

第46回自然史標本情報発信に関する研究会
—S-Net提供データ作成に向けたPC使い方講座—

OpenRefineを使って データクリーニングをやる (基礎編)

2026-02-28

国立科学博物館 標本資料センター
柿添翔太郎・太田藍乃

データを標準化する意味



サイエンスミュージアム ネット
Science Museum Net S-Net

[ホーム](#) > 検索結果

レコード種別

Preserved Specimen (7,820,325)

Fossil Specimen (178,466)

[もっと見る](#)

自然史標本情報

該当件数：8,173,128 件 (1-20件目)

20 件ずつ表示 [絞り込み検索](#)

機関名 (日本語)

国立科学博物館 (1,861,689)

神奈川県立生命の星・地球博物館 (717,996)

1 2 3 4 5 6 >

国 (日本語)

[日本国](#) (735,342) [台湾](#) (45,988) [フィリピン](#) (34,974) [ネパール](#) (27,456) [インドネシア](#) (25,042) [タイ](#) (20,659) [中国](#) (19,285) [マレーシア](#) (18,195) [インドネシア共和国](#) (16,523) [Japan](#) (13,231) [ロシア](#) (9,126) [韓国](#) (7,998) [アメリカ合衆国](#) (7,727) [フィリピン共和国](#) (7,626) [ベトナム](#) (6,918) [インド](#) (6,587) [アメリカ](#) (5,710) [不明](#) (5,443) [オーストラリア](#) (5,360) [ペルー](#) (5,171) [タイ王国](#) (5,143) [フィリピン](#) (4,717) [カンボジア](#) (4,343) [ブラジル](#) (3,644) [パラオ共和国](#) (3,526) [パプアニューギニア](#) (3,426) [日本?](#) (3,175) [パラオ](#) (2,948) [米国](#) (2,841) [フランス](#) (2,804) [コロンビア](#) (2,738) [フィジー](#) (2,319) [ザンビア](#) (2,282) [ラオス](#) (2,278) [ドイツ](#) (2,236) [メキシコ](#) (1,992) [シンガポール](#) (1,984) [スペイン](#) (1,901) [カナダ](#) (1,880) [マレシア](#) (1,756) [Vietnam](#) (1,733) [イラク](#) (1,677) [ニュージーランド](#) (1,667) [大韓民国](#) (1,634) [ニューカレドニア](#) (1,493) [アフガニスタン](#) (1,467) [パキスタン](#) (1,464) [イタリア](#) (1,456) [チリ](#) (1,396) [ボリビア](#) (1,381) [北マリアナ諸島自治連邦区](#) (1,242) [コスタリカ](#) (1,157) [モンゴル](#) (1,109) [イギリス](#) (1,044) [エクアドル](#) (1,033) [パプアニューギニア](#) (976) [南アフリカ](#) (970) [オーストリア](#) (965) [トルコ](#) (957) [Indonesia](#) (955) [スリランカ](#) (941) [イラン](#) (917) [マダガスカル](#) (906) [インド共和国](#) (887) [ソロモン諸島](#) (859) [北朝鮮](#) (815) [トンガ](#) (812) [Malaysia](#) (804) [中華人民共和国](#) (776) [香港](#) (773) [コンゴ](#) (750) [エチオピア](#) (702) [スイス](#) (684) [Philippines](#) (680) [キルギスタン](#) (658) [タンザニア](#) (651) [unknown](#) (643) [パプアニューギニア独立国](#) (624) [フレンチポリネシア](#) (593) [スウェーデン](#) (584) [日本?](#) (572) [タジキスタン](#) (564) [ミャンマー](#) (562) [ウクライナ](#) (529) [Thailand](#) (521) [ミクロネシア](#) (521) [コスタリカ共和国](#) (503) [モルディブ](#) (498) [ベネズエラ](#) (490) [ギリシャ](#) (485) [中央アフリカ](#) (481) [カザフスタン](#) (472) [サウジアラビア王国](#) (465) [スリナム](#) (460) [モロッコ](#) (422) [グアム](#) (419) [中華民国](#) (418) [日本およびその周辺地域](#) (415) [チェコ共和国](#) (411) [ウズベキスタン](#) (402)

同じ国でも表記がバラバラ

→特定の国のデータだけ調べるのも大変

[もっと見る](#)

[閉じる](#)

Paratype (37,380)

Holotype (10,131)

[もっと見る](#)

8

[Didemnum moseleyi](#)

シロウスボヤ

19840615

国立科学博物館

データ活用のためには標準化する必要がある

S-Net形式

57.	国 ※2021年6月から必須。	○	文字列・半	例：Japan 公海上や採集国が不明の場合は「-（半角ハイフン）」を指定。
58.	国（日本語） ※2021年6月から必須。	○	文字列・日	例：日本 公海上や採集国が不明の場合は「-（半角ハイフン）」を指定。
59.	国地域コード		文字列・半	国名のID。日本はJP。ISO 3166-1-alpha-2 country codes (https://www.iso.org/iso/country_codes) に従うことが推奨される（入力しなくてもいい）
60.	都道府県		文字列・半	例：Tokyo, Ibaraki, Kyoto 複数都道府県をまたぐ場合は「/（半角スラッシュ）」で区切って指定。例：Yamanashi/Kanagawa
61.	都道府県（日本語） ※2021年6月から63. 郡・市区町村（日本語）が入力されている場合は必須。		文字列・日	例：東京都、茨城県、京都府 複数都道府県をまたぐ場合は「/（全角スラッシュ）」で区切って指定。例：山梨県/神奈川県
62.	郡・市区町村		文字列・半	例：Saitama
63.	郡・市区町村（日本語）		文字列・日	例：さいたま市
64.	詳細地名		文字列・半	
65.	詳細地名（日本語）		文字列・日	例：南区辻一丁目

85.	学名	○	文字列・半	該当標本の学名のフルスペル。属も不明の場合は、同定できた一番下の階級（目名、科名など）を書く（例：Aspergillus oryzae、Fagus sp.、Cornaceae、Coleoptera、Anas zonorhyncha x A. platyrhynchos）。雑種式は半角小文字のx（エックス）を使用
86.	カテゴリー ※変換ツールでは非表示の項目		文字列・日	当面使用しない。
87.	界名（学名）	○	文字列・半	Viruses、Bacteria、Archaea、Protozoa、Chromista、Plantae、Fungi、Animaliaのいずれか。大文字・小文字を区別して正確に入力
88.	界名（日本語名）	○	文字列・日	ウイルス界、真正細菌界、古細菌界、原生物界、クロミスタ界、植物界、菌界、動物界のいずれか。全角文字で正確に入力
89.	門名（学名）		文字列・半	
90.	門名（日本語名）		文字列・日	
91.	綱名（学名）		文字列・半	
92.	綱名（日本語名）		文字列・日	
93.	目名（学名）		文字列・半	

GBIFのデータ形式（DwC-A）に準拠

GBIFで失われてしまう日本語表記などの情報にも対応

データ形式を揃えている（標準化）ので

多くの機関の膨大な標本情報を横断検索できる！

データクリーニングとは

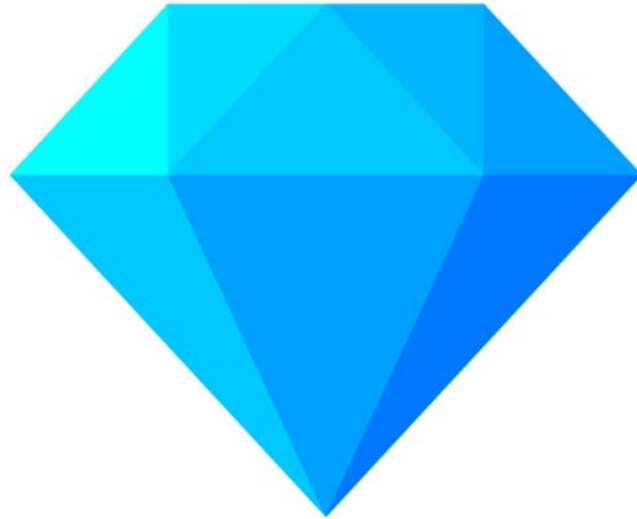
データを正しく整える作業

例

- ・ 不要文字の除去 「茨城 県」 「茨城県 」 → 「茨城県」
- ・ 表記の統一 「茨城」 「いばらきけん」 → 「茨城県」
- ・ 標準化 「2026/2/28」 「28 Feb. 2026」 → 「20260228」
- ・ 整合性チェック 採集日が「30260228」など
- ・ データの補完 学名から高次分類群を埋めるなど

Excelの関数ではすごく手間がかかる、データが膨大でExcelでは扱えない、データのゆらぎを探すのが大変...など

OpenRefineを活用してみませんか？



OpenRefine

データクリーニングを行うための
オープンソースアプリケーション

OpenRefineを使ってできること

- データの俯瞰・異常値の発見と一発クリーニング

データのばらつきや空欄を瞬時にリスト化し、全体像を直感的に把握
関数不要で、不要な空白や見えない文字をワンクリック除去

- 表記ゆれの自動名寄せ

学名や地名などの微妙な違いを自動検知して一括修正

- データの自由な分割・結合

セル内の混在した情報を別々の列に分割したり、複数列を1つにまとめる作業が瞬時に完了

- 正規表現による高度な抽出・置換

「特定のパターンに一致する文字だけを抽出・削除」など複雑な処理が可能

- 全操作の履歴保存と手順の使いまわし

元データを上書きせず、間違えてもいつでも1クリックで復元可能
マクロ不要で操作手順を保存し、別のデータへ一瞬で適用（使い回し）

OpenRefineを使ってできること

Excel

- セル単位の編集
- 関数を用いた計算
- 不可逆的編集
- 表記ゆらぎ検出には不向き
- 膨大なデータには不向き

OpenRefine

- 列単位の編集
- 可逆的編集（全編集履歴が保存される・レシピ保存）
- 表記ゆらぎ検出向き（ファセット・クラスタリング）
- 膨大なデータもOK

CSVファイルをOpenRefineで読み込む

OpenRefine 乱雑なデータを処理するためのパワーツール。

新規プロジェクト

既存プロジェクト

インポート

言語設定

拡張機能

インポートしたデータからプロジェクトを新規作成します。どんな種類のデータですか？

TSV、CSV、*SV、Excel(xls/xlsx)、JSON、XML、RDF as XML、OpenDocument(ods)とGoogle Dataがサポートされています。他のフォーマットも拡張機能（エクステンション）として追加できます。

データを取得する

コンピューター上のファイルを選択してください:

このコンピューター

ファイル選択 選択されていません

ここにファイルをドラッグ

次へ »

ここにCSVをドラッグ
※Excelファイルも読み込み可能

コンピューター上のファイルを選択してください:

ファイル選択 dummy_sp...snet_30.csv

ここにファイルをドラッグ

次へ »

CSVをドラッグすると
このように表示される

CSVファイルをOpenRefineで読み込む

ここをクリック

OpenRefine 乱雑なデータを処理するためのパワーツール。

新規プロジェクト [« 最初からやり直す](#) [解析オプションを設定](#) 名 dummy specimen snet 30 csv タグ [プロジェクトを作成 »](#)

