

2016年5月4日

# デジタルアーカイブの構築の歩み 【詳細】

同志社大学大学院総合政策科学研究科  
嘱託講師 中山正樹

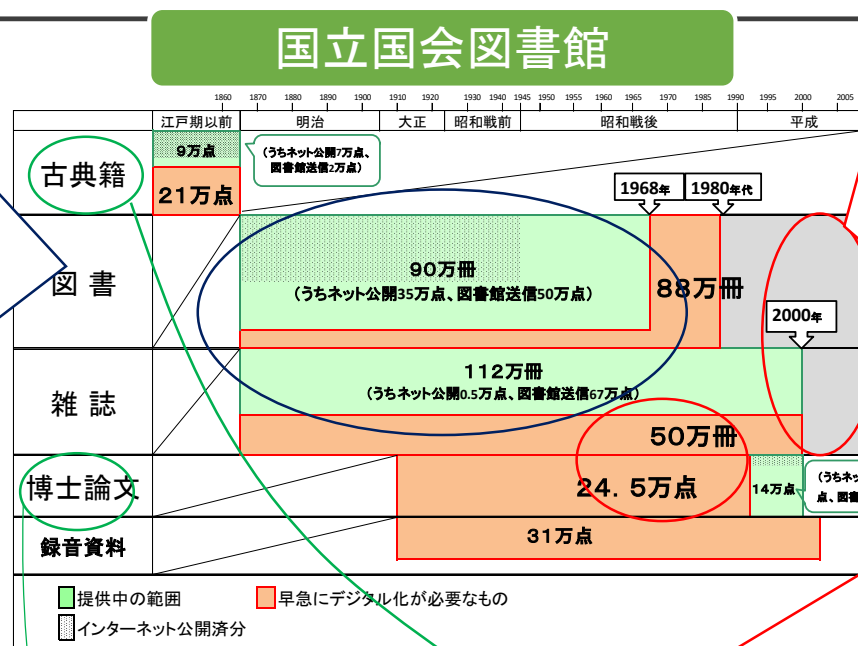
# デジタルアーカイブ構築【概要】

# ① 資料デジタル化

- ・【現状】
- ・246.5万点（図緑色）のデジタル化を実施
- ・48万点をインターネット提供、131万点（絶版等資料）を図書館等へ送信
- ・復刻・翻刻目的に限ったデジタル化データの試行提供

## 【利活用の課題】

- ・学協会等の学術情報のインターネット公開の拡大
- ・デジタル化データの二次利用の促進＝教育目的、出版目的←孤児著作物の権利処理
- ・全文テキスト化とその活用（検索目的、視覚障害者等の利用目的）



## 【デジタル化の課題①】

- ・現行NDLデジタル化方針における対象は赤色部分。デジタル化の優先順位の見直し
- 国内刊行欧文誌
- 官庁出版物
- (古い)博士論文
- (新しい)学協会出版物/紀要類

## 【デジタル化の課題②】

- ・NDL以外の図書館等のデジタル化の促進
- ・他の図書館等が作成したデジタル化データのNDLからの送信

## 公共図書館

- ・地域資料のデジタル化

## 大学図書館

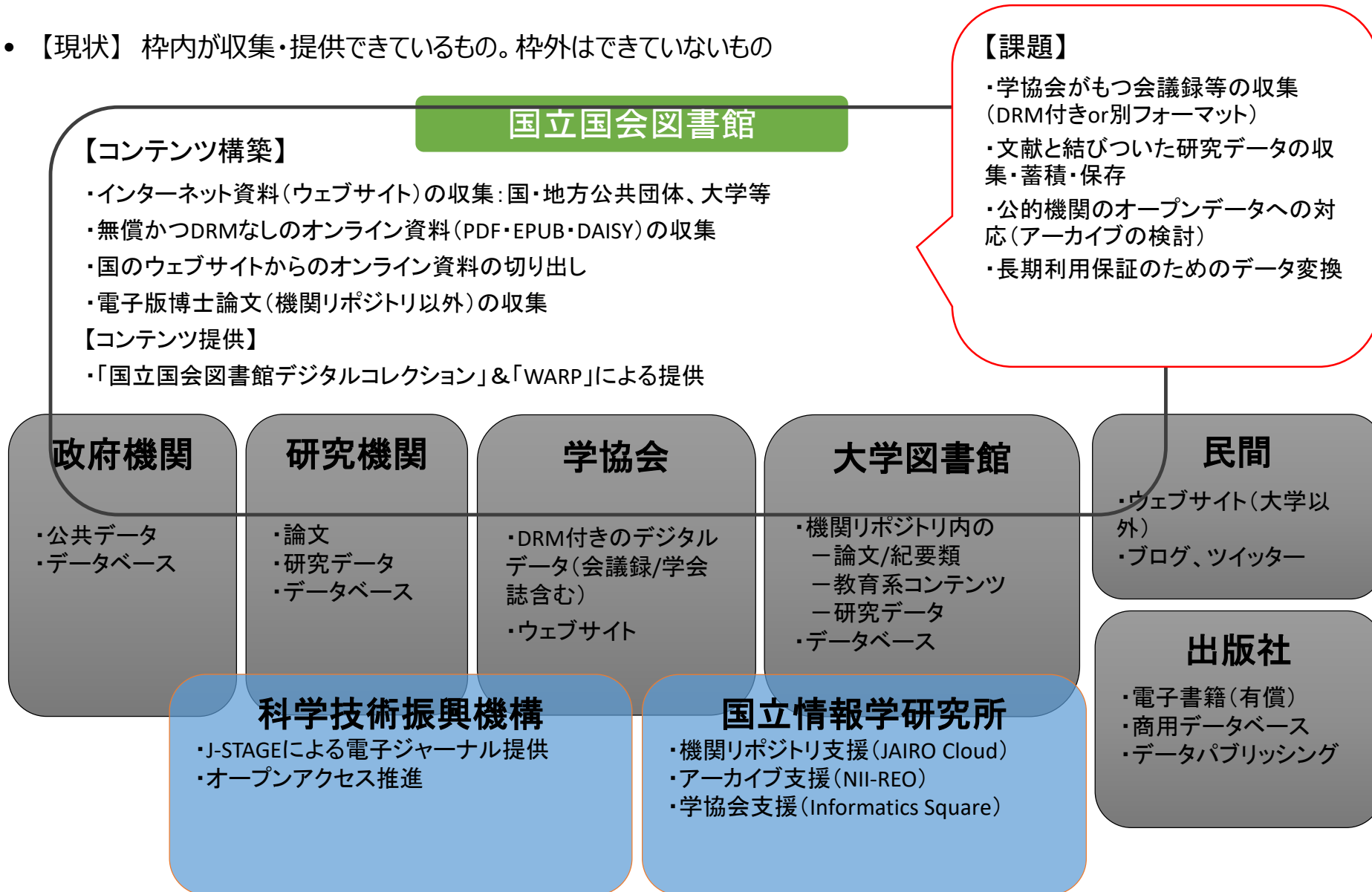
- ・博士論文(H13～)のデジタル化
- ・論文/紀要類のデジタル化

## 文書館・美術館・博物館

- ・「日本語の歴史的典籍」プロジェクト(国文学研究資料館)
- ・所蔵品のデジタル化

## ② インターネット資料・オンライン資料等

- 【現状】 枠内が収集・提供できているもの。枠外はできていないもの



# 資料デジタル化と提供状況（インターネット及び図書館送信）

「いつでもどこでもだれでも」が電子図書館サービスの理念

館内限定  
**62万点**

185万点（48万＋137万点）は、公共図書館の蔵書と言える  
しかし、インターネット公開していない200万点は、ILLもできない。

国立国会図書館は所蔵資料をデジタル化できる  
（著作権法第31条2項）

全**248.5**  
万点

国立国会図書館  
館内利用者



アクセス月19万件

市場で入手困難な資料は図書館送信できる  
（著作権法第31条3項）

デジタル化資料

国立国会図書館  
デジタルアーカイブ

著作権  
存続

権利者・詳細不明

権利者許諾

絶版等

137万点

公共・大学  
図書館等  
館内利用者



アクセス月1万3千件

※著作権関連  
の慎重調査

著作権  
消滅

文化庁  
長官裁定

NDLデジ  
タルコレク  
ション等

48万点  
アクセス  
月490万件

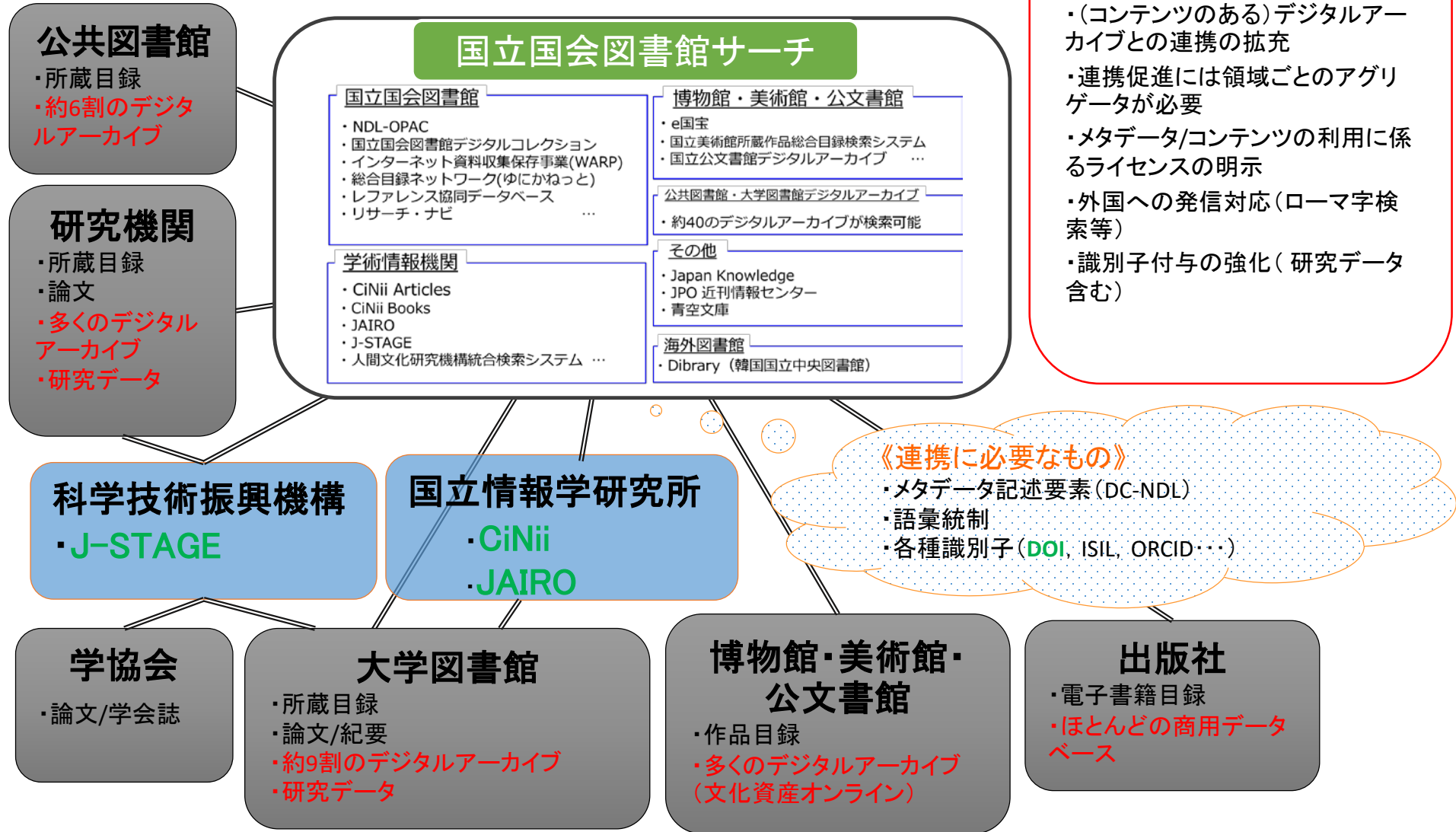


インター  
ネット  
利用者

現在500館程度  
ILL加入館の1/5

# ③ 電子情報資源の組織化と連携

【現状】 国立国会図書館サーチによる連携先データベースは約100  
赤字が検索対象になっていないもの。





# 資料のデジタル化



# NDL所蔵資料のデジタル化

- 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。



# 図書館資料の保存と利用

資料は、利用によって、劣化損傷する

～紙質や利用頻度から～

資料保存対策

媒体変換

～保存と利用の両立～

マイクロ化

(長年の実績)

