

はじめに  
この研究でやること  
やったこと  
おわりに

# WebスクレイピングしたデータをQGISに適したフォーマットに変換するツールの開発

平松 楓也

富山県立大学 情報基盤工学講座

November 19, 2020

## 背景

昨今、人々の生活に多大な影響を及ぼしているコロナウイルスの伝播する様子を視覚的に表現するツールとして QGIS(Quantum Geographic Information System) が用いられている。一般的に QGIS は国や地方が公開しているビックデータを使い地図上に視覚的に表示することに使われている。しかし、QGIS に適した形式のデータを探すことや変換するには時間や手間がかかる問題がある。

## 目的

ある事柄に対した情報を Web から自動で収集し QGIS に対応したデータに変換するツールはまだ開発されていない。そこで、本研究では、Web スクレイピングしてきたデータを QGIS に適した形式のデータに変換するツールの開発を行う。

はじめに

この研究でやること

やったこと

おわりに

はじめに  
この研究でやること  
やったこと  
おわりに

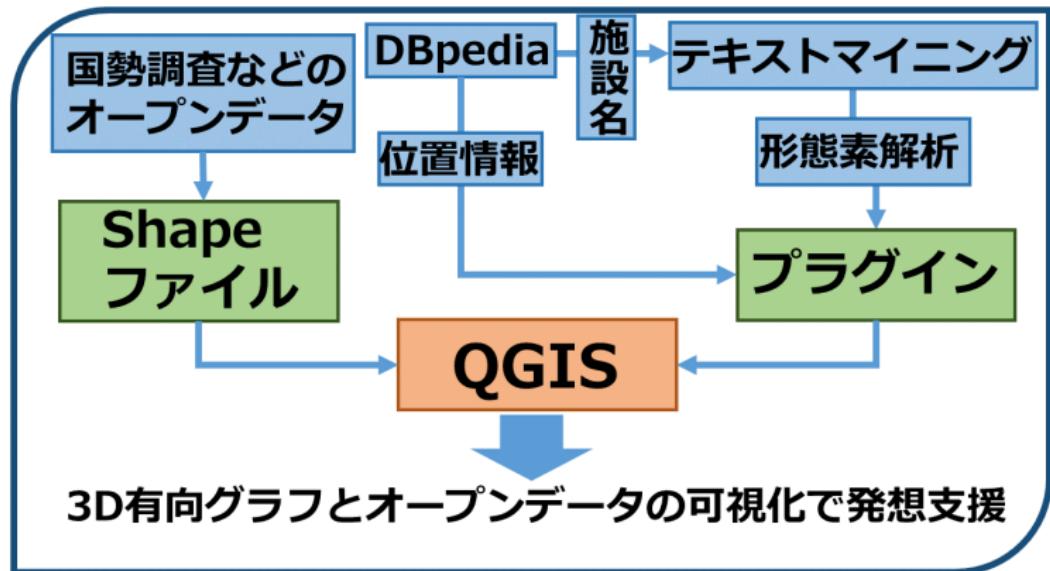


図 1: 全体イメージ

# 目標とできている事と課題

4/9

QGIS のクリックした地名と高齢者などのキーワードで web 上での情報で 3 D グラフを作成と、Twitter で地名でスクレイピングして 3 D グラフを作成をして過去の蓄積されたデータとトレンドデータで比較できる発想支援

## できている事

- ① 地図をクリックしたら住所を取得すること
- ② Twitter でキーワードで一週間分の情報をスクレイピングすること

## 課題

- ① プログラムがバラバラなのでくっつける必要がある
- ② 地名ともう一つのキーワードをどう決めるか検討中

Twitter のスクレイピングは原則禁止されているので、Twitter のデベロッパー登録をし、TwitterAPI を取得してスクレイピングを行う  
この TwitterAPI は課金することで取得できるツイートの期間を増やすことができるが、今回は無料で取得できる一週間分を扱う