既存プロジェクト
インポート
言語設定
拡張機能

レコード種別	記録年月日 (始め)	記録年月日 (終わり)	記録年月日オリジナル表記	国	国 (日本語)	都道府県	都道府県 (日本語)	郡・市区町村	郡・市区町村 (日本語)	地名オリジナル表記	学名
1. PreservedSpecimen			12-Jan-54	Japan	日本	Osaka	大阪府	Kishiwada City		大阪府 岸和田市	Rostanga sp.
2. PreservedSpecimen			1959年5月15日	Japan	日本	Gunma Pref.	群馬県	Midori-shi		群馬県 みどり市	Lacon sp.
3. PreservedSpecimen			1985/11/12	Japan	日本国	Ibaraki Prefecture	茨城県	Tsukuba		茨城県つくば市天王台	Polydactylus plebeius
4. PreservedSpecimen			1978年3月10日	Japan	日本	Ibaraki Pref.	茨城県	Tsukuba-shi		茨城県 つくば市	Hibiscus hamabo Sieb
5. PreservedSpecimen			8-May-55	Japan	日本	Gunma Prefecture	群馬県	Midori-shi		群馬県みどり市	Decapterus macarellus
6. PreservedSpecimen			1994/1/30	Japan	日本	群馬県	群馬県	Midori		みどり市	Scorpaenodes guamei
7. PreservedSpecimen			1983年5月7日	Japan	日本	Gunma Prefecture	群馬県	Midori-shi		みどり市	Rhinogobius similis
8. PreservedSpecimen			18-Nov-06	Japan	日本	Ibaraki Prefecture	茨城県	Tsukuba-shi		つくば市	Narcotica niveosparsa
9. PreservedSpecimen			2007年2月26日	Japan	日本	Osaka	大阪府	Kishiwada-shi		大阪府岸和田市岸之浦町	Linnaea borealis L.
10. PreservedSpecimen			30-Mar-53	Japan	日本国	Gunma Prefecture	群馬県	Midori City		群馬県みどり市	Vaccinium uliginosum
11. PreservedSpecimen			1983/2/17	Japan	日本	Ibaraki Prefecture	茨城県	Tsukuba-shi		つくば市	Trimma annosum
12. PreservedSpecimen			1978年5月3日	Japan	日本	群馬県	群馬県	Midori-shi		みどり市	Gyraulus soritai Habe,
13. PreservedSpecimen			1986年8月22日	Japan	日本	Gunma P.	群馬県	Midori-shi		群馬県 みどり市	Setaria viridis (L.) P.Be
14. PreservedSpecimen			1976年3月4日	Japan	日本	Osaka Prefecture	大阪府	Kishiwada-shi		大阪府 岸和田市	Pyrola japonica Klenz
15. PreservedSpecimen			1961/6/9	Japan	日本国	Gunma Prefecture	群馬県	Midori-shi		群馬県 みどり市	Girella punctata
16. PreservedSpecimen			1951/5/8	Japan	日本	Osaka	大阪府	Kishiwada-shi		大阪府岸和田市岸之浦町	Alkekengi officinarum I R.J.Wang
17. PreservedSpecimen			1999年4月9日	Japan	日本	Osaka Pref.	大阪府	Kishiwada City		大阪府岸和田市岸之浦町	Trifolium repens L.
18. PreservedSpecimen			1979.7.9	Japan	日本	Osaka	大阪府	Kishiwada-shi		大阪府岸和田市岸之浦町	Lespedeza virgata (Th
19. PreservedSpecimen			2010/10/27	Japan	日本	Ibaraki Pref.	茨城県	Tsukuba City		茨城県 つくば市	Chenopodium acuminata
20. PreservedSpecimen			1991年10月12日	Japan	日本	Ibaraki Prefecture	茨城県	Tsukuba-shi		茨城県つくば市天王台	Hemerocallis citrina Ba M.Hotta
21. PreservedSpecimen			1973/5/3	Japan	日本	Ibaraki Pref.	茨城県	Tsukuba		茨城県つくば市天王台	Eufentonia nihonica
22. PreservedSpecimen			1983/6/30	Japan	日本	Ibaraki Prefecture	茨城県	Tsukuba City		茨城県つくば市天王台	Rhynchocypris lagows

データの解析形式

文字区切りテキスト(CSV/TSV)

行単位のテキスト

固定長テキスト

PC-Axis text files

JSONファイル

MARCファイル

JSON-LDファイル

RDF/N3ファイル

文字コード

UTF-8

カラムの区切り文字

カンマ区切り (CSV)

タブ (TSV)

カスタム, _____

文字 " " でセルを囲む

前後の空白文字を削除する

特殊文字はバックスラッシュでエスケープしてください

先頭を無視 0 行分 (先頭から)

カラム名として 1 行分
解釈

カラム名 (カンマ区切り)

先頭データを破棄 0 行分

最後の行を破棄 0 行分

[プレビューを更新](#)

自動プレビューしない

セルを解析
文字から数字

空白行も保存

空白カラムも保存

nullとして空白セルを保存

ファイルのソース

アーカイブファイルの保存



バージョン 3.9.5 (TRUNK)

設定
ヘルプ
About

ファセット/フィルタ

取り消す/やり直す 0/9

ファセットとフィルタを使う

ファセットやフィルタでデータのサブセットを選択できます。各カラムのメニューからファセットとフィルタメソッドを選択してください。

使い方がわからない?

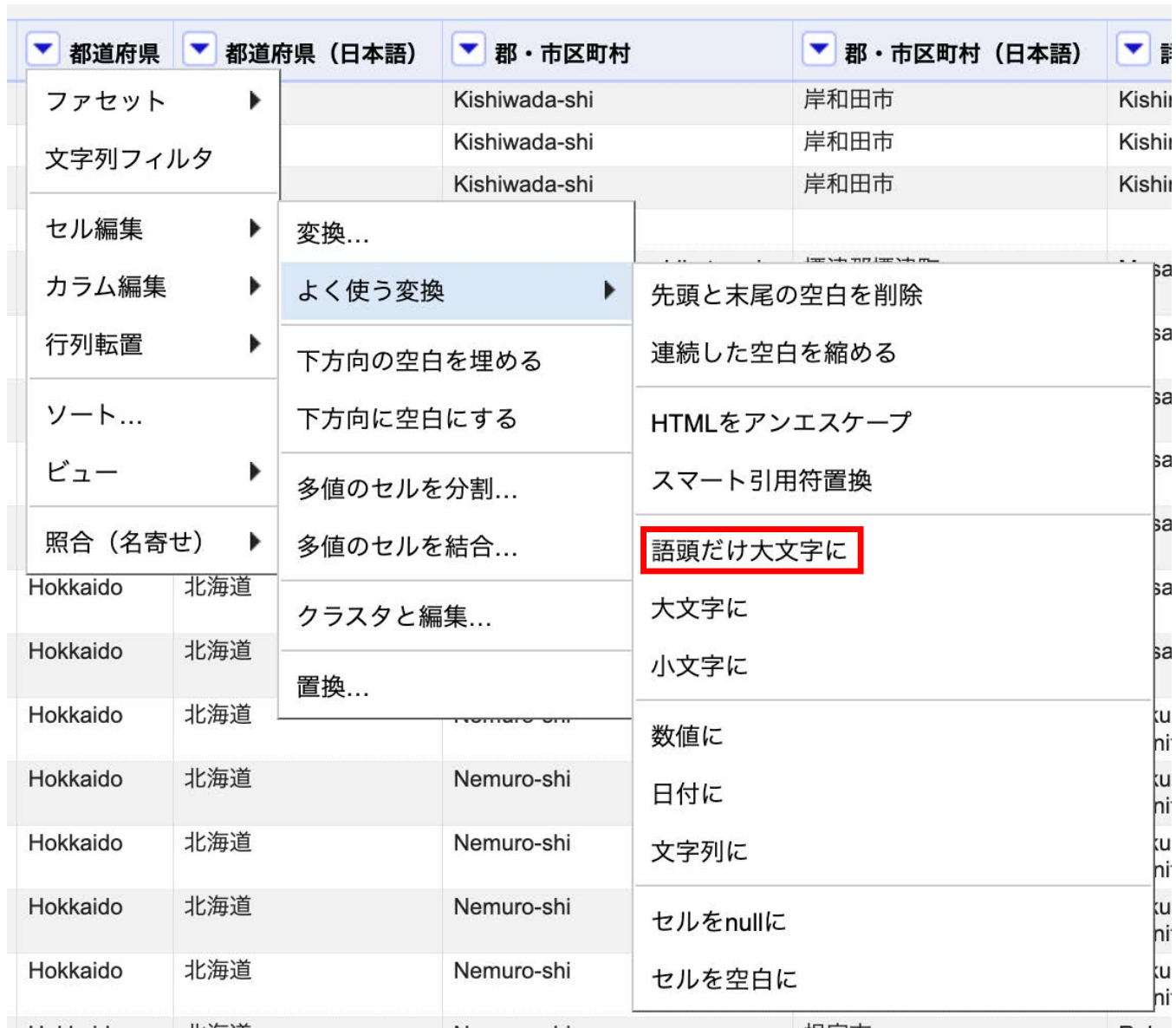
[動画をご覧ください \(ただし英語\)](#)

30 行

モード: 行 レコード 表示: 5 10 25 50 100 500 1000 行

全て	レコード種別	記録年月日 (始め)	記録年月日 (終わり)	記録年月日オリジナル表記	国	国 (日本語)	都道府県	都道府県 (日本語)
☆	1.	PreservedSpecimen		12-Jan-54	Japan	日本	Osaka	大阪府
☆	2.	PreservedSpecimen		1959年5月15日	Japan	日本	Gunma Pref.	群馬県
☆	3.	PreservedSpecimen		1985/11/12	Japan	日本国	Ibaraki Prefecture	茨城県
☆	4.	PreservedSpecimen		1978年3月10日	Japan	日本	Ibaraki Pref.	茨城県
☆	5.	PreservedSpecimen		8-May-55	Japan	日本	Gunma Prefecture	群馬県
☆	6.	PreservedSpecimen		1994/1/30	Japan	日本		群馬県
☆	7.	PreservedSpecimen		1983年5月7日	Japan	日本	Gunma Prefecture	群馬県
☆	8.	PreservedSpecimen		18-Nov-06	Japan	日本	Ibaraki Prefecture	茨城県
☆	9.	PreservedSpecimen		2007年2月26日	Japan	日本	Osaka	大阪府
☆	10.	PreservedSpecimen		30-Mar-53	Japan	日本国	Gunma Prefecture	群馬県
☆	11.	PreservedSpecimen		1983/2/17	Japan	日本	Ibaraki Prefecture	茨城県
☆	12.	PreservedSpecimen		1978年5月3日	Japan	日本		群馬県
☆	13.	PreservedSpecimen		1986年8月22日	Japan	日本	Gunma P.	群馬県
☆	14.	PreservedSpecimen		1976年3月4日	Japan	日本	Osaka Prefecture	大阪府
☆	15.	PreservedSpecimen		1961/6/9	Japan	日本国	Gunma Prefecture	群馬県
☆	16.	PreservedSpecimen		1951/5/8	Japan	日本	Osaka	大阪府
☆	17.	PreservedSpecimen		1999年4月9日	Japan	日本	Osaka Pref.	
☆	18.	PreservedSpecimen		1979.7.9	Japan	日本	Osaka	大阪府
☆	19.	PreservedSpecimen		2010/10/27	Japan	日本	Ibaraki Pref.	茨城県
☆	20.	PreservedSpecimen		1991年10月12日	Japan	日本	Ibaraki Prefecture	茨城県
☆	21.	PreservedSpecimen		1973/5/3	Japan	日本	Ibaraki Pref.	
☆	22.	PreservedSpecimen		1983/6/30	Japan	日本	Ibaraki Prefecture	茨城県
☆	23.	PreservedSpecimen		2021年12月17日	Japan	日本	Gunma Prefecture	群馬県
☆	24.	PreservedSpecimen		1997年9月20日	Japan	日本	Osaka	大阪府
☆	25.	PreservedSpecimen		2-Nov-96	Japan	日本	Gunma Prefecture	
☆	26.	PreservedSpecimen		1973/11/24	Japan	日本	Osaka	大阪府
☆	27.	PreservedSpecimen		1984年6月30日	Japan	日本		茨城県
☆	28.	PreservedSpecimen		1952/10/5	Japan	日本国	Ibaraki	茨城県
☆	29.	PreservedSpecimen		1955/3/18	Japan	日本	Osaka P.	大阪府
☆	30.	PreservedSpecimen		1965年5月24日	Japan	日本	Osaka	大阪府

