



一般社団法人
不動産テック協会

不動産オープンID

～ 基礎知識、政府・行政による取り組み、不動産IDについて ～

企画: 不動産テック協会, Geolonia
執筆: 片岡義明, Geolonia

公開日: 2023年3月
更新日: 2023年3月

TABLE OF CONTENTS

日本の住所	2
住所の基礎知識	2
2つの住所	3
① 住居表示住所	3
② 地番住所	4
住居表示住所と地番住所の関係	4
住所の構造	5
① 都道府県	5
② 市区町村	5
③ 町・丁目（ちょう・ちょうもく、ちょう・ちょうめ）	6
④ 番地・号	6
⑤ ビル名	7
⑥ 部屋番号	7
住居表示住所は、自治体が管理している	7
住所コード	7
住所の課題とソリューション	8
住所の表記ゆれ	8
住所の表記ゆれの原因	8
住所の正規化	11
座標の取得 = ジオコーディング	12
その他の住所課題	14
地名変更	14
建築途中の建物に住所がない問題	14
建物名の表記ゆれと建物名変更	15
ふりがなと英語表記	15
政府・行政による関連の取り組み	15
不動産オープンID API	18
基本的な仕組み	18
処理の方法	18
処理内容の概要	18
不動産オープンIDのソースコード	19
不動産オープンID API	19
Geolonia 住所データ	19
住所正規化エンジン	20
コミュニティージオコーダー	21
今後の展開について	21
本件に関するお問い合わせ	21

日本の住所

住所は場所を指し示す方法のひとつです。日本の住所は、太閤検地、明治期の住所制度整備、近代化を推し進めた時代の制度の高度化をへて現在に至ります。人々の暮らしや伝統から生まれてきた地名と、各時代の為政者による工夫によって整備、運用されてきました。

社会、企業のDXが急速に進む今、「位置」を指し示す重要な方法のひとつである住所をデジタル処理する必要に、私たちは歴史上初めて直面しています。「住所をデジタル化してコンピュータを活用したい」という一見するとシンプルな要請ですが、制度、地域の多様性由来する課題が多くあることが分かっています。

この資料では、住所の構造、課題をまとめて、現在利用できるソリューションをご紹介します。住所を使うすべての人が共通してぶつかる問題の全体像が理解できます。あなたの業務を進めるための一助となれることを狙っています。

住所の基礎知識

2つの住所

日本には主に2つの住所体系があります。建物につけられる「**住居表示住所**」と、土地につけられている「**地番住所**」です。

① 住居表示住所

山川里1丁目2番3号

住居表示住所は、住居表示法に基づいて地方自治体が定めたもので、戸建て、ビル、集合住宅などの建物が対象です。

住居表示住所は町名（字名）・街区符号・住居番号の組み合わせで表す街区方式と、道路名称と住居番号を組み合わせた道路方式の2通りがあり、ほとんどの市区町村では街区方式を採用しています。

街区方式では「山川里1丁目2番3号」や「山川里1-2-3」のような表記をされます。また、道路方式では「東根市板垣中通り18」のような表記をされます。いずれも直感的で分かりやすい仕組みで、行政サービス、郵便など、普段遣いをするのに分かりやすいメリットがあります。

日本全国のすべての建物に住居表示住所があるのではなく、現在も市区町村が主体となって整備が進行中です（[街区整備事業](#)といいます）。また、建物に付与される番号のため、山林、更地、農地、駐車場などに住居表示住所はありません。

② 地番住所

山川里1234番地

地番住所は、法務局が土地に振っている住所です。

「山川里1234番地」や「山川里1234」のような表記をされます。「山川里1234-1」のように枝番がつくこともあります。不動産の登記、税金の計算などに利用されるほか、住居表示住所がない地域では、地番住所が普段使われています。

土地に付与される番号であるため、数十、数百の建物が同じ地番住所を使っていたり、隣の建物であっても極端に離れた番号がついていたりして、荷物の配達や訪問といった日常生活、災害時の連絡・被害確認といった非常時の活動で不便があります。