今回は任意のキーワードで検索しスクレイピングし、その結果を  
JSON ファイルで出力する

取得する情報は、ツイートされたテキストとツイートした人の ID,  
名前である。

はじめに

この研究でやること

やったこと

おわりに

## twitter スクレイピングした出力

1 「text': {"e": "A.太閤山ランド\n昭和58年(1983)開園\nにっぽん新世界博覧会\n平成4年(1992)年1回\nサンキスボ富山'92\n\n花の名所:アジサイ\nヘメカリクス", "i": "「はるちゃん( @maru19978684 )」\n院内 missan #silvia #silvia14 #14 #14kouki #nancan #canonphotography #canon\_photos\n#toyama, https://t.co/XUC0SDNEz", "l": "【寺院紹介】 太閤山 常寂寺 - Tsuchiyama-Jousenji temple ->https://t.co/8p5KxW7wDE#愛知県の「日本」を感じさせる美富をご紹介するリース!\"#Japan\"inAiichi\"#First上屏浦-https://t.co/MviLkwd8U", "s": "「太閤山ランド」を感じさせる富山は太閤山ランドでした!「みなさま あやすみなさい」 https://t.co/veUu1q19", "t": "「太閤山ランド...は違う」, "u": "「みひとつばこ」, "v": "@makaron\_neu 太閤山ランド?", "w": "「太閤山ランド https://t.co/MoE6VeYoei」, "x": "「豊田店 ワッショイ音無簫子です。」\n院内スタッフから情報されていた開園日長が「ハイパー富山」を真似ました。院内スタッフの音「音」を聞いて、みんなもそれに沿って乗って残る長でした。」\n院内太閤山店に行かれましたも、頃頃でてくださいね! https://t.co/4she2leId", "y": "「#kurokiro6546 太閱山も良いですね♪\nめちゃくちゃ迷いました。コロナもあまですし。お泊りはちょっと難しいですか?」, "z": "「#shuni\_humi\_aya 肺炎は任せたでない! 『#n』がありがとうございます! 1人少人数で太閱山でももらしてもよし...春が来ます!」, "aa": "「#東京 岩国公園太閱山ランドに富山を真似ました: https://t.co/pnMdC15C」, "ab": "「#紅葉 #公園 #太閱山ランド https://t.co/laew9q58B」, "ac": "「はなのはいくんが太閱山ランドに来てるすごいけどなぜ」, "ad": "「#ソララント太閱山店 【#東京朝9時開店】! 」\n当店では桜満開のヨコナリス風を実感しております。」\nマスク着用、消毒の徹底などへのご協力を宜しくお願いします! https://t.co/DkWpx8Gdx9y", "ae": "「#うちはん勝ったアルム。いつたってたワットラ。2012年太閱山ランドのフェス、チャーチのラップ終了後の音がさんざん「#ワットラ」と最高でした!」\nたって伝えたからと照ねながら「うわったってワットラ」と言っただって、僕はおっしゃる瀧らしさも。https://t.co/HyCj30v0n», "af": "「#北陸街道 #毛越寺 #新潟市。耐水木石造(うち木)。明治35年(1902年)。清酒の醸造を始め、「太閱山」や「杉正宮」の銘柄で販売した。水は金胎堂の湧水を使ったと伝わる。 https://t.co/4z2yJ5q6tr」, "ag": "「太閱山ランドを行ってきたたけん」\nhttps://t.co/ASqZetq4zGZ», "ah": "「太閱山ランド https://t.co/MoE6VeYoei」, "ai": "「明日太閱山ランドいこ」, "aj": "「「ブル」といえば「太閱山ランド」」, "ak": "「太閱山ランド」といわば「太閱山ランド」」, "al": "「太閱山ランド https://t.co/MoE6VeYoei」, "am": "「やっぱり太閱山の見景、高麗、新潟の温泉までのくと華やかだら、のり渡りで走るんだる。アラビアさんもそうだね。」\n#新潟の海王や太閱山の温泉もそんなんだよ。なまなま温泉までくと華やかだら、のり渡りで走るんだる。」, "an": "「#新潟市太閱山、元気な空き家用相談。 https://t.co/AMyHjRDSy」, "ao": "「#時水市」\n小林、大門、三ツ・太閱山、南大関山、貴島金属高貴直毅り#nryu貴金属、金・ブチナ、シルバー・ダイヤ 貴金属専門店「e-shops」イニシエーション https://t.co/PSuQw0x6Dw」, "ap": "「温泉に入るか~(太閱山天然温泉 太閱の湯 在#新潟市 富山県) https://t.co/7iuMoS2X0D」, "aq": "「#時水、太閱山の温泉を楽しめぬるでさうした前海海岸\ne行けてたかだり~」\nhttps://t.co/hJ31av1yG5」, "ar": "「#なんか太閱山ランドとかいうのさ」, "as": "「#13292121235465336」, "at": "「#13292121235465336」, "au": "「#1328681313458932577」, "av": "「#132868131458932577」, "aw": "「#132868131458932577」, "ax": "「#132868131458932577」, "ay": "「#132868131458932577」, "az": "「#132868131458932577」, "ba": "「#132868131458932577」, "bc": "「#132868131458932577」, "bd": "「#132868131458932577」, "be": "「#132868131458932577」, "bf": "「#132868131458932577」, "bg": "「#132868131458932577」, "bh": "「#132868131458932577」, "bi": "「#132868131458932577」, "bj": "「#132868131458932577」, "bk": "「#132868131458932577」, "bl": "「#132868131458932577」, "bm": "「#132868131458932577」, "bn": "「#132868131458932577」, "bo": "「#132868131458932577」, "bp": "「#132868131458932577」, "bq": "「#132868131458932577」, "br": "「#132868131458932577」, "bs": "「#132868131458932577」, "bt": "「#132868131458932577」, "bu": "「#132