無事に読み込めるとこのように表示される



「よく使う変換」だけでも結構便利

ファセットを使ってゆらぎを探す

表記	▼ 国	▼ 国 (日本語)	▼ 都道府県	▼ 都道府県 (日本語)
	Japan	日本		群馬県

Context Menu for '国':

- ファセット
 - 文字列ファセット
 - 数値ファセット
 - タイムラインファセット
 - 散布図ファセット...
 - カスタム文字列ファセット...
 - カスタム数値ファセット...
 - カスタムファセット
- 文字列フィルタ
- セル編集
- カラム編集
- 行列転置
- ソート...
- ビュー
- 照合 (名寄せ)

ファセット/フィルタ

取り消す/やり直す 8/8

再読み込み すべてリセット すべて削除

国 変更

2 件 ソート基準: 名前 カウント クラスタ

Japan 26

Japan 4

選択肢のカウントによるファセット

「Japan」と
「Japan」が混在

カラムの「▼」をクリック
ファセット→文字列ファセット

ファセットから表記ゆらぎを修正

国 変更

2件 ソート基準: 名前 カウント クラスタ

Japan 26

Japan 4

編集 include

選択肢のカウントによるファセット

「編集」を押す

国 変更

2件 ソート基準: 名前 カウント クラスタ

Japan 26

Japan 4

選択肢のカウントによるファセット

全て	レコード種別	記録年月日 (始め)	記録年月日 (
☆ ↶	1. PreservedSpecimen		
☆ ↶	2. PreservedSpecimen		
☆ ↶	3. PreservedSpecimen		

Japan

適用 キャンセル

Enter Esc

Japanの後の空白を削除 → 適用

国 変更

1件 ソート基準: 名前 カウント クラスタ

Japan 30

選択肢のカウントによるファセット

ゆらぎが無事に修正された！

クラスタリングによる表記ゆらぎの修正

国 変更

2件 ソート基準: 名前 カウント

クラスタ

Japan 26

Japan 4 編集 include

選択肢のカウントによるファセット

「クラスタ」を押す

クラスタリングと編集 "国"

同じものの別の名前かもしれないセル値のグループを探します。たとえば、"New York" と "new york" は似た概念ですが表記の違いがあります。また、"Gödel" と "Godel/Goedel" は多分同じ人を示します。 [詳細はこちら...](#)

方法 キー関数

自動更新 1 cluster が見つかりました

マージしますか?	クラスタの値	新しいセルの値	クラスタサイズ	行数
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> Japan (26行) <input checked="" type="checkbox"/> Japan (4行)	<input type="text" value="Japan"/>	2	30

マージしますか?
にチェック

新しいセルの値を「Japan」に

クラスタリングによる表記ゆらぎの修正



ゆらぎが無事に修正された！

カラムのコピー

▼ 記録年月日オリジナル表記 ▼ 国 ▼ 国 (日本語) ▼ 都道府県

ファセット ▶	Japan	日本	Osaka
文字列フィルタ	Japan	日本	Gunma Pref.
セル編集 ▶	Japan	日本国	Ibaraki Prefectu
カラム編集 ▶	Japan	日本	Ibaraki Pref.

複数のカラムに分割...
カラムの結合...
このカラムに基づいてカラムを追加...
URLでカラムを追加...
照合値からカラムを作る...
カラム名を変更する...
このカラムを取り除く
先頭にカラムを移動
末尾にカラムを移動
左にカラムを移動
右にカラムを移動

1978年5月3日
1986年8月22日
1976年3月4日
1961/6/9
1951/5/8
1999年4月9日
1979.7.9

このカラムに基づき新しいカラムを追加 記録年月日オリジナル表記

新しいカラム名

エラー 空白をセット エラーを保存 カラムから値をコピー

式 (expression)

言語

文法エラーはありません。

プレビュー

履歴

星付き

ヘルプ

row	value	value
1.	12-Jan-54	12-Jan-54
2.	1959年5月15日	1959年5月15日
3.	1985/11/12	1985/11/12
4.	1978年3月10日	1978年3月10日
5.	8-May-55	8-May-55
6.	1994/1/30	1994/1/30

OK

キャンセル

コピー元のカラム

→カラム編集

→このカラムに基づいてカラムを追加

新しいカラム名を設定してOKを押す

カラムのコピー

▼ 記録年月日 (始め) 2	▼ 国	▼ 国 (日本語)	▼ 都道府県
ファセット ▶	Japan	日本	Osaka
文字列フィルタ	Japan	日本	Gunma Pref.
セル編集 ▶	Japan	日本国	Ibaraki Prefecture
カラム編集 ▶	Japan	日本	Ibaraki Pref.
行列転置 ▶	複数のカラムに分割...		
ソート...	カラムの結合...		
ビュー ▶	このカラムに基づいてカラムを追加...		
照合 (名寄せ) ▶	URLでカラムを追加...		
	照合値からカラムを作る...		
1978年5月3日	カラム名を変更する...		
1986年8月22日	このカラムを取り除く		
1976年3月4日	先頭にカラムを移動		
1961/6/9	末尾にカラムを移動		
1951/5/8	左にカラムを移動		
1999年4月9日	右にカラムを移動		
1979.7.9			

**カラムの移動、名称変更、削除などは
「カラム編集」から可能**

記録年月日の表記ゆらぎ修正

28-Feb-26
2026年2月28日
2026/2/28
2026-2-28
2026-02-28



20260228

様々な表記ゆらぎ

S-Net形式の日付表記

今回行うアプローチ

- ・ 余計な空白の削除
- ・ 年月日をハイフンつながぎに置換
- ・ 文字型データを日付型データに変換
- ・ 日付データを8桁表記に変換

▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日 (記録4)
ファセット ▶		12-Jan-54
文字列フィルタ		1959年5月
		1985/11/1
セル編集 ▶	変換...	1985年3月
カラム編集 ▶	よく使う変換 ▶	1985年5月
行列転置 ▶	下方向の空白を埋める	1985年5月
ソート...	下方向に空白にする	1985年5月
ビュー ▶	多値のセルを分割...	1985年5月
照合 (名寄せ) ▶	多値のセルを結合...	1985年5月
1978年5月3日	クラスタと編集...	1985年5月
1986年8月22日	置換...	1985年5月
1976年3月4日		1985年3月
1984/1/10		1984/1/10

置換

検索

大文字・小文字を区別しない 単語全体 正規表現として解釈

各文字に置換文字を追加する場合は空白のまま。

改行などの特殊文字や複雑なパターンを検索する場合は正規表現が使えます。

この文字列に置換

改行文字には\n、タブ文字には\tを使います。「\n」や「\t」を検索するには「\\n」や「\\t」を使います (エスケープ文字)。

正規表現に()を使ったグループがある場合、\$0がマッチ全体を、\$1,\$2...が各グループを返します。

置換機能を使って1つ1つ置換することもできる

「年」「月」「.」 → 「-」

「」 「日」 → 「」

年月日や「.」をハイフンつながりに置換

	▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年
e	ファセット ▶		12-Jan-54
e	文字列フィルタ		1959年5月
e	セル編集 ▶	変換...	5/11/12
e	カラム編集 ▶	よく使う変換 ▶	3年3月
e	行列転置 ▶	下方向の空白を埋める	ay-55
e	ソート...	下方向に空白にする	/1/30
e	ビュー ▶	多値のセルを分割...	3年5月
e	照合 (名寄せ) ▶	多値のセルを結合...	ov-06
e	2007年2月26日	クラスタと編集...	7年2月
e	30-Mar-53	置換...	lar-53
e	1983/2/17		1983/2/17

「変換」を使うと一気に置換可能

年月日や「.」をハイフンつながぎに置換

式による値の変換：記録年月日（始め）

式 (expression) 言語 General Refine Expression Language (GREL) ▾

value 文法エラーはありません。

プレビュー 履歴 星付き ヘルプ

row	value	value
1.	12-Jan-54	12-Jan-54
2.	1959年5月15日	1959年5月15日
3.	1985/11/12	1985/11/12
4.	1978年3月10日	1978年3月10日
5.	8-May-55	8-May-55
6.	1994/1/30	1994/1/30

- エラー オリジナルを保持 再帰的な変換： 回（変化しなくなるまで）
- 空白をセット
- エラーを保存

OK

キャンセル

「変換」の「式」に指示内容を入れる
年月日や「.」をハイフンつなぎに置換

```
value.trim().replace("年","-").replace("月","-").replace("日","").replace(".", "-")
```

`value`

値を引用

`.trim()`

空白の削除

`.replace("年","-")`

「年」を「-（ハイフン）」に置換

`.replace("月","-")`

「月」を「-」に置換

`.replace("日","")`

「日」を削除

`.replace(".", "-")`

「.（ピリオド）」を「-」に置換

年月日や「.」をハイフンつなぎに置換

30 行

モード: 行 レコード 表示: 5 10 25 50 100 500 1000 行

<input type="checkbox"/> 全て	<input type="checkbox"/> レコード種別	<input type="checkbox"/> 記録年月日 (始め)	<input type="checkbox"/> 記録年月日 (終わり)	<input type="checkbox"/> 記録年月日オリジナル表記	<input type="checkbox"/>
 	1.	PreservedSpecimen	12-Jan-54	12-Jan-54	Ja
 	2.	PreservedSpecimen	1959-5-15	1959年5月15日	Ja
 	3.	PreservedSpecimen	1985/11/12	1985/11/12	Ja
 	4.	PreservedSpecimen	1978-3-10	1978年3月10日	Ja
 	5.	PreservedSpecimen	8-May-55	8-May-55	Ja
 	6.	PreservedSpecimen	1994/1/30	1994/1/30	Ja
 	7.	PreservedSpecimen	1983-5-7	1983年5月7日	Ja
 	8.	PreservedSpecimen	18-Nov-06	18-Nov-06	Ja
 	9.	PreservedSpecimen	2007-2-26	2007年2月26日	Ja
 	10.	PreservedSpecimen	30-Mar-53	30-Mar-53	Ja
 	11.	PreservedSpecimen	1983/2/17	1983/2/17	Ja
 	12.	PreservedSpecimen	1978-5-3	1978年5月3日	Ja
 	13.	PreservedSpecimen	1986-8-22	1986年8月22日	Ja
 	14.	PreservedSpecimen	1976-2-4	1976年2月4日	Ja

年月日や「.」をハイフンつながりに置換

▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)
ファセット ▶		12-Jan-54	Japan	日本
文字列フィルタ		1959年5月15日	Japan	日本
セル編集 ▶	変換...	5/11/12	Japan	日本国
カラム編集 ▶	よく使う変換 ▶			
行列転置 ▶	下方向の空白を埋める			
ソート...	下方向に空白にする			
ビュー ▶	多値のセルを分割...			
照合 (名寄せ) ▶	多値のセルを結合...			
2007-2-26	クラスタと編集...			
30-Mar-53	置換...			
1983/2/17		1983		
1978-5-3		1978		
1986-8-22		1986		
1976-3-4		1976		
1961/6/9		1961		

- 先頭と末尾の空白を削除
- 連続した空白を縮める
- HTMLをアンエスケープ
- スマート引用符置換
- 語頭だけ大文字に
- 大文字に
- 小文字に
- 数値に
- 日付に
- 文字列に
- セルをnullに
- セルを空白に

