

# 日本の気象観測データの (オープンなようでオープンでない) 話

27<sup>th</sup> June 2020, Tokyo.R #86  
Yuta Kanzawa @yutakanzawa



Data Science Senior Analyst at Janssen Pharmaceutical K.K., Tokyo  
A Family Company of Johnson & Johnson



# I am...

- 神沢雄大 Yuta Kanzawa (twitter: [@yutakanzawa](https://twitter.com/yutakanzawa))
- Data scientist at Janssen Japan, Tokyo  | PHARMACEUTICAL COMPANIES OF 
- A pharmaceutical company of J&J
- Opera & wine lover
  - Wagner
  - Bourgogne
- 7 languages
  - Human: Japanese, English, German
  - Computer: R, Python, SAS, SQL



# アジェンダ

- 今日話すこと
  - オープンデータ
  - 気象庁の気象観測データのダウンロードサービス
  - 国土数値情報API（悲報）
- 今日話ないこと
  - 気象観測データ
  - 国土数値情報
  - Python

# 前回までの話： 国土数値情報APIのススメ

Previously on Tokyo.R #84

# (再掲) 始まり：日本の1キロメッシュの年平均気温図を作りたい。

 **R-bloggers**  
R news and tutorials contributed by hundreds of R bloggers

Home | About | RSS | add your blog! | Learn R | R jobs | Contact us

**WELCOME!**

 Follow @rbloggers 81.9t

Here you will find daily news and tutorials about R, contributed by hundreds of bloggers.

There are many ways to follow us -  
By e-mail:  
  
  
52.206 readers  
BY FEEDBURNER

On Facebook:  
 **R blog...**  
78K likes

## Temperature changes in Germany visualized in R

November 5, 2019  
By Yongfu, Liao

 Like 263  Share  Tweet  Share

[This article was first published on [Jan Knappe - Personal Blog](#), and kindly contributed to R-bloggers]. (You can report issue about the content on this page [here](#))

Want to share your content on R-bloggers? [click here](#) if you have a blog, or [here](#) if you don't.

 Share  Tweet

Recently ZEIT online, a leading German online newspaper, published an article on long-term changes in temperature throughout the country based on publicly available data. The article included some very striking and informative visualization that I wanted to recreate using R and the tidyverse.

MOST VISITED ARTICLES OF THE WEEK

1. Tidying the John Hopkins Covid-19 data
2. Tidying the new Johns Hopkins Covid-19 time-series datasets
3. 5 Ways to Subset a Data Frame in R
4. How to create a simple Coronavirus dashboard specific to your country in R
5. How to write the first for loop in R
6. Date Formats in R
7. What to study if you're under quarantine
8. R – Sorting a data frame by the contents of a column
9. Installing R packages

**ZEIT ONLINE**

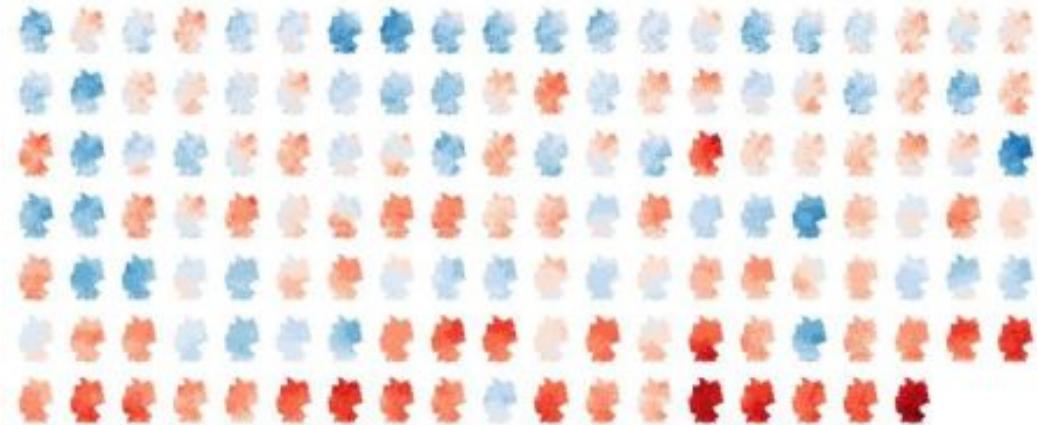
Suche

Politik Gesellschaft Wirtschaft Kultur Wissen Digital Campus Arbeit Entdecken Sport ZEITmagazin Podcasts mehr... 