デジタル化

(可能性大の新技术)

<留意点>

・著作権の尊重    ・利用者利便性の向上    ・出版文化の隆盛

# 資料デジタル化予算の推移

※平成21、22両年度の補正予算にて大規模に予算措置、  
著作権法改正と相まって一気にデジタル化が進展

平成12～26年度 総額 153億円

年度	12	13	14	15	16	17	18	19	20	21当初	22当初	23	24	25	26
予算	1.0	1.5	2.2	2.4	1.2	0.4	2.2	0.8	1.3	1.3	1.3	0	0	0.2	0.2

(単位：億円)

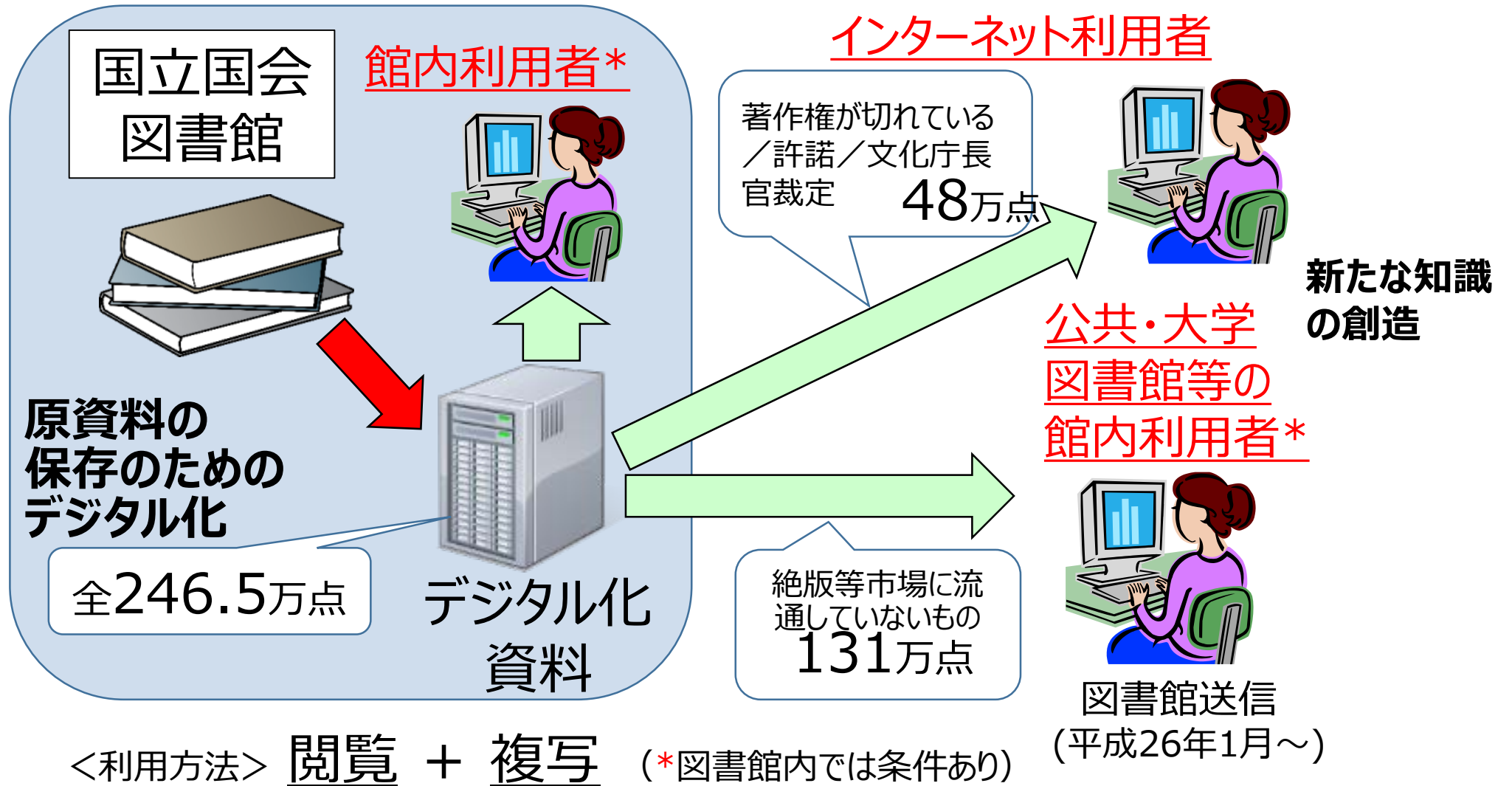
10年間で 14億円

※平成12年度  
明治期刊行図書デジタル化予算からデ  
ジタル化を開始（2～4万冊/年）

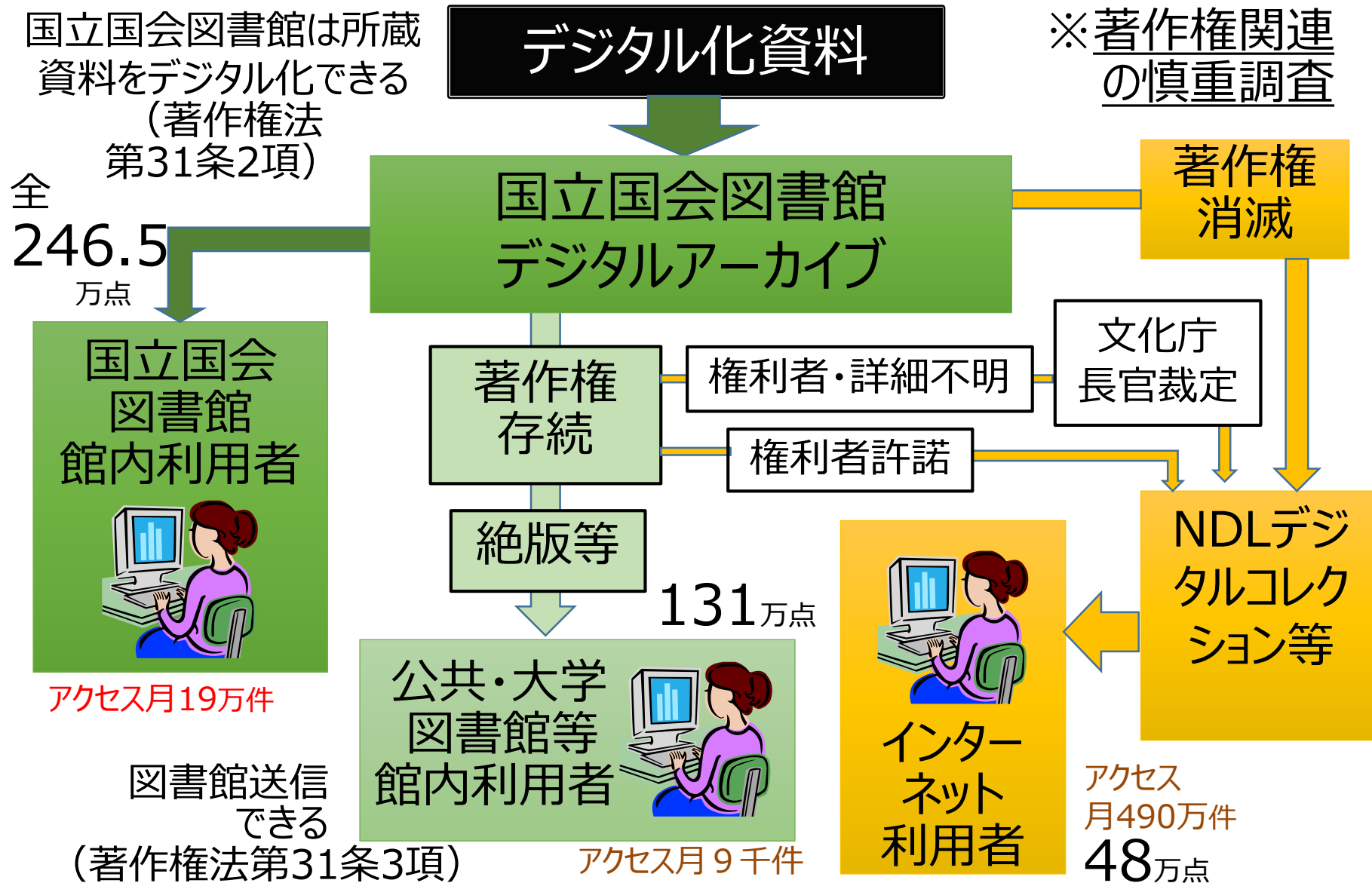


21・22年度  
補正予算

# デジタル化資料の提供状況～利用者別



# デジタル化資料の提供と著作権保護





# デジタル化及びデジタル資料の収集の歩み

デジタル化	内容
～2009年	著作権処理を行い、デジタル化、インターネット提供
2010年1月（改正著作権法施行）	原資料保存のため、著作権者の許諾なく、図書館資料をデジタル化し館内提供
2013年1月（改正著作権法施行） 2014年1月（正式運用開始）	デジタル化した資料のうち、絶版になっているものを、公共図書館へ送信開始

デジタル資料の収集	内容
2010年10月	パッケージ系電子出版物の 納本制度による収集開始
2002年11月	WARP（インターネット資料収集保存事業） 開始（許諾による収集）
2010年4月	公的機関のインターネット資料の 制度的収集開始
2013年7月	民間のオンライン資料（電子書籍、電子雑誌等）の制度的収集 開始（当面、無料かつDRMのないもの）

# 資料デジタル化の現下の方針 (1)

「国立国会図書館の資料デジタル化に係る基本方針」  
(2013年5月27日国図電1305201号)

