

全文テキスト化・実証実験

■ 全文テキスト化・検索実証実験

(目的)

- (1) 視覚障がい者等の利用向上のために、デジタル化資料のテキスト化(画像を文字列へ変換)、デジタル化資料及び電子書籍等のテキストデータを用いた視覚障がい者等への読上げサービスに係る実証実験。
- (2) デジタル化資料及び電子書籍等のテキストデータを用いて、従来のメタデータ(データの作成者、作成日時、属性等の記録)検索にとどまらない、全文検索サービスの実証実験。

○出版社等に広く協力を呼びかけ、電子書籍データなどのデジタル出版データを基にした全文テキスト検索に関する技術的課題の検証と、過去から現在までの資料の統合的な全文テキストデータ検索の実証実験を行う。

大規模デジタル化の方法

【原資料から】

- オーバヘッド方式のスキャナ
- 資料に対して光学解像度400dpi
- 24ビットフルカラー

【マイクロフィルムから】

- A3サイズ400dpi
- 8ビットグレイスケール

【共通】

- JPEG2000(保存用、提供用)
- サムネイル画像
- 目次情報の入力
- 本文のテキスト化(OCR)は実施しない



原本保存のためのデジタル化(平成22年1月施行)

□ 国立国会図書館における所蔵資料の電子化

(平成21年著作権法の一部改正の内容)

国立国会図書館においては、所蔵資料を納本後直ちに電子化することができる。

⇒ 従来は、劣化・損傷している場合に限定(31条2号)

(改正の趣旨)

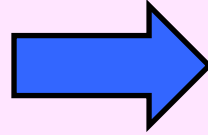
所蔵資料が損傷・劣化する前に電子化し、原資料を文化的遺産として保存できることが重要

デジタル化による資料保存のイメージ

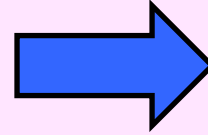
デジタル化
しない場合



図書館資料



閲覧



資料の劣化
利用不能

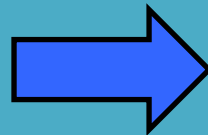
デジタル化
した場合



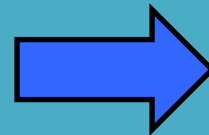
図書館資料



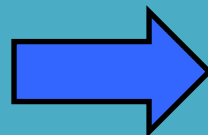
デジタル化 = 原資料の代替



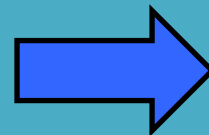
書庫で保管



原資料の保存



閲覧



将来の閲覧

「関係者協議会」での検討

□ 資料デジタル化及び利用に係る関係者協議会 (平成20年9月～)

- 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設置
- デジタル化資料の利用提供について、権利者の経済的利益や出版ビジネスへ配慮する必要
- 館内利用の基本要件、実施手順等について合意を得る

基本方針(「関係者協議会」第一次合意 平成21年3月)

□ 基本方針

- (1) 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は、画像データの作成を当面の範囲とする。
- (2) 検索利用等を目的とした資料の「テキスト化」の実施については、今後の検証事業等の結果を踏まえて、あらためて、関係者との協議により方針を定める。
- (3) デジタル化の実施に際しては、権利者を始めとする関係者の理解と協力を得るように努め、民間の市場経済活動を阻害することがないように十分に留意する。

館内提供(「関係者協議会」第一次合意)

□ 館内提供の基本要件

- 来館利用者は、特定の端末からデジタル化資料を利用。
- 同一文献に対する同時利用は、資料の所蔵部数を超えない範囲とする。(1冊であれば、同時には1人)
- 複写は、プリントアウトのみ(デジタル複製は行わない。)
- 著作物の適正利用の注意喚起のため、プリントアウトに際して、出所等を印字する。
- 端末に一時複製されたファイルは利用終了後、すみやかに破棄する。
- コンテンツの流出(違法複製等)防止のため、外部のネットワークと完全に遮断する。

利活用に向けた検討課題

デジタル出版物についての図書館サービスと民間ビジネスの境界範囲、利用ルールの明確化が必要

- ① 図書館間の資料相互貸借(市場入手困難資料)に相当する利用の方策(公共図書館館内への限定送信等の検討)
- ② 「画像」からテキスト形式データの作成(視覚障がい者等への対応、全文検索の利用)
- ③ 遠隔地での利用モデル(有償配信)の検討