SCHWERPUNKTE DONALD TRUMP KLIMAGIPFEL UKRAINE-KONFLIKT

Klimawandel

## Viel zu warm hier



Durch den Klimawandel hat sich Deutschland deutlich erwärmt. Wir zeigen, wie sich die Temperatur in Ihrer Gemeinde in den letzten 137 Jahren entwickelt hat.

Von Paul Bickel; Elena Erdmann; Flavio Gortana u.a.  
+ 745 Kommentare

\*1 <https://www.r-bloggers.com/temperature-changes-in-germany-visualized-in-r/>

\*2 <https://www.zeit.de/wissen/umwelt/2019-12/klimawandel-globale-erwaermung-warming-stripes-wohnort>

# (再掲) Next steps

- 国土数値情報API
  - ダウンロード機能をライブラリ化してみたい (RやPython)。
  - To 中の人：正式化と対応データ拡充を是非お願いします！
- 日本の年平均気温図
  - 気象庁サイトからの気象データダウンロード：大変そう
    - 意図的にAPIを用意していない。
    - 繙時的な観測所のリストがない。
      - サイトのHTMLからマイニングできそう<sup>\*1</sup>。



\*1 [http://washitake.com/weather/amedas/obs\\_stations.md](http://washitake.com/weather/amedas/obs_stations.md)

# 気象データダウンロードサービス

**Not so open data**

# その前に、オープンデータとは？

- ・日本政府の定義：『オープンデータ基本指針』\*1
  1. 営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの
  2. 機械判読に適したもの
  3. 無償で利用できるもの
- ・同指針のオープンデータに関する基本的ルール：公開環境\*1
  - ・「各府省庁は、ウェブサイトで容易に検索・利用できる形でデータを公開する。特にニーズが高いと想定されるデータについては、利用者の利便性に加え、システムの負荷及び効率性の観点から、一括ダウンロードを可能とする仕組みの導入や、APIを通じた提供を推進する。」

\*1 [https://cio.go.jp/sites/default/files/uploads/documents/data\\_shishin.pdf](https://cio.go.jp/sites/default/files/uploads/documents/data_shishin.pdf)

# 参考：日本政府の公式オープンデータカタログサイト\*1

日本語 | English

## DATA ● GO.JP データカタログサイト

お知らせ ▾ データ ▾ データベースサイト一覧 ▾ 公共データ活用事例 コミュニケーション ▾ 開発者向け情報 ▾

データセットを検索...

検索

[意見受付コーナー](#) にて、オープンデータの掲載に関する御要望等を受け付けております（匿名可）

データ

データセット 組織 グループ タグ

オープンデータの取組（リンク集）

オープンデータに関する方針・決定  
公共データ活用事例一覧  
データベースサイト一覧

コミュニケーション

意見受付コーナー 意見・回答公開コーナー  
掲載データ利用の御連絡

一億総活躍社会の実現

子育て支援 安心につながる社会保障 希望を生み出す強い経済

2020年東京オリンピック・パラリンピック競技大会

大会の円滑な準備及び運営 大会を通じた新しい日本の創造

地方公共団体データベースサイト一覧

\*1 <https://www.data.go.jp/>

# 気象庁「過去の気象データ・ダウンロード」\*1

The screenshot shows the 'Various Data & Materials' section of the JMA website. At the top, there are links for Home, Disaster Prevention Information, Various Data & Materials (highlighted in dark blue), Knowledge & Explanations, About JMA, and Inquiries & Applications. On the right, there are English language and Google Custom Search options, along with a search button.