## 1. 国立国会図書館の取組の考え方

- 人々が将来にわたり利活用できるように、我が国の豊かな文化的な資産を広く収集・保存する
- 関係機関と連携して社会全体で知識・文化の基盤を構築する
- これらの取組によって、人々の創造的な活動に貢献する
- この取組の一環として
  - 財源の確保に努め、所蔵資料のデジタル化を進める
  - デジタル化資料の提供により、原資料を保存する
  - 検索の利便性や障害者を含むあらゆる人々の利用可能性を高める
  - 関係機関等との有機的な連携により、知識・文化の基盤を構築する

# 資料デジタル化の現下の方針 (2)

## 2. 対象資料の範囲

- ①1980年頃までの国内刊行図書
- ②歴史的価値の高い外国刊行の日本関係資料等

## 3. 対象資料選定の際の評価要素

- ①唯一性・希少性 ②資料の劣化状況・保存の緊急性 ③利用機会の拡大 ④社会的ニーズ ⑤内外の体系的デジタルコレクション構築への貢献

## 4. デジタル化の方法

- ①当面、画像データ作成 ②ただし、検索の便のため、目次情報はテキストデータ化
- ③視覚障害者等に対する提供、戦前期議会会議録の活用を視野に、本文テキストデータ化を段階的に取り組む

