



デジタル化資料の全文テキストデータの 視覚障害者等への提供について 【事業の概要】



本事業の目的

- 国立国会図書館のデジタル化資料は、画像データで提供しています。
- 画像データの形では、画面の情報を音声で読み上げて利用する**視覚障害者等の方は利用できません**。
- 一方、デジタル化資料は、全文検索用にOCR処理によって「全文テキストデータ」を作成しています。

この全文テキストデータを、国立国会図書館が2014年から行っている**視覚障害者等の方向けのデータ送信サービス**で**配信する**という事業になります。

1. 全文テキストデータの概要

全文テキストデータとは

- 当館が提供するデジタル化資料（画像データ）から、OCR（光学的文字認識）処理により作成したテキストデータです。
- 現在提供しているのは、2020年12月時点でデジタル化資料を提供していた図書、雑誌、博士論文等の**約247万冊分**です。

※当館の資料デジタル化については、以下の当館ホームページをご参照ください。

資料デジタル化について

<https://www.ndl.go.jp/jp/preservation/digitization/index.html>

現在の全文テキストデータ提供件数

種別	内容	点数（概数）
図書	<ul style="list-style-type: none">・ 1968年までに受け入れた図書・ 震災・災害関係資料の一部（1969年以降受入分も含む）	97万点
雑誌	<ul style="list-style-type: none">・ 明治期以降に刊行された雑誌 （刊行後5年以上経過したもの）	132万点
博士論文	<ul style="list-style-type: none">・ 1990～2000年に送付を受けた論文	15万点
その他	<ul style="list-style-type: none">・ 官報等	2万点
計		247万点

全文テキストデータを提供している資料

- 1990年以降に刊行された図書は約3,000点、雑誌は約30万点です。
- 全体で見ても、約247万点のうち、1990年以降に刊行された資料は約44万点（17.8%）です。
- 逆に、1969年までに刊行された資料が全体の6割以上を占めています。

現在、全文テキストデータを提供している資料は古い年代が中心です。

2. 全文テキストデータの利用

視覚障害者等用データ送信サービスと「みなサーチ」

- 国立国会図書館は、点字データ、DAISYデータ等を、インターネットを通じて配信するサービス「視覚障害者等用データ送信サービス」を、2014年から実施しています。
- この「視覚障害者等用データ送信サービス」を通じて、著作権法第37条第3項に基づき、視覚障害者等の方に限って全文テキストデータを提供しています。
- 全文テキストデータは、OCRをかけたままの形で提供しています。
- サービスの窓口となるのが、2024年1月に正式版を公開した「**みなサーチ**」（国立国会図書館障害者用資料検索）です。



みなサーチ

国立国会図書館障害者用資料検索

(参考) 著作権法第37条第3項

(視覚障害者等のための複製等)

第37条

1・2 (略)

3 視覚障害その他の障害により視覚による表現の認識が困難な者（以下この項及び第百二条第四項において「視覚障害者等」という。）の福祉に関する事業を行う者で政令で定めるものは、公表された著作物であつて、視覚によりその表現が認識される方式（視覚及び他の知覚により認識される方式を含む。）により公衆に提供され、又は提示されているもの（当該著作物以外の著作物で、当該著作物において複製されているものその他当該著作物と一体として公衆に提供され、又は提示されているものを含む。以下この項及び同条第四項において「視覚著作物」という。）について、専ら視覚障害者等で当該方式によつては当該視覚著作物を利用することが困難な者の用に供するために必要と認められる限度において、当該視覚著作物に係る文字を音声にすることその他当該視覚障害者等が利用するために必要な方式により、複製し、又は公衆送信を行うことができる。ただし、当該視覚著作物について、著作権者又はその許諾を得た者若しくは第七十九条の出版権の設定を受けた者若しくはその複製許諾若しくは公衆送信許諾を得た者により、当該方式による公衆への提供又は提示が行われている場合は、この限りでない。

全文テキストデータの利用方法



ヘルプ ログイン

 **みなサーチ**
国立国会図書館障害者用資料検索

簡易検索 すべての資料を検索します



校正済のデータに絞る ダウンロードできるデータに絞る

いろいろな検索

詳細検索 >
細かい条件を指定して検索

全文検索 >
本の内容をキーワードで検索

ジャンル検索 >
気になるジャンルを探そう

全文テキストデータの利用方法

みなサーチ
国立国会図書館障害者利用資料検索

ヘルプ ログイン

全文検索

テキストデータのある本の本文を対象にして、キーワードで検索します。

キーワード 本のタイトルや著者

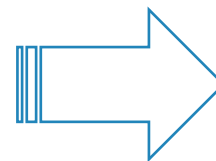
光る君

キーワード「光る君」の検索結果2,253 件中 1~20件 (タイトル:昇順)