868131458932577」, "bv": "「#132868131458932577」, "bw": "「#132868131458932577」, "bx": "「#132868131458932577」, "by": "「#132868131458932577」, "bz": "「#132868131458932577」, "ca": "「#132868131458932577」, "cb": "「#132868131458932577」, "cc": "「#132868131458932577」, "cd": "「#132868131458932577」, "ce": "「#132868131458932577」, "cf": "「#132868131458932577」, "cg": "「#132868131458932577」, "ch": "「#132868131458932577」, "ci": "「#132868131458932577」, "cj": "「#132868131458932577」, "ck": "「#132868131458932577」, "cl": "「#132868131458932577」, "cm": "「#132868131458932577」, "cn": "「#132868131458932577」, "co": "「#132868131458932577」, "cp": "「#132868131458932577」, "cq": "「#132868131458932577」, "cr": "「#132868131458932577」, "cs": "「#132868131458932577」, "ct": "「#132868131458932577」, "cu": "「#132868131458932577」, "cv": "「#132868131458932577」, "cw": "「#132868131458932577」, "cx": "「#132868131458932577」, "cy": "「#132868131458932577」, "cz": "「#132868131458932577」, "da": "「#132868131458932577」, "db": "「#132868131458932577」, "dc": "「#132868131458932577」, "de": "「#132868131458932577」, "df": "「#132868131458932577」, "dg": "「#132868131458932577」, "dh": "「#132868131458932577」, "di": "「#132868131458932577」, "dj": "「#132868131458932577」, "dk": "「#132868131458932577」, "dl": "「#132868131458932577」, "dm": "「#132868131458932577」, "dn": "「#132868131458932577」, "do": "「#132868131458932577」, "dp": "「#132868131458932577」, "dq": "「#132868131458932577」, "dr": "「#132868131458932577」, "ds": "「#132868131458932577」, "dt": "「#132868131458932577」, "du": "「#132868131458932577」, "dv": "「#132868131458932577」, "dw": "「#132868131458932577」, "dx": "「#132868131458932577」, "dy": "「#132868131458932577」, "dz": "「#132868131458932577」, "ea": "「#132868131458932577」, "eb": "「#132868131458932577」, "ec": "「#132868131458932577」, "ed": "「#132868131458932577」, "ef": "「#132868131458932577」, "eg": "「#132868131458932577」, "eh": "「#132868131458932577」, "ei": "「#132868131458932577」, "ej": "「#132868131458932577」, "ek": "「#132868131458932577」, "el": "「#132868131458932577」, "em": "「#132868131458932577」, "en": "「#132868131458932577」, "eo": "「#132868131458932577」, "ep": "「#132868131458932577」, "eq": "「#132868131458932577」, "er": "「#132868131458932577」, "es": "「#132868131458932577」, "et": "「#132868131458932577」, "eu": "「#132868131458932577」, "ev": "「#132868131458932577」, "ew": "「#132868131458932577」, "ex": "「#132868131458932577」, "ey": "「#132868131458932577」, "ez": "「#132868131458932577」, "fa": "「#132868131458932577」, "fb": "「#132868131458932577」, "fc": "「#132868131458932577」, "fd": "「#132868131458932577」, "fe": "「#132868131458932577」, "fg": "「#132868131458932577」, "fh": "「#132868131458932577」, "fi": "「#132868131458932577」, "fj": "「#132868131458932577」, "fk": "「#132868131458932577」, "fl": "「#132868131458932577」, "fm": "「#132868131458932577」, "fn": "「#132868131458932577」, "fo": "「#132868131458932577」, "fp": "「#132868131458932577」, "fq": "「#132868131458932577」, "fr": "「#132868131458932577」, "fs": "「#132868131458932577」, "ft": "「#132868131458932577」, "fu": "「#132868131458932577」, "fv": "「#