近代デジタルライブラリー

- 国立国会図書館所蔵の明治期・大正期・昭和前期刊行図書をデジタル化して収録
- 著作権処理を行ってインターネット提供
 - 著作権保護期間満了
 - － 著作権者の許諾
 - － 文化庁長官裁定
- インターネット提供できない資料を館内提供

提供資料数

- インターネット提供資料 約17万冊
 - 明治期約13万冊(所蔵数の8割)、
大正期約4万冊(所蔵数の4割)
 - 館内提供資料 約22万冊
 - 明治期約3万冊、大正期約5万冊、
昭和前期約14万冊(大規模デジタル化の成果)
- ※館内では、インターネット提供資料と館内提供資料を合わせた約39万冊を利用可能

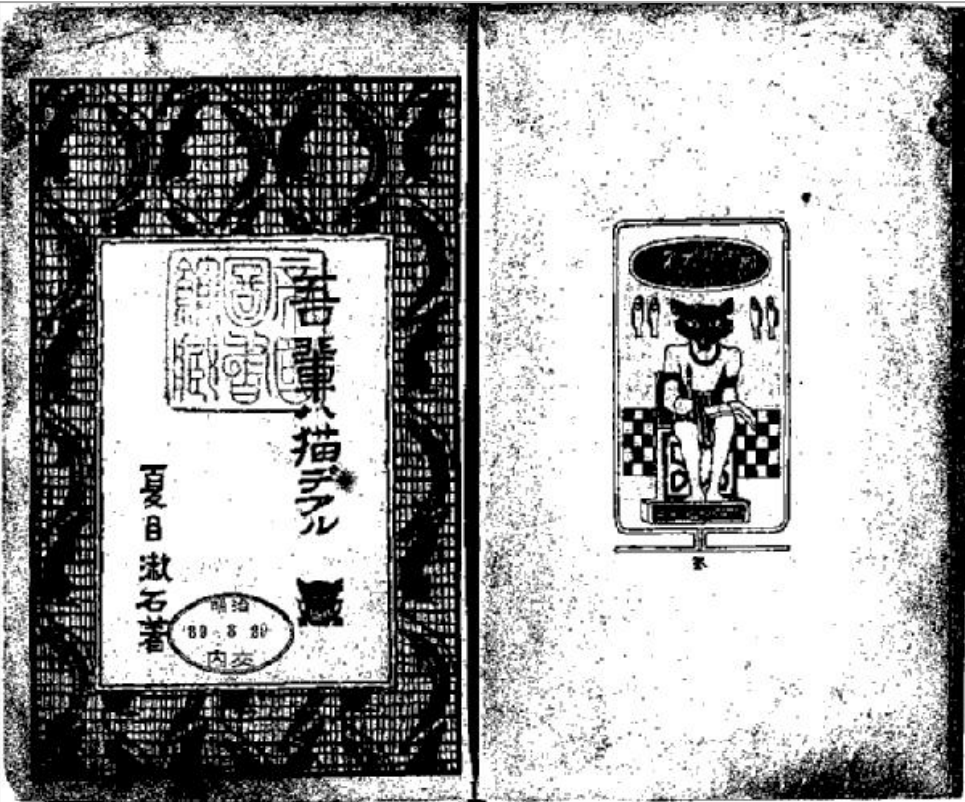
平成22年9月現在

閲覧できる資料

『吾輩ハ猫デアル 上』

夏目漱石著

大倉書店 明38-40



『ハムレット』

シェークスピア著，坪内逍遙訳

早稲田大学出版部 明42

利用方法

- アクセス

国立国会図書館HP (<http://www.ndl.go.jp/>)

→電子図書館

→近代デジタルライブラリー (<http://kindai.ndl.go.jp/>)

The screenshot shows the homepage of the National Diet Library (NDL). The header includes the logo and name '国立国会図書館 National Diet Library'. Navigation links include 'トップ', 'よくあるご質問', 'ヘルプ', 'サイトマップ', and a search box. A secondary navigation bar lists various services like '利用案内', '東京本館', '関西館', etc. The main navigation bar has tabs for '国会サービス関連情報', '資料の検索', '調べ方案内', '電子図書館', '電子展示会', and '刊行物'. The '電子図書館' tab is highlighted with a red box, and a dropdown menu is visible below it, listing '近代デジタルライブラリー', '貴重書画像データベース', '児童書デジタルライブラリー', 'インターネット資料', 'デジタルアーカイブポータル', and '国立国会図書館サーチ(開発版)'. The page also features a '新着情報' (New Information) section with an RSS icon and a 'スポットライト' (Spotlight) section.