文字型データを日付型データに変換

▼ 記録年月日 (始め)
12-Jan-54
1959-5-15
1985/11/12
1978-3-10
8-May-55
1994/1/30
1983-5-7
18-Nov-06
2007-2-26
30-Mar-53
1983/2/17
1978-5-3
1986-8-22
1976-3-4

文字型 → 日付型



文字型ではないデータは
緑色の文字色になる

▼ 記録年月日 (始め)
1954-01-12T00:00:00Z
1959-05-15T00:00:00Z
1985-11-12T00:00:00Z
1978-03-10T00:00:00Z
1955-05-08T00:00:00Z
1994-01-30T00:00:00Z
1983-05-07T00:00:00Z
2006-11-18T00:00:00Z
2007-02-26T00:00:00Z
1953-03-30T00:00:00Z
1983-02-17T00:00:00Z
1978-05-03T00:00:00Z
1986-08-22T00:00:00Z
1976-03-04T00:00:00Z

文字型データを日付型データに変換

value.toString("yyyyMMdd")

toString **文字型へ変換**という意味
yyyyは年、MMは月、ddは日

記録年月日 (始め)	記録年月日 (終わり)	記
12-Jan		
1959年		
2007-02-26T00:00:00Z		
1953-03-30T00:00:00Z		
1983-02-17T00:00:00Z	1983/2	

式による値の変換： 記録年月日 (始め)

式 (expression)

言語 General Refine Expression Language (GREL) ▾

value.toString("yyyyMMdd")

文法エラーはありません。

プレビュー

履歴

星付き

ヘルプ

row	value	value.toString("yyyyMMdd")
1.	1954-01-12T00:00:00Z	19540112
2.	1959-05-15T00:00:00Z	19590515
3.	1985-11-12T00:00:00Z	19851112
4.	1978-03-10T00:00:00Z	19780310
5.	1955-05-08T00:00:00Z	19550508
6.	1994-01-30T00:00:00Z	19940130

エラー オリジナルを保持 再帰的な変換： 回 (変化しなくなるまで)
 空白をセット
 エラーを保存

OK

キャンセル

日付型データを8桁表記に変換

30行

モード: 行 レコード 表示: 5 10 25 50 100 500 1000 行

▼ 全て	▼ レコード種別	▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)	▼ 都道府県	▼ 都道	
☆	1.	PreservedSpecimen	19540112		12-Jan-54	Japan	日本	Osaka	大阪府
☆	2.	PreservedSpecimen	19590515		1959年5月15日	Japan	日本	Gunma Pref.	群馬県
☆	3.	PreservedSpecimen	19851112		1985/11/12	Japan	日本国	Ibaraki Prefecture	茨城県
☆	4.	PreservedSpecimen	19780310		1978年3月10日	Japan	日本	Ibaraki Pref.	茨城県
☆	5.	PreservedSpecimen	19550508		8-May-55	Japan	日本	Gunma Prefecture	群馬県
☆	6.	PreservedSpecimen	19940130		1994/1/30	Japan	日本		群馬県
☆	7.	PreservedSpecimen	19830507		1983年5月7日	Japan	日本	Gunma Prefecture	群馬県
☆	8.	PreservedSpecimen	20061118		18-Nov-06	Japan	日本	Ibaraki Prefecture	茨城県
☆	9.	PreservedSpecimen	20070226		2007年2月26日	Japan	日本	Osaka	大阪府
☆	10.	PreservedSpecimen	19530330		30-Mar-53	Japan	日本国	Gunma Prefecture	群馬県
☆	11.	PreservedSpecimen	19830217		1983/2/17	Japan	日本	Ibaraki Prefecture	茨城県
☆	12.	PreservedSpecimen	19780503		1978年5月3日	Japan	日本		群馬県
☆	13.	PreservedSpecimen	19860822		1986年8月22日	Japan	日本	Gunma P.	群馬県
☆	14.	PreservedSpecimen	19760304		1976年3月4日	Japan	日本	Osaka Prefecture	大阪府
☆	15.	PreservedSpecimen	19610609		1961/6/9	Japan	日本国	Gunma Prefecture	群馬県
☆	16.	PreservedSpecimen	19510508		1951/5/8	Japan	日本	Osaka	大阪府
☆	17.	PreservedSpecimen	19990409		1999年4月9日	Japan	日本	Osaka Pref.	
☆	18.	PreservedSpecimen	19790709		1979.7.9	Japan	日本	Osaka	大阪府

8桁表記に変換された結果

日付型データを8桁表記に変換

▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)
ファセット ▶		12-Jan-54	Japan	日本
文字列フィルタ		1959年5月15日	Japan	日本
セル編集 ▶	変換...	1959/11/12	Japan	日本国
カラム編集 ▶	よく使う変換 ▶			
行列転置 ▶	下方向の空白を埋める			
ソート...	下方向に空白にする			
ビュー ▶	多値のセルを分割...			
照合 (名寄せ) ▶	多値のセルを結合...			
20070226	クラスタと編集...			
19530330	置換...			
19830217		1983		
19780503		1978		
19860822		1986		
19760304		1976		
19610609		1961		

- 先頭と末尾の空白を削除
- 連続した空白を縮める
- HTMLをアンエスケープ
- スマート引用符置換
- 語頭だけ大文字に
- 大文字に
- 小文字に
- 数値に**
- 日付に
- 文字列に
- セルをnullに
- セルを空白に

文字型 → 数値型 への変換

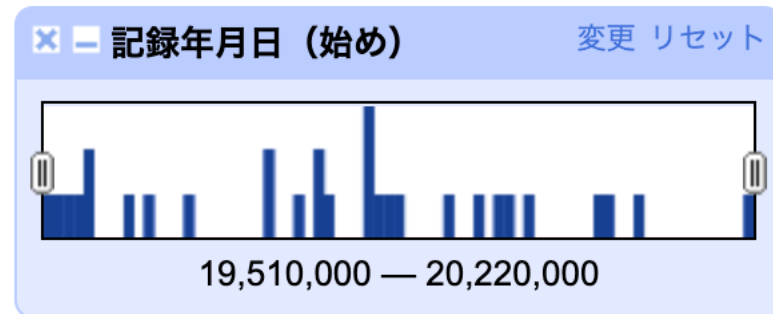
▼ 全て	▼ レコード種別	▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)	▼ 都道府県
☆	1. PreservedSpecimen	19540112		12-Jan-54	Japan	日本	Osaka
☆	2. PreservedSpecimen	19590515		1959年5月15日	Japan	日本	Gunma Pref.
☆	3. PreservedSpecimen	19851112		1985/11/12	Japan	日本国	Ibaraki Prefecture
☆	4. PreservedSpecimen	19780310		1978年3月10日	Japan	日本	Ibaraki Pref.
☆	5. PreservedSpecimen	19550508		8-May-55	Japan	日本	Gunma Prefecture

数値型になると文字色が緑色になる

▼ 全て	▼ レコード種別	▼ 記録年月日 (始め)	▼ 記録年月日 (終わり)	▼ 記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)	▼ 都道府県
☆	1. PreservedSpecimen	19540112		12-Jan-54	Japan	日本	Osaka
☆	2. PreservedSpecimen	19590515		1959年5月15日	Japan	日本	Gunma Pref.
☆	3. PreservedSpecimen	19851112		1985/11/12	Japan	日本国	Ibaraki Prefecture
☆	4. PreservedSpecimen <small>編集</small>	19780310		1978年3月10日	Japan	日本	Ibaraki Pref.
☆	5. PreservedSpecimen	19550508		8-May-55	Japan	日本	Gunma Prefecture

数値型に変換した結果

文字型 → 数値型への変換



数値型にすると

範囲検索ができるようになる！

1951年から2021年までのデータが含まれていると分かる

数値ファセットの活用

学名から属だけ切り出し（分割）

学名	門名 (学名)	門名
ファセット	MOLLUSCA	軟体動物
文字列フィルタ	Arthropoda	節足動物
セル編集	et Zucc.	
カラム編集		
行列転置		
ソート...		
ビュー		
照合 (名寄せ)		
Trimma annosum		
Gyraulus soritai Habe, 19		
Setaria viridis (L.) P.Beau		
Pyrola japonica Klentze e		
(Maxim.) Andres		
Girella punctata		
Alkekengi officinarum Mc		
(Mast.) R.J.Wang		
Trifolium repens L.		
Lespedeza virgata (Thunb.) DC.	Tracheophyta	

カラム"学名"を複数カラムに分割

カラムの分割するには

文字で分割

区切り文字 正規表現として解釈

分割先 最大のカラム (空白はそのまま)

フィールド長で分割

カンマ区切りの数値をリストに (例: 5,7,15)

分割後

セルタイプを推測

このカラムを取り除く

OK

キャンセル

特定の区切り文字や長さで分割できる
区切り文字には「（半角スペース）」を利用

学名から属だけ切り出し（分割）

▼ 学名	▼ 学名 1	▼ 学名 2	
Rostanga sp.	Rostanga	sp.	M
Lacon sp.	Lacon	sp.	A
Polydactylus plebeius	Polydactylus	plebeius	
Hibiscus hamabo Siebold et Zucc.	Hibiscus	hamabo Siebold et Zucc.	
Decapterus macarellus	Decapterus	macarellus	
Scorpaenodes guamensis	Scorpaenodes	guamensis	
Rhinogobius similis	Rhinogobius	similis	
Narcotica niveosparsa	Narcotica	niveosparsa	A
Linnaea borealis L.	Linnaea	borealis L.	T
Vaccinium uliginosum L. var. japonicum T.Yamaz.	Vaccinium	uliginosum L. var. japonicum T.Yamaz.	T
Trimma annosum	Trimma	annosum	
Gyraulus soritai Habe, 1976	Gyraulus	soritai Habe, 1976	M
Setaria viridis (L.) P.Beauv.	Setaria	viridis (L.) P.Beauv.	T
Pyrola japonica Klenze ex Alef. var. subaphylla (Maxim.) Andres	Pyrola	japonica Klenze ex Alef. var. subaphylla (Maxim.) Andres	T
Cirilla punctata	Cirilla	punctata	C

特定の区切り文字や長さで分割できる
区切り文字には「（半角スペース）」を利用

正規表現を使った変換

パターン1

```
value.replace(/^(*?)(都|道|府|県)/, "")
```

パターン2

```
value.replace(/^([都道府県]+[都道府県])/, "")
```

パターン3

```
value.replace(/^¥s*.{2,3}?¥s*[都道府県]¥s*/, "")
```

**最近はこのGREL指示文は
生成AIに書かせたほうが楽な場合も多い**

例:

OpenRefineのGRELを使用して、住所データから「都道府県名」だけを削除（空文字に置換）するコードを書いてください。

以下の条件をすべて満たす、最も安全で確実な1行のコードを提案してください。

【条件】

- ・ 行の先頭にある都道府県名（都・道・府・県で終わる言葉）を消したい。
- ・ 先頭に全角や半角のスペースが入っていることがあるので、それも一緒に消したい。
- ・ 「群馬 県」のように、文字の間にスペースが混ざっているケースにも対応したい。

OpenRefineのGRELを使用して、住所データから「都道府県名」だけを削除(空文字に置換)するコードを書いてください。

以下の条件をすべて満たす、最も安全で確実な1行のコードを提案してください。

【条件】

- ・行の先頭にある都道府県名(都・道・府・県で終わる言葉)を消したい。
- ・先頭に全角や半角のスペースが入っていることがあるので、それも一緒に消したい。
- ・「群馬 県」のように、文字の間にスペースが混ざっているケースにも対応したい。

思考時間: 6s >

時間がかかってもいいので正確に推論してください。こちらの思考力を考慮せず、計算資源を最大限利用してください。一方針どおり、仕様を満たす“安全寄り”の1行を提示します。

```
grel
```

📄 コピーする

```
value.replace(/^[\s ]*(^[^\s ]+[\s ]*){0,4}[都道府県](?=[^\s ])/, "")
```

- ・ 行頭の全角/半角スペース ([\s]*) をまとめて除去
- ・ **「群馬 県」**のように都道府県名の途中にスペースが混ざるのを許容 ((^[^\s]+[\s]*){0,4} で「数文字+任意スペース」を最大4回まで、= “だいたい1~4文字県名” を想定)
- ・ **末尾の「都/道/府/県」**で確定して削除 ([都道府県])
- ・ 後ろに住所が続くときだけ削るようにして誤爆を減らす ((?=[^\s]))
 - ・ 例: 値が「東京都」だけのように都道府県名しか無い場合は削らず残しやすい(※要件次第でここは外せます)

このまま「Transform cells → Replace / transform」等で使えます。

📄 🔊 🗨️ 🔍 …

OpenRefineのGRELを使用して、住所データから「都道府県名」だけを削除（空文字に置換）するコードを書いてください。
以下の条件をすべて満たす、最も安全で確実な1行のコードを提案してください。...



