

SquareFree Seminar #115, Sep 24 2020 18:40-19:10

COVID-19とマーケティング 経験や勘を超えたデータ分析

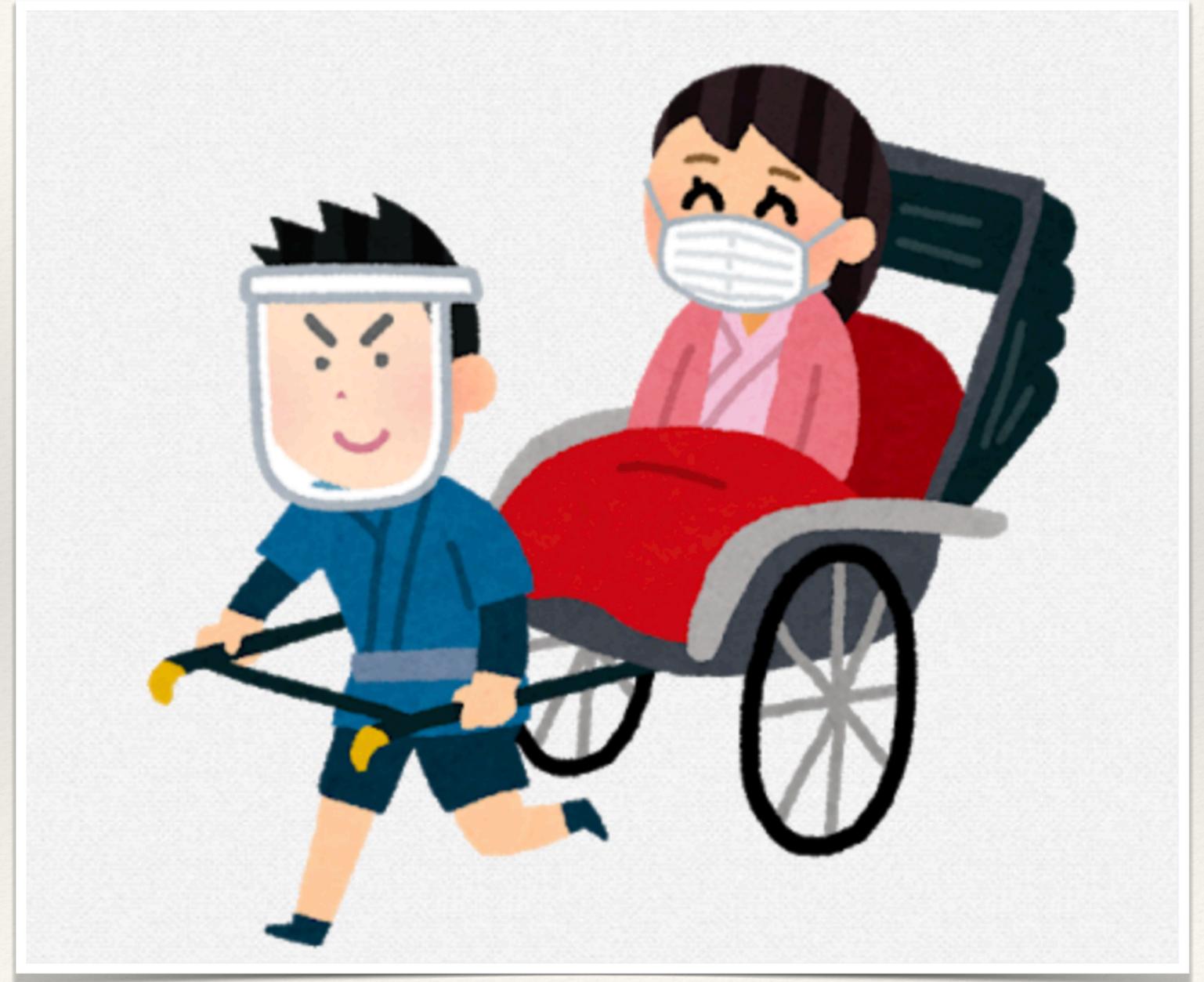
Koji Annoura

案浦浩二

koji@annoura.com

COVID-19の苦労話とマーケティング

- ❖ COVID-19の苦労話
 - ❖ **全国**で**80,000**件ものデータを集めてどのような問題があったのか?
 - ❖ 問題点はどのようなものか
- ❖ コロナ禍でのマーケティング
 - ❖ Webから情報を得る



