

高校生の作文検索インターフェース

このデータについて

1. 高校生の作文、約 500 作文が掲載されています。岐阜県の複数の高校から約 300 作文、大阪府の複数の高校から 100 作文、神奈川県複数の高校から 100 作文で構成されています。
2. 調査は高校三学年にわたって行われましたが、このインターフェースに掲載したのは全て高校 1 年生のデータです。
3. 基本的に調査に協力してくださった全ての高校から均等に作文を選択しています。
4. テキストサイズが均等に分散するように作文を選択しました。文字数は最大 1688 文字、最小 42 文字、平均 610 文字です。以下の図 1 は作文の文字数の分布状況です。

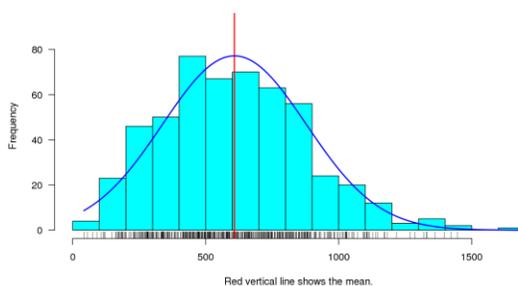


図 1 作文の文字数の分布

5. 5 つの課題について、約 100 作文ずつ収録されています。
6. 個人が特定されない作文だけを選択しました。個人が特定される情報が含まれている作文については、特定される恐れのある表現を★印に変換しました。
7. 実際に書いた作文を見ていただけるように、最も個人の特定性が低い描写文だけを pdf ファイルでダウンロードできるようにしました。

作文課題について

以下の5つのジャンルについて、一つずつ課題を設定しました。

表1 ジャンルと課題

ジャンル	課題指示
課題1 描写文 もの・人・状況・場所がどのようなものか、どのような印象を与えるか述べる	下のグラフを説明し、それが示すことについての分析・考察を述べなさい(グラフのタイトル:世界各地のCO ₂ 排出量の推移(1990-2017))。
課題2 課題文 書き手の主張が正しいこと、真実であることを、納得できるものであることを、与えられたテキストを引用したりしながら論じる	この4人の生徒は制服の廃止についていろいろな意見を述べています。ここに示された様々な意見のうちいくつかを取りあげて説明しながら、自分の意見を述べてください。その際、自分の意見とは異なる立場の意見を持つ人にも納得してもらえるように書いてください。タイトルも考えて作文を書いてください。
課題3 説明文 手順・定義・例・比較対照の結果・因果関係・解決方法などがどのようなものか説明する	あなたがよく知っている地域(日本、その他の国、あなたが住んでいる/いた町や都道府県など)を一つ選び、そこに行ったことがない人達にその地域の魅力を紹介する文章を書いてください。タイトルも考えて作文を書いてください。
課題4 意見文 書き手の主張が正しいこと、真実であることを、納得できるものであることを論じる	あなたは「高校生の意見」という作文コンクールに応募することになりました。テーマは、世の中(例えば、世界の国々や日本の政府)、学校、地域などに対して「変わってほしい・変えたいこと」です。タイトルも考えて作文を書いてください。
課題5 ナラティブ 何が起こったか、何・誰がどうしたか述べる	インターネットでエッセイを募集しています。あなたは、ペンネームで応募します。テーマは「これまでの人生で自分を変えたできごと」です。小さなことでもいいので、自分の体験をもとに書いてください。タイトルも考えて作文を書いてください。

検索インターフェースでできること

1. アンケート結果による作文の絞り込み機能

作文のアンケートの設問の回答で作文の絞り込みをすることができます。CTRL キーを押しながら項目をクリックすると、複数の項目での絞り込みができます。

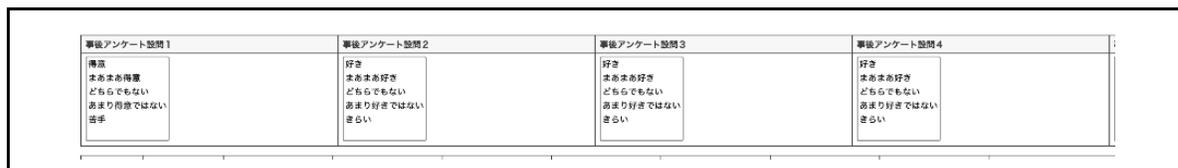


図2 アンケート結果による絞り込み検索画面

表2 アンケート内容

	内容	回答選択肢
設問1	作文は得意ですか、苦手ですか	得意、まあまあ得意、どちらでもない、あまり得意ではない、苦手
設問2	マンガを読むのが好きですか、嫌いですか	好き、まあまあ好き、どちらでもない、あまり好きではない、きらい
設問3	マンガ以外の本を読むのが好きですか、嫌いですか	好き、まあまあ好き、どちらでもない、あまり好きではない、きらい
設問4	人前で発表するのは好きですか、嫌いですか	好き、まあまあ好き、どちらでもない、あまり好きではない、きらい
設問5	作文は練習すると上手になると思いますか	思う、まあまあ思う、どちらでもない、あまり思わない、思わない
設問6	作文がもともと上手な人がいると思いますか	思う、まあまあ思う、どちらでもない、あまり思わない、思わない
設問7	この課題は簡単でしたか、難しかったですか	簡単だった、まあまあ簡単だった、どちらでもない、まあまあ難しかった、難しかった