OpenRefineで住所データから都道府県名を削除する場合、正規表現を用いた `replace` 関数を使うのが最も確実です。

ご要望の条件（先頭の空白、文字間の空白、4種類の末尾文字）をすべて網羅したコードは以下の通りです。

提案コード

JavaScript



```
value.replace(/^\
```

現在位置： [ホーム](#) > [政策](#) > アドレス・ベース・レジストリ

アドレス・ベース・レジストリ

最終更新日:2025年10月16日

アドレス・ベース・レジストリは、ベース・レジストリ※として整備する住所・所在地関係データベースです。

※法律上におけるベース・レジストリの正式名称は「公的基礎情報データベース」と定義されています。

お知らせ

- 2025年10月16日
 - [町字データに関する留意事項](#)について、対応が完了し、[関連情報](#)に掲載しました。

アドレス・ベース・レポジトリ

「[全国町字マスター](#)」など（※かなり重い）

https://www.digital.go.jp/policies/base_registry_address

市町村、詳細地名のマッチングに使える辞書

「学名」 → 「scientificName」 に

府県	都道府県 (日本語)	郡・市区町村	郡・市区町村 (日本語)	地名オリジナル表記	scientificName	門名 (学名)	門名 (日本語名)	綱名 (学名)	綱名 (日本語名)	目名 (学名)
Pref.	大阪府	Kishiwada City	大阪府 岸和田市	岸和田市	Notostingu sp.	MOLLUSCA	軟体動物門	GASTROPODA	腹足綱	NUDIBRANCHIA
Pref.	群馬県	Midori-shi	群馬県 みどり市	みどり市	Lacon sp.	Arthropoda	節足動物門	Insecta	昆虫綱	Coleoptera
Prefecture	茨城県	Tsukuba	茨城県つくば市天王台	天王台	Polydactylus plebeius					
Pref.	茨城県	Tsukuba-shi	茨城県 つくば市	つくば市	Hibiscus hamabo Siebold et Zucc.					
Prefecture	群馬県	Midori-shi	群馬県みどり市	みどり市	Decapterus macarellus					
Prefecture	群馬県	Midori	みどり市	みどり市	Scorpaenodes guamensis					
Prefecture	群馬県	Midori-shi	みどり市	みどり市	Rhinogobius similis					
Prefecture	茨城県	Tsukuba-shi	つくば市	つくば市	Narctica niveosparsa	Arthropoda	節足動物門	Insecta	昆虫綱	
Prefecture	大阪府	Kishiwada-shi	大阪府岸和田市岸之浦町	岸之浦町	Linnaea borealis L.	Tracheophyta		Magnoliopsida		Dipsacales
Prefecture	群馬県	Midori City	群馬県みどり市	みどり市	Vaccinium uliginosum L. var. japonicum T.Yamaz.	Tracheophyta		Magnoliopsida		Ericales
Prefecture	茨城県	Tsukuba-shi	つくば市	つくば市	Trimma annosum					
Prefecture	群馬県	Midori-shi	みどり市	みどり市	Gyraulus soritai Habe, 1976	Mollusca	軟体動物門	Gastropoda	腹足綱	Lymnaeida
Prefecture	群馬県	Midori city	群馬県 みどり市	みどり市	Setaria viridis (L.) P.Beauv.	Tracheophyta		Liliopsida		Poales

ツール | コミュニティ | GBIFについて

出版者	ユーザー	GBIFラボ
IPT	GBIFがホストするポータル	種名マッチングツール
データ検証ツール	Scientific collections	学名解析ツール
GeoPick ⇄	データ処理	配列ID
New data model	派生データセット	関連した観察傾向
自然史コレクション (GRSciColl) ⇄	rgbif	開発者ブログ ⇄
掲載したいデータセットを提案する	pygbif	
Metabarcoding data toolkit ⇄	SQLダウンロード	
	ツールカタログ	

TOOLS | LOOK UP

Normalize species names from a csv file against the GBIF backbone.
The file is expected to be have a column called 'scientificName' and an optional column 'kingdom' and 'id'.

SIMPLEEXAMPLE.CSV | ADVANCEDEXAMPLE.CSV

SELECT FILE

DROP HERE

ここにCSVをドラッグ

項目名を「scientificName」にしたCSVを
GBIF「種名マッチングツール」で分析



TOOLS | LOOK UP

If no kingdom specified then prefer

- animalia
- plantae
- fungi
- chromista
- bacteria
- protozoa
- viruses
- archaea

MATCH TO GBIF BACKBONE

scientificName	preferred kingdom
Rostanga sp.	
Lacon sp.	
Polydactylus plebeius	
Hibiscus hamabo Siebold et Zucc.	
Decapterus macarellus	
Scorpaenodes guamensis	
Rhinogobius similis	
Narcotica niveosparsa	
Linnaea borealis L.	
Vaccinium uliginosum L. var. japonicum T.Yamaz.	
Trimma annosum	
Gyraulus soritai Habe, 1976	

「scientificName」の列だけ読み込まれる

種	Plantae	Tracheophyta	Liliopsida	Fabales	Fabaceae	Se
変種	Plantae	Tracheophyta	Magnoliopsida	Ericales	Ericaceae	Py
種	Animalia	Chordata		Perciformes	Kyphosidae	Gi
変種	Plantae	Tracheophyta	Magnoliopsida	Solanales	Solanaceae	All
種	Plantae	Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Tri
種	Plantae	Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Le
種	Plantae	Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	Ch
変種	Plantae	Tracheophyta	Liliopsida	Asparagales	Asphodelaceae	He

Exclude unmatched

GENERATE CSV

CSVに出力可能

高次分類群を埋めるとき便利

30行 エクステンション Wikidata

モード: 行 レコード 表示: 5 10 25 50 100 500 1000 行

1	verbatimScientificName	scientificName	key	matchType	confidence	status	rank	kingdom	phylum	class	order	family	genus	species	canonicalName
	Rostanga sp.	Rostanga Bergh, 1879	2291985	HIGHERRANK	94	ACCEPTED	GENUS	Animalia	Mollusca	Gastropoda	Mulibranchia	Dicopodidae	Rostanga		Rostanga
	Lacon sp.	Lacon Laporte, 1838	1162067	HIGHERRANK	92	ACCEPTED	GENUS	Animalia	Arthropoda	Insecta	Coleoptera	Elateridae	Lacon		Lacon
	Polydactylus plebeius	Polydactylus plebeius (Broussonet, 1782)	2393274	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Polynemidae	Polydactylus	Polydactylus plebeius	Polydactylus pleb
	Hibiscus hamabo Siebold et Zucc.	Hibiscus hamabo Siebold & Zucc.	7279115	EXACT	100	SYNONYM	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Malvales	Malvaceae	Talipariti	Talipariti hamabo	Hibiscus hamabo
	Decapterus macarellus	Decapterus macarellus (Cuvier, 1833)	2391064	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Carangidae	Decapterus	Decapterus macarellus	Decapterus maca
	Scorpaenodes guamensis	Scorpaenodes guamensis (Quoy & Gaimard, 1824)	2334382	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Scorpaeniformes	Scorpaenidae	Scorpaenodes	Scorpaenodes guamensis	Scorpaenodes gu
	Rhinogobius similis	Rhinogobius similis Gill, 1859	2377652	EXACT	97	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Gobiidae	Rhinogobius	Rhinogobius similis	Rhinogobius simil
	Narcotica niveosparsa	Narcotica niveosparsa	10339497	EXACT	99	ACCEPTED	SPECIES	Animalia	Arthropoda	Insecta	Lepidoptera	Noctuidae	Narcotica	Narcotica niveosparsa	Narcotica niveosp
	Linnaea borealis L.	Linnaea borealis L.	5334220	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Dipscales	Caprifoliaceae	Linnaea	Linnaea borealis	Linnaea borealis
	Vaccinium uliginosum L. var. japonicum T.Yamaz.	Vaccinium uliginosum var. japonicum T.Yamaz.	4167183	EXACT	100	ACCEPTED	VARIETY	Plantae	Tracheophyta	Magnoliopsida	Ericales	Ericaceae	Vaccinium	Vaccinium uliginosum	Vaccinium uliginos
	Trimma annosum	Trimma annosum Winterbottom, 2003	2378632	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Gobiidae	Trimma	Trimma annosum	Trimma annosum
	Gyraulus soritai Habe, 1976	Gyraulus soritai Habe, 1976	9225472	EXACT	100	SYNONYM	SPECIES	Animalia	Mollusca	Gastropoda		Planorbidae	Choanophthalus	Choanophthalus soritai	Gyraulus soritai
	Setaria viridis (L.) P.Beauv.	Setaria viridis (L.) P.Beauv.	5289684	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Liliopsida	Poales	Poaceae	Setaria	Setaria viridis	Setaria viridis
	Pyrola japonica Klienze ex Alef. var. subaphylla (Maxim.) Andres	Pyrola japonica var. subaphylla (Maxim.) Andres	7738475	EXACT	100	SYNONYM	VARIETY	Plantae	Tracheophyta	Magnoliopsida	Ericales	Ericaceae	Pyrola	Pyrola subaphylla	Pyrola japonica su
	Girella punctata	Girella punctata Gray, 1835	2381231	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Kyphosidae	Girella	Girella punctata	Girella punctata
	Alkekengi officinarum Moench var. franchetii (Mast.) R. J. Wang	Alkekengi officinarum var. franchetii (Mast.) R. J. Wang	9537212	EXACT	100	ACCEPTED	VARIETY	Plantae	Tracheophyta	Magnoliopsida	Solanales	Solanaceae	Alkekengi	Alkekengi officinarum	Alkekengi officina
	Trifolium repens L.	Trifolium repens L.	5358748	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Trifolium	Trifolium repens	Trifolium repens
	Lespedeza virgata (Thunb.) DC.	Lespedeza virgata (Thunb.) DC.	2955629	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Lespedeza	Lespedeza virgata	Lespedeza virgata
	Chenopodium acuminatum Willd.	Chenopodium acuminatum Willd.	3760713	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	Chenopodium	Chenopodium acuminatum	Chenopodium ac
	Hemerocallis citrina Baroni var. vespertina (H.Hara) M.Hotta	Hemerocallis citrina var. vespertina (H.Hara) M.Hotta	2781039	EXACT	100	ACCEPTED	VARIETY	Plantae	Tracheophyta	Liliopsida	Asparagales	Asphodelaceae	Hemerocallis	Hemerocallis citrina	Hemerocallis citri
	Eufentonia nihonica	Eufentonia nihonica	1822924	EXACT	99	ACCEPTED	SPECIES	Animalia	Arthropoda	Insecta	Lepidoptera	Notodontidae	Eufentonia	Eufentonia nihonica	Eufentonia nihoni
	Rhynchocypris lagowskii steindachneri	Rhynchocypris lagowskii (Dybowski, 1869)	2365606	HIGHERRANK	99	ACCEPTED	SPECIES	Animalia	Chordata		Cypriniformes	Cyprinidae	Rhynchocypris	Rhynchocypris lagowskii	Rhynchocypris lag
	Prunella vulgaris L. subsp. asiatica (Nakai) H.Hara	Prunella vulgaris subsp. asiatica (Nakai) H.Hara	5829992	EXACT	100	ACCEPTED	SUBSPECIES	Plantae	Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	Prunella	Prunella vulgaris	Prunella vulgaris a
	Pseudanthias rubrizonatus	Pseudanthias rubrizonatus (Randall, 1963)	2387622	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Perciformes	Serranidae	Pseudanthias	Pseudanthias rubrizonatus	Pseudanthias rubi
	Cantherhines pardalis	Cantherhines pardalis (Rüppell, 1837)	2406982	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Tetraodontiformes	Monacanthidae	Cantherhines	Cantherhines pardalis	Cantherhines parc
	Achellognathus rhombeus	Achellognathus rhombeus (Temminck & Schlegel, 1846)	2359434	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Cypriniformes	Cyprinidae	Achellognathus	Achellognathus rhombeus	Achellognathus rh
	Maianthemum japonicum (A.Gray) LaFrankie	Maianthemum japonicum (A.Gray) LaFrankie	2769420	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Liliopsida	Asparagales	Asparagaceae	Maianthemum	Maianthemum japonicum	Maianthemum jap
	Comibaena ingrata	Comibaena ingrata (Wileman, 1911)	11701672	EXACT	99	ACCEPTED	SPECIES	Animalia	Arthropoda	Insecta	Lepidoptera	Geometridae	Comibaena	Comibaena ingrata	Comibaena ingrat
	Chelonodon patoca	Chelonodon patoca (Hamilton, 1822)	2407317	EXACT	99	ACCEPTED	SPECIES	Animalia	Chordata		Tetraodontiformes	Tetraodontidae	Chelonodon	Chelonodon patoca	Chelonodon patoc
	Ribes latifolium Jancz.	Ribes latifolium Jancz.	4194367	EXACT	100	ACCEPTED	SPECIES	Plantae	Tracheophyta	Magnoliopsida	Saxifragales	Grossulariaceae	Ribes	Ribes latifolium	Ribes latifolium

