のシナジー効果の発揮
- ⑦その他電子出版の制作・流通の促進に向けた環境整備

デジタルアーカイブ関連での連携協力

● 国内連携

- 国立情報学研究所（NII）、科学技術推進機構（JST）
- 国立公文書館、国立美術館、東京国立博物館、人間文化研究機構
- MLA連携
 - デジタル情報資源ラウンドテーブル」を設置

● 国際連携

- 日中韓電子図書館イニシアティブ
 - 2010年8月、日本と中国、韓国の3つの国立図書館は、デジタルアーカイブを連携して進めていくことに合意。
 - メタデータスキーマの標準化、情報サービスの相互運用性確保及び電子情報への長期アクセスの保証
- ワールドデジタルライブラリー
 - LCとユネスコが中心となって実施しているワールドデジタルライブラリーに対しても、国立国会図書館は協力。特に翻訳と言語の委員会で協力する。

デジタルアーカイブのポータル

● PORTA（国立国会図書館デジタルアーカイブポータル）

- Google等の検索エンジンはノイズが多く、多くのデータは深層ウェブの情報として存在するので、探せない。
- 国立国会図書館では、日本のデジタル情報資源を案内し、一元的に検索可能とするポータルシステムを開発し、提供。
- 当館のデジタル情報資源の他、公文書、博物館資料、美術作品等、広範な分野のデジタルコンテンツが利用可能。
- 162種類のデジタルアーカイブ、コンテンツ総数は約2800万件である。
- 簡易検索の他、詳細検索、連想検索、分類検索等、複数の検索手法を用意。デザインや検索方法のカスタマイズが可能。

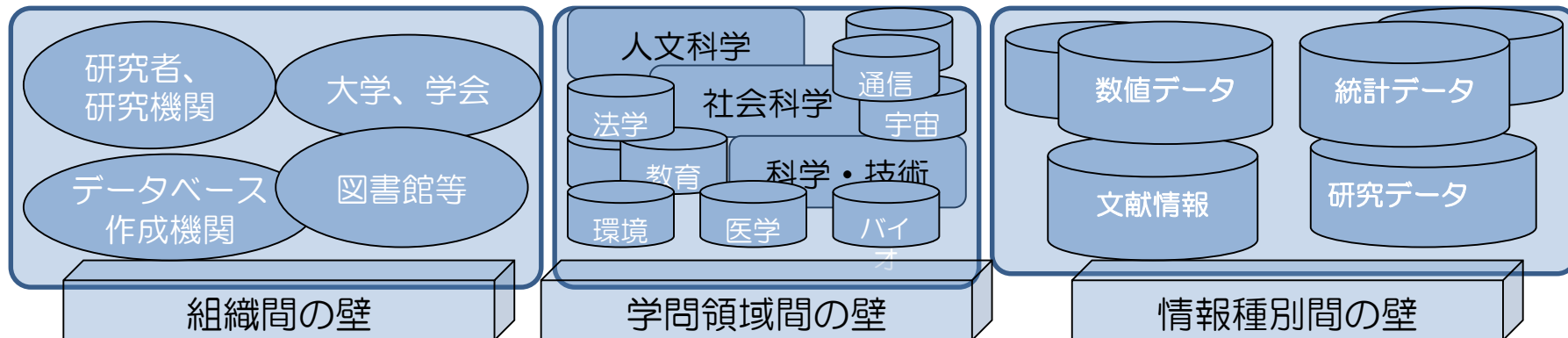
学術デジタルコミュニケーション

-知識インフラの構築に向けて-

- 知識インフラの必要性
- 知識インフラの構築に向けて
- 知識インフラの課題と当館の役割
- 近い将来に取り組むべき事項
- 当館が取り組んできたこと

知識インフラの必要性

■ 科学技術情報流通の「壁」 ～現状～



■ 「科学技術基本政策策定の基本方針」* (デジタル時代の科学技術情報の流通)