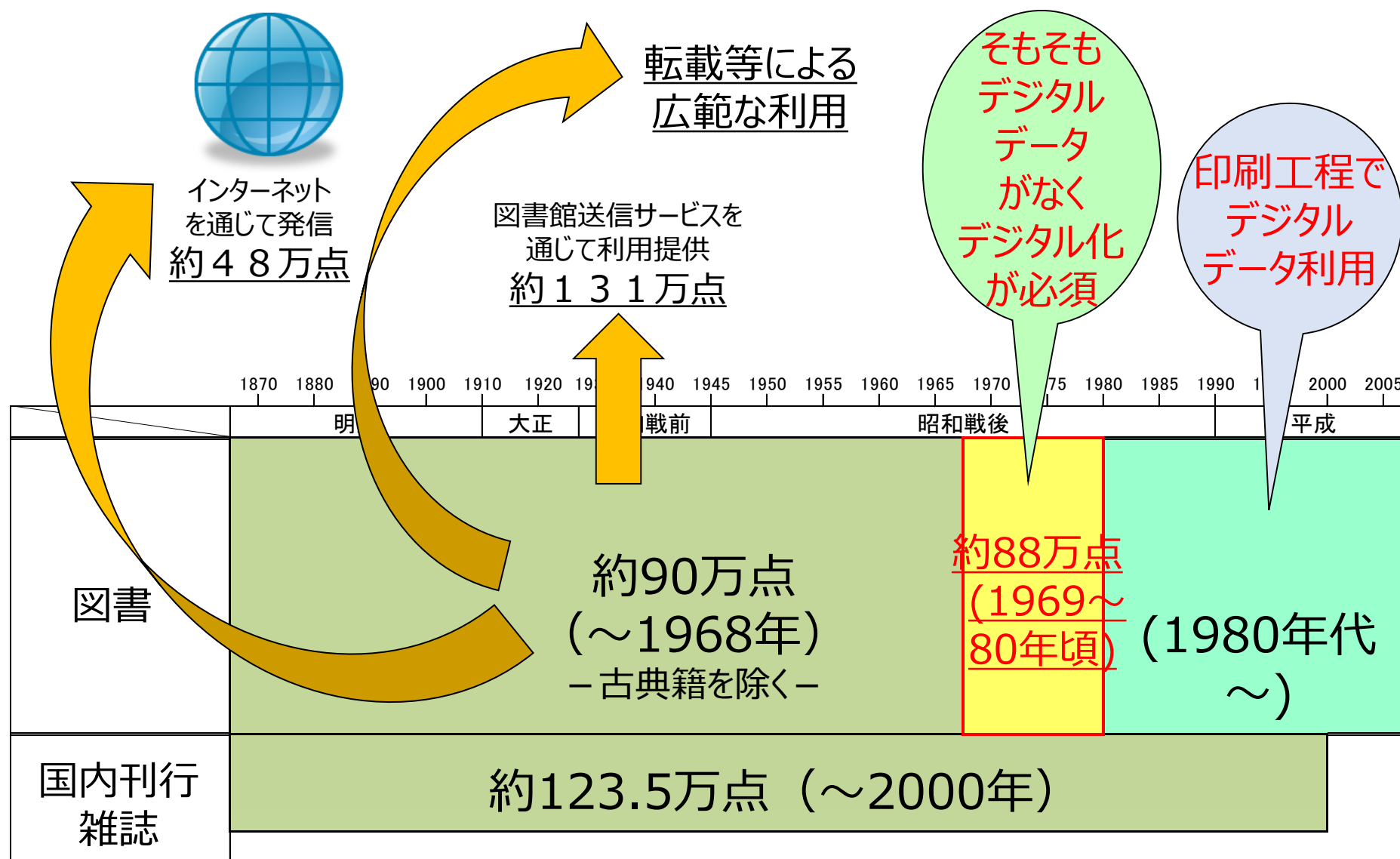
## 5. 資料デジタル化の成果を利活用する努力

- ①国の文化的資産として、資料デジタル化の成果を社会的に利活用できること
- ②そのため、著作権等に留意しつつ社会的な理解を得ること

## 6. 東日本大震災アーカイブ構築のために

関係機関と調整しながら関係記録を後世に伝えるため、デジタル化に取り組む

# 現下の最優先のデジタル化対象



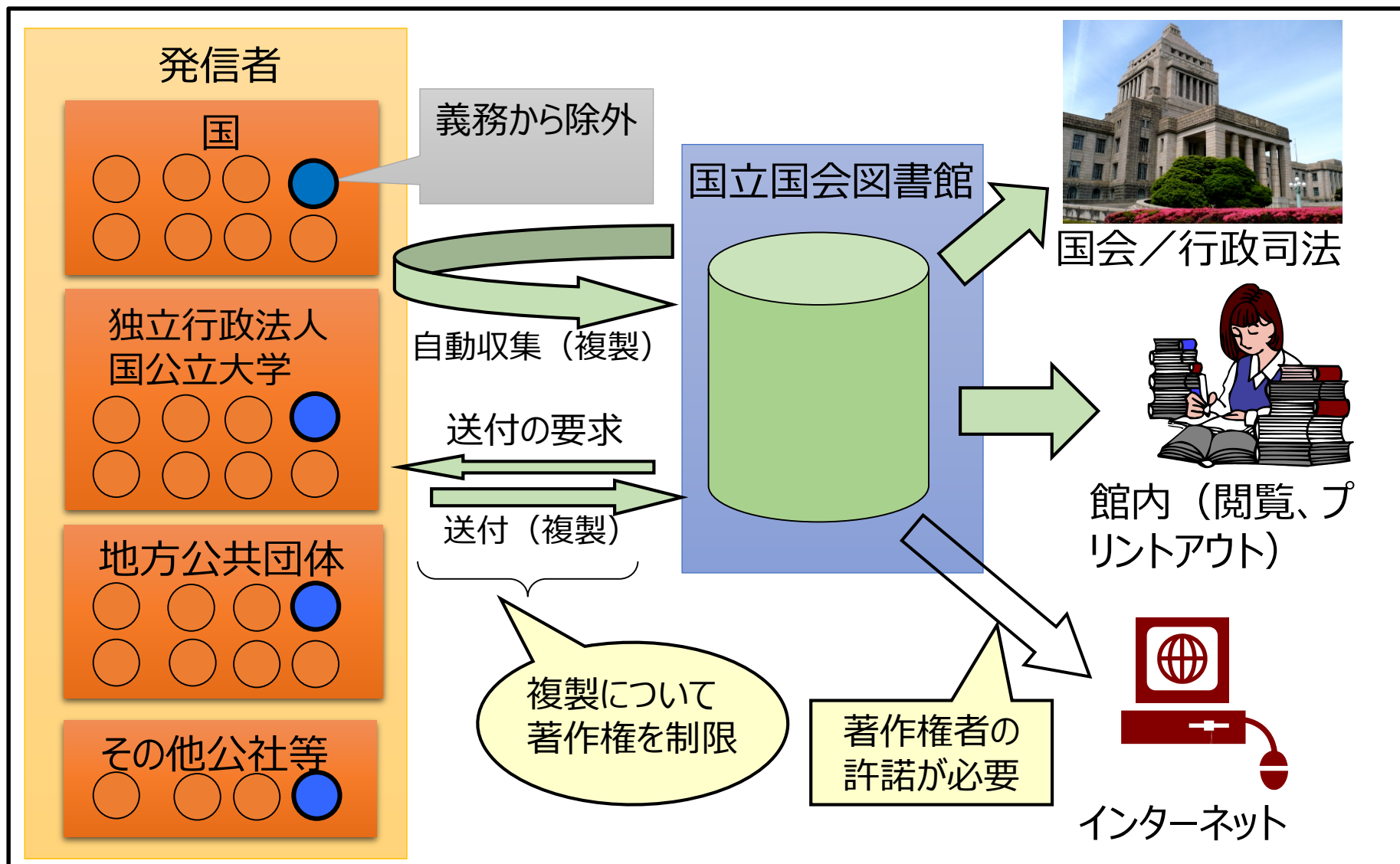




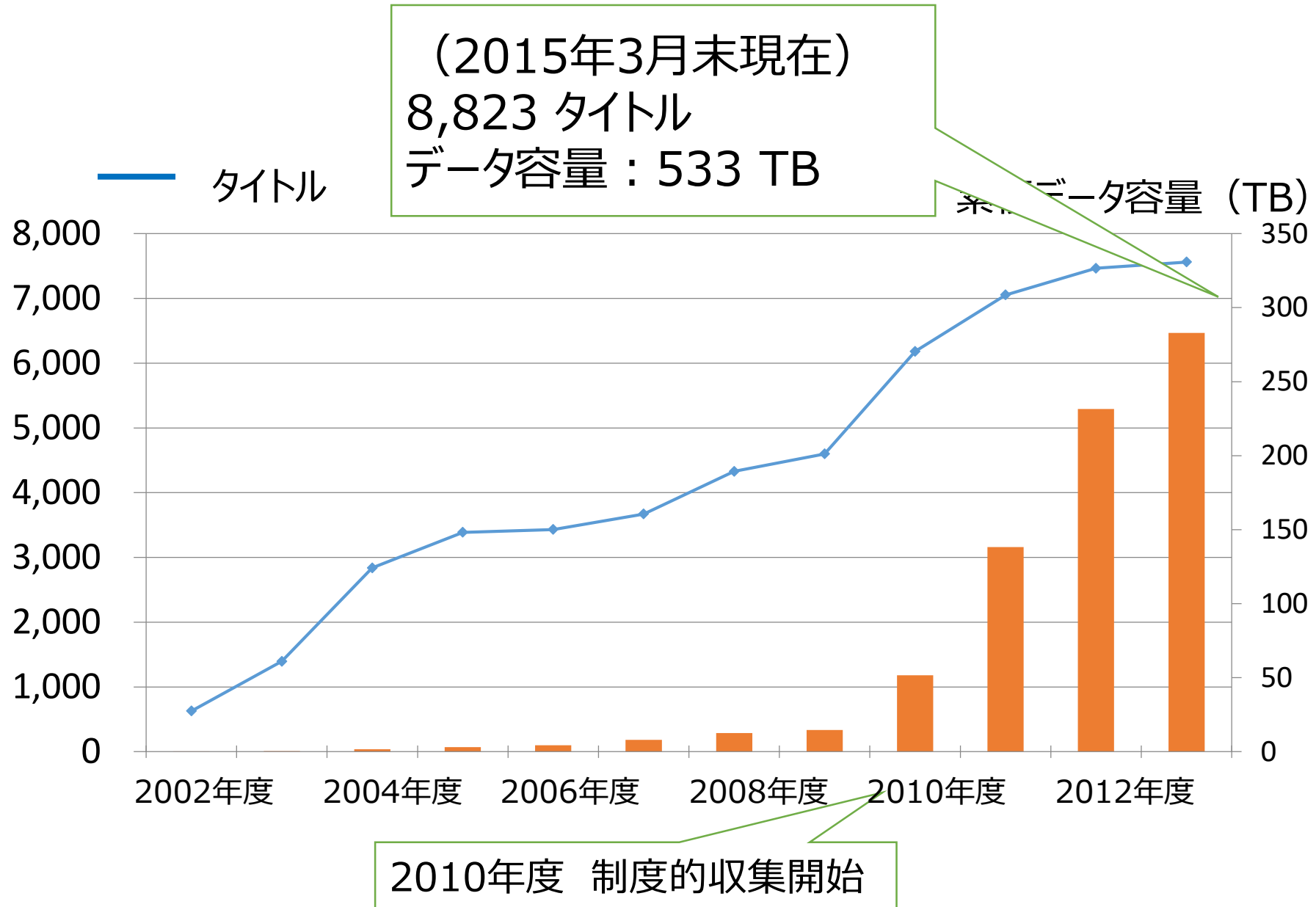
# インターネット資料収集




# インターネット資料の制度的収集



# 収集したウェブサイトのデータ量の変遷





# オンライン資料収集



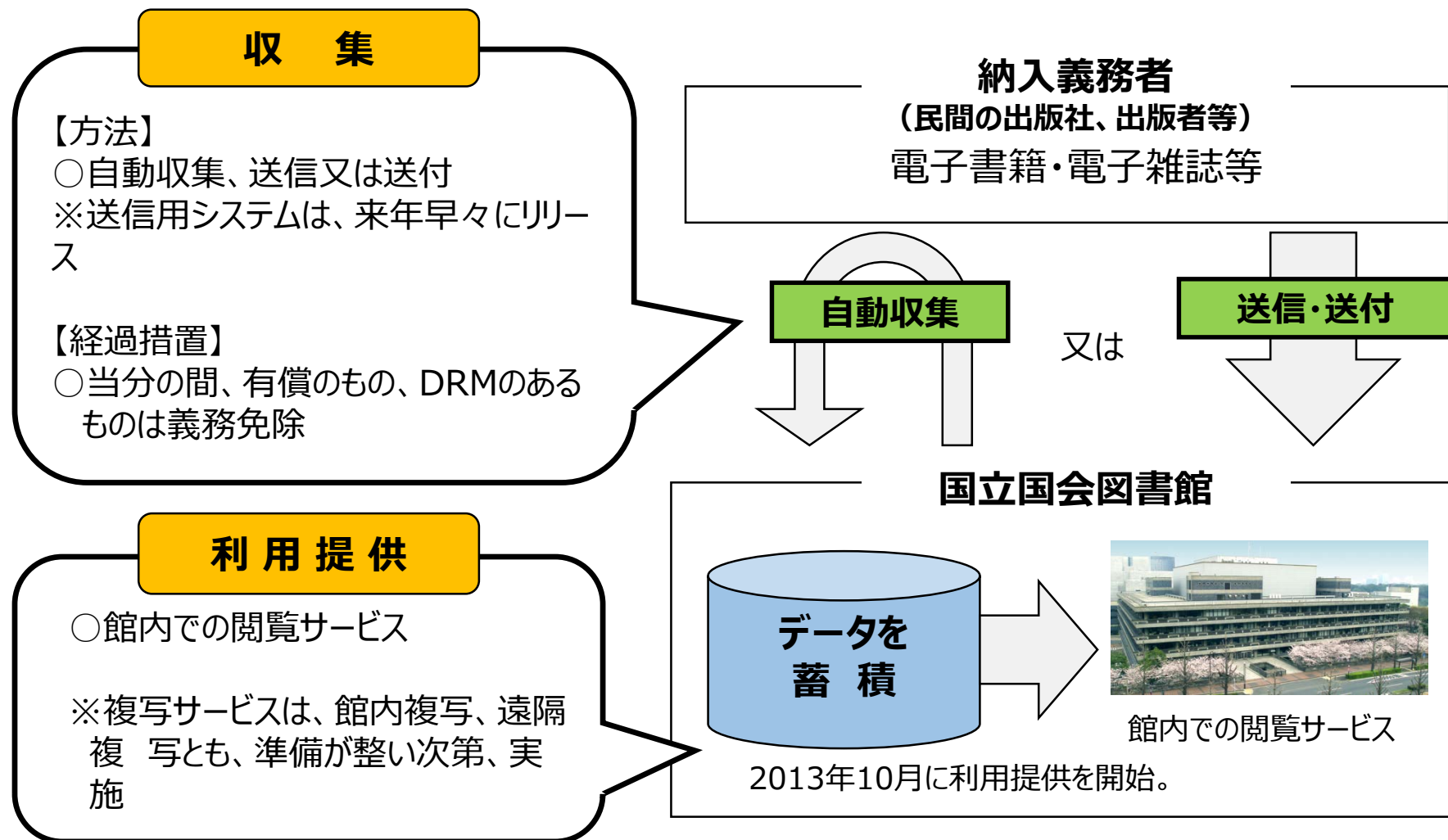
# オンライン資料の収集

2013年7月に、制度に基づく民間の「オンライン資料」の収集を開始した。

- ◆ オンライン資料とは  
インターネット等により出版（公開）される電子情報で、図書又は逐次刊行物に相当するものであり、電子書籍・電子雑誌等を指す。（次ページ参照）
- ◆ 納入義務対象は、2013年7月1日以降に、私人がインターネット等で出版（公開）した電子書籍・電子雑誌等のうち、  
⇒特定のコード（ISBN、ISSN、DOI）が付与されたもの  
⇒特定のフォーマット（PDF、EPUB、DAISY）で作成されたもの  
のいずれかであって、無償かつDRMのないもの\*
- ◆ 具体例  
年鑑、要覧、機関誌、調査報告書、事業報告書、学術論文、紀要、技報、ニュースレター、小説、実用書、児童書等

\*当分の間、有償のもの、DRMのあるものは義務免除。

# オンライン資料の収集から提供までの流れ

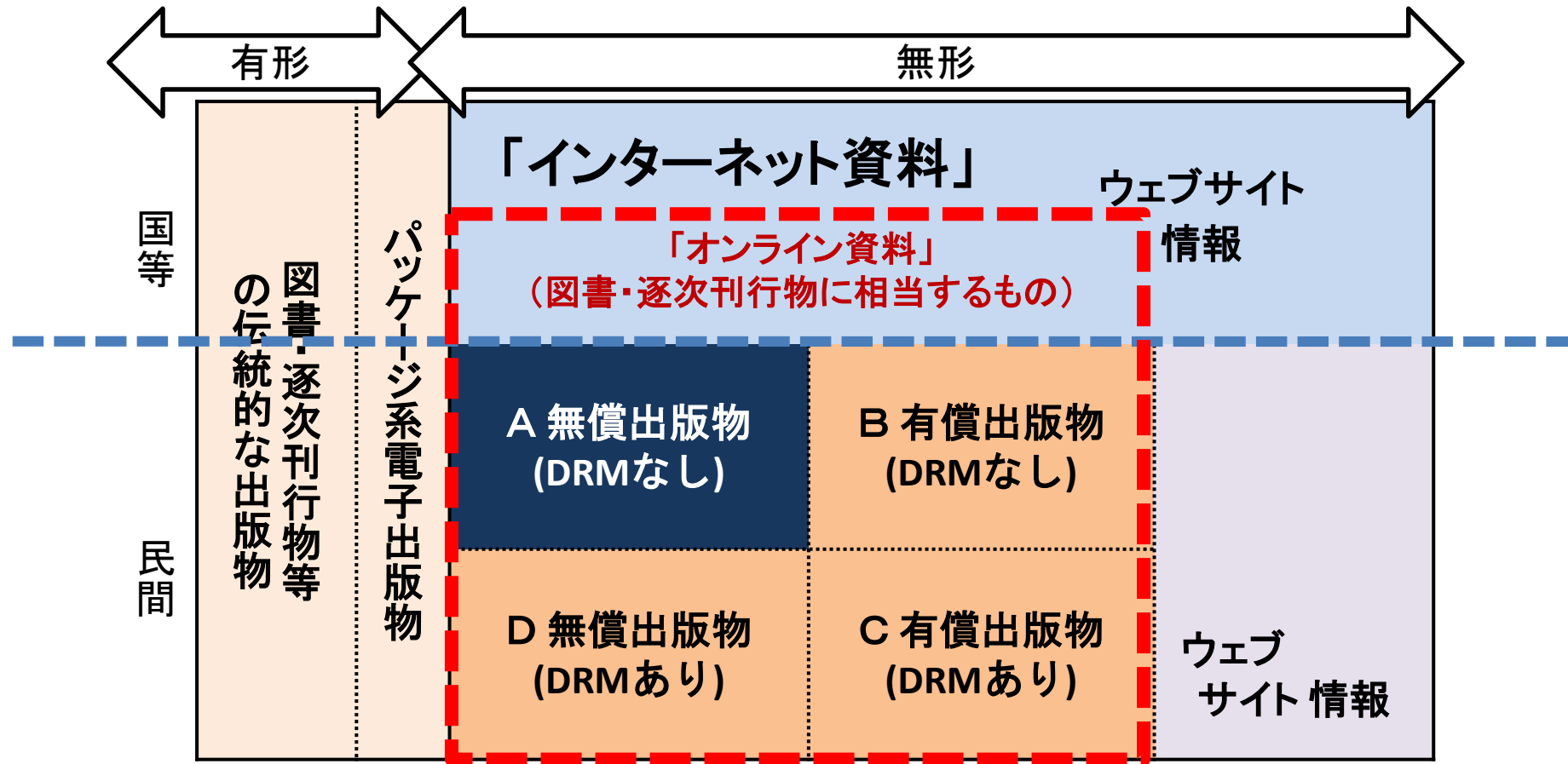


# 電子版博士論文の収集

2013年4月1日に、学位規則が改正・施行されたことに伴い、博士論文は、インターネットの利用により公表されることになった。それに伴い、NDLでは、従来紙媒体を収集していたところ、電子データを収集することになった。