住居表示住所と地番住所の関係



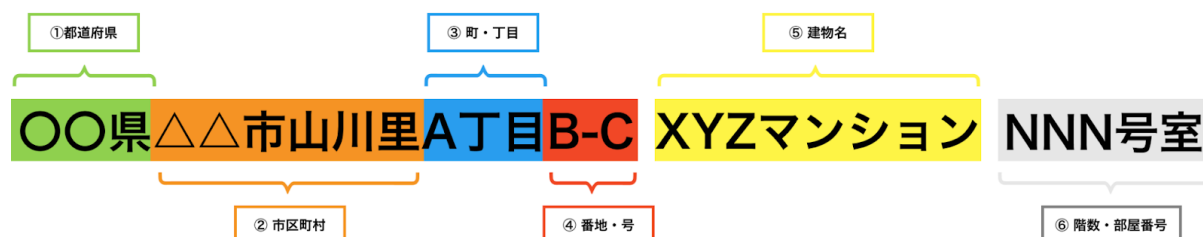
一方は土地に、もう一方は建物についており、管理の体系もまったく別です。住居表示住所がある建物にも、地番があります。住居表示住所を利用している人は、家やオフィスの地番を知らないことも普通です。

1つの地番住所の中に複数の建物が立っていたり、逆に一つの建物が複数の地番住所にまたがって立っていたり、1対1に対応するものではありません。お互いを機械的に変換する方法はありません。

住所の構造

住所の構造を、住居表示住所を使ってまとめています。

街区方式の住所は、おおむね以下のような構成でできています。



① 都道府県

47都道府県です。

② 市区町村

都道府県の下には、市、町、村があります。

「市」とは、町や村が人口、市街地に住んでいる人の割合、都市的な業態に就く人の割合などの要件を満たす場合に昇格したものです。

市、町、村の要件の主な違い

市	<ul style="list-style-type: none">● 人口5万人以上● 市街地にある戸数が、全体の6割以上● 商工業その他の都市的業態に従事する者及びその者と同一世帯に属する者の数が、全人口の6割以上
町	市街地要件、商工業従事者の割合などを都道府県の条例で決めている
村	要件はない

「区」は、政令指定都市だけにある下位区分です。例としては、神奈川県横浜市青葉区、愛知県名古屋市中区、福岡県福岡市博多区などがあります。東京23区は「市」をはさみず区がありますが、これは昔の東京市がなくなったときの名残です。

「郡」は、町や村の上にある行政区画のことで、住所に含めて書かれますが、行政の仕組み上の意味はありません。

「大字（おおあざ）」と「字（あざ）」は、もともと農山村だった地域で使われています。江戸時代の村を単位としているのが大字、さらに小さい集落が字です。現在では、合併を繰り返すうちに複数がかっついたり、消えてしまったりしています。

この他、北海道では条丁目という独自の仕組みが碁盤の目のような市街地において使われることがあります。

③ 町・丁目（ちょう・ちょうもく、ちょう・ちょうめ）

「〇〇町一丁目」などを指します。なお、「丁目」は漢数字を使うのが正しく、「番地」と「号」は算用数字を使うのが正式な表記となります。

④ 番地・号

街区（通常「番」と呼ばれる）と、住居番号（「号」と呼ばれる）で構成されます。

「街区」とは、町や字を道路や線路、河川や水路で分けた区分です。道路や川などを渡らずに歩いて一回りできるエリアと理解できます。海外ではストリートやアベニューといった、一本の道を1つのまとまりとして、番号をつけていく道路方式が多くあります。日本では、多くの地域で街区方式が採用されています。

「号」は住居の番号です。街区の中の建物に順番に数字を振っていきます。建物の増減や密集などの影響もあり、番号が飛んだりすることもあります。一般には、号までが分かっているれば建物が決定しますが、1つの住所に複数の建物が立つこともあります。

⑤ ビル名

雑居ビルの名前や集合住宅の名前です。「方書」と呼ばれることも多く、本来は住所の文字列とは切り離れたデータとして管理されるべきものです。

⑥ 部屋番号

集合住宅の部屋の番号や、雑居ビルなどの部屋の番号です。階数が記されていることも多いです。また、お店の名前などが入ってくることもあります。なお、集合住宅において戸数が20戸以上の場合は部屋番号までが正式な住居表示となります。

住所の構造についての参考リンク:

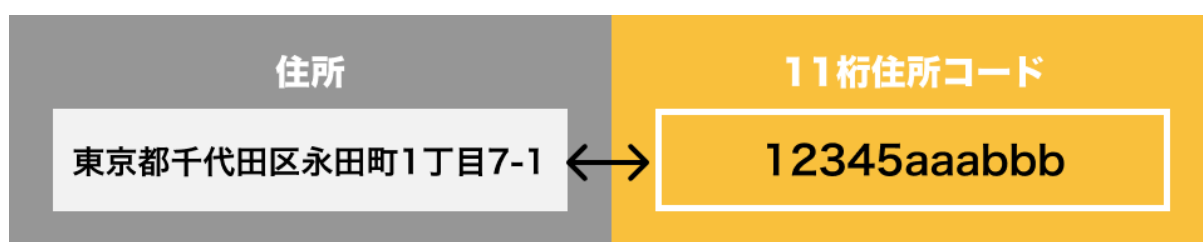
- [市とはなにか](#)（総務省）
- [政令指定都市について](#)（総務省）
- [街区方式と道路方式について](#)（Increment P, MAP WORLD+）
- [大字、字について](#)（ZENRIN）

住居表示住所は、自治体が管理している

住居表示住所の整備や地名の変更など、住所の管理は日本の政府や官庁が行っているのではなく、1,724の市区町村が行っています。中央で一元的に管理されていない理由は、地名がその土地に住む人々の暮らし、歴史、文化から生まれたものだからです。住居表示住所の整備に時間差があり、土地の分け方や表記のルールに地域差があるのはこのためです。

一方、デジタル庁の誕生の前後から、デジタル化の基盤を整えようとする議論や取り組みも始まっています。

住所コード



住所コードとは、長い文字列を数字に置き換えて管理を容易にするためのソリューションです。代表的なものが4つあります。官公庁や自治体で主に使われているのは国土地理協会の11桁住所コードです。4つの住所コードはそれぞれ仕様が異なり、互換性也没有ありません。

- [国土地理協会の住所コード](#)

国土地理協会（JGDC）が提供する住所マスターデータ「全国町・字ファイル」で使われている11桁の住所コードで、JISで定められている5桁の住所コード（JIS住所コード）に町・字コード（6～8桁目）と丁目・字コード（9～11桁目）が追加されています。

- [日本加除出版の各種コード](#)

日本加除出版が提供するデータベース「日本行政区画便覧データファイル」に収録されている住所コードで、統計で使用されている全国地方公共団体コード（3桁・5桁・6桁）に加えて、市区町村以下の地名固有のコードを付与しており、町・丁目レベルまでのコード管理が可能です。

- [国土交通省 大字・町丁目レベル位置参照情報](#)

国土交通省が整備するデータで、「大字」および「町丁目」（自治体によっては「町字」）を示す住所代表点と、その地点が示す位置座標を紐づけた情報です。全国の都市計画区域相当範囲を対象に街区単位の位置座標を整備した「街区レベル位置参照情報」を補完するものとして提供されています。

- [国勢調査町丁・字等別境界データ](#)

総務省統計局によって発行され、国勢調査で使用されている住所コードで、国土地理協会の住所コードと同じく11桁で構成されています。

住所の課題とソリューション

行政、ビジネスにおいて、住所を扱う開発者が直面する課題とソリューションについて基本的なところをまとめます。

課題	課題の内容	対応するソリューション	ソリューションの内容
表記ゆれ	人やシステムにより住所の分割方法、記載方法が違うことにより、同じ住所が違う住所として認識されてしまう	正規化	どのようなフォーマットの記載方法であっても、ひとつの定義された正しい記述方法に変換すること
位置の特定	住所が指し示す位置が分からない	ジオコーディング	住所を入力とし、座標を出力とするシステムで、住所を緯度経度の座標に変換すること。

住所の表記ゆれ

住所のコンピュータ処理、不動産情報の交換、物流における行き先の決定、会社内でのリスト管理などで問題になるのが住所の表記ゆれです。

住所の表記ゆれの原因

表記ゆれが起こってしまうのは、住所の書き方が人それぞれ、組織それぞれだからです。ユーザーがウェブサイトのフォームなどで入力する際に発生する場合もあれば、データを管理する組織内で表記方法に関するルールが統一されていないことで発生する場合もあります。

同一住所でも人によって表記は様々

都道府県の有無	
東京都文京区千石4丁目15-7	文京区千石4丁目15-7
大字、字の省略	
福島県会津若松市 大戸町大字高川甲 1324	福島県会津若松市 大戸町高川甲 1324
漢数字と英数字	
一丁目	1丁目
全角と半角	
1丁目	1丁目
漢字表記かハイフンか	
一丁目2番5号	1-2-5
建物名の有無	
1-1-1-101	1-1-1 山田マンション101