出力されたCSVファイル

ファセット/フィルタ

取り消す/やり直す 23 / 23

再読み込み すべてリセット すべて削除

× 記録年月日 (始め) 変更 リセット

2,541 マッチした行: (2,616 t

モード: 行 レコード 表示: 5

全て レコード種別

☆	↶	4.	PreservedSpecimen
---	---	----	-------------------

編集履歴を開く

ファセット/フィルタ

取り消す/やり直す 23 / 23

抜き出し... 適用...

フィルタ

11. カラム 界名 (学名) を削除

2,541 マッチした行: (2,616 t

モード: 行 レコード 表示: 5

全て レコード種別

☆	↶	4.	PreservedSpecimen
☆	↶	5.	PreservedSpecimen

「抜き出し...」から別のファイルに適用したい
操作履歴を取り出す

操作履歴を出力し次回も使えるレシピに

操作履歴を抜き出し

操作履歴を抜き出し、他のプロジェクトなどで適用できるようにJSONに保存します。

- カラム カテゴリ を削除
- カラム オリジナルの学名 を削除
- カラム 界名 (学名) を削除
- カラム 界名 (日本語名) を削除
- 式 value.toNumber() でカラム 緯度 (十進数表記) のセルを変換
- 式 value.toNumber() でカラム 経度 (十進数表記) のセルを変換
- 式 value.toNumber() でカラム 最低海拔 のセルを変換
- 式 value.toNumber() でカラム 最高海拔 のセルを変換
- 式 value.toNumber() でカラム 緯度経度誤差半径 のセルを変換
- 式 value.toNumber() でカラム 記録年月日 (終わり) のセルを変換
- 式 value.toNumber() でカラム 記録年月日 (始め) のセルを変換
- 式 value.replace(" Pref.", "") でカラム 都道府県 のセルを変換
- 式 value.replace("Prefecture", "") でカラム 都道府県 のセルを変換
- 式 value.replace("-ken", "") でカラム 都道府県 のセルを変換
- 式 value.replace(" ", "") でカラム 都道府県 のセルを変換

全選択 全選択解除

出力

```
{
  "type": "range",
  "name": "記録年月日 (始め)",
  "expression": "value",
  "columnName": "記録年月日 (始め)",
  "from": 20240200,
  "to": 20241300,
  "selectNumeric": true,
  "selectNonNumeric": true,
  "selectBlank": true,
  "selectError": true
}
],
"mode": "row-based"
},
"columnName": "都道府県",
"expression": "value.replace(\" \", \"\")",
"onError": "keep-original",
"repeat": false,
"repeatCount": 10,
"description": "式 value.replace(\" \", \"\") でカラム 都道府県 のセ
```

JSONファイルとして保存可能

閉じる

ファセット/フィルタ

取り消す/やり直す 23 / 23

抜き出し... 適用...

2,541 マッチした行: (2,616 t

モード: 行 レコード 表示: 5

全て	レコード種別
4.	PreservedSpecimen
5.	PreservedSpecimen

JSONファイルを「適用」から読み込むとそのファイルにも同じ操作が適用される

操作履歴を出力し次回も使えるレシピに

000 行

記録年月日オリジナル表記	▼ 国	▼ 国 (日本語)	▼ 都道府県	▼ 郡・市・町
Jan-54	Japan	日本	Osaka	
19年5月15日	Japan	日本	Gunma Pref	
15/11/12	Japan	日本国	Ibaraki Pref	
'8年3月10日	Japan	日本	Ibaraki Pref	
lay-55	Japan	日本	Gunma Pref	
14/1/30	Japan	日本		
13年5月7日	Japan	日本	Gunma Pref	
Nov-06	Japan	日本	Ibaraki Pref	
17年2月26日	Japan	日本	Osaka	
Mar-53	Japan	日本国	Gunma Pref	
13/2/17	Japan	日本	Ibaraki Pref	
'8年5月3日	Japan	日本		
16年8月22日	Japan	日本	Gunma P.	
'6年3月4日	Japan	日本	Osaka Prefecture	大阪府 Kishiwada-shi
11/6/9	Japan	日本国	Gunma Prefecture	群馬県 Midori-shi

プロジェクトをエクスポート

タブ区切り (TSV) で出力

カンマ区切り (CSV) で出力

HTMLテーブル

Excel (.xls)

Excel 2007+ (.xlsx)

ODFスプレッドシート

カスタム表形式...

SQL...

テンプレート...

Wikidata編集...

クイックステートメント・ファイル

Wikidataスキーマ

次回様々な出力形式に対応

新規プロジェクト

既存プロジェクト

インポート

言語設定

拡張機能

🔍 All

		最終更新日時	プロジェクト名	タグ	作成者	題名	内容	行数
×	文書情報	2026/02/28 13:59:03	dummy specimen snet 30 csv					30
×	文書情報	2026/02/27 10:17:28	Book 2026 02 27 xlsx					119
×	文書情報	2026/02/02 15:59:05	rm presentations20260202 csv					32



バージョン 3.9.5 [TRUNK]

設定
ヘルプ
About

作業ディレクトリを閲覧

補足：高速化のためのテクニック

Windowsの場合：

OpenRefineのフォルダ内にある `openrefine.l4j.ini` というファイルをメモ帳で開き、`-Xmx1024M`（または`2048M`）となっている数字を、PCの空きメモリに合わせて `-Xmx4096M`（4GB）などに書き換えて上書き保存

Macの場合：

アプリケーションフォルダのOpenRefineを右クリックし「パッケージの内容を表示」> **Contents** > **Info.plist** をテキストエディタで開き、`-Xmx1024M` の部分を書き換え

搭載メモリの1/4ぐらいにすると良い
16GBなら「`-Xmx4096M`」
64GBなら「`-Xmx16384M`」

```
<string>-Xdock:name=OpenRefine</string>  
<string>-Xms512M</string>  
<string>-Xmx2048M</string>  
<string>-Drefine.version=3.9.5</string>  
<string>-Drefine.version=${APP_ROOT}/Co
```

実習前の質疑応答

OpenRefineのインストール

Google

openrefine



AI モード すべて 画像 動画 ショッピング ニュース ショート動画 もっと見る ツール



OpenRefine

<https://openrefine.org> · このページを訳す

OpenRefine

OpenRefine is a powerful free, open source tool for working with messy data: cleaning it; transforming it from one format into another; and extending it.



関西大学

<https://www.dh.ku-orcas.kansai-u.ac.jp> · ...

OpenRefineを使ってデータをきれいにする方法

2020/03/31 — OpenRefineというプログラムを使えば、空白セル、重複レコード、表記ゆれなどのシステムティックなエラーを簡単に特定できるようになります。OpenRefineは ...

レッスンの目標

OpenRefineの概要

演習の概要: パワーハウスミュ...



Wikipedia

<https://ja.wikipedia.org/wiki/OpenRefine> · ...

OpenRefine

OpenRefine (旧称: Google Refine、さらに以前はFreebase Gridworks) は、スタンドアロンでオープンソースのデスクトップアプリケーションで、データのクリーンナップ、他の ...



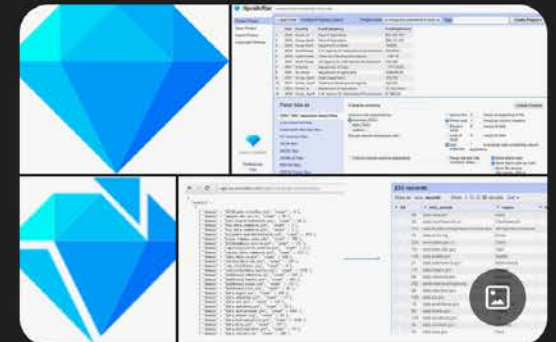
Qiita

<https://qiita.com/openrefine> · ...

OpenRefineで神エクセルと戦う

2020/05/31 — 「▽市町村」の▽をクリックしてセル編集>下方向の空白を埋めると選びます。すると全ての行に市町村名がコピーされ、空白セルはなくなりました。

OpenRefine



OpenRefineは、スタンドアロンでオープンソースのデスクトップアプリケーションで、データのクリーンナップ、他の形式への変換、データラングリングと呼ばれる処理を行うことができる。見た目はスプレッドシートアプリケーションに似ているが、よりデータベースに近い動作をする。

出典: ウィキペディア

ライセンス: BSD License

開発元: Freebase。後にGoogle。現在はオープンソースコミュニティ

OpenRefineのインストール

OpenRefine

OpenRefine is a powerful free, open source tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data.

Our goal is to empower everyone to meaningfully engage with data by providing an accessible open source tool and nurturing a diverse, supportive community.



[Download](#)

ここからダウンロード

Main features



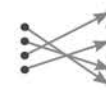
Faceting

Drill through large datasets using facets and apply operations on filtered views of your dataset.



Clustering

Fix inconsistencies by merging similar values thanks to powerful heuristics.



Reconciliation

Match your dataset to external databases via reconciliation services.