平成22年6月 総合科学技術会議基本政策専門調査会

● デジタル情報資源のネットワーク化、データの標準化を進める。

● 学問領域横断的な統合検索、構造化、知識抽出の自動化に向けた研究開発を国全体として推進する。

● 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。

知識インフラの課題と当館の役割

- 知識インフラ構築における課題
 - － 電子情報資源を取り込むためには、継続的な研究開発が必要
 - － 研究プロセスで生じる研究データや中間成果物をも対象として提供することが必要
 - － 現存する個別のデータベース及びシステムを有機的に関連させることが必要
- 知識インフラ構築における当館の役割
 - － 知識インフラ構築に積極的に関与して、その「ノード」としての役割を果たす
 - － 出発点は、関係する機関と協議の場を形成すること
 - － 「知識インフラ」のサイクルに組み込む体制についても継続的研究を続ける
 - － 各国の国立図書館等との連携を進め、学術情報の国際的なオープンアクセス推進

近い将来に取り組むべき事項

- 国内学術出版物のデジタル化と電子情報資源の収集
- デジタル化のための環境整備
- 電子情報資源の管理・保存
- 電子情報資源の利活用の促進
- 従来在所蔵資料・サービスと電子情報資源との有機的連携
- 利用情報の解析と利活用
- 知識インフラのノードとしての社会的な機能の展開

近い将来に取り組むべき事項No.1

国内学術出版物のデジタル化と電子情報資源の収集

- 印刷物のデジタル化を推進し、国民各層の求めに応じた提供
- 電子情報資源に関する収集と保存
- 電子情報資源へ永続的にアクセスしやすい状況を早期に実現
- 出版社、学会、大学等の関連機関の動向に十分留意した上で、持続可能な体制を整備

デジタル化のための環境整備

- 電子情報資源の書誌データ基準の普及、電子情報資源に関する識別子の付与・登録機関の設立推進等を行う
- 国レベルの協議の場の形成に向け、関係機関と協議を進める

電子情報資源の管理・保存

- 電子情報資源についても、「最後の拠り所」として広く国民からのアクセスに応えるために、長期に管理、保存するシステムを構築
- 研究開発や応用、実践に関して、常に最新状況を把握
- 海外の国立図書館等長期保存に取り組む機関との情報共有や、共同研究や調査の連携を実施

近い将来に取り組むべき事項No.2

電子情報資源の利活用の促進

- 個別システムやデータベースへのリンクや一括して検索を行うシステム
- 科学技術振興機構、国立情報学研究所等関連機関との棲み分けと連携
- 政府が保有する各種統計データを中心とした電子情報資源へのナビゲーションに優先的に取り組む
- 保有する各種電子情報資源のAPI 提供等を推進

従来所蔵資料・サービスと電子情報資源との有機的連携

- 従来の印刷物を中心とする所蔵資料と電子情報資源との有機的な連携を図り、利用者が資源の種別に関係なく一括して検索でき、その違いを意識することなくシームレスに情報本体へと案内することを目指す。
- 情報検索に習熟していない利用者への支援を進める
- より複雑な情報ニーズを持つ利用者に対しては、そのニーズを的確に判断し、国立国会図書館の所蔵にとどまらず、また、媒体を問わず、適切な資料・情報へと導くレファレンスサービス等の整備を行う。

近い将来に取り組むべき事項No.3

利用情報の解析と利活用

- 検索語、利用された資料名とその頻度、利用した資料・サービスの経路といった利用動向の把握が容易
- 特定の主題分野で利用頻度の多い資料や類似の資料に関する情報を提示
- 個人情報を除く等の統計上の加工を行い利用情報として活用及び提供が可能か、検討

知識インフラのノードとしての社会的な機能の展開

- 国会議員と研究者コミュニティをつなぎ、最新の科学技術の動向を政策立案に組み込むなどの立法府の活動を補佐することを通じて、最終的には科学技術の成果の国民への還元に寄与する。
- 国民が学術情報へ容易にアクセスできることを保障することで、サイエンスコミュニケーションの基盤形成に寄与する。