The main content area is titled 'Past Meteorological Data Download'. It features a navigation bar with links for Important Notices, How to Use This Page, Frequently Asked Questions, and CSV File Format. Below this is a search condition section with tabs for 'Select Location' (selected), 'Select Item', 'Select Period', and 'Select Options'. A note states: 'One request per day can be made for data download. If many requests are made at once, it may take time to process.' To the right, there is a map of Japan showing various regions and specific locations like Sapporo, Hakodate, and Otaru. Buttons for 'Display on Screen' and 'Download CSV File' are also present.

On the left side, there are arrows for navigating through the data. A note at the bottom left says: 'Please select a prefecture first'.

At the bottom right, there is a note: 'Select the location you want to download data for'.

On the far right, there is a note: 'Select the items you want to download'.

At the very bottom right, there is a note: 'Select the period you want to download data for'.

\*1 <https://www.data.jma.go.jp/gmd/risk/obndl/index.php>

# お気づきだろうか？

地点を選ぶ

項目を選ぶ

期間を選ぶ

表

す

一回のリクエストで表示・ダウンロードできるデータ量には上限があります  
(右上棒グラフ参照)。また、このページへのアクセスが集中したり、リクエストのデータ量が多い場合には、表示・ダウンロードまで時間がかかる場合があります。

後志

檜山

# ダウンロードしてみた：東京気象台の過去30年間の日次平均気温

過去の気象データ・ダウンロード

! 重要なお知らせ

このページの使い方

よくある質問

CSVファイルの形式

検索条件

地点を選ぶ

項目を選ぶ

期間を選ぶ

表示オプションを選ぶ

選択済みのデータ量 0% 100% (上限)