Infinite undo/redo

Revert to any previous state of your



Privacy

Your data is stored on your machine



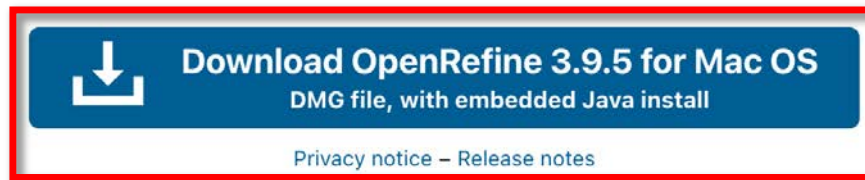
Wikibase

Connect to Wikibase, the free

OpenRefineのインストール

Download OpenRefine

OpenRefine is free software released under the [BSD 3-clause license](#), brought to you by [our contributors](#).



自分の環境が表示されていないならば
下から選ぶ

You also can install OpenRefine via Homebrew with: `brew install --cask openrefine`

Other platforms and versions

Version	Windows (installer)	Windows (zip file)	Mac OS	Linux
3.9.5 (2025-09-02)	↓ EXE	↓ ZIP	↓ DMG	↓ TAR.GZ
3.8.7 (2024-11-21)	↓ EXE	↓ ZIP	↓ DMG	↓ TAR.GZ
3.7.9 (2024-02-10)		↓ ZIP	↓ DMG	↓ TAR.GZ

See the [full list of releases on GitHub](#) for other versions.

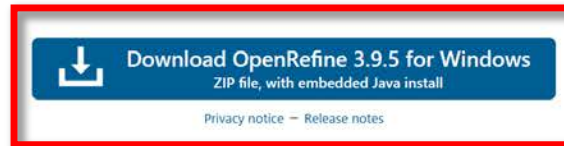
Extensions

See the [extensions](#) for additional ways to use and extend OpenRefine.

OpenRefineのインストール (Windows)

Download OpenRefine

OpenRefine is free software released under the [BSD 3-clause license](#), brought to you by our contributors.



You can also install OpenRefine via [package managers](#) of various Linux distributions.

Other platforms and versions

Version	Windows (installer)	Windows (zip file)	Mac OS	Linux
3.9.5 (2025-09-02)	↓ EXE	↓ ZIP	↓ DMG	↓ TAR.GZ
3.8.7 (2024-11-21)	↓ EXE	↓ ZIP	↓ DMG	↓ TAR.GZ
3.7.9 (2024-02-10)		↓ ZIP	↓ DMG	↓ TAR.GZ

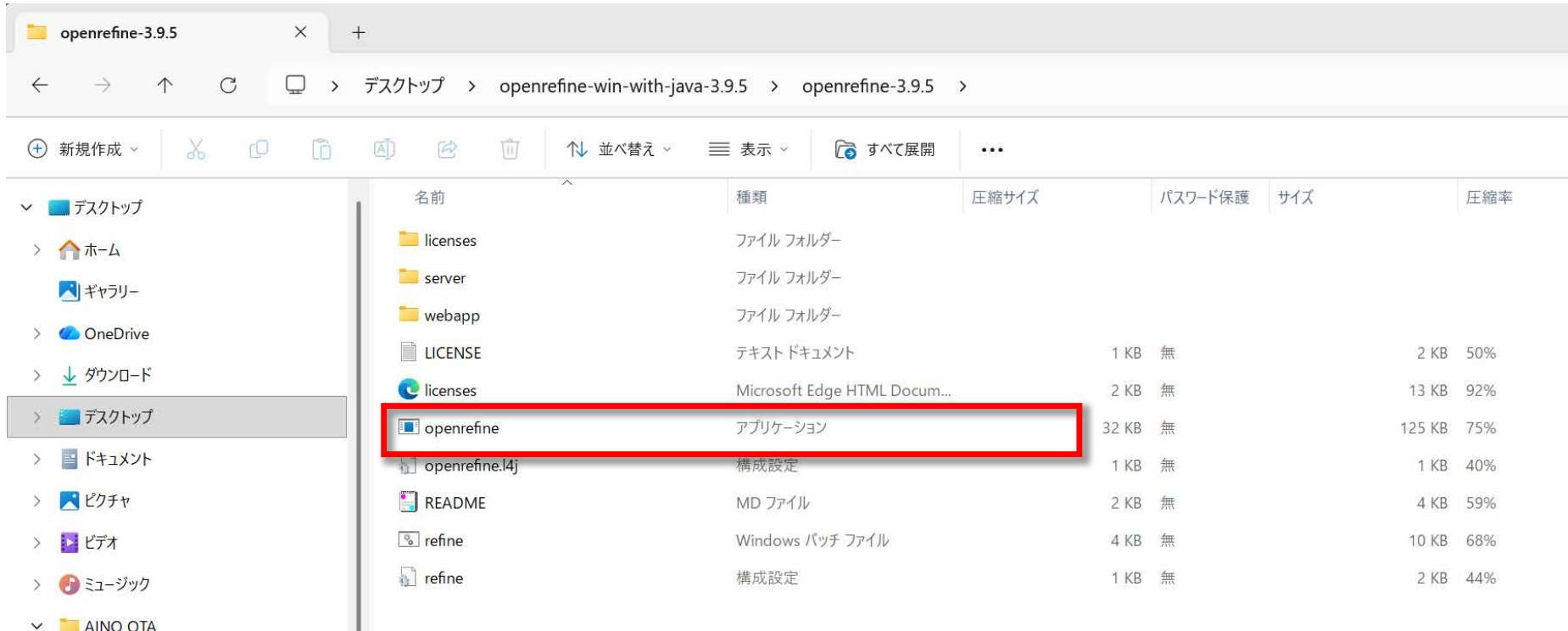
See the [full list of releases on GitHub](#) for other versions.

Extensions

See the [extensions](#) for additional ways to use and extend OpenRefine.

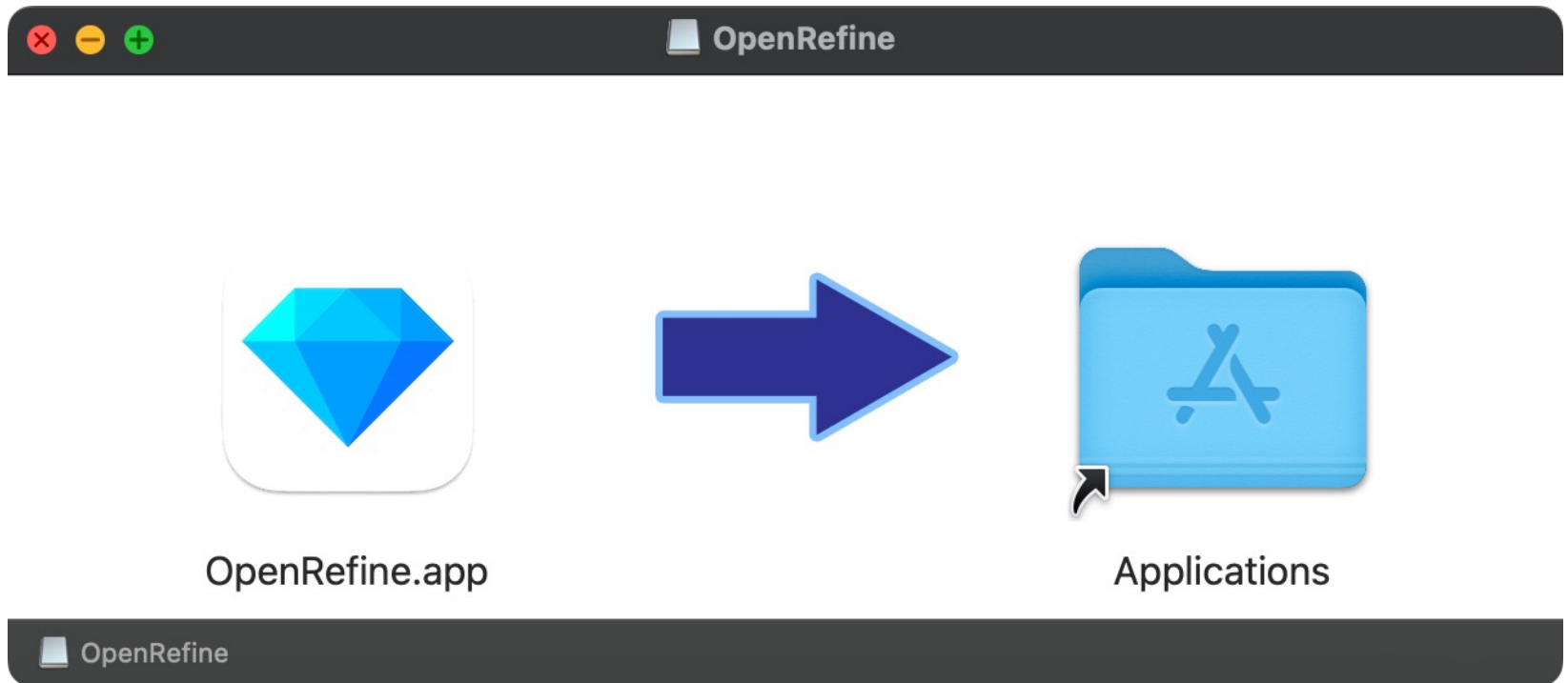
zipファイルをダウンロード・解凍

OpenRefineのインストール (Windows)



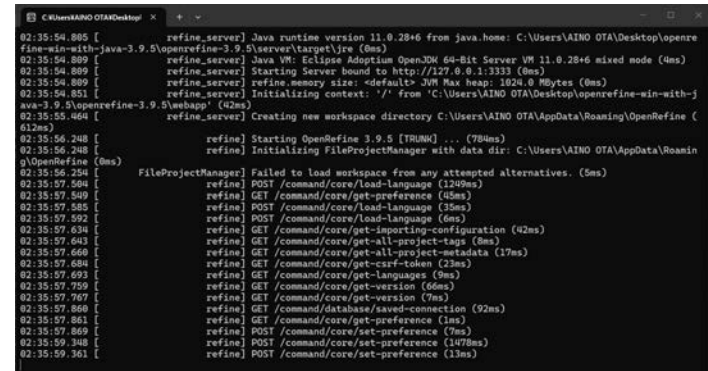
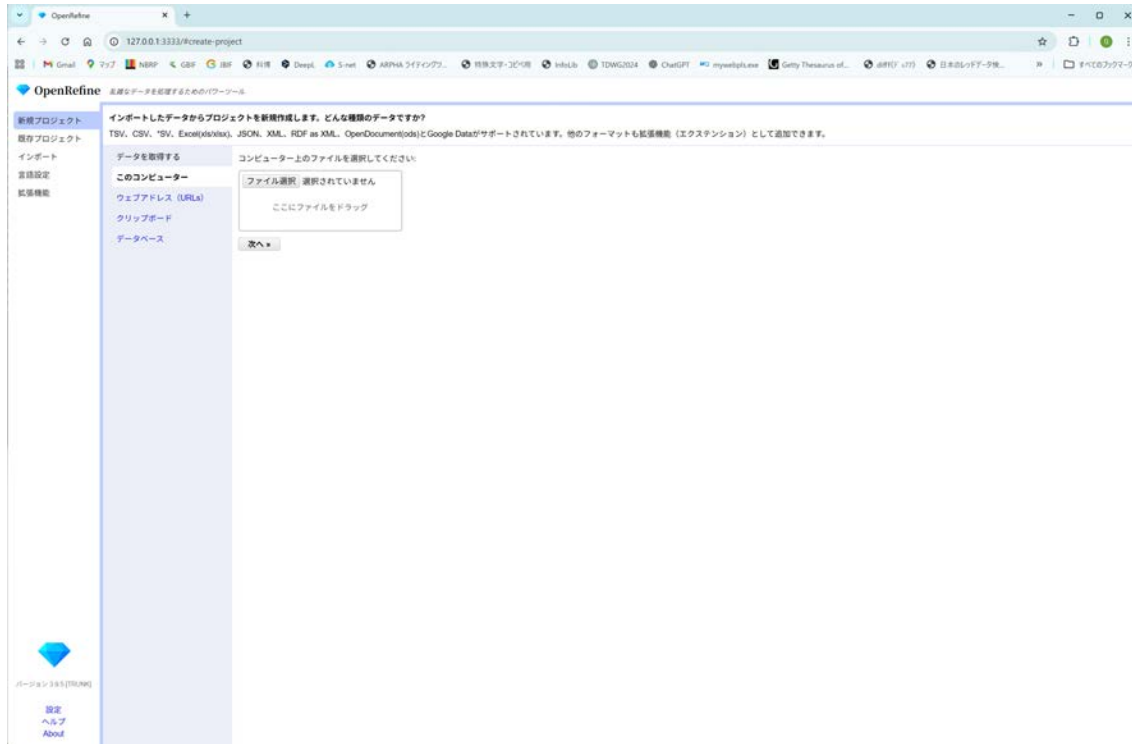
openrefineというアプリケーションを開く