- ◆ 博士論文の送付方法は、主に以下の2種類。
  - ① 学位授与大学が、機関リポジトリで公表し、メタデータを国立情報学研究所（NII）の「学術機関リポジトリデータベース（IRDB）」に提供している場合は、IRDBのメタデータを利用して当館が自動収集。
  - ② ①以外の機関リポジトリやホームページ等で公表する場合は、送信用システムにより送付。
- ◆ 収集した博士論文は当館施設内で閲覧に供する。また、許諾を得られたものはインターネット公開。

# オンライン資料の制度収集



＝国立国会図書館法(館法)24条、24条の2、25条に基づく収集

＝館法25条の3に基づく収集

＝現在の収集の対象(2013年6月22日公布の館法25条の4)





# ☆デジタル情報の長期保存

- OAIS, 情報パッケージ(SIP, AIP, DIP)
- 

# 再生技術情報の管理（論理保存）の課題

- 媒体、ストレージの問題だけでは解決しない。
- 電子情報のコンテンツの再生可能性を長期にわたって保障する仕組みが必要。
  - 技術・環境の変化：旧式化  
→長期保存の対応策を含む必要がある
- 利用保証の2つのレベル
  - 情報の欠損／変化がないことを保障する  
→ビット保存（bit preservation）
  - 情報の内容を理解できることを保障する  
→論理保存（logical preservation）

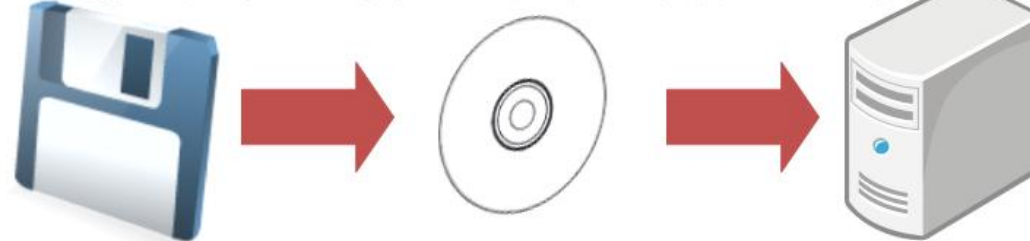
# 再生技術情報の管理の必要性

- 参考：マイグレーション

① 同種の新しい記録媒体への移行



② 異種の新しい記録媒体への移行



実施例：  
5.25インチFD  
→HDD

PhotoCD, TIFF  
→JPEG2000

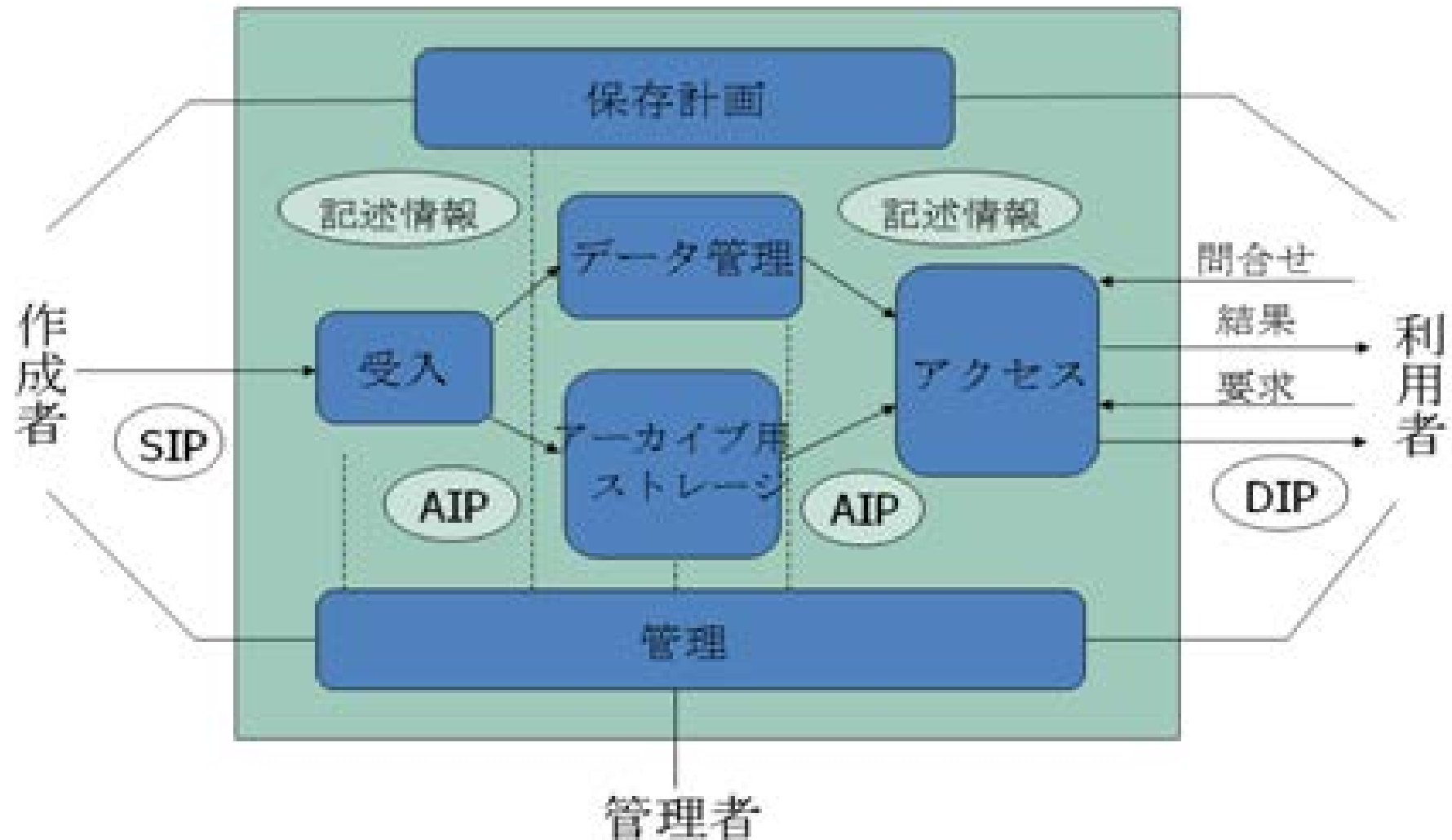
③ 新しい記録フォーマット(ファイル形式)への変換



# ☆ 長期保存システムの概念モデル

## • OAIS参照モデル

- Space data and information transfer systems -- Open Archival Information System -- Reference model (ISO 14721:2003)




# 電子情報の長期保存の課題（まとめ）

- 大容量で長期間監視不要な記録媒体
- ストレージ技術のさらなる進歩
  - コスト、データ移行、遠隔地バックアップ等
- 再生技術情報の管理（論理保存）の実現



# NDL所蔵資料の二次利用

- デジタル化資料の利活用に向けて
- 関係者協議による調整
- デジタル化資料の転載利用



# デジタル化資料の利活用に向けて

- 2000年度から資料デジタル化実施（法第31条1項2号）
  - － 著作権処理を行いインターネットで公開
- 文化審議会「**過去の著作物等の利用の円滑化のための方策について（中間総括）**」（2008年5月）－ 国立国会図書館のデジタル化資料の利用につき、関係者間の協議を提言
- 2009年**著作権法改正**（法第31条2項新設）
  - 国立国会図書館での原本保存目的のデジタル化が可能に
- 2009,2010年度に大規模デジタル化事業を実施
- 2012年**著作権法改正**（法第31条3項新設）
  - デジタル化資料のうち絶版等資料を全国の図書館等で利用可能に  
⇒ 2014年1月から図書館送信サービス開始



# 関係者協議による調整

- 国立国会図書館と権利者、出版者等の関係者間の協議の場として、「**資料デジタル化及び利用に係る関係者協議会**」を設置（2008年9月～）
- 図書館への限定送信に関するワーキングチームを設置（2011年11月～）
- 「**国立国会図書館のデジタル化資料の図書館等への限定送信に関する合意事項**」（2012年12月）

[http://www.ndl.go.jp/jp/aboutus/digitization/digitization\\_agreement02.pdf](http://www.ndl.go.jp/jp/aboutus/digitization/digitization_agreement02.pdf)

⇔ この範囲において「図書館送信サービス」を実施

- デジタル化する**録音資料**の利用に係る協議のために、「録音資料のデジタル化及び利用に係る関係者協議会」を設置（2014年5月～）





# デジタル化資料の転載利用

～デジタル化資料を広く社会の利用に供する～

- インターネット提供のデジタル化資料（著作権保護期間満了分）に係る転載利用は、**転載申込み手続が不要**

平成26年5月1日から。ただし、許諾または文化庁長官裁定分の転載（復刻、翻刻、掲載、放映又は展示等）を行う場合には、これまで同様、国立国会図書館への照会が必要

- デジタル化資料の画像データの試行提供中

平成26年8月1日から1年間、図書館送信及び国立国会図書館内限定公開の資料について、復刻・翻刻を目的とした利用に限って、画像データの試行提供（ニーズを把握、手続等を検証し、サービス体制を構築・整備）