期間選択の使い方

期間

連続した期間で表示する

最近1年 最近1か月

1990 年 6 月 25 日から

2020 年 6 月 25 日までの日別値を表示

特定の期間を複数年分、表示する

1 月 1 日から 1 月 1 日 の値を

2020 年から 2020 年まで表示

画面に表示 ▶

CSVファイルをダウンロード ▶

選択地点・項目をクリア

選択された地点 観測項目

東京



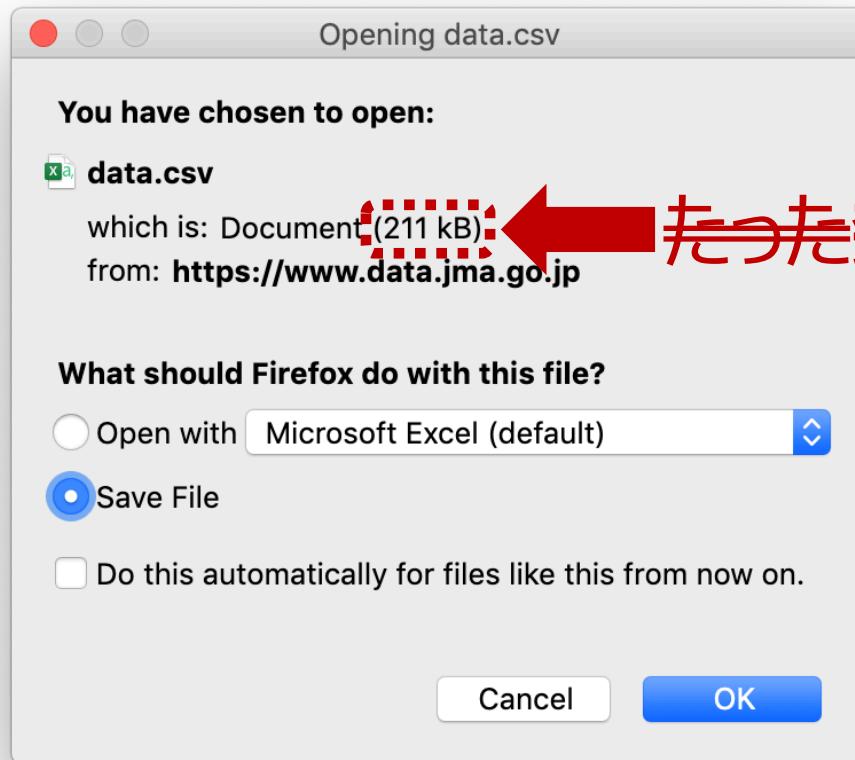
削除

選択された項目

日平均気温

削除

# 結果：イマイチ...



A screenshot of a CSV file in a text editor. The file contains 10,965 lines of data. A red dashed box highlights the first five lines, which include the header and some initial data. A large red arrow points upwards from the bottom of this box towards the top of the page. Red Japanese text is overlaid on the right side of the editor window.

1	ダウンロードした時刻 : 2020/06/26 20:54:52
2	
3	,東京,東京,東京
4	年月日,平均気温(°C),平均気温(°C),平均気温(°C)
5	品質情報,均質番号
6	1990/6/25,28.0,8,1
7	1990/6/26,24.5,8,1
8	1990/6/27,21.2,8,1
9	1990/6/28,19.5,8,1
10	1990/6/29,19.0,8,1
11	1990/6/30,20.2,8,1
12	1990/7/1,20.0,8,1
13	1990/7/2,22.3,8,1
14	1990/7/3,23.2,8,1
15	1990/7/4,23.6,8,1
16	1990/7/5,26.5,8,1
17	1990/7/6,26.2,8,1
18	1990/7/7,23.9,8,1
19	1990/7/8,22.5,8,1
20	1990/7/9,22.4,8,1
21	1990/7/10,21.5,8,1
22	1990/7/11,24.1,8,1
23	1990/7/12,19.8,8,1
24	1990/7/13,20.5,8,1
25	1990/7/14,24.9,8,1
26	1990/7/15,28.0,8,1
27	1990/7/16,29.4,8,1
28	1990/7/17,30.5,8,1
29	1990/7/18,30.0,8,1
30	1990/7/19,30.1,8,1

Lines: 10,965 Characters: 204,829 Location: 1 Line: 1 216 KB

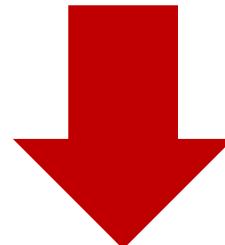
# なぜ上限がある？？？

- 「よくある質問」\*1へのご回答

取得できるデータ量に制限があって使いづらいのですが、まとめて取得する方法はありますか？

本ツールは、多くの方々が同時に利用しても円滑に動作するよう、一度に取得できるデータ量に制限を設けています。  
大容量のデータが必要な場合は、これらを販売している(一財)気象業務支援センターへお問い合わせください。

~~たった~~数百KB



~~天下り先へ利益誘導~~

気象ビッグデータのプロが有償対応！



\*1 <https://www.data.jma.go.jp/gmd/risk/obsdl/top/faq.html>

# (再掲) その前に、オープンデータとは？

- ・日本政府の定義：『オープンデータ基本指針』\*1
  1. 営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの
  2. 機械判読に適したもの
  3. 無償で利用できるもの
- ・『オープンデータに関する基本的ルール』：公開環境\*1
  - ・「各府省庁は、ウェブサイトで容易に検索・利用できる形でデータを公開する。特にニーズが高いと想定されるデータについては、利用者の利便性に加え、システムの負荷及び効率性の観点から、一括ダウンロードを可能とする仕組みの導入や、APIを通じた提供を推進する。」

\*1 [https://cio.go.jp/sites/default/files/uploads/documents/data\\_shishin.pdf](https://cio.go.jp/sites/default/files/uploads/documents/data_shishin.pdf)