課題による作文の絞り込み機能

上で示した5つの課題で作文を絞り込むことができます。「課題番号」の検索窓で選択します。

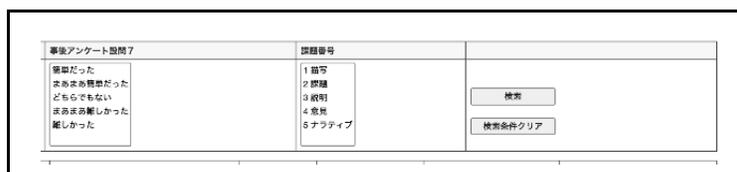


図3 作文課題による絞り込み検索画面

ダウンロード機能

作文データは、全てダウンロードできます。検索結果画面右側に番号.txtと表示されているリンクをクリックするとダウンロードできます。全てダウンロードできるのはデジタルテキストですが、課題1の描写文については、手書きの作文の pdf データもダウンロードすることができます。pdf をダウンロードしたい時は、図 X 左側の pdf アイコンをクリックしてください。

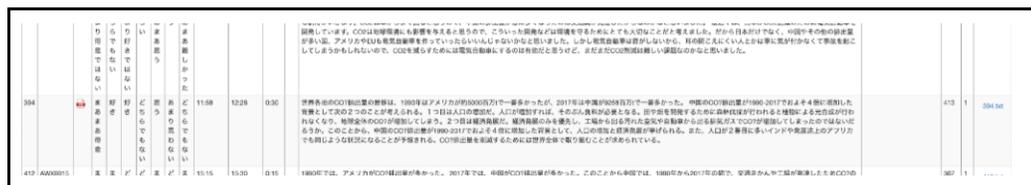


図 4 テキストダウンロード機能

文字数による検索機能

一定の文字数を設定し、その作文だけを抽出することができます。

文字数での絞り込み

文字数下限: ~ 文字数上限:

図 5 文字数での絞り込み機能

語検索機能

作文の中でどんな単語が使われているかを調べることができます。次の二つの方法があります。

文字列検索

一番上の検索窓です。文字列検索機能では、作文テキストを検索し、文字列が含まれる作文を一覧表示することができます。

日本人高校生作文データベース検索

【文部科学省委託】高等学校等における日本語能力評価に関する学術的調査研究事業

文字列検索：作文テキストを検索し、文字列が含まれる作文を表示します。

特別な意味を持つ記号が使用できます。記号の前

図 6 文字列検索画面

この検索では、以下の、特別な意味を持つ記号が使用できます。

[1] 半角のアンダーバー「_」は1文字の文字列を表します。

(例:新潟県_市と入力すると、「新潟県燕市」など、任意の1文字の市町村名が検索できます。「新潟県新潟市」「新潟県阿賀野市」は対象外になります)

[2] 半角のパーセント「%」は任意の文字列を表します。

(例:新潟県%市とすると文字数に関係なく「新潟県新潟市」、「新潟県燕市」、「新潟県阿賀野市」などが検索できます)

[3]半角のカギ括弧「[」`]」で囲まれた文字は、そのうちのどれか1文字を表します。(例:新潟県[長新]「新潟県新潟市」「新潟県長岡市」など、「新潟県燕市」などが検索できません。「新潟県阿賀野市」は対象外になります)

[4]半角のカギ括弧「[」`]」にcaret「^」をつけると、そのうちのどれか1文字以外を表します。(例:新潟県[^長新]「新潟県燕市」「新潟県阿賀野市」などが検索できます。「新潟県新潟市」「新潟県長岡市」が対象外になります)

短単位検索

作文テキストを UNIDIC という辞書と Mecab という形態素解析器を使用して形態素解析し、絞り込みます。UNIDIC について知りたい方は[こちら](#)、Mecab については[こちら](#)を、短単位について知りたい方は[こちら](#)をご覧ください。

キー項目	検索語
書字形 読み形 語彙素読み	<input type="text"/>
書字形 読み形 語彙素読み	<input type="text"/>
書字形 読み形 語彙素読み	<input type="text"/>

図7 短単位検索画面

品詞のとき前方一致で比較します。(例:「品詞」=「名詞-普通名詞」のとき、「名詞-普通名詞-副詞可能」も「名詞-普通名詞-一般」も一致します)

それ以外は完全一致で検索します。この検索で書字形を検索したとき、下段に KWIC 形式(前後 10ワード)でも表示します。キー項目は3個指定でき、絞り込み検索します。

指定できる項目は以下の UNIDIC の出力項目となります。それぞれの出力項目が何を表すのか知りたい方は[こちら](#)をご覧ください。

[書字形、発音形、語彙素読み、語彙素、品詞、活用型、活用形、語形、書字形基本形、語種]