

手書き資料の テキストデータ化のご案内

手書きで作成された資料は、分析・集計に利用しようと思うと、タイピング（入力）が必要なため、死蔵されてしまうことがあります。学術情報メディアセンターでは、深層学習による文字認識技術を活用し、手書き資料のテキストデータ化の事業を試行的に実施します。

募集期間： **2018年11月末** まで

テキストデータ化試行を希望される手書き資料がありましたら、わかる範囲で結構ですので、次の点についてお知らせ下さい。

- 1 資料が紙媒体か電子化画像か
- 2 1ページあたりの記入欄の数
- 3 資料の全数
- 4 記入欄内に複数行の記入があるか

◇資料が紙媒体であれば

- 5-1 紙のサイズ
- 5-2 紙の状態（紙の種類、劣化状態）

◇資料が電子化画像であれば

- 6-1 画像ファイル形式（JPEG/TIFF/PNGなど）
- 6-2 画素数、1ページあたりのデータ容量

裏面もご確認下さい



お問い合わせ
ご相談先

E-mail: mf-info@stat.media.kyoto-u.ac.jp / 内線：87464
学術情報メディアセンター 食料・農業統計情報開発研究分野
(担当 吉川、岡本)

手書き資料のテキストデータ化 対象資料についてのお知らせ

本事業が行う手書き文字認識によるテキストデータ化では、次のような資料を対象といたします。

本事業のテキストデータ化に適した手書き資料

- ・横書き資料
- ・日本語、ローマ字
- ・フォーマットが定まっており、決まった位置に文字がある
【イメージ】

	<p>このように、だめを完全に塗りつぶしてください</p>	<p>コンピュータの利用状況や「情報」に関する知識を問うて下さい。 回答結果は回答以降の授業に役立ちます。成績には関係ないので、深く考えずに思ったままを記入して下さい。</p>	<p>わ か ら な い</p>	<p>わ か ら な い</p>
インターネットがよく利用していますか。	<input type="radio"/>	CDの音楽をパソコンに保存したことがありますか。	<input type="radio"/>	<input type="radio"/>
パソコンを使う授業は好きでしたか。	<input type="radio"/>	DVDの映画やアニメをパソコンで見たことがありますか。	<input type="radio"/>	<input type="radio"/>
自分はパソコンの操作ができるほうだと思いますか。	<input type="radio"/>	インターネットの動画サイト（YouTubeやニコニコ動画などの動画を 見たことがありますか。	<input type="radio"/>	<input type="radio"/>
「ワープロ」を使って文章を作ったことがありますか。	<input type="radio"/>	インターネットのフリッターや掲示板に何か書き込んだことがありますか。	<input type="radio"/>	<input type="radio"/>
「表計算ソフト（エクセルや三四郎）で作成したことがありますか。	<input type="radio"/>	インターネットを通じて知り合いになった人がいますか。	<input type="radio"/>	<input type="radio"/>
コンピュータを使って絵を描いたことがありますか。	<input type="radio"/>	インターネットを通じて何か買ったことがありますか。 買ったことがある商品などいろいろの気楽に書いて下さい。	<input type="radio"/>	<input type="radio"/>
コンピュータを使って音楽を作ったことがありますか。	<input type="radio"/>	コンピュータファイルについて学んだことがありますか。	<input type="radio"/>	<input type="radio"/>
コンピュータを使ってメールを送ったり受け取ったりしたことがありますか。	<input type="radio"/>	自分で作ったホームページをインターネットに公開したことがありますか。	<input type="radio"/>	<input type="radio"/>
インターネットから絵や写真などもらったことがありますか。	<input type="radio"/>	自分のブログ（プロフィール）を作ったことがありますか。	<input type="radio"/>	<input type="radio"/>
絵や写真を入れた文章を作ったことがありますか。	<input type="radio"/>	自分も含めて、インターネットを使用して何かの練習があった人を知っていますか。	<input type="radio"/>	<input type="radio"/>
コンピュータを用いた簿記ソフトや表計算ソフトを作ったことがありますか。	<input type="radio"/>	自分の名前や写真をインターネットに掲載したことがありますか。	<input type="radio"/>	<input type="radio"/>
メールに絵や写真をつけてメールに送ったことがありますか。	<input type="radio"/>	「著作権」という言葉を学んだことがありますか。	<input type="radio"/>	<input type="radio"/>
音楽CDをパソコンで聴いたことがありますか。	<input type="radio"/>	これからの社会ではコンピュータの知識は必要だと思いますか。	<input type="radio"/>	<input type="radio"/>

初級申込書 【 】

1. 本人の住所を記入してください。
2. 所属する学校を記入してください。
3. この申込書は、保護者・指導員・担任先生の手紙と一緒に送って下さい。
4. 記入した内容は、保護者等に知らせません。
5. 届いた後、必ず返書が来るまで待ちます。
6. 届いた後、必ず返書が来るまで待ちます。

受講年月日 _____ 学年 _____ 年 _____ 月 _____ 日

氏名 _____ 性別 _____ 年齢 _____ 年 _____ 月 _____ 日 ()歳

現住所 〒 _____ 市 _____ 区 _____ 丁目 _____ 番 _____ 号

連絡先(自宅) 自宅 ☎ _____ 番 _____ 号

連絡先(学校) _____

※この申込書は、必ず本人の住所を記入してください。
※この申込書は、必ず本人の住所を記入してください。
※この申込書は、必ず本人の住所を記入してください。

※この申込書は、必ず本人の住所を記入してください。
※この申込書は、必ず本人の住所を記入してください。
※この申込書は、必ず本人の住所を記入してください。

テキストデータ化が難しい資料

- ・縦書き資料
- ・日本語、ローマ字以外の文字
- ・文字の位置が資料ごとに異なる

本事業は、株式会社 Cogent Labsの画像認識技術「Tegaki」により実施しております。コージェンラボの開発したエンジンを使用して下記の文章を認識した場合の認識率は99.22%となります。



なお、本取組みは、平成30年度総長裁量経費の支援を受けて実施いたしております。本学ご所属の方にご利用いただけます。

**不明点などございましたら、
表面のお問い合わせ・ご相談先までご連絡下さい。**