当館が取り組んできたこと

電子情報資源収集のための法改正を含む制度整備

- 許諾を得た国内のウェブサイトを選択的に収集
- 国立国会図書館法改正により、平成22年4月から新たに、国等の機関のウェブサイトを許諾なしに収集することを開始
- 平成17年度末の3.1TBから平成21年度末の14.7TBに増加
- 平成22年8月末で総提供量は4,661サイト、18TB

所蔵資料のデジタル化の進展

- 「近代デジタルライブラリー」約29万タイトル（39万冊）このうち、科学技術分野は約2割
- 国内刊行図書、国内刊行雑誌、古典籍資料等所蔵資料の大規模デジタル化
- 平成21年度には、国立国会図書館が保存のために所蔵資料のデジタル化を著作権者の許諾なしに行うことを可能とする著作権法の改正

書誌データの整備

- OCLCを通じて図書に関する書誌データの提供を開始
- 440万件の図書に関する基本的な情報が全世界からアクセスできる

情報探索サービス（NDL Search）の構築・運用

知識インフラのノードとしての 社会的な機能の展開

- NDLSearchとは
- NDL Searchが当面目指す方向性
- NDLSearchの将来像
- 24年1月のサービスイメージ
- 22年8月公開の開発版の到達点
- 22年度機能強化
- NDL Searchでの連携協力イメージ
- 知識の利活用の促進を目指して