- 申請者による著作権処理又は著作権保護期間満了の確認が必要
- 出版の形態等により、使用料若しくは製品の一定数の寄贈、又はその双方が条件となる場合がある



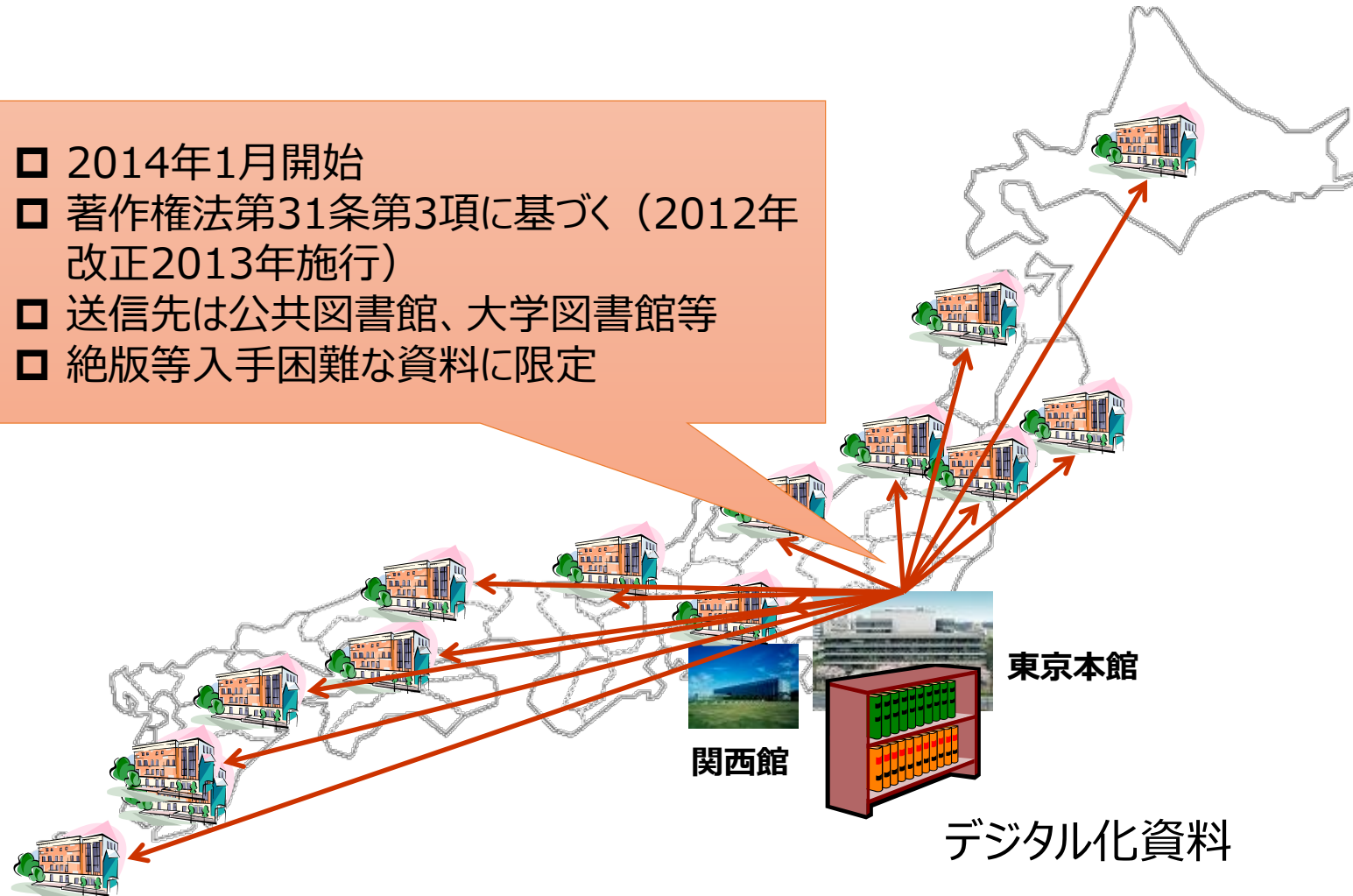
# 図書館送信サービス

「国立国会図書館のデジタル化資料の図書館等への限定送信」  
サービス



# 図書館向けデジタル化資料送信サービス

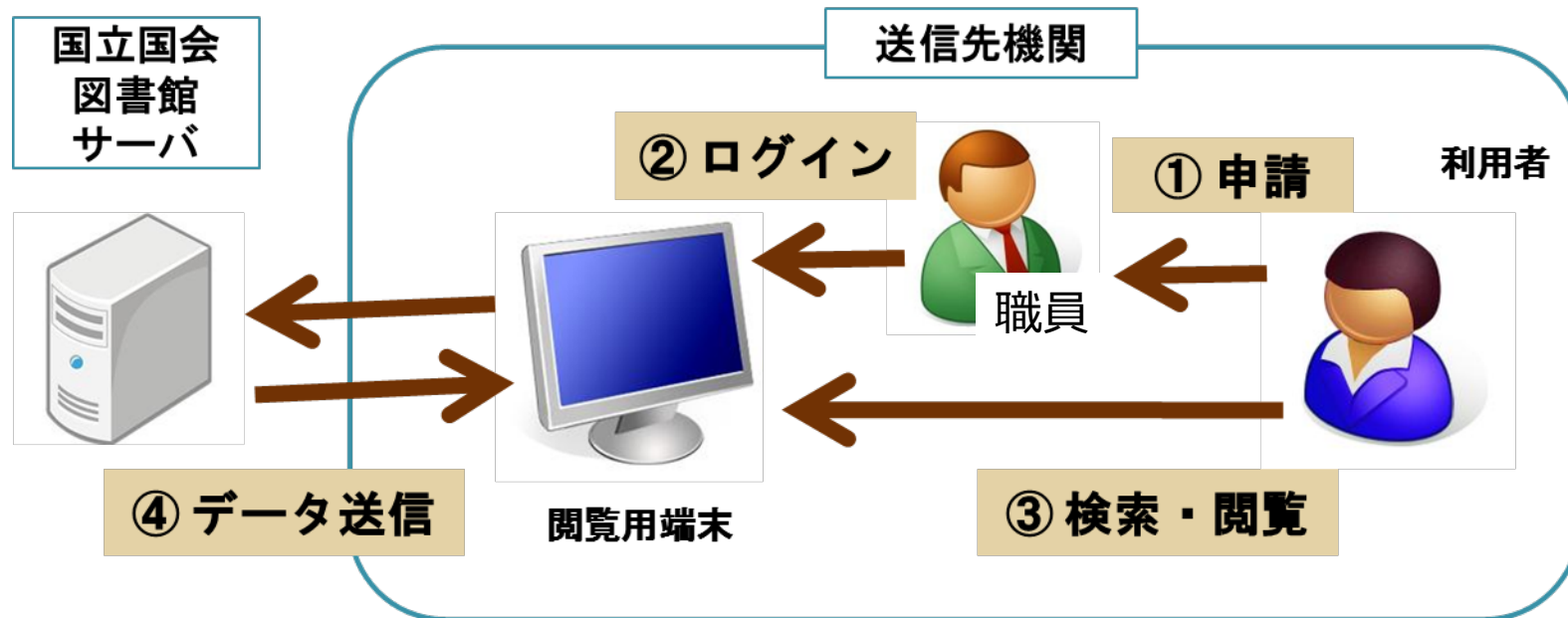
- ❑ 2014年1月開始
- ❑ 著作権法第31条第3項に基づく（2012年改正2013年施行）
- ❑ 送信先は公共図書館、大学図書館等
- ❑ 絶版等入手困難な資料に限定



[http://www.ndl.go.jp/jp/library/service\\_digi/index.html](http://www.ndl.go.jp/jp/library/service_digi/index.html)

# 閲覧利用

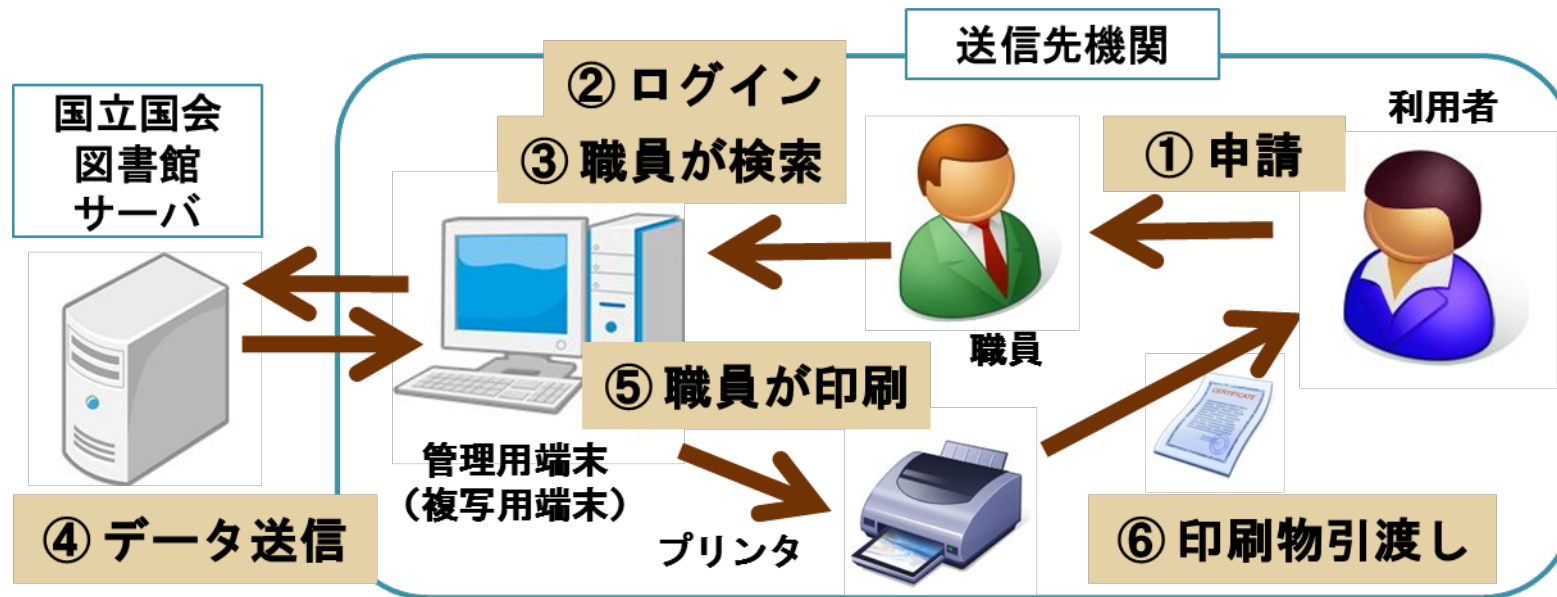
「国立国会図書館のデジタル化資料の図書館等への限定送信に関する合意事項」に基づき利用提供



- ・端末は職員の目の届く場所に設置
- ／
- ・利用は送信先機関の「登録利用者」のみ
- ／
- ・閲覧申込の都度、職員がログイン
- ／
- ・利用後、ブラウザを終了
- ／
- ・不正の監視・注意喚起

# 複写利用

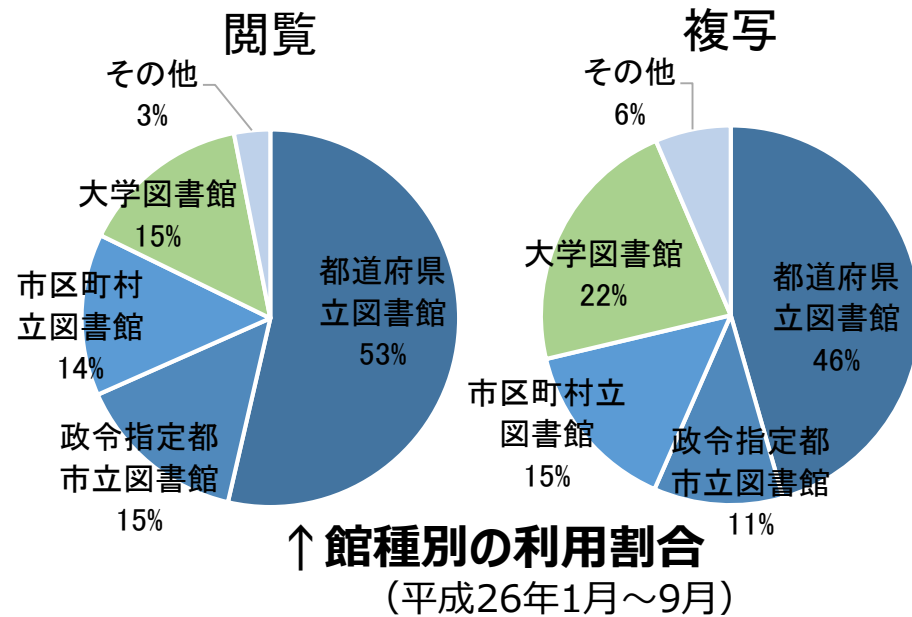
「国立国会図書館のデジタル化資料の図書館等への限定送信に関する合意事項」に基づき利用提供



