

人文学のオープンデータ事例

東京大学史料編纂所 久留島典子

事例1

国立国会図書館 NDL

古典籍資料（貴重書等）

(97887点)、憲政資料(3981点)・・・

図書資料に準じた記載情報
→種々の資料に最適化した
情報付加の必要

<http://dl.ndl.go.jp/>



事例 2 - 1

国立公文書館

<https://www.digital.archives.go.jp/>

階層構造を示すなど、資料の特性に応じた情報を持つ



The screenshot shows the National Archives of Japan Digital Archive website. The search results for '公文録' are displayed as follows:

タイトル	第一編 公文録
年次	明治12年度、明治13年度、明治14年度、明治15年度
件数	1,000件 (明治12年度) - 1,000件 (明治15年度)
作成機関	内閣府(内閣府)
形式	本館蔵に於いて複製した公文書(複製) - 本館蔵に複製した公文書(複製)、「公文録(複製)」...

共有権 第一編 1,000件 (1 - 100)

検索結果一覧

No.	公文録
1	公文録 - 明治12年
2	公文録 - 明治13年
3	公文録 - 明治14年
4	公文録 - 明治15年
5	公文録 - 明治16年

事例 2 - 2

アジア歴史資料センター <https://www.jacar.go.jp/>



2019/6/9

事例 3

京都府立京都学・歴彩館

<http://hyakugo.kyoto.jp/>

『東寺百合文書目録』(1976-1979年刊行)のデータをもとに画像公開

原本24,067通の保管と
目録・画像データの公開

2019/6/9

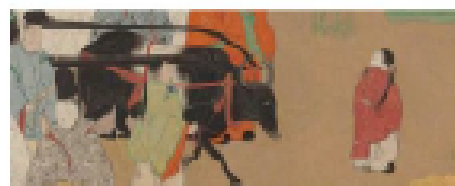
事例 5

情報・システム研究機構人文学オープンデータ共同利用センター

<http://codh.rois.ac.jp/>

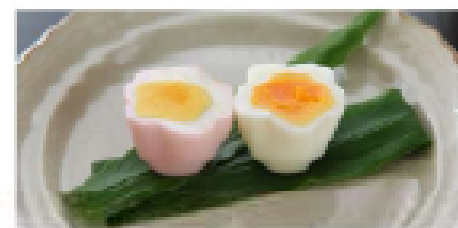
データセット一覧

データセットの提供



日本古典データセット

歴史的典拠データベースにおいてデジタル化された古典籍のうち、主に国文学研究館蔵本を対象に、画像データと音読データセットで公開しています。さらに一部の古典籍には作品紹介や複製テキストデータ、タグ情報なども付与しています。



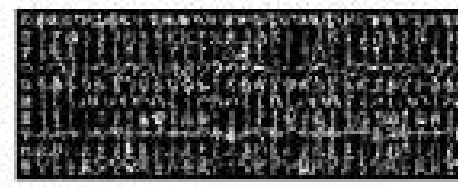
江戸料理レシピデータセット

日本古典籍データセットに含まれる江戸の料理本を対象に、江戸の料理文化に関するデータとして、書籍・地域・部材・レシピ化という作業を加えたレシピデータを提供します。



日本古典縦くずし字データセット

日本古典籍データセットで公開するデジタル化された古典籍を対象に、複製テキストの制作過程で生じるくずし字のなり出された字形と書体情報などを、複製の人間



KMNISTデータセット

日本古典縦くずし字データセットを元に、書体学研究室で有名なMNISTデータセット互換のくずし字データセットKMNI1を作成しました。目下において、Mandarin-MNIST、MNIST-10、Mandarin-MNISTの3種類のデータ

事例 6

東京大学史料編纂所

<https://www.hi.u-tokyo.ac.jp/index-j.html>

- ・ 他機関所蔵資料
の目録公開
リンクして画像公開
- ・ 種々のデータ
ベース
→ 整理の必要



日本における人文学系データ公開の特徴と課題 1

• 資料

所蔵機関による公開が中心 事例 1, 2, 3

画像が先行、全文データは限られている

テキスト作成の努力→データセット（事例 5）

→電子辞典（事例 6）

• 研究成果

オープンな公開は機関リポジトリ中心

各機関、IRD (<https://irdb.nii.ac.jp/>) 収載などは少数

雑誌論文・著書は課題

（目録はCinii (<https://ci.nii.ac.jp/>) でほぼオープンな公開実現）

辞典類等の一部は商業的に電子的公開 ジャパンナレッジ Lib

日本における人文学系データ公開の特徴と課題 2

- 研究経過の公開？ 公開以前の保存・記録？



- 目録・画像・テキストに関連知識・文献等がリンク
データを分析できるデータベース 事例4 分野が限定

【参考】海外の人文学系研究成果のオープン化事例

Open Library of Humanities. (OLH) <https://www.openlibhums.org/>

2013年に英国・米国の大学研究者3名が立ち上げ、APCではなく、図書館コンソーシアムからの出資金で運営。2017年現在14種のオープンアクセスジャーナルを発行。900超論文掲載

Knowledge Unlatched (KU) <http://www.knowledgeunlatched.org/>

2012年に英国で設立、図書館からの出資金で、人文社会科学系学術書のオープンアクセス化に取り組む

*参考文献：東京大学附属図書館編『オープンアクセスハンドブック』2017年

<http://hdl.handle.net/2261/72694>