NDL Search(国立国会図書館サーチ) とは

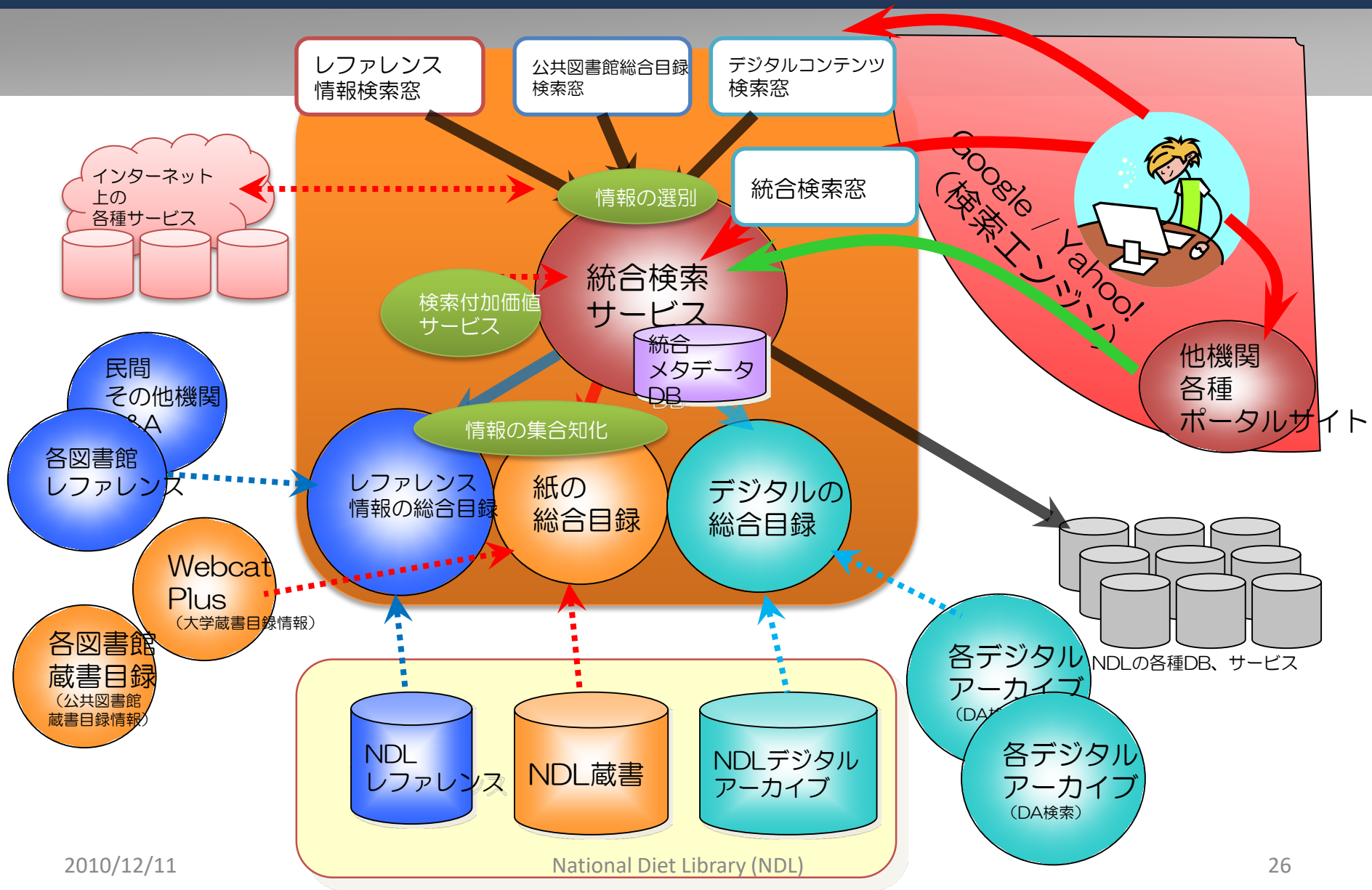
● 目的

- 「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」

● 概要

- PORTAの後継システムを開発中。
- 今年8月に公開し、2012年からは本格システムとして稼働予定。
- 試行運用の段階であるが、34個のデータベースから収集した約5,500万件を検索でき、9種のお機関の大規模データベースと横断検索できる。
- PORTAとの最大の相違点は、紙とデジタルを統合的に検索できること。
- 「日韓・日中・日英翻訳機能」を追加するなど、機能拡張を図っている。
- ユーザの評価等に基づく改善を重ね、我が国を代表するデジタル情報検索の玄関となるよう取り組む。