OpenRefineのインストール (Mac)



**OpenRefine.appを
アプリケーションフォルダにドラッグ
アプリケーションフォルダからOpenRefine.appを開く**

OpenRefineのインストール

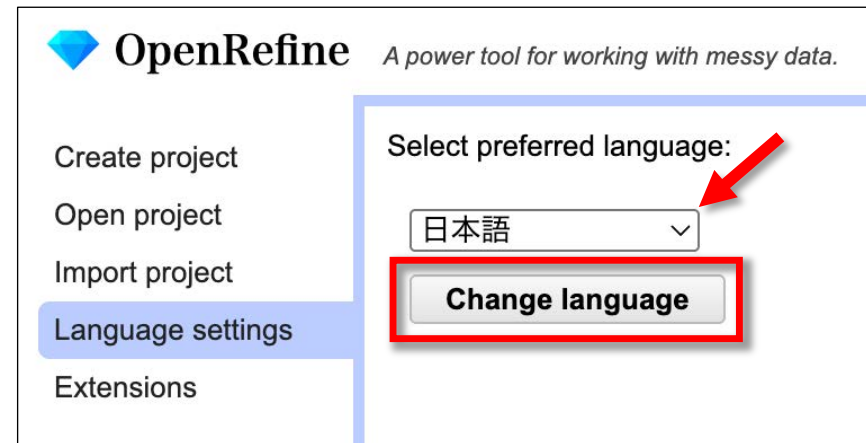
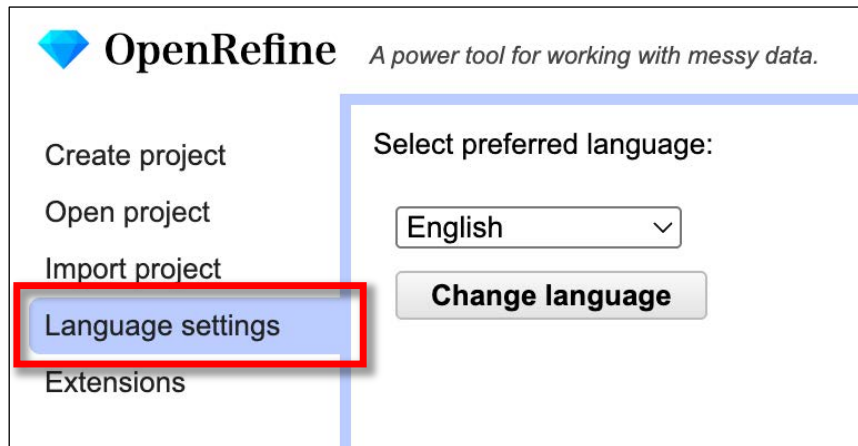


黒い画面は無視する
(Windows)

起動画面の例 (共通)

ブラウザ (Edge、Chrome、Safariなど) で
OpenRefineが起動する
インターネット接続は不要

OpenRefineのインストール



使用言語を日本語に設定可能

左タブ→Language settings→「日本語」を選択→Change language

デモデータが終わった人むけ: specimen_20260224.csv



サイエンスミュージアム ネット
Science Museum Net S-Net

「詳しく検索」では学名、和名、採集した年月日、採集場所の地名など「絞り込んだ検索」ができます。

一覧で並び替えを行うには10万件以下、採集場所の地図表示を行うには2000件以下、

検索結果のダウンロードを行うには5000件以下になるように検索結果を絞り込んでください。

English

検索 簡易検索

学名	<input type="text"/>	部分一致 ▾
和名 (全角カナ)	<input type="text"/>	部分一致 ▾
分類	<input type="text"/>	部分一致 ▾
界名、門名、綱名、目名、科名、属名、亜属名の英語と日本語の情報から検索します。		
記録年月日	<input type="text" value="2022"/> 年 <input type="text" value="7"/> 月 <input type="text" value="1"/> 日 ~ <input type="text" value="2022"/> 年 <input type="text" value="7"/> 月 <input type="text" value="31"/> 日	
地名	<input type="text"/>	部分一致 ▾
国、都道府県、郡・市区町村、詳細地名の英語と日本語の情報から検索します。		
海拔	<input type="text"/> ~ <input type="text"/> m	
水深	<input type="text"/> ~ <input type="text"/> m	
緯度 (十進数表記)	<input type="text"/> ~ <input type="text"/>	-90~90の値を指定。正の値は北半球 (北緯)、負の値は南半球 (南緯)。十進の数値は度+分/60+秒/3600で求められます。
経度 (十進数表記)	<input type="text"/> ~ <input type="text"/>	-180~180の値を指定。正の値は東半球 (東経)、負の値は西半球 (西経)。十進の数値は度+分/60+秒/3600で求められます。

S-Netから2022年7月のデータを抽出

デモデータが終わった人むけ: specimen_20260224.csv



サイエンスミュージアム ネット
Science Museum Net S-Net

レコード種別

PreservedSpecimen (2,861)

機関名 (日本語)

鹿児島大学総合研究博物館 (794)

筑波大学山岳科学センター菅平高原実験所
(502)

[もっと見る](#)

コレクションコード

I (920)

Insect (437)

[もっと見る](#)

国 (日本語)

日本 (2,677)

- (119)

[もっと見る](#)

都道府県 (日本語)

鹿児島県 (670)

長野県 (505)

[もっと見る](#)

タイプ標本種別

Paratype (4)

Holotype (1)

[もっと見る](#)

学名

[ホーム](#) > 検索結果

自然史標本情報

該当件数: 2,861 件 (1-20件目)

20

件ずつ表示

ソート順



ダウンロード



絞り込み検索

1 2 3 4 5 6 > >>

No	学名	和名	記録年月日 (始め)	国 (日本語)	都道府県 (日本語)	機関名 (日本語)
1	Favonigobius gymnauchen	ヒメハゼ	20220728	日本	大阪府	きしわだ自然資料館
2	Acanthogobius flavimanus	マハゼ	20220728	日本	大阪府	きしわだ自然資料館
3	Trdentiger obscurus	チチブ	20220728	日本	大阪府	きしわだ自然資料館
4	Gabbia kiusiuensis (Hirase, 1927)	ヒメマルマメタニシ	20220723	日本	京都府	きしわだ自然資料館
5	Myrmica jessensis Forel	エゾクシケアリ	20220704	日本	北海道	釧路市立博物館
6	Myrmica jessensis Forel	エゾクシケアリ	20220704	日本	北海道	釧路市立博物館
7	Myrmica kurokii Forel	クロキクシケアリ	20220704	日本	北海道	釧路市立博物館
8	Myrmica kurokii Forel	クロキクシケアリ	20220704	日本	北海道	釧路市立博物館

S-Netから2022年7月のデータを抽出

デモデータが終わった人むけ: specimen_20260224.csv

The screenshot shows the S-Net search interface. On the left is a sidebar with filters for Record Type, Institution Name, Collection Code, Country, Prefecture, Type, and Scientific Name. The main area contains a '実行' (Execute) button and a list of search criteria with checkboxes. The criteria are organized into three columns. The first column includes 'グローバルユニーク番号', 'レコード種別', '機関名', '機関名 (日本語)', '機関コード', 'コレクションコード', 'カタログ番号', '採集者番号', 'オカレンス備考', 'オカレンス備考 (日本語)', and '性別'. The second column includes '島 (日本語)', '国', '国 (日本語)', '国地域コード', '都道府県', '都道府県 (日本語)', '郡・市区町村', '郡・市区町村 (日本語)', '詳細地名', '詳細地名 (日本語)', and '地名オリジナル表記'. The third column includes '和名', '同定に関する補足情報', '同定に関する補足情報 (日本語)', '分類群に関する備考', '分類群に関する備考 (日本語)', '備考1 (公開,日本語)', '上限の推定年代 (累代・累界レベル)', '上限の推定年代 (累代・累界レベル) (日本語)', '下限の推定年代 (累代・累界レベル)', '下限の推定年代 (累代・累界レベル) (日本語)', and '上限の推定年代 (代・界レベル)'. A '閉じる' (Close) button is in the top right corner.

サイエンスミュージアム
Science Museum

レコード種別
PreservedSpecimen (2,861)

機関名 (日本語)
鹿児島大学総合研究博物館 (7)
筑波大学山岳科学センター管
(502)
[もっと見る](#)

コレクションコード
I (920)
Insect (437)
[もっと見る](#)

国 (日本語)
日本 (2,677)
- (119)
[もっと見る](#)

都道府県 (日本語)
鹿児島県 (670)
長野県 (505)
[もっと見る](#)

タイプ標本種別
Paratype (4)
Holotype (1)
[もっと見る](#)

学名
Dendrophilus (Dendrophilus) xax
Myrmica ruginodis Nylander (57)
[もっと見る](#)

界名 (日本語名)
動物界 (2,030)
植物界 (760)
[もっと見る](#)

実行

閉じる

ファイル形式: Excel形式 CSV形式(カンマ区切りテキスト形式) TSV形式(タブ区切りテキスト形式)
※Excel形式は、検索結果が1000件以下の場合のみ利用できます。

文字コード: S-JIS UTF-8

項目名
 全て

グローバルユニーク番号

島 (日本語)

レコード種別

国

機関名

国 (日本語)

機関名 (日本語)

国地域コード

機関コード

都道府県

コレクションコード

都道府県 (日本語)

カatalog番号

郡・市区町村

採集者番号

郡・市区町村 (日本語)

オカレンス備考

詳細地名

オカレンス備考 (日本語)

詳細地名 (日本語)

性別

地名オリジナル表記

和名

同定に関する補足情報

同定に関する補足情報 (日本語)

分類群に関する備考

分類群に関する備考 (日本語)

備考1 (公開,日本語)

上限の推定年代 (累代・累界レベル)

上限の推定年代 (累代・累界レベル) (日本語)

下限の推定年代 (累代・累界レベル)

下限の推定年代 (累代・累界レベル) (日本語)

上限の推定年代 (代・界レベル)

S-Netから2022年7月のデータを抽出

デモデータが終わった人むけ: specimen_20260224.csv

余裕があったらチャレンジ

・ **オカレンス最多の都道府県（日本語）
は？**

文字列ファセット→ソート基準でカウントに

・ **緯度、経度の範囲は合ってる？**

よく使う変換→数値型に変換→数値ファセット

・ **都道府県の表記ゆらぎを修正**

置換を活用（正規表現を使ってみてもOK）