- ・端末は利用者が操作できない場所に設置
- ／
- ・利用は送信先機関の「登録利用者」のみ
- ／
- ・複写申込の都度、職員がプリントアウト（セルフプリントアウト不可）
- ／
- ・著作権法上の要件の確認
- ／
- ・利用後、ブラウザを終了（キャッシュの自動削除をONに）
- ／
- ・複写記録の作成と1年間保存（個人情報を除く）

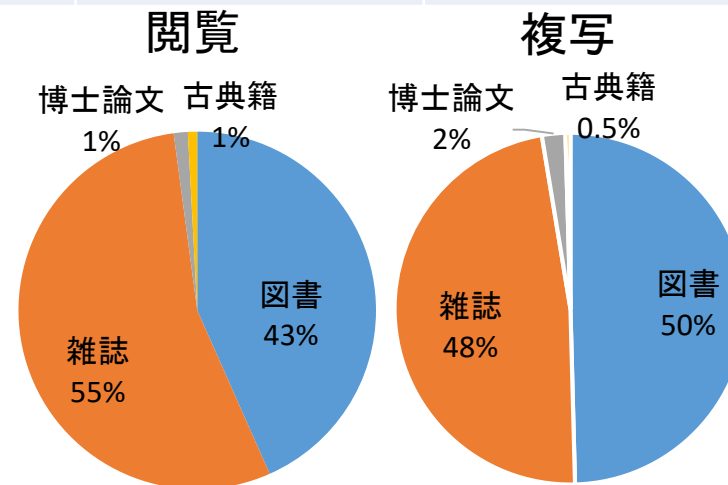
# (参考) 図書館送信サービスの利用統計

- 参加館は2014年10月10日現在で**333館**（都道府県立図書館53館、政令指定都市立図書館33館、市町村立図書館93館、大学図書館141館、その他13館）  
参加館リスト：[http://dl.ndl.go.jp/ja/soshin\\_librarylist.html](http://dl.ndl.go.jp/ja/soshin_librarylist.html)
- 利用統計：[http://dl.ndl.go.jp/ja/soshin\\_library\\_stats.html](http://dl.ndl.go.jp/ja/soshin_library_stats.html)



	平成26年1月～9月	平成26年9月
閲覧	56,489 回	10,342回 (1日平均345回)
複写	24,941回	4,224回 (1日平均141回)

資料種別の利用割合→  
(平成26年1月～9月)



# 歴史的音源の公共図書館等への配信提供

## 歴史的音源とは

- ◆ 1900年初頭～1950年頃に日本国内で製造されたSP盤及び金属原盤等に収録された音楽・演説等の音源
- ◆ 歴史的音盤アーカイブ推進協議会（HiRAC）がデジタル化。
- ◆ 2011年5月に提供開始。
- ◆ 提供音源は約5万音源。
- ◆ 著作権・著作隣接権保護期間を満了したものはインターネット公開（約1,000音源）。  
⇒つまり、大部分は国立国会図書館の施設内での提供のみ。



## 2012年7月に、公共図書館への配信を開始

- ◆ 配信参加館は、館内限定音源を含む全音源を図書館利用者に提供可能。
- ◆ 現在（2013年11月時点）、126館に配信中。以後、継続的に配信参加館募集

ナレッジデータベース



# はじめに

2005年

- 「電子図書館中期計画2004」
  - 情報資源に関する情報の充実
    - 情報に到達するための手段や情報の充実
    - 検索手段の充実
    - 主題情報の充実



# 何故、ナレッジDBか


- 「電子図書館中期計画2004」に基づき、
  - 情報に到達するための手段や情報の充実
  - 検索手段の充実
  - 主題情報の充実
- **NDL全職員を対象とする**
  - ナレッジデータベースを構築する。
- 必要な人（利用者）が自由に利用可能



NDLの強みを活かす

# もう少し詳しく・・・

- 情報資源に関する情報の更なる充実を図る
- 個人的な主題知識（ナレッジ）を継承する

NDL組織全体として、職員↓の主題知識や  
情報収集能力などを出し合うことが必要。

- システム的には、
  - メインシステム「ナレッジデータベース」
  - サブシステム「ナレッジコミュニティ」

# NDLナレッジDBの定義

- ナレッジデータベース

- 知識の集まりだけでなく、データ、情報、知恵、ノウハウなどNDL組織で発生した知的資産全体を指す。

- ナレッジコミュニティ（サブシステム）

- ユーザー同士の半同期的な質問と回答、ニュース等の収集により形成される共同知識の集約サイトを指す。
- [Yahoo!知恵袋](#)、[はてな](#)、[NAVER - 知識PLUS](#)、[教えて! Goo](#)、[OKWebコミュニティ](#)

# ナレッジDBの目的

- 当館で発生したすべての知的情報
- ナレッジコミュニティによる関連情報

媒体に関わらず全館的に収集し、蓄積する。

- ナレッジDBの目的
  - 利用者のニーズ、日常の調査活動
  - レファレンスへの効率化、知的価値創造に貢献

# ナレッジDB対象コンテンツ

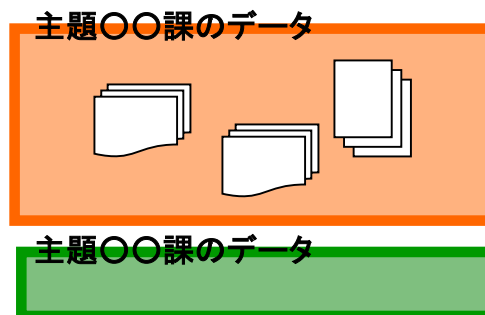
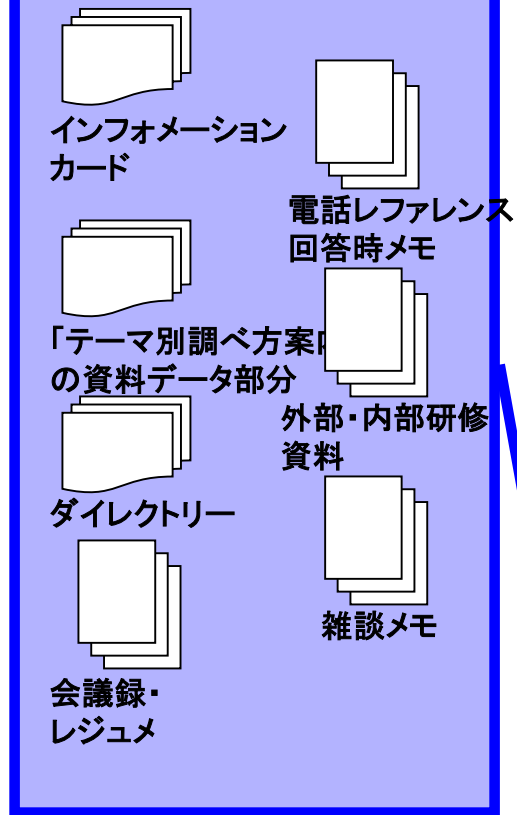
## ナレッジデータベース

- ・既存のデータ
- ・当館刊行物／統制語辞書類／調査研究レポート類
- ・個人的資料（業務上作成資料を含む）
- ・非公式参考情報メモ（他者の参考になる情報）
- ・外部発表資料（論文、報告書、ppt等講演資料類）
- ・当館作成マニュアル、教材、レジメ
- ・有用な情報源の所在情報Webサイト Dnavi
- ・貴重書や電子展示会等での解題情報等

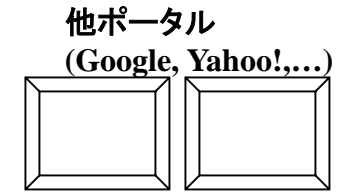
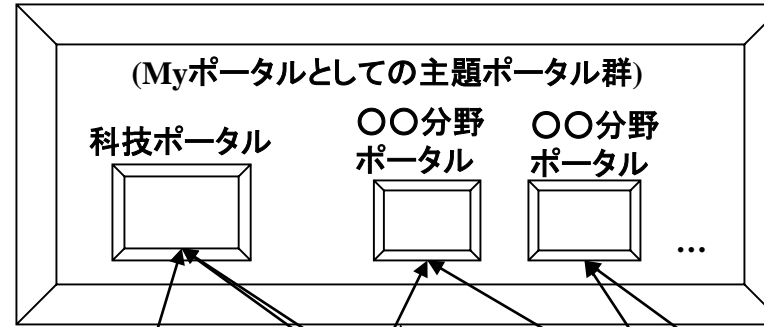
### ナレッジ コミュニティ

レファレンス事例情報  
／Q&A／支援情報  
メール情報  
KnowHow  
KnowWho  
ウォッチング収集情報  
(ニュース、ブログ…)

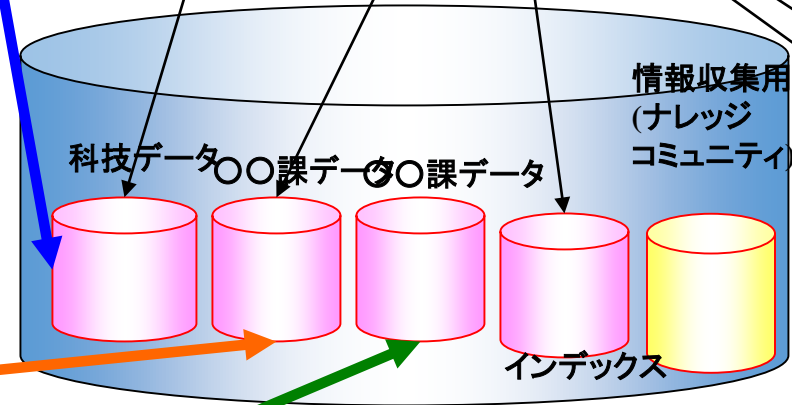
科技で発生する資料・インフォーマルなメモ等すべて



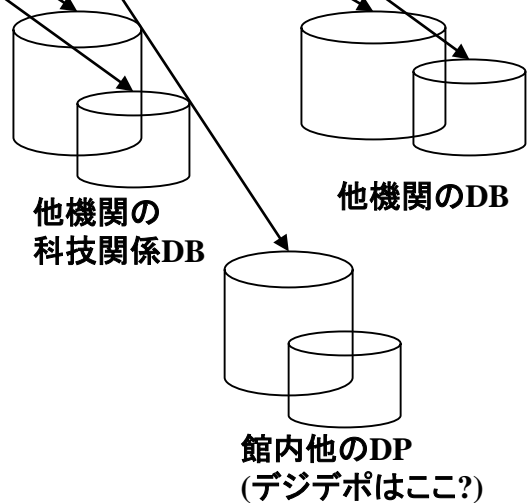
NDLデジタルアーカイブポータル  
(→将来は「日本のデジタルアーカイブポータル」へ)