- 都道府県を書くか、省略するか
- 大字、字を書くか、省略するか
- 漢数字と英数字（一丁目と1丁目）
- 全角と半角（1丁目と1丁目）
- 丁目、番、号を漢字で書くか、ハイフンにするか（一丁目2番5号と1-2-5）
- 建物名を含めるか含めないか（1-1-1と1-1-1 山田マンション）
- 建物名をカタカナ表記にするか、英語表記にするか（SHINEとシャイン）

この他にも、次のような事情に起因する表記ゆれもあります。

- 京都の通り名は、同じ場所を示すのに複数の書き方が可能
- OCR（コンピュータによる手書き文字などの読み取り）で結果が意図と異なってしまう
 - 漢数字の「一」とハイフン
 - カタカナの「八」と漢数字の「八」、カタカナの「二」と漢数字の「二」

- OSや言語設定、IME（パソコン等で日本語入力をする際にアルファベットからかなや漢字への変換などをするプログラム）の設定によってハイフンなどが異なっている

人間が住所の表記を見るとき、1-1-1と一丁目1番1号は同じものであると、特に意識をすることなく分かりますが、機械にはそのように認識できず、別のデータとして登録されてしまいます。

Excel、顧客管理システム、データベースなどで同じ住所がデータが重複して登録されてしまったり、他社のデータやオープンデータと突き合わせたときに、同じ住所が同じであると分からず、出来上がったデータのクオリティーが上がらないこととなります。結果として、扱う人の手間が増えたり、同じ人に複数の手紙を送ってしまったり、統計処理をしたときの正確性が損なわれたり、といった問題が発生してしまいます。

住所の正規化

住所の正規化とは、住所の表記ゆれなどの問題を吸収し、ビジネス、特にコンピュータやソフトウェアで住所を扱う際の問題を解決するためのソリューションです。



いくつもの形に表記が揺れている住所を、1つのルールに沿って変更することにより、結果的に重複を廃して問題の解決に繋げることができます。正規化を正確に行うためには地名のリストが必要です。ルールが二重であったり（地番住所と住居表示住所）、場所によってルールが異なっていたりするため、正規表現（多様な文字列のパターンを記述するプログラミングの方法）だけでは正規化ができないことが分かっています。

座標の取得 = ジオコーディング

ジオコーディングとは、住所をもとに緯度・経度などの地理座標に変換する処理のことで、「アドレスマッチング」とも呼ばれます。

地図アプリやウェブ地図サービスでは、住所や地名、施設名、駅名などで検索すると該当する施設や場所が地図上に表示されます。このとき住所や施設名などの文字列をもとに、該当する地図上の特定の位置を示すためにジオコーディングが必要となります。ジオコーディングで示されるのは代表点が1点のみの場合もあれば、市区町村単位など、特定の領域が面で示される場合もあります。

ジオコーディングを行うには、住所や施設名と地理座標を紐付けたデータベースが必要となります。住所から地理座標に変換する場合に、住所データを都道府県や市区町村など行政区分に依りてレベルごとに分解した上で、データベースを照合して座標を導き出します。

なお、逆に緯度・経度などの座標から住所や地名などの情報に変換することは「逆ジオコーディング（リバースジオコーディング）」といいます。

現在、国内で提供されている主なジオコーディングのサービスは以下の通りです。

- [CSV アドレスマッチングサービス](#)

東京大学空間情報科学研究センター（CSIS）が提供する無料のサービスで、商用利用や自治体などの日常業務にも利用可能です。ただしメンテナンスなどのため無断で停止する可能性があるため、防災システムなどサービス停止が人命に関わる目的で利用することは推奨されていません。

- [Community Geocoder](#)

Geoloniaが開発したオープンソースのジオコーディングAPIで、国土交通省の位置参照情報の「街区レベル位置参照情報」および「大字・町丁目レベル位置参照情報」を利用しています。Geoloniaの住所正規化ライブラリを利用して、緯度経度を取得したい住所を正規化（都道府県、市区町村、町丁目などに分解する）した上で緯度経度を取得します。

- [Google Maps Platform](#)

Googleが提供するサービスで、Geocoding APIまたはMaps JavaScript APIのGeocodingサービスへのリクエストが発生した場合に課金されます。

- [Yahoo!ジオコードAPI](#)