# 自分でオープンにしてみた：

→気象庁のデータ検索ページ<sup>\*1</sup>からスクレイピング

ホーム > 各種データ・資料 > 過去の気象データ検索

過去の気象データ検索

各地の気温、降水量、風など 高層の気温、風など

最新の気象データ 過去の気象データ ダウンロード 地域平均 気象データ

地点と年月日時を選択して、表示するデータの種類を選択してください。検索条件を全てクリア

データの種類

年月日の選択

年月日の選択をクリア

都道府県 選択

都道府県・地方を選択

（未選択）

（未選択）

年月日の選択

年月日の選択をクリア

1月 1日 16日  
2月 2日 17日  
3月 3日 18日  
4月 4日 19日  
5月 5日 20日  
6月 6日 21日  
7月 7日 22日  
8月 8日 23日  
9月 9日 24日  
10月 10日 25日  
11月 11日 26日  
12月 12日 27日  
13日 28日  
14日 29日  
15日 30日  
31日

年ごとの値を表示  
(地点を指定してください)

3か月ごとの値を表示  
(地点、年を指定してください)

観測開始からの月ごとの値を表示  
(地点を指定してください)

月ごとの値を表示  
(地点、年を指定してください)

旬ごとの値を表示  
(地点を指定してください)

半旬ごとの平年値を表示  
(地点を指定してください)

日ごとの平年値を表示  
(地点、月を指定してください)

霜・雪・結氷の初終日と初冠雪日の平年値を表示  
(気象台、測候所などのみのデータです)

要素別データの公開期間  
(気象台、測候所などのみのデータです)

メッシュ平年値図

歴代全国ランキング

年月日は1981-2010年の30年間の観測値の平均をもとに算出しています。

```
library(tidyverse)
library(rvest)
library(lubridate)

obs <- readRDS("r_data/obs_stations.rdf")

obs_list <- mapply(list, obs$type, obs$prec_no, obs$block_no, SIMPLIFY=F)

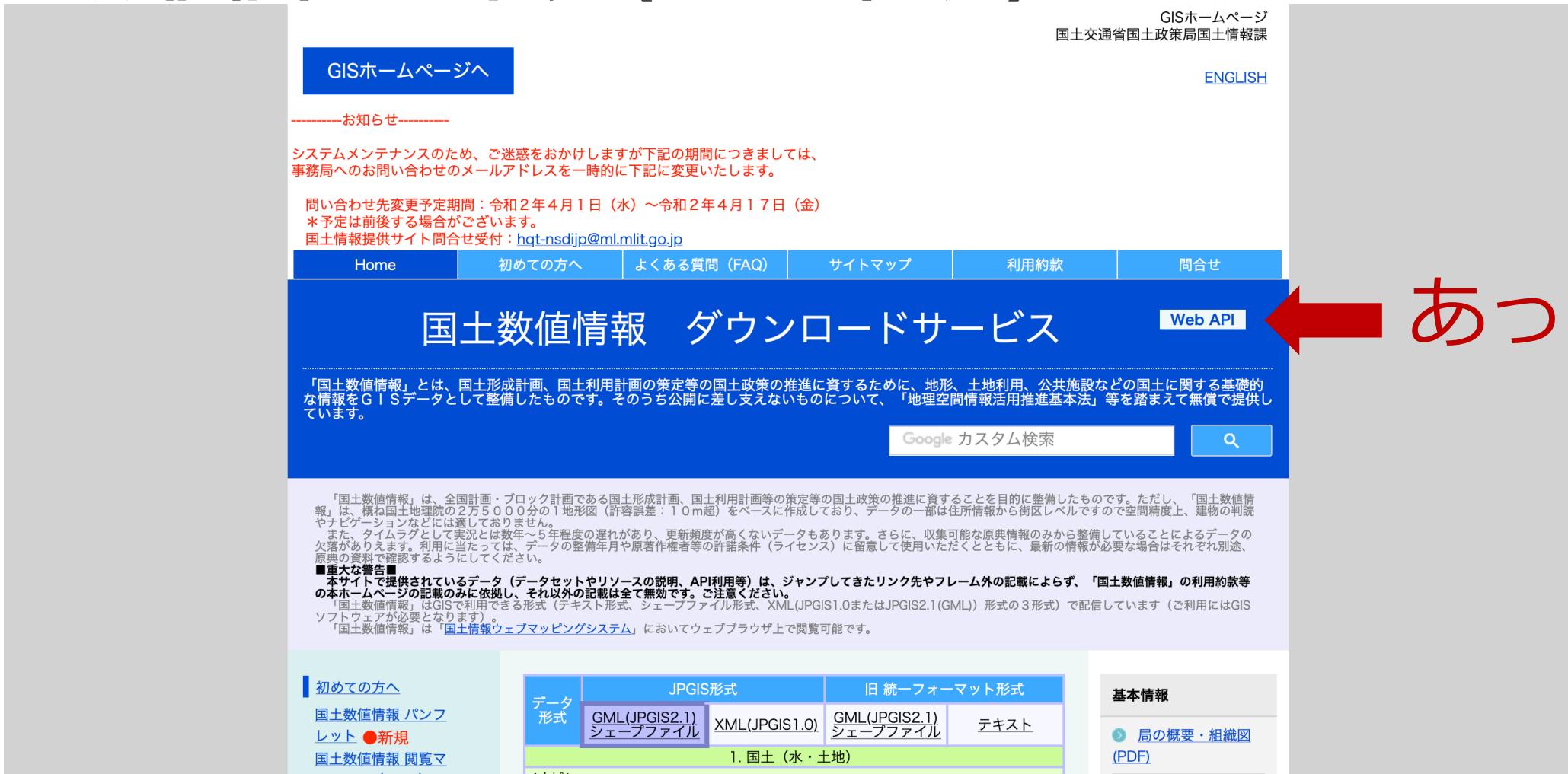
base_url <- "https://www.data.jma.go.jp/obd/stats/etrn/view/"
df <- data.frame()
for (e in obs_list){
  type <- e[[1]]
  prec_no <- e[[2]]
  block_no <- e[[3]]
  print(block_no)
  page_url <- paste0("annually_", type, ".php?prec_no=", prec_no,
  "&block_no=", block_no,
  "&year=&month=&day=&view=a2")
  block <- paste0(base_url, page_url) %>% read_html()
  tbl <- block %>% html_node("[id='tablefix1']") %>%
    html_table(header = FALSE, fill = TRUE)
```