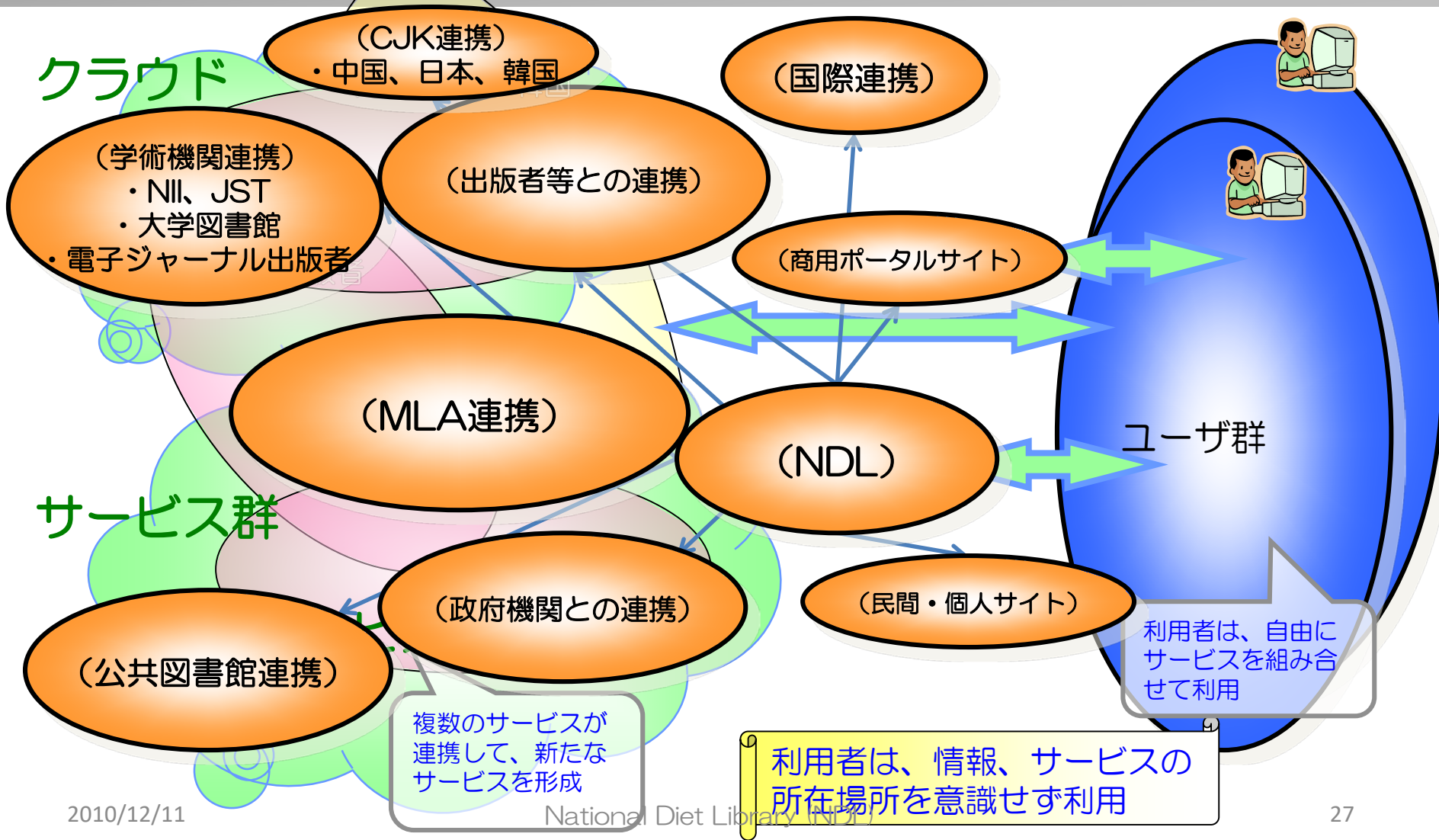
NDL Searchが当面目指す方向性



NDL Searchの将来像

- 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、中核的なサービスプロバイダとなる

(クラウドの世界でのサービスの連携)



平成24年1月のサービスイメージ

- 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。
- 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）
- 情報の再利用（APIでの利用）
 - NDL Search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）
 - NDL Search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）
- サービス・機能の再利用
 - NDL Search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）
 - NDL Search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）
 - NDL Search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）

平成22年8月公開の開発版の到達点

- 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。
 - 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。

平成22年度機能強化

■ 組織化(Converter)

- メタデータへの読みの自動生成・付与
- 書誌同定：精度向上（同定キーの見直し、正規化等）
- グループ化：精度向上（グループ化キーの見直し、正規化等）

■ 検索インデックスの見直し

- 検索コンテキストを考慮したランキングロジックの実装
- 類義語辞書の追加
 - ✓ カタカナ語、外来語などの基本的な類義語、件名典拠、人名典拠
- 日本語のため処理ロジック追加
- スニペットとハイライト機能
- クラスタリング機能
 - ⇒ 検索結果の書誌データから特徴的な単語を使った自動facetデータを提示