近代デジタルライブラリー

Digital Library from the Meiji Era

国立国会図書館

国立国会図書館が所蔵する明治・大正期刊行図書をデジタル画像で閲覧できるサービスです。

お探しの作品、作者、テーマなどを入力してください

フランダースの犬

検索

インターネットで画像閲覧可能な資料に限定する

▶ 詳しく検索

項目ごとに条件を
指定して検索します

▶ 分野で検索

日本十進分類法
(NDC)で検索します

お知らせ



NEW 2010年8月2日

アンケート(全16問)にご協力ください(平成22年8月2日~9月30日)。

- 利用状況を調査し、今後のサービス改善に役立てることを目的としています。
- アンケート結果は統計的に処理します。アンケートの目的以外に使用することはありません。
- NEC社のCSVIEW/Webアンケートサービスを利用しています。

NEW 2010年7月27日

約14,000冊(約11,000タイトル)の資料を追加公開しました。(追加資料の紹介)

- ・大正期の資料は約14,000冊(約11,000タイトル)を追加しました。大正期追加資料のリストはCSVファイルをご覧ください。
 - ・明治期の資料は約100冊(約70タイトル)を追加しました。明治期追加資料のリストはCSVファイルをご覧ください。
- 今回の追加公開により、近代デジタルライブラリーでの提供総資料数は約170,000冊(約119,000タイトル)となりました。

▶ お知らせ一覧へ

資料あれこれ



近代デジタルライブラリーで提供している本の一部をテーマごとに紹介します

▶ 資料あれこれへ

▶ このデータベースについて ▶ よくある質問 ▶ お問い合わせ ▶ ヘルプ

国立国会図書館



Adobe Readerのダウンロードは左のアイコンからできます。

検索結果 (Hit)

検索結果:[1件中 1 - 1件を表示]

件ずつ表示

1. フランダースの犬 / ウイダ(ルイス・デ・レミイ) 著他, 内外出版協会, 明41. 11

[↑このページの先頭へ](#)

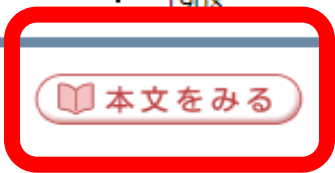
書誌情報

ページの中ほどにある「本文をみる」ボタンをクリックすると、この書誌の本文画像を見ることができます。

検索結果： [1件中 1件目を表示]

[検索結果に戻る](#)

タイトル : フランダースの犬
タイトルよみ : フランダース ノ イヌ
責任表示 : ウイダ(ルイス・デ・レミイ) 著, 日高善一(柿軒) 訳
出版事項 : 東京:内外出版協会, 明41.11
形態 : 109p;20cm
NDC分類 : 930
著者標目 : Ouida(1839-1908);日高, 善一(1879-1956)
著者標目よみ : Ouida(1839-1908) ヒダカ, ゼンイチ(1879-1956)
全国書誌番号 : 41002991
請求記号 : YDM101356
西暦年 : 1908



この図書の著作権情報

| 著者名 | 著作権 | 裁定年月日 |
|-----|-----|-------|
|-----|-----|-------|

サムネイル



1/67

22%

765

× 600

変更

リセット

設定

ヘルプ



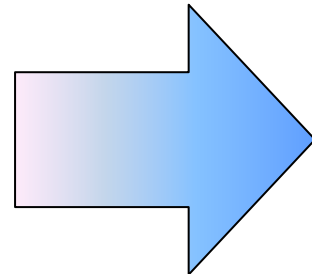
表示オプション

インターネット資料収集保存事業

- 紙資料では入手できない情報がある
 - インターネットだけの情報
 - 紙から移行した電子版資料
- インターネットは情報が失われやすい
 - 内容変更が頻繁（更新、消去）
 - サイト自体の閉鎖