\*1 <https://www.data.jma.go.jp/obd/stats/etrn/index.php>

# (悲報) 国土数値情報API

Goodbye

# 国土数値情報APIわざ（2020年4月）



GISホームページ  
国土交通省国土政策局国土情報課

ENGLISH

GISホームページへ

お知らせ

システムメンテナンスのため、ご迷惑をおかけしますが下記の期間につきましては、事務局へのお問い合わせのメールアドレスを一時的に下記に変更いたします。

問い合わせ先変更予定期間：令和2年4月1日（水）～令和2年4月17日（金）  
＊予定は前後する場合がございます。  
国土情報提供サイト問合せ受付：[hqt-nsdijp@ml.mlit.go.jp](mailto:hqt-nsdijp@ml.mlit.go.jp)

Home 初めての方へ よくある質問（FAQ） サイトマップ 利用約款 問合せ

## 国土数値情報 ダウンロードサービス

Web API

「国土数値情報」とは、国土形成計画、国土利用計画の策定等の国土政策の推進に資するために、地形、土地利用、公共施設などの国土に関する基礎的な情報をG I Sデータとして整備したものです。そのうち公開に差し支えないものについて、「地理空間情報活用推進基本法」等を踏まえて無償で提供しています。

Google カスタム検索

「国土数値情報」は、全国計画・ブロック計画である国土形成計画、国土利用計画等の策定等の国土政策の推進に資することを目的に整備したもので、ただし、「国土数値情報」は、概ね国土地理院の2万5000分の1地形図（許容誤差：100m超）をベースに作成しており、データの一部は住所情報から街区レベルまで空間精度上、建物の判読やナビゲーションなどには適しておりません。