結果を絞り込んだり並び替えたりするには [詳細検索](#)

◀ ◁ ▷ ▶

番号	タイトル	資料形態	著者	原本の出版者	原本の出版年	製作館	検索にヒットした箇所
1	 十五世市村羽左衛門舞台写真集 未校正	プレーンテキスト	菱田正男 編	舞台すがた社	1951	国立国会図書館	ログインしてダウンロード
本文 松王丸「増補忠臣蔵」の桃井若狭之助「新版歌祭文」の丁稚久松「箱書附魚屋茶碗」の娘の次郎吉三語一の 光る君 「柳巷春着薊色縫一の極楽寺所化■山婦女庭訓」月歌舞伎座「時近江春桃」の三上太郎「妹背(吉野川)の久...郎吉三月歌舞伎座「時近江春桃」の三上太郎「妹背山婦女庭訓」(吉野川)の久我之助■船「源氏物語」の 光る君 「柳巷春着薊色縫一の極楽寺所化■心後に鬼薊■吉四月東京劇場(初開場)「壽式三番叟」の千歳「義経千本...							
2	 6年の学習 9(7) 未校正	プレーンテキスト	学習研究社	1954-10	国立国会図書館	ログインしてダウンロード	
本文 都のだれ知らぬものもありませんでした。cat 光る君 は、時の天皇の二番めの王子で、源といつかうみょうじ...ありました。かれ「あれは、だれの住む家だろう。」ひか丘と、 光る君 がききますと、おともの人は、ほ「あれは、光る君さまもごぞんじ...なみだをぬぐっています。あま君は、ひめのおばあさまなのだ、 光る君 は思いうかべました。かなひめも、おばあさまのなくさまを見ると...							



ログイン

利用者ID

パスワード

ログインしたままにする

ログイン

[登録利用者IDをお忘れですか？](#)

[パスワードをお忘れですか？](#)

視覚障害者等利用者としてのIDとパスワードを入力する。

全文テキストデータの使用方法

 みなサーチ
国立国会図書館障害者用資料検索

ヘルプ 利用者情報

視覚障害者等用データ送信サービスの利用に係る注意事項

視覚障害者等用データ送信サービスは、国立国会図書館が各機関から収集した視覚障害者等用データと国立国会図書館が製作した視覚障害者等用データを、インターネット経由でダウンロードしてご利用いただけるサービスです。以下の注意事項をご確認の上、サービスをご利用ください。

- 1 本サービスをご利用いただける方は、あらかじめ当館から利用の承認を受けた、視覚障害その他の理由で通常の活字の印刷物の読書が困難な方及び図書館等です。
- 2 視覚障害者等用データは、視覚障害その他の理由で通常の活字の印刷物の読書が困難な方のために製作されたデータです。著作権法に定められた権利制限規定に該当する場合を除き、複製等による第三者への提供を固く禁じます。
- 3 本サービスの利用にあたって使用するID及びパスワードは厳重に管理し、ID及びパスワードを第三者に利用させたり知られたりしないようご注意ください。また、登録情報に変更があった場合は、速やかに変更を行ってください。
- 4 業務上、特に必要があると認めるときは、臨時に、本サービスの一部又は全部を休止する場合があります。

OK



検索結果 2,253 件中 1~20 件 (タイトル:昇順)

キーワード「光る君」の検索結果2,253 件中 1~20件 (タイトル:昇順)
結果を絞り込んだり並び替えたりするには [詳細検索](#)

番号	タイトル	資料形態	著者	原本の出版者	原本の出版年	製作館	検索にヒットした箇所
1	十五世市村羽左衛門舞台写真集	プレーンテキスト	菱田正男 編	舞台すがた社	1951	国立国会図書館	ダウンロード

視覚障害者等として初めてログインまたは初めてダウンロードする場合は、サービスの利用にあたっての注意事項の画面が表示され、確認ボタンを押下しなければダウンロードできない。

対象利用者①

- 全文テキストデータは、以下の視覚障害者等個人の方、国立国会図書館に登録している図書館等に配信します。
- 各図書館では、障害者手帳等により視覚障害者等に該当することを確認した方にのみ、配信を受けた全文テキストデータを提供します。

①視覚障害者等個人

障害者手帳や医師の診断書等の障害を示す文書、又は直接本人を目視で障害状況を確認する等により、視覚障害者等に該当する者として当館が確認した上で、サービスの利用者登録をした個人

②図書館等

視覚障害者等へのサービス提供体制が整備されていることを書面により当館が確認し、サービスを利用することを承認した図書館等（=送信承認館）

対象利用者②

視覚障害者等
登録利用者数

877名
(2024年8月末現在)