132868131458932577」, "fw": "「#132868131458932577」, "fx": "「#132868131458932577」, "fy": "「#132868131458932577」, "fz": "「#132868131458932577」, "ga": "「#132868131458932577」, "gb": "「#132868131458932577」, "gc": "「#132868131458932577」, "gd": "「#132868131458932577」, "ge": "「#132868131458932577」, "gf": "「#132868131458932577」, "gh": "「#132868131458932577」, "gi": "「#132868131458932577」, "gj": "「#132868131458932577」, "gk": "「#132868131458932577」, "gl": "「#132868131458932577」, "gm": "「#132868131458932577」, "gn": "「#132868131458932577」, "go": "「#132868131458932577」, "gp": "「#132868131458932577」, "gq": "「#132868131458932577」, "gr": "「#132868131458932577」, "gs": "「#132868131458932577」, "gt": "「#132868131458932577」, "gu": "「#132868131458932577」, "gv": "「#132868131458932577」, "gw": "「#132868131458932577」, "gx": "「#132868131458932577」, "gy": "「#132868131458932577」, "gz": "「#132868131458932577」, "ha": "「#132868131458932577」, "hb": "「#132868131458932577」, "hc": "「#132868131458932577」, "hd": "「#132868131458932577」, "he": "「#132868131458932577」, "hf": "「#132868131458932577」, "hg": "「#132868131458932577」, "hi": "「#132868131458932577」, "hj": "「#132868131458932577」, "hk": "「#132868131458932577」, "hl": "「#132868131458932577」, "hm": "「#132868131458932577」, "hn": "「#132868131458932577」, "ho": "「#132868131458932577」, "hp": "「#132868131458932577」, "hq": "「#132868131458932577」, "hr": "「#132868131458932577」, "hs": "「#132868131458932577」, "ht": "「#132868131458932577」, "hu": "「#132868131458932577」, "hv": "「#132868131458932577」, "hw": "「#132868131458932577」, "hx": "「#132868131458932577」, "hy": "「#132868131458932577」, "hz": "「#132868131458932577」, "ia": "「#132868131458932577」, "ib": "「#132868131458932577」, "ic": "「#132868131458932577」, "id": "「#132868131458932577」, "ie": "「#132868131458932577」, "if": "「#132868131458932577」, "ig": "「#132868131458932577」, "ih": "「#132868131458932577」, "ii": "「#132868131458932577」, "ij": "「#132868131458932577」, "ik": "「#132868131458932577」, "il": "「#132868131458932577」, "im": "「#132868131458932577」, "in": "「#132868131458932577」, "io": "「#132868131458932577」, "ip": "「#132868131458932577」, "iq": "「#132868131458932577」, "ir": "「#132868131458932577」, "is": "「#132868131458932577」, "it": "「#132868131458932577」, "iu": "「#132868131458932577」, "iv": "「#132868131458932577」, "iw": "「#132868131458932577」, "ix": "「#132868131458932577」, "iy": "「#132868131458932577」, "iz": "「#132868131458932577」, "ja": "「#132868131458932577」, "jb": "「#132868131458932577」, "jc": "「#132868131458932577」, "jd": "「#132868131458932577」, "je": "「#132868131458932577」, "jf": "「#132868131458932577」, "jh": "「#132868131458932577」, "ji": "「#132868131458932577」, "jj": "「#132868131458932577」, "jk": "「#132868131458932577」, "jl": "「#132868131458932577」, "jm": "「#132868131458932577」, "jn": "「#132868131458932577」, "jo": "「#132868131458932577」, "jp": "「#132868131458932577」, "jq": "「#132868131458932577」, "jr": "「#132868131458932577」, "js": "「#132868131458932577」, "jt": "「#132868131458932577」, "ju": "「#132868131458932577」, "jv": "「#132868131458932577」, "jw": "「