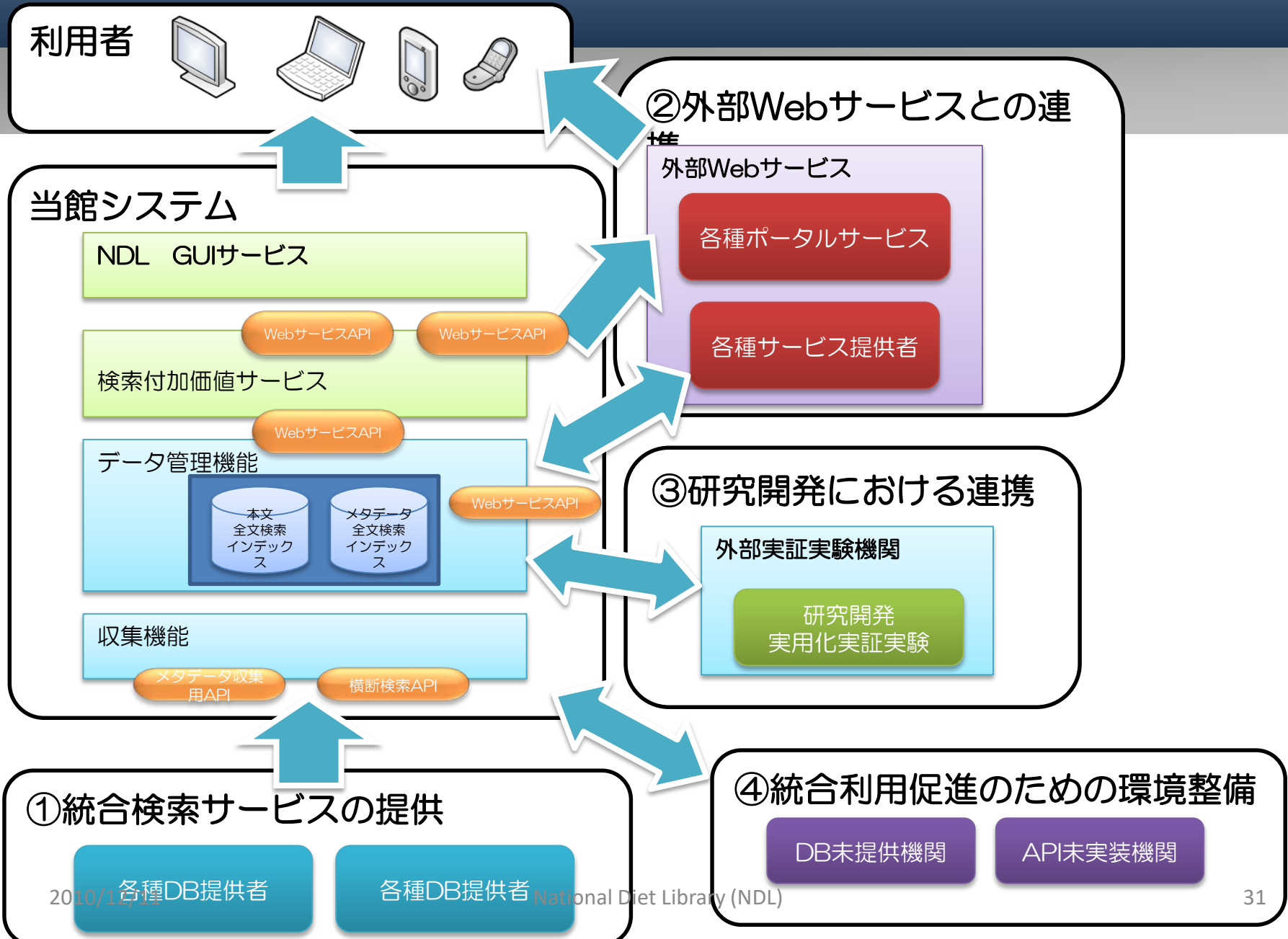
■ 集合知関連

- キーワードサジェスト：
入力されている途中の文字列を元に、関連が深いキーワードを列挙
- 再検索キーワード・サジェスト：
検索結果や0件ヒット画面にて、新たな検索軸を使って再検索するために利用できるキーワードを列挙
- キーワード・レコメンド：書誌に関連する検索用キーワードを提示
- 書誌・レコメンド：利用者の行動履歴を元にして、書誌との関連が深い書誌を提示
- 利用者プロフィールによる検索ランキング
- 検索結果でのクラスタリング