住所を指定することで地理座標を出力するAPIです。施設名などでの検索に対応した「Yahoo!ローカルサーチAPI」や、緯度・経度から住所を検索できる「Yahoo!リバースジオコードAPI」も提供されています。

- [座標付与サービス](#)

ゼンリンが提供するサービスで、顧客のデータに含まれる住所や名称、電話番号などをキーとして、住宅地図データや電話帳データなどをもとに対応する地理座標を付与します。住宅地図データを使用しているため、建物ごとに紐付いた座標値を取得することが可能です。

- [住所ジオコーディングサービス](#)

ジオテクノロジーズが提供するWeb APIで、住所を1件ずつ処理できるだけでなく、一括処理することも可能です。住所表記ゆれの解消や、古い住所表記の更新などの処理を行うため、管理が重複している物件を見つけやすいのが特徴です。Excelのマクロ形式でも提供しています。

その他の住所課題

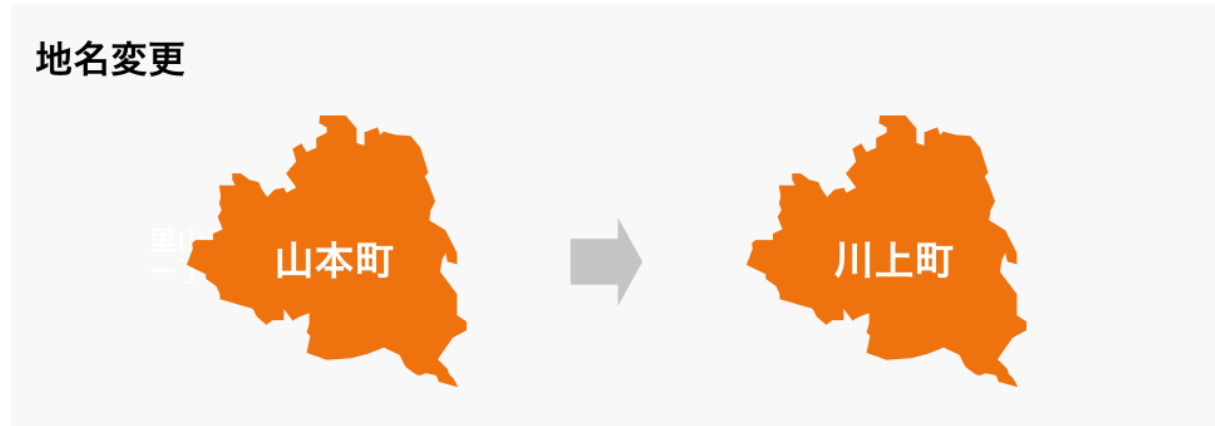
地名変更

街区整備などで地名が変わることで、名寄せをする際などに、本当は同じ建物であるのに、見かけ上は異なる文字列になってしまう問題です。地名の変更は、毎月全国のどこかで数件は発生しています。その情報は各自治体がバラバラの方法で公開しており、国などによって

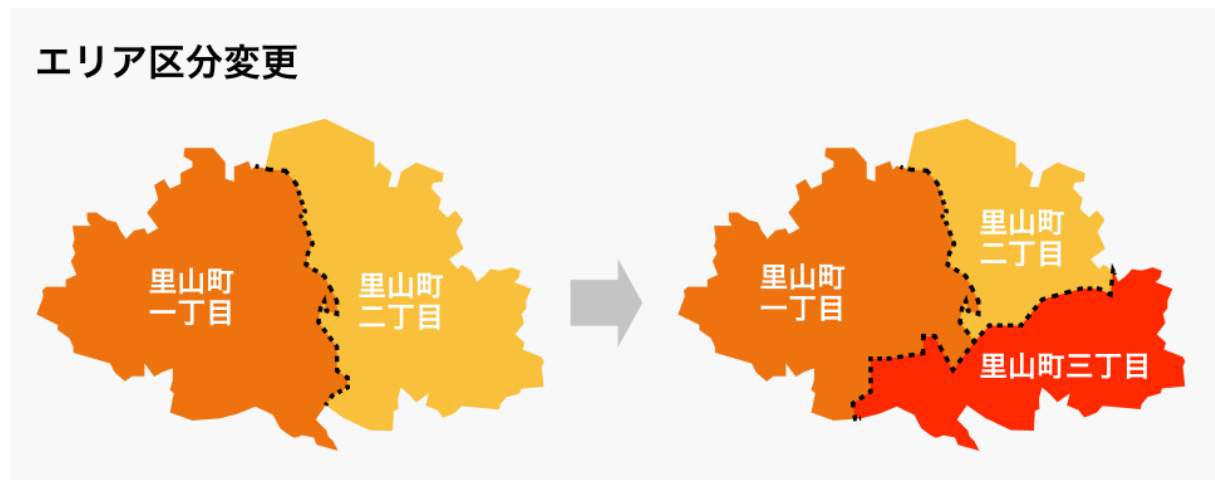
その詳細が統一的に公開されることがなく、変更された事実や変更の詳細を把握することはとても困難です。