ウェブサイト



インターネット資料収集保存事業とは



- 定期的にサイトを収集し、変更前の情報を保存していく
- 失われやすいインターネット情報を、紙の資料と同様に後世に残していく

失われたウェブサイトを当事業で保存

- 2002年FIFAワールドカップ（日本組織委員会）
- アジア野球選手権2003
- 世界陸上2007大阪大会
- 防衛庁（現・防衛省）
- 加茂町（現・木津川市 京都府）
- 高松町（現・かほく市 石川県）

これまでの収集

- 平成14年～平成22年3月（WARP時代）は
収集や提供の許諾を個別に文書で依頼

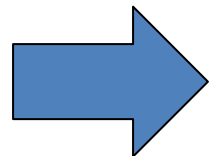
- ・公的機関
- ・電子雑誌
- ・大学
- ・イベントのウェブサイト

応諾率 約60%

法律改正後の収集 平成22年4月～

- 公的機関のウェブサイトは、当館が著作権者の許諾なく収集できる
- 公的機関以外（イベント等）のサイトは引き続き許諾を得て収集する
- 収集頻度の増加

年1回



国の機関は月1回

公的機関は年4回

その他機関は年1回

データ量

■ H14年 ~H22年3月まで ~H22年7月まで

| | | |
|--------|----------|----------|
| タイトル | 4,599 | 4,661 |
| 総件数 | 22,000 | 22,400 |
| ファイル数 | 1億6,000万 | 1億8,100万 |
| 容量(GB) | 15,000 | 18,000 |

■ H22年度末予測(前年比)

| | |
|--------|--------------------|
| タイトル | 4,785 (+186) |
| ファイル数 | 2億3,000万 (+7,000万) |
| 容量(GB) | 25,000 (+10,000) |

収集データの利用

- 書誌情報にあたるものを手動で付与
- 当館内のパソコンで閲覧可
- 著作権者の許諾が得られれば
 - 当館内での複写(プリントアウト)利用
 - 当館ホームページからインターネット公開

NDLのデジタル化関連事業：

現状のまとめ

- 著作権処理の可能な範囲までの資料
- 関係者との調整をはかりながらデジタル化

◎過去の出版物をもとにした「調査」のインフラとしては整備されつつある

☆市場性のある領域の資料の提供（著作権処理）には限界

☆「読書」＝「本との出会い」のインフラ足り得るために、今後必要なことは何か??

「読書」のイメージ

(2010年10月20日国民読書年記念シンポジウム資料より)

本に書かれている
「情報」「内容」

本の
「形」「物質性」



著者の「意図」

理解・思考

想像的世界の
創造・構築

楽しみ・娯楽
心の豊かさ

子どもと読書を取りまく要素

(イメージ図)

第一回調査会キーワードによるまとめ
平成16年12月15日
(第二回調査会資料から転載)

社会生活を営むための能力

自己決定

大人の情報行動

情報リテラシー

自立した大人

職業・社会生活

キャリアディベロップメント

「子どもの本」の
読書

美術館、博物館

知性

科学する心・おどろき・感動

異文化理解(異なる文化との接触)

感性

楽しさ・遊び

リアルな体験(現実、日常の体験)と時間をこえた本の中の体験(現在、過去、外国)

体験・経験

心の豊かさ
読書の楽しみ

伝える

コミュニケーション

主体としての子ども

調べる力

理解する

ハンディキャップ児童

読む力

創造性

中高生の読書

専門性確保

読み聞かせ・口承・音読

読書

バリアフリー(障害・地域)

地域格差

バリアフリー資料

様々なメディア(映像、動画等)

雑誌

本

著作権

財政難

公共図書館

文庫

学校図書館

インターネット

研究・評論・教育関係

国際連携
学術情報の生産と流通、研究発表
司書教育、専門職の養成

資料・情報の保存

デジタルで発表される作品の保存

子どもの文化

情報のつくり手

出版・編集
作家・ライター
画家・イラストレーター・写真家
翻訳家・通訳など

書店

手渡す人(仲介者)