■ 性能改善

- 書誌同定、グルーピングの効率化
- 横断検索の効率化、適正化

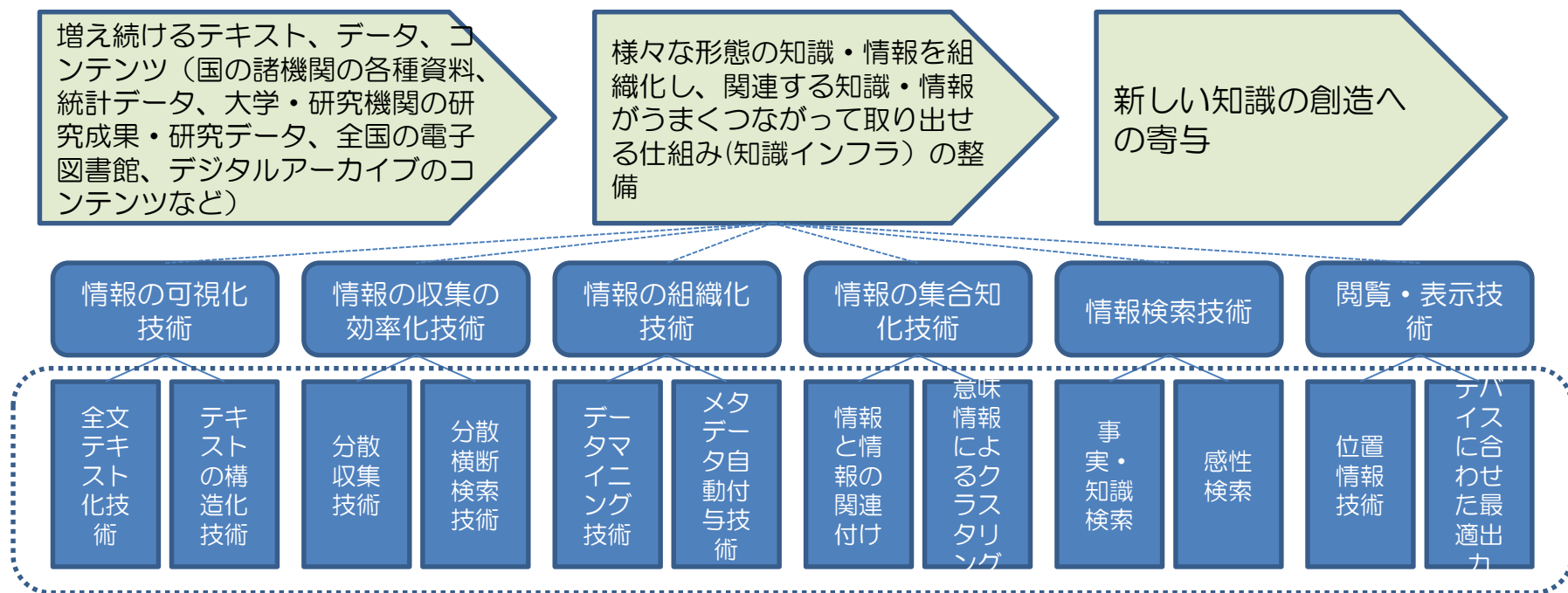
NDL Searchでの連携協力のイメージ



-知識の利活用の促進を目指して-

次世代図書館システムが目指す方向性

図書館資料のデジタル化の進展を踏まえ、次の世代の図書館システムは、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。



**知識の利活用の促進を目指した「知識インフラ」の構築
という目標達成の一翼を担う**

終わりに

膨大な情報を知識として活用して、新たな知識の
創造をめざす
→知識はわれらを豊かにする」の実現手段の一つ

国立国会図書館: <http://www.ndl.go.jp>
e-mail: m-nakaya@ndl.go.jp

ご静聴ありがとうございました