また、地名の変更と同時にエリアの区分けも変更されることが多くあります。

シンプルな地名変更では、「山本町」が「川上町」に変わったり、「ABC町」が「ABC市」に変わったりすることがあります。



一方、「里山町一丁目、二丁目」が存在していた場所に「里山町三丁目」が加わるケースでは、一丁目と二丁目の面積が減って、三丁目が新たに誕生します。



その結果、新しく見える建物の住所が、地名が変更された既存の建物なのか、新しい建物が立ったのか、あるいは住所の一覧データにその地名が抜けていたのか、判別できなくなります。

建築途中の建物に住所がない問題

住居表示住所は建物につけられる住所であるため、建築の計画中、工事中のマンションやビルには住居表示住所がありません。完成してから割り当てられます。一方、建物のオーナーは、建築の完了前に購入者や借り主の募集を開始します。このとき、隣の建物の住所がつけられたり、「〇〇町1-2-(3)」のような仮の住所が流通することがあります。建物が完成すると、同じ建物が正式な住所（仮の住所とは全く違うもの）となって再度流通します。

結果として、同じ建物を別の物件と認識してしまう問題が起こります。

建物名の表記ゆれと建物名変更

マンションなどの集合住宅や、オフィスビル、雑居ビル等のビルの場合、住所の末尾に建物名が付きますが、この建物名は住所そのものよりも複雑に表記ゆれします。たとえば、以下の建物名がすべて同じ建物を指していることがあります。

- AB-Garden Mansion THE SOUTH
- A B ガーデンマンション S棟
- AB ガーデン マンション 南棟



同一建物

この例では、ABの全角・半角、ハイフンの有無、スペースの有無、ガーデンがアルファベットかカタカナか、棟の名前の英語、アルファベット、漢字といういくつかのパターンで表記が揺れます。

また、特に賃貸物件では、印象を新しくするなどの目的で、「丸山コーポ」を「Residence 丸山」に変更し、看板を付け替えるということも行われます。

ふりがなと英語表記

現在の法律には、住所にある地名をどう読むのか、また、アルファベットで表記する際にどう書くのか、決まりはありません。

政府・行政による関連の取り組み

- [国土数値情報／位置参照情報](#)

国土数値情報および位置参照情報は、国土交通省が無償で提供しているデータです。国土数値情報は、地形や土地利用、公共施設、地価などの基礎的な情報をGISデータとして整備したものです。データ形式はGML形式やXML形式で、データごとに商用化・非商用などの利用条件が異なります。

位置参照情報は、全国の都市計画区域相当範囲を対象としたデータで、街区レベル（「○○町△丁目□番」）、大字・町丁目レベル（大字・町丁目の住所代表点）の位置座標を整備しています。出典を明記することにより再配布も可能です。

- [アドレス・ベース・レジストリ](#)

ベース・レジストリとは、国や地方自治体といった公的機関が管理しており、さまざまな場面において必要なときに国民が参照できるデータベースのことを意味します。このうち住所や地番などのデータベースがアドレス・ベース・レジストリで、デジタル庁より提供されています。

日本の住所表記や地番表記は表記ゆれが大きく、官公庁や民間企業など住所・地番を扱う組織によって管理する場合の表記ルールに違いがあるため、データを標準化した上で一元的に管理するために整備が進められています。これにより、行政コストの削減が図れるとともに、民間も含めた幅広い利活用が期待できます。

- [PLATEAU（プラトー）](#)

現実世界（フィジカル空間）をサイバー（仮想）空間上に双子（ツイン）のように再現し、そこに多種多様なデータを組み合わせることで解析・シミュレーションを行う「都市のデジタルツイン」。この取り組みを推進するために国土交通省が推進しているプロジェクトが「PLATEAU（プラトー）」です。