送信承認館数

353館
(2024年8月末現在)

公共図書館	192
(都道府県立)	(40)
(市区町村立)	(152)
大学図書館	80
視覚障害者情報提供施設	27
学校図書館	49
その他	5

※送信承認館の一覧は「各サービスの承認館・参加館一覧」

(https://www.ndl.go.jp/jp/library/supportvisual/supportvisual_partic_1.html) を参照。

利用者に対するアナウンス①

- 利用者には、全文テキストデータがOCRをかけたままのデータであることを、きちんと理解した上で利用していただきたいと考えています。
- 図書館等には、全文テキストデータを素材として新たな視覚障害者等用資料を製作する場合は、きちんと校正を行った上で利用者に提供することを推奨しています。

利用者に対するアナウンス②

全文テキストデータの冒頭には、以下の趣旨のアナウンスを挿入し、注意事項を理解した上での利用を求めます。

- 全文テキストデータが著作権法第37条第3項の規定に基づき、視覚障害者等に限定して提供するためのものであり、視覚障害者等以外は利用できないこと
- 校正されていないテキストデータであり、OCRで読み込んだ後、人手による確認を経していないこと
- 未校正のデータであるため、校正を行った上での新たな視覚障害者等用データの製作を推奨すること

全文テキストデータの利用数

- 全文テキストデータのアクセス件数は、2024年1月～8月の総件数で約2,000件程度でした（月平均約250件）。
- DAISYや点字データなど、その他の視覚障害者等用データへのアクセス件数は、約71,000件／月であることを考慮すると、全文テキストデータの利用が全体に占める割合は、まだごくわずかです。
- ただ、実際に利用した当事者の方からは、アクセスできる資料が飛躍的に増えたことを歓迎する声が寄せられています。
- また、視覚障害者情報提供施設や学校図書館の送信サービスへの参加が増えつつあるなど、膨大な全文テキストデータを、視覚障害者等に提供する選択肢の一つとして活用していこうとする動きが見えています。

全文テキストデータに寄せられた感想等

- これまでは調べものをする際に古い資料はあきらめていたが、全文テキストデータによって、内容を確認して、自分で新たな資料を見つけることができた。
- 全文テキストデータとして提供されているものは、点字やDAISYの形式では提供されていない学術論文や雑誌が多いので、利用価値が高い。
- 自宅にいながらアクセシブルな資料を多く利用できるようになり、ありがたい。
- （視覚障害者情報提供施設から）利用者から全文テキストデータを校正して、校正済みのテキストデータまたはテキストDAISYを作成してほしいとの要望があった。
- （学校図書館から）発達に特性がある児童が、自分の興味のある分野について学びを深めていく際に、音声で利用できる資料となると限られたものしかなかった。全文テキストデータであれば、様々な資料が提供されているので可能性が広がる。

3. 新たな全文テキストデータの提供

新たな全文テキストデータの作成

- これまでに全文テキストデータを作成した資料（約247万点）は、2020年12月までに当館がデジタル化していた資料でした。
- 2021年から2023年3月末までにデジタル化した資料を対象として、2024年1月から新たにテキスト化を開始しました。
- 今回のテキスト化対象資料は、1969年～1987年（一部1995年までのものも含む）に出版された図書のほか、雑誌、博士論文なども含んだ約75万点です。

コレクション	対象資料数(点)
図書	656,000
雑誌	72,000
博士論文	21,000
その他	100

- 約75万点のテキスト化は、2025年3月末までに完了する予定です。

視覚障害者等の方への全文テキストデータの提供

- 新たに作成した全文テキストデータは、視覚障害者等への提供に係る確認手続を実施した上で、2025年1月から順次、みなサーチにおいて提供を開始する予定です。
- 確認手続では、当館だけではなく、出版社様の御協力を賜りまして、後述の除外基準に合致するデータがないかを確認いたします。

4. 提供対象から除外されるデータ

提供対象から除外されるもの

著作権法第37条第3項ただし書では、既に同じ方式の視覚障害者等用資料が市場に流通している場合には、著作権者等の許諾を得ずに複製等を行うことは許容されない旨を規定しています。

⇒全文テキストデータのうち、**ただし書に該当するものを提供対象から除外します。**

(参考) 著作権法第37条第3項 (下線部がただし書)

(視覚障害者等のための複製等)

第37条

1・2 (略)