#132868131458932577」, "jx": "「#132868131458932577」, "jy": "「#132868131458932577」, "jz": "「#132868131458932577」, "ka": "「#132868131458932577」, "kb": "「#132868131458932577」, "kc": "「#132868131458932577」, "kd": "「#132868131458932577」, "ke": "「#132868131458932577」, "kf": "「#132868131458932577」, "kg": "「#132868131458932577」, "kh": "「#132868131458932577」, "ki": "「#132868131458932577」, "kj": "「#132868131458932577」, "kk": "「#132868131458932577」, "kl": "「#132868131458932577」, "km": "「#132868131458932577」, "kn": "「#132868131458932577」, "ko": "「#132868131458932577」, "kp": "「#132868131458932577」, "kq": "「#132868131458932577」, "kr": "「#132868131458932577」, "ks": "「#132868131458932577」, "kt": "「#132868131458932577」, "ku": "「#132868131458932577」, "kv": "「#132868131458932577」, "kw": "「#132868131458932577」, "kx": "「#132868131458932577」, "ky": "「#132868131458932577」, "kz": "「#132868131458932577」, "la": "「#132868131458932577」, "lb": "「#132868131458932577」, "lc": "「#132868131458932577」, "ld": "「#132868131458932577」, "le": "「#132868131458932577」, "lf": "「#132868131458932577」, "lg": "「#132868131458932577」, "lh": "「#132868131458932577」, "li": "「#132868131458932577」, "lj": "「#132868131458932577」, "lk": "「#132868131458932577」, "ll": "「#132868131458932577」, "lm": "「#132868131458932577」, "ln": "「#132868131458932577」, "lo": "「#132868131458932577」, "lp": "「#132868131458932577」, "lq": "「#132868131458932577」, "lr": "「#132868131458932577」, "ls": "「#132868131458932577」, "lt": "「#132868131458932577」, "lu": "「#132868131458932577」, "lv": "「#132868131458932577」, "lw": "「#132868131458932577」, "lx": "「#132868131458932577」, "ly": "「#132868131458932577」, "lz": "「#132868131458932577」, "ra": "「#132868131458932577」, "rb": "「#132868131458932577」, "rc": "「#132868131458932577」, "rd": "「#132868131458932577」, "re": "「#132868131458932577」, "rf": "「#132868131458932577」, "rg": "「#132868131458932577」, "rh": "「#132868131458932577」, "ri": "「#132868131458932577」, "rj": "「#132868131458932577」, "rk": "「#132868131458932577」, "rl": "「#132868131458932577」, "rm": "「#132868131458932577」, "rn": "「#132868131458932577」, "ro": "「#132868131458932577」, "rp": "「#132868131458932577」, "rq": "「#132868131458932577」, "rr": "「#132868131458932577」, "rs": "「#132868131458932577」, "rt": "「#132868131458932577」, "ru": "「#132868131458932577」, "rv": "「#132868131458932577」, "rw": "「#132868131458932577」, "rx": "「#132868131458932577」, "ry": "「#132868131458932577」, "rz": "「#132868131458932577」, "ta": "「#132868131458932577」, "tb": "「#132868131458932577」, "tc": "「#132868131458932577」, "td": "「#132868131458932577」, "te": "「#132868131458932577」, "tf": "「#132868131458932577」, "tg": "「#132868131458932577」, "th": "「#132868131458932577」, "ti": "「#132868131458932577」, "tr": "「#132868131458932577」, "ts": "「#132868131458932577」, "tu": "「#132868131458932577」, "tv": "「#132868131458932577」, "tw": "「#132868131458932577」, "tx": "「#132868131458932577」, "ty": "「#132868131458932577」, "tz": "「#132868131458932577」, "ua": "「#132868131458932577」, "ub": "「#132868131458932577」, "uc": "「#132868131458932577」, "ud": "「#132868131458932577」, "ue": "「#132868131458932577」, "uf": "「#132868131458932577」, "ug": "「#132868131458932577」, "uh": "「#132868131458932577」, "ui": "「#132868131458932577」, "uj": "「#132868131458932577」, "uk": "「#132868131458932577」, "ul": "「#132868131458932577」, "um": "「#132868131458932577」, "un": "「#132868131458932577」, "uo": "「#132868131458932577」, "up": "「#132868131458932577」, "ur": "「#132868131458932577」, "us": "「#132868131458932577」, "ut": "「#132868131458932577」, "uv": "「#132868131458932577」, "uw": "「#132868131458932577」, "ux": "「#132868131458932577」, "uy": "「#132868131458932577」, "uz": "「#132868131458932577」, "va": "「#132868131458932577」, "vb": "「#132868131458932577」, "vc": "「#132868131458932577」, "vd": "「#132868131458932577」, "ve": "「#132868131458932577」, "vf": "「#132868131458932577」, "vg": "「#132868131458932577」, "vh": "「#132868131458932577」, "vi": "「#132868131458932577」, "vj": "「#132868131458932577」, "vk": "「#132868131458932577」, "vl": "「#132868131458932577」, "vm": "「#132868131458932577」, "vn": "「#132868131458932577」, "vo": "「#132868131458932577」, "vp": "「#132868131458932577」, "vr": "「#132868131458932577」, "vs": "「#132868131458932577」, "vt": "「#132868131458932577」, "vu": "「#132868131458932577」, "vv": "「#132868131458932577」, "vw": "「#132868131458932577」, "vx": "「#132868131458932577」, "vy": "「#132868131458932577」, "vz": "「#132868131458932577」, "wa": "「#132868131458932577」, "wb": "「#132868131458932577」, "wc": "「#132868131458932577」, "wd": "「#132868131458932577」, "we": "「#132868131458932577」, "wf": "「#132868131458932577」, "wg": "「#132868131458932577」, "wh": "「#132868131458932577」, "wi": "「#132868131458932577」, "wj": "「#132868131458932577」, "wk": "「#132868131458932577」, "wl": "「#132868131458932577」, "wm": "「#132868131458932577」, "wn": "「#132868131458932577」, "wo": "「#132868131458932577」, "wp": "「#132868131458932577」, "wr": "「#132868131458932577」, "ws": "「#132868131458932577」, "wt": "「#132868131458932577」, "wu": "「#132868131458932577」, "wv": "「#132868131458932577」, "wx": "「#132868131458932577」, "wy": "「#132868131458932577」, "wz": "「#132868131458932577」, "xa": "「#132868131458932577」, "xb": "「#132868131458932577」, "xc": "「#132868131458932577」, "xd": "「#132868131458932577」, "xe": "「#132868131458932577」, "xf": "「#132868131458932577」, "xg": "「#132868131458932577」, "xh": "「#132868131458932577」, "xi": "「#132868131458932577」, "xj": "「#132868131458932577」, "xk": "「#132868131458932577」, "xl": "「#132868131458932577」, "xm": "「#132868131458932577」, "xn": "「#132868131458932577」, "xo": "「#132868131458932577」, "

図 2: スクレイピングした JSON ファイル

# クリックした箇所の住所を出す

7/9

はじめに  
この研究でやること  
やったこと  
おわりに



図 3: QGIS クリック処理

# reverse-geocording

reverse-geocording(逆ジオコーディング)とは、緯度経度から住所を出力することである。

QGIS の地図上で左クリックすることで、その地点の緯度経度を取得する。

その情報をを使って googlemapAPI を使い、逆ジオコーディングを行っている。

まだ完成していないがこの取得してきた住所の一部(例えば太閤山)などをキーワードにして 3D グラフを作成する予定である。

このキーワードは、施設名、町名、市名ぐらいの優先順位で決定したいと考えている

はじめに

この研究でやること

やったこと

おわりに

## まとめ

- ① バラバラのプログラムになってしまっているが基礎部分のプラグインは順調に進んでいる
- ② 来週にはプログラムをまとめて1つにし1つのプラグインにしたい

はじめに  
この研究でやること

やったこと  
おわりに