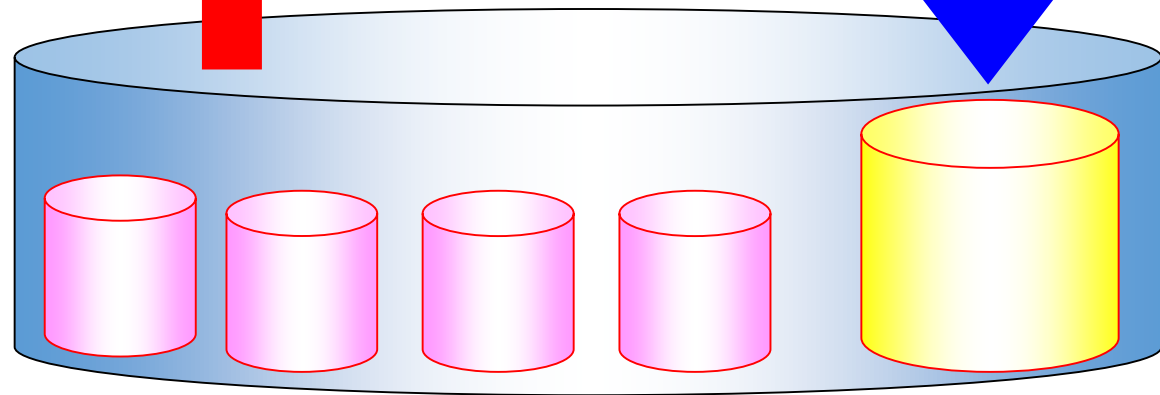
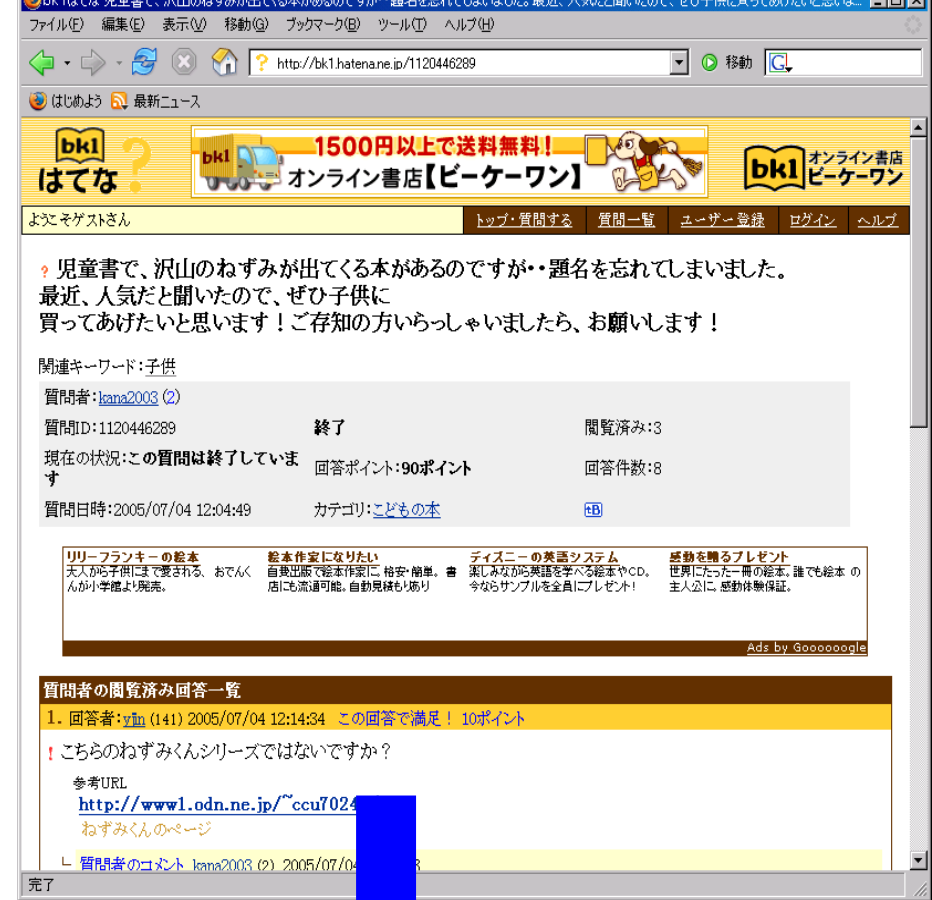
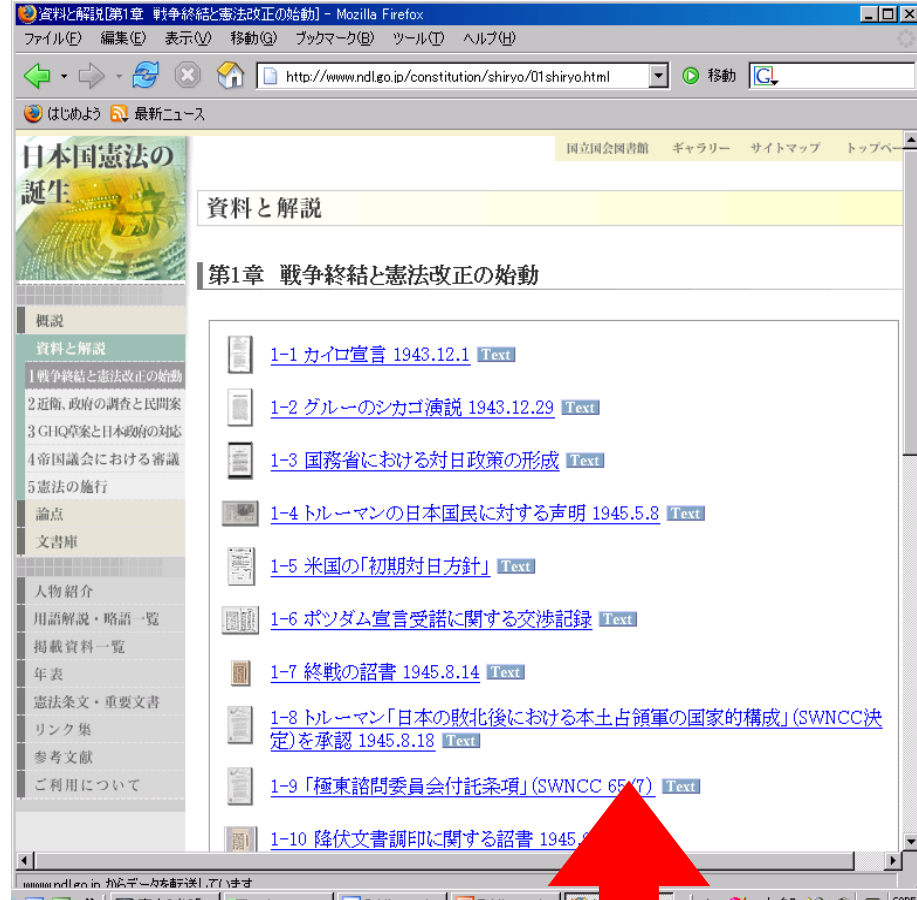


(統合検索等サービス動的にページ生成)



NDLナレッジDB

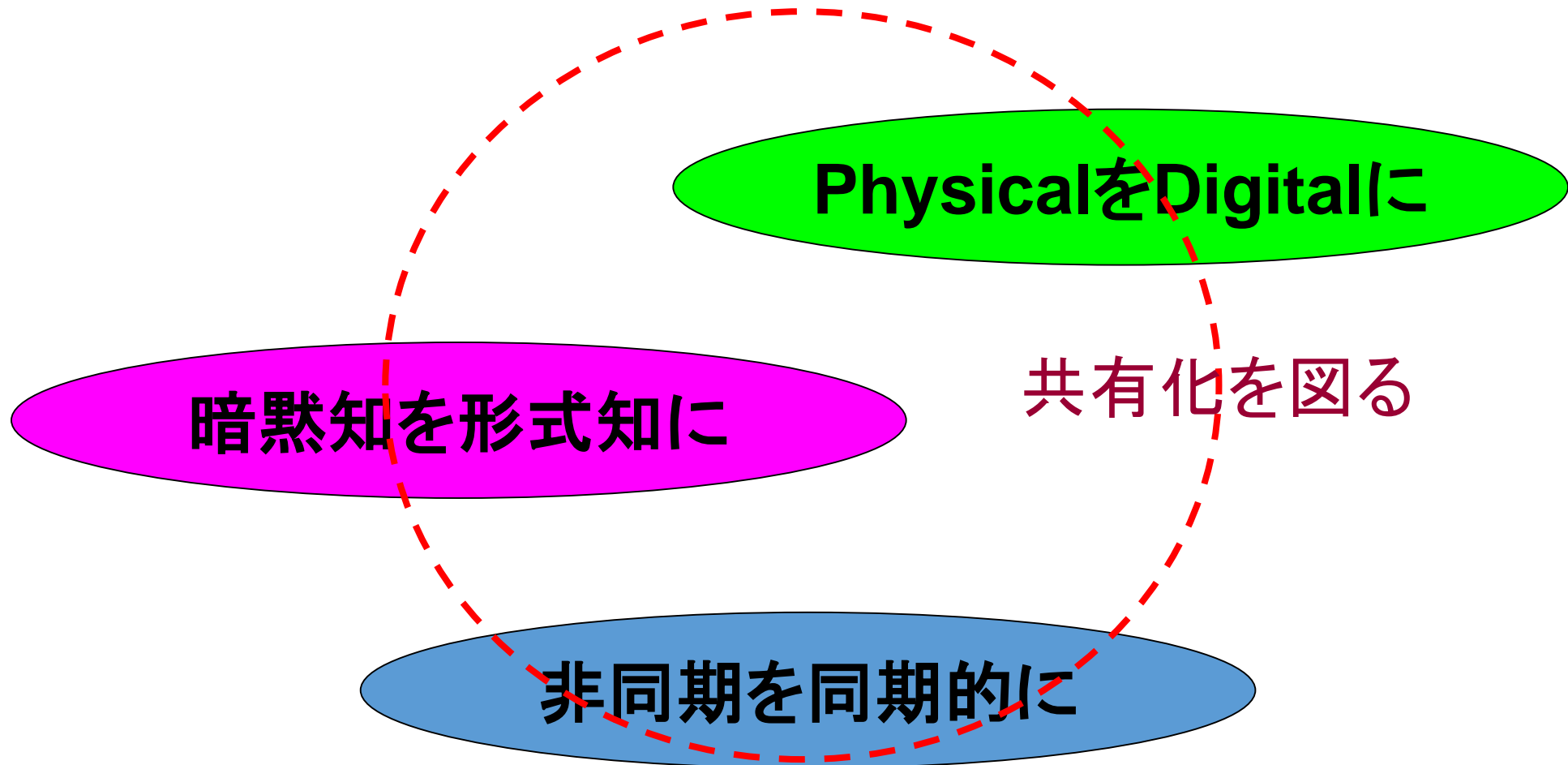






# ナレッジDB構築のためには

- 知的関連情報として再利用しやすくすること。



# 求められる機能

- 収集・組織化
  - メタデータの入力、自動収集、自動作成機能
  - RSS機能
- 登録・蓄積
  - メタデータ登録機能
  - 一次情報自動アップロード機能
  - リポジトリ機能
- 提供
  - メタデータ検索機能
  - 全文検索機能
  - 連想検索機能
  - 自然言語群検索機能
  - サブジェクトゲートウェイ機能
  - パスファインダー機能
  - ILL機能
  - オンラインチュートリアル機能（動画、音声ヘルプ機能）
  - 他システムとのシステム連携機能（Webサービス機能）