同プロジェクトでは都市のデジタルツインの基礎となる全国各地の3D都市モデルを整備してオープンデータとして公開するとともに、そのユースケースを創出しており、2021年度にかけて56都市以上の3Dモデルを整備しました。

PLATEAUが提供する3D都市モデルは、都道府県や市区町村が整備する「都市計画基本図」と呼ばれる2Dの地図をもとに航空測量による高さ情報や都市計画基礎調査などの情報を組み合わせられており、その3D都市モデルは国際標準のオープンフォーマット「CityGML」で記述されたデータとなっています。

- [法務省登記所備付地図データ](#)

不動産の登記申請業務を行う法務省の登記所において使用されている「登記所備付地図データ」が2023年1月23日から、法務省より加工可能な形で無償公開されています。このデータ

は利用規約に抵触しない限り商用利用も含めて誰でも自由に利用可能で、再配布も可能です。

データは地図XML形式を採用しており、その中には地図名や市町村名、市町村コード、座標系などの属性情報が含まれています。なお、登記所備付地図データには、地籍調査などに基づいた公共座標系の地図と、旧土地台帳附属地図から引き継がれた公図に基づく「任意座標系」の図面があり、任意座標系の地図は、そのままでは地図の上に重ねることはできません。

地図XMLファイルはそのままではGISソフトウェアなどで読み込むことができないため、ファイル変換が必要となります。ファイル変換ツールとしては、デジタル庁が公開したGeoJSON形式への変換ツールをはじめ、さまざまなツールが公開されています。

- [国土交通省 不動産ID](#)

国土交通省が整備中の不動産コードで、不動産情報の収集・名寄せを容易にすることで事業者の負担を軽減させるとともに、不動産関連情報の連携・蓄積・活用や消費者への的確な情報発信などを促進させることを目的としています。

2022年3月にルールガイドラインが発表されており、このルールに従って不動産関連情報の保有者などが、自ら保有する不動産関連情報に不動産IDを紐付けて連携・蓄積・活用することを想定しており、国が一元的なデータベースを作成してIDを発番したり、不動産情報を収集・蓄積して提供したりすることは意図していません。

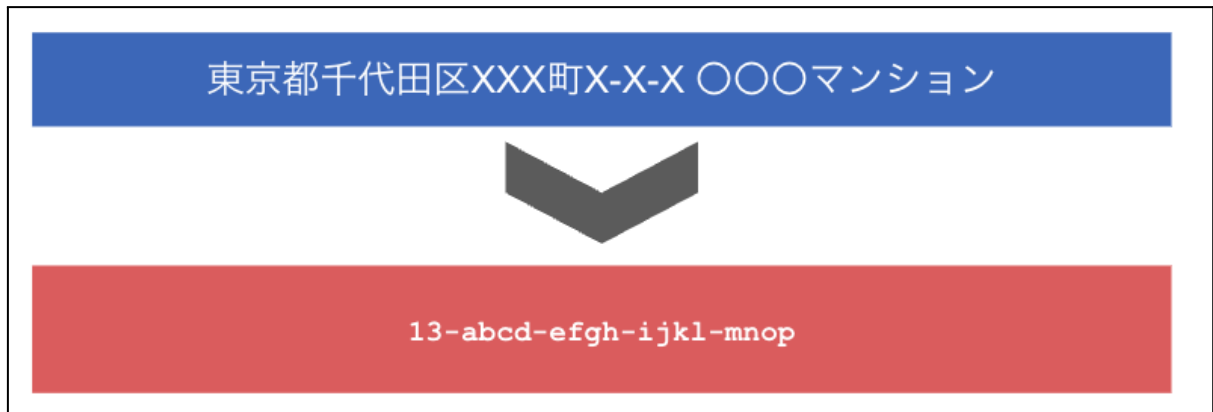
不動産IDは登記簿謄本に記載されている13桁の不動産番号を基本として、ビル内のテナントやマンションの部屋番号など、13桁だけでは特定できない物件には、さらに4桁のコードが付与されます。

不動産オープンID API

不動産オープンIDは、不動産テック協会と株式会社 Geolonia が開発・運用するIDのAPIです。

基本的な仕組み

住所の文字列を入力として受け付け、一意の不動産オープンIDを返却します。



「住所をIDに変換するAPI」が基本のコンセプト

以下は、API に対してリクエストを送っている例です。

```
curl -v -G \  
  -H "x-access-token: <アクセストークン>" \  
  --data-urlencode "q=東京都千代田区永田町1丁目7-1" \  
  --data-urlencode "api-key=<APIキー>" \  
  "https://api.propid.jp/v1/"
```