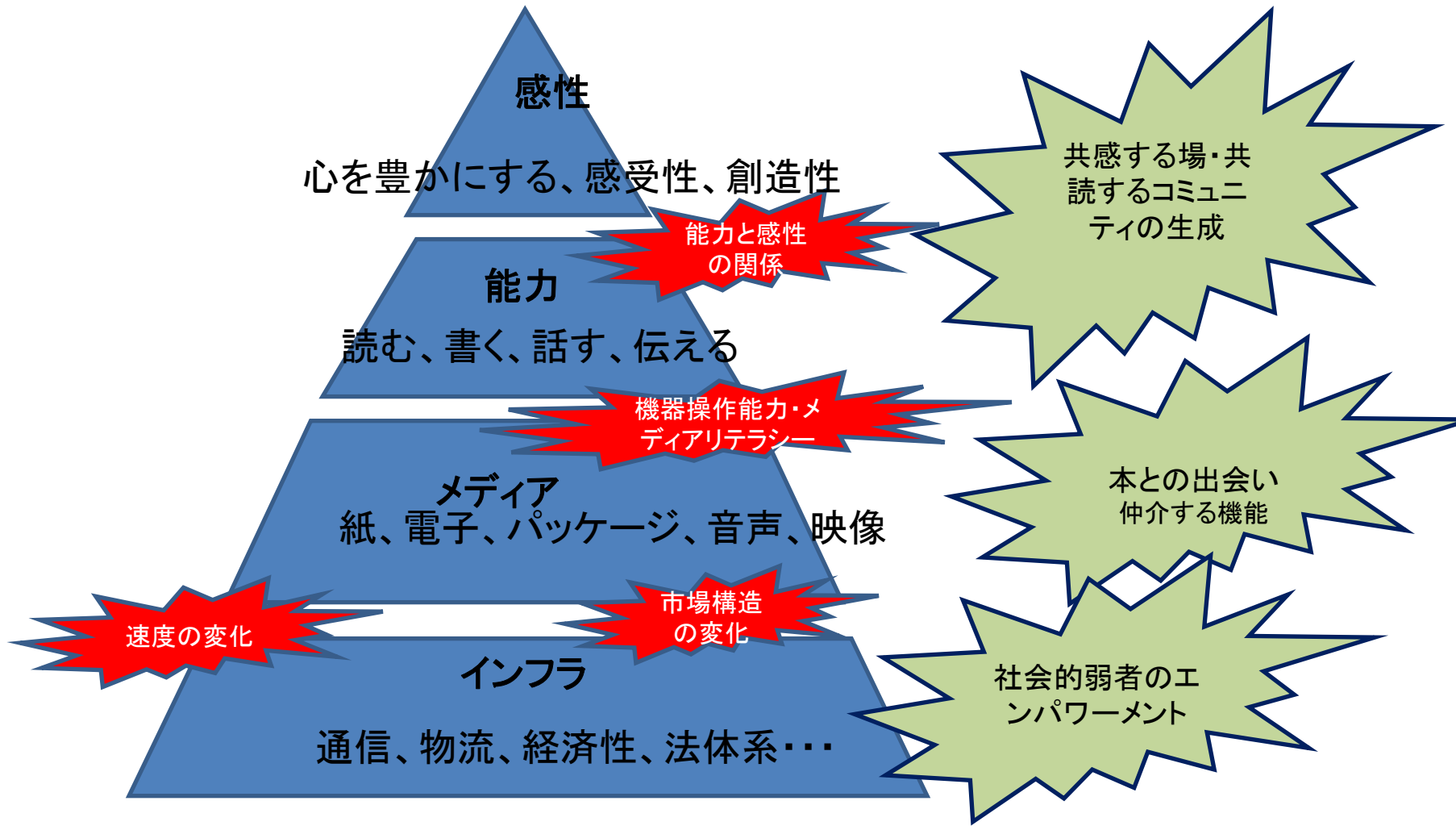
保育者、司書、教諭
専門ではないが媒介する人
ボランティアなど

図書館が担える分野

読書のレイヤー

(大人・一般)

「読書とは何か」シンポメモに加筆
2010年10月20日(岸)



読書推進と図書館の役割

- 資料への仲介
おすすりめ本、おすすりめコーナー
(リアルでも、web上でも)
- 共読コミュニティの発生・持続を支援
- 国民に等しく開かれた場としての公共性