3 視覚障害その他の障害により視覚による表現の認識が困難な者（以下この項及び第百二条第四項において「視覚障害者等」という。）の福祉に関する事業を行う者で政令で定めるものは、公表された著作物であつて、視覚によりその表現が認識される方式（視覚及び他の知覚により認識される方式を含む。）により公衆に提供され、又は提示されているもの（当該著作物以外の著作物で、当該著作物において複製されているものその他当該著作物と一体として公衆に提供され、又は提示されているものを含む。以下この項及び同条第四項において「視覚著作物」という。）について、専ら視覚障害者等で当該方式によつては当該視覚著作物を利用することが困難な者の用に供するために必要と認められる限度において、当該視覚著作物に係る文字を音声にすることその他当該視覚障害者等が利用するために必要な方式により、複製し、又は公衆送信を行うことができる。ただし、当該視覚著作物について、著作権者又はその許諾を得た者若しくは第七十九条の出版権の設定を受けた者若しくはその複製許諾若しくは公衆送信許諾を得た者により、当該方式による公衆への提供又は提示が行われている場合は、この限りでない。

除外基準

①原本を購入した視覚障害者等が出版者等に申し出た場合に、当該資料又は同内容の著作物のテキストデータの提供を当該出版者等が行う旨を、原本や出版者等が運営するホームページ等において公表している場合

②当該資料又は同内容の著作物が、スクリーンリーダーによる読み上げ（Text to Speech（TTS））に対応している等、視覚障害者等が支援技術を通じて利用できる方式により、電子書籍市場や出版者のホームページ等で流通している場合（おおむね1か月を目安として流通予定であることを公開情報により確認できた場合を含む）

※「同内容の著作物」とは、全文テキストデータの内容を全て含む著作物で、文庫版、作品集・全集、他出版社からの再版等を含みます。

※当該資料又は同内容の著作物のオーディオブックが流通している場合は、暫定的に全文テキストデータの提供を留保します。

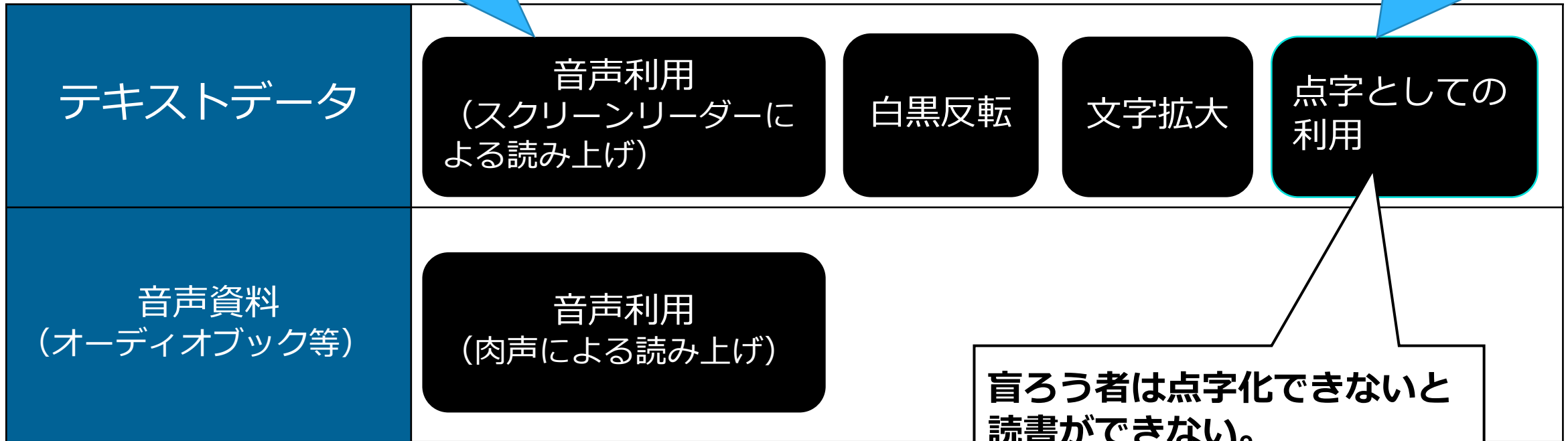
※この除外基準は、各図書館等で視覚障害者等用資料を製作する際の、著作権法第37条第3項ただし書該当資料を確認するための基準ではありません。

なお、経済的利益以外の正当な理由（人権侵害、個人情報保護等）により、提供停止のお申出があった場合は、別途対応いたします。

(参考) テキストデータと音声資料 (オーディオブック等)

詳細読みで1文字単位の漢字確認が可能

点字ディスプレイへの出力及び自動点訳ソフトによる点訳



確認手続のプロセス

四段階の確認手続

- ① スクリーンリーダーによる読み上げ対応の電子書籍の入手可能性調査
- ② 事前確認手続
- ③ テキストデータ提供の原本記載有無の確認
---①～③の確認手続を経て提供対象を決定し、提供を開始---
- ④ 事後確認手続（随時受付）

出版社の皆様には、特に提供開始前に行う②について、御協力を賜りますようお願い申し上げます。