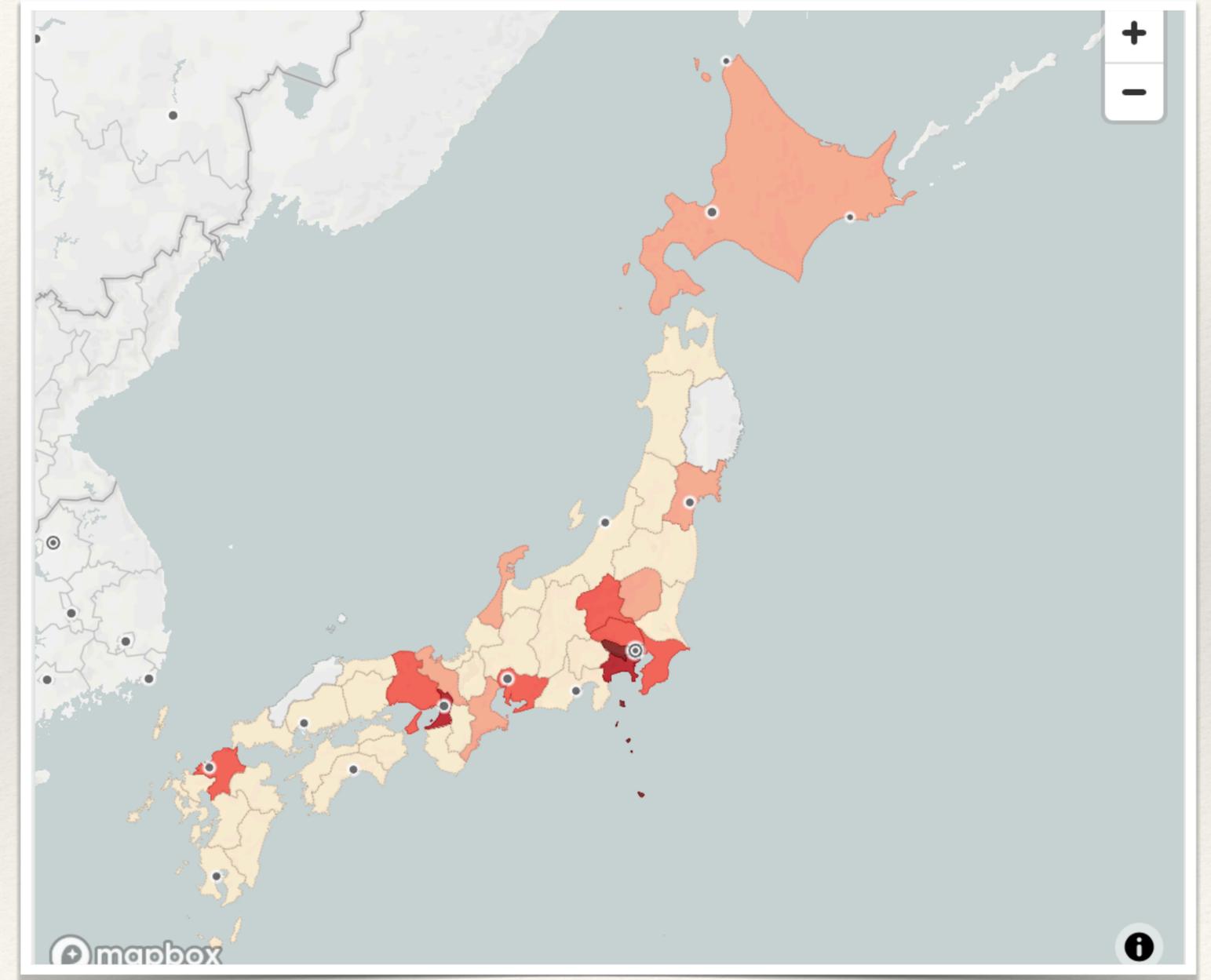
About Me

- ❖ **福岡県** から参加しています
- ❖ グラフ理論が大好きです
- ❖ Apple IIが好きすぎてソフトウェア業界に参入しました
- ❖ Neo4j
 - ❖ Neo4j ユーザーグループ設立者の1人 (2013)
 - ❖ Neo4j Ambassador
 - ❖ Neo4j Speaker, Neo4j Ninja, Neo4j Developer



COVID-19 in Japan

- ❖ Sep 23 2020
- ❖ Japan 80,004 (+219)
 - ❖ 全ての都道府県で陽性者が出ています
 - ❖ 右のマップは現在感染者数
 - ❖ 累計感染者 - 回復者 - 死亡者



COVID-19 in Japan

❖ Sep 24 2020 (+yesterday)

❖ Japan 80,004 (+219)