API として提供することで、既存のシステムとの連携が可能になります。リクエストの送り方、エンドポイント、レスポンス、エラーコードなど[不動産オープンIDの詳細な仕様は、ドキュメンテーションで確認](#)することができます。

また、住所変更への対応方針、信頼性、履歴情報等について、「[よくある質問](#)」ページも用意していますので参考にしてください。

処理の方法

入力された住所の文字列から、不動産オープンIDを出力するまでの処理は、おおむね以下のようになっています。



処理内容の概要

- 住所の入力を受け付ける
- 正規化を行う
- 正規化した結果を用いてジオコーディングを行い、座標を得る
- ウェブ地図のタイル番号に変換する
 - ※ タイル番号については、[国土地理院による地図タイルの概念説明のページ](#)を参照してください
- タイル番号、正規化済みの住所文字列、既存のIDデータをもとに不動産オープンIDを生成して返却

不動産オープンIDのソースコード

不動産テック協会では、できる限りオープンに仕様やデータを取り扱うこととしています。以下は、Geolonia が公開しているデータとソースコードです。

不動産オープンID API

不動産オープンID API の運用に利用しているソースコードです。誰でも同じように不動産オープンIDの発行プロセスを再現することが可能です。

<https://github.com/geolonia/prop-id-api>

Geolonia 住所データ

不動産オープンID 発行のプロセスには、大量の住所文字列が必要です。国土交通省位置参照情報ダウンロードサービスで公開されている住所をもとに郵便番号データや、自治体による更新情報等のオープンなデータを組み合わせて文字列を整備しています。

また、このデータを用いたAPIも公開しています。

たとえば、<https://geolonia.github.io/japanese-addresses/api/ja/<都道府県名>/<市区町村名>.json> というURLにアクセスすることで、該当する市区町村にある町丁目とその代表点座標のレスポンスを取得することができます。（[長野県長野市のレスポンス](#)）

<https://github.com/geolonia/japanese-addresses>

住所正規化エンジン

住所文字列を入力として受け付け、住所を正規化するロジックを公開しています。正規化の内容や、利用方法については、リポジトリで公開しており、エンジニアに渡すことですぐに利用が可能です。不動産オープンIDにおける正規化では、上記のGeolonia 住所データの他、アドレスベースレジストリに公開されたデータも利用しています。

<https://github.com/geolonia/normalize-japanese-addresses>

他のエンジンにない特徴として、ユーザーからのフィードバックが蓄積されており、正規化のエラーを吸収するロジックが、2年にわたって追加され続けている点があります。これまでに追加してきた修正は、[リポジトリのIssueページ](#)で確認できます。

コミュニティジオコーダー

前述の正規化エンジンによる正規化、住所データに含まれる座標情報を参照して、町丁目レベルの座標を返却するジオコーダーです。

<https://github.com/geolonia/community-geocoder>

今後の展開について

これまでの不動産オープンIDの開発の歩みは、データが揃うことで精度が上がることを確認する作業でもありました。国交省の27万件の住所文字列だけを利用して必要最小限の機能を持ったソリューションとして開発をスタートしたのち、郵便局のデータ、ユーザーからのフィードバックを取り込み、一定の精度を確保してサービスを開始しました。

その後、2022年に公開されたアドレスベースレジストリのデータ2000万件強を追加することで、住居表示住所に関してはさらに精度を上げることができました。

現在は、2023年に公開された登記所備付地図、いわゆる公図XMLデータを合わせると2億件を超えるデータ量となり、さらなる精度向上のために統合を検討しています。

このようにして、一定のユースケースにはサービスとして価値を提供できる状態になっています。

一方で、たとえば住居表示と地番を相互に参照したり、一つの住所に複数の建物がある問題を解決したりといったことが十分にできる状態には至っておりません。これらの問題の解決には、データの量と質（情報の内容）の向上が必要です。公図XMLの取り込みの後、国、自治体による取り組みが進む際、国が用意する不動産IDとの連携も含め、開発者に使いやすく、不動産情報の流通を高速回転させるためのソリューションとして、発展させていく予定です。

本件に関するお問い合わせ

不動産テック協会

info@retechjapan.org