また、タイムラグとして実況とは数年～5年程度の遅れがあり、更新頻度が高くなないデータもあります。さらに、収集可能な原典情報のみから整備していることによるデータの欠落があります。利用に当たっては、データの整備年月や原著作権者等の許諾条件（ライセンス）に留意して使用いただくとともに、最新の情報が必要な場合はそれぞれ別途、原典の資料で確認するようにしてください。

**重大な警告** ■ 本サイトで提供されているデータ（データセットやリソースの説明、API利用等）は、ジャンプしてきたリンク先やフレーム外の記載によらず、「国土数値情報」の利用約款等の本ホームページの記載のみに依拠し、それ以外の記載は全て無効です。ご注意ください。

「国土数値情報」はGISで利用できる形式（テキスト形式、シェーファイル形式、XML(JPGIS1.0またはJPGIS2.1(GML))形式の3形式）で配信しています（ご利用にはGISソフトウェアが必要となります）。

「国土数値情報」は「[国土情報ウェブマッピングシステム](#)」においてウェブブラウザ上で閲覧可能です。

初めての方へ

国土数値情報 パンフレット ●新規  
国土数値情報 閲覧マニュアル (PDF)

データ形式

	JPGIS形式	旧統一フォーマット形式
GML(JPGIS2.1) シェーファイル	XML(JPGIS1.0)	GML(JPGIS2.1) シェーファイル
1. 国土（水・土地）		

基本情報

局の概要・組織図(PDF)

あった

\*1 <http://nlftp.mlit.go.jp/ksj/index.html>

# 国土数値情報APIなう（2020年6月）

The screenshot shows the 'National Digital Land Information Download' page. The top navigation bar includes links for 'Top Page' and 'National Digital Land Information Download'. The main content area features a search bar with 'ENHANCED BY Google' and a search button. A large red text overlay 'もうない！' (No longer available!) is positioned on the right side. The left sidebar lists download categories: 'National Digital Land Information Download', 'Location Reference Information Download', 'National Survey (Land Classification Survey · Water Survey)', and 'Web Mapping'. Under 'GML(JPGIS2.1) Shapefile', there are five options: 1. 国土 (Water · Land), 2. Policy Area, 3. Region, 4. Transportation, and 5. Various Statistics. Below these are three other download formats: XML(JPGIS1.0), GML(JPGIS2.1) Shapefile, and Text. The central part of the page shows data formats: 'JPGIS Format' and 'Old Unified Format'. Under 'JPGIS Format', 'GML (JPGIS2.1) Shapefile' is highlighted. Under 'Old Unified Format', 'XML (JPGIS1.0)' and 'GML (JPGIS2.1) Shapefile' are listed, along with 'Text'. A dropdown menu for '1. 国土 (Water · Land)' is open, showing 'Waters' (水域) with sub-options: 'Coastline (Line)', 'Lake/Bay (Polygon)', 'Dam (Point)', 'Coastal Protection Facilities (Line) (Point)', 'Basin Mesh', and 'River (Line) (Point)'. A blue circular arrow icon with an upward arrow is located in the bottom right corner of the main content area.

\*1 <http://nlftp.mlit.go.jp/ksj/index.html>

# RIP...

- ・「API」でサイト内検索をしても0件。
- ・kokudosuuchiライブラリーはエラーが返ってくる。

```
> library(kokudosuuchi)
```

このサービスは、「国土交通省 国土数値情報（カテゴリ名）」をもとに加工者が作成  
以下の国土数値情報ダウンロードサービスの利用約款をご確認の上ご利用ください：

<http://nlftp.mlit.go.jp/ksj/other/yakkan.html>

```
> getKSJSummary()
```

read\_xml.raw(charToRaw(enc2utf8(x)), "UTF-8", ..., as\_html = as\_html, でエラー：  
EntityRef: expecting ';' [23]

```
> |
```

# まとめ

**Long story short**

# Next steps

- 気象庁の気象データダウンロードサービス
  - To 中の人：もっとオープンにして下さい！
- 国土数値情報API
  - To 中の人：復活させて下さい！！！
- 日本の年平均気温図
  - モデリング
    - まずは単純な重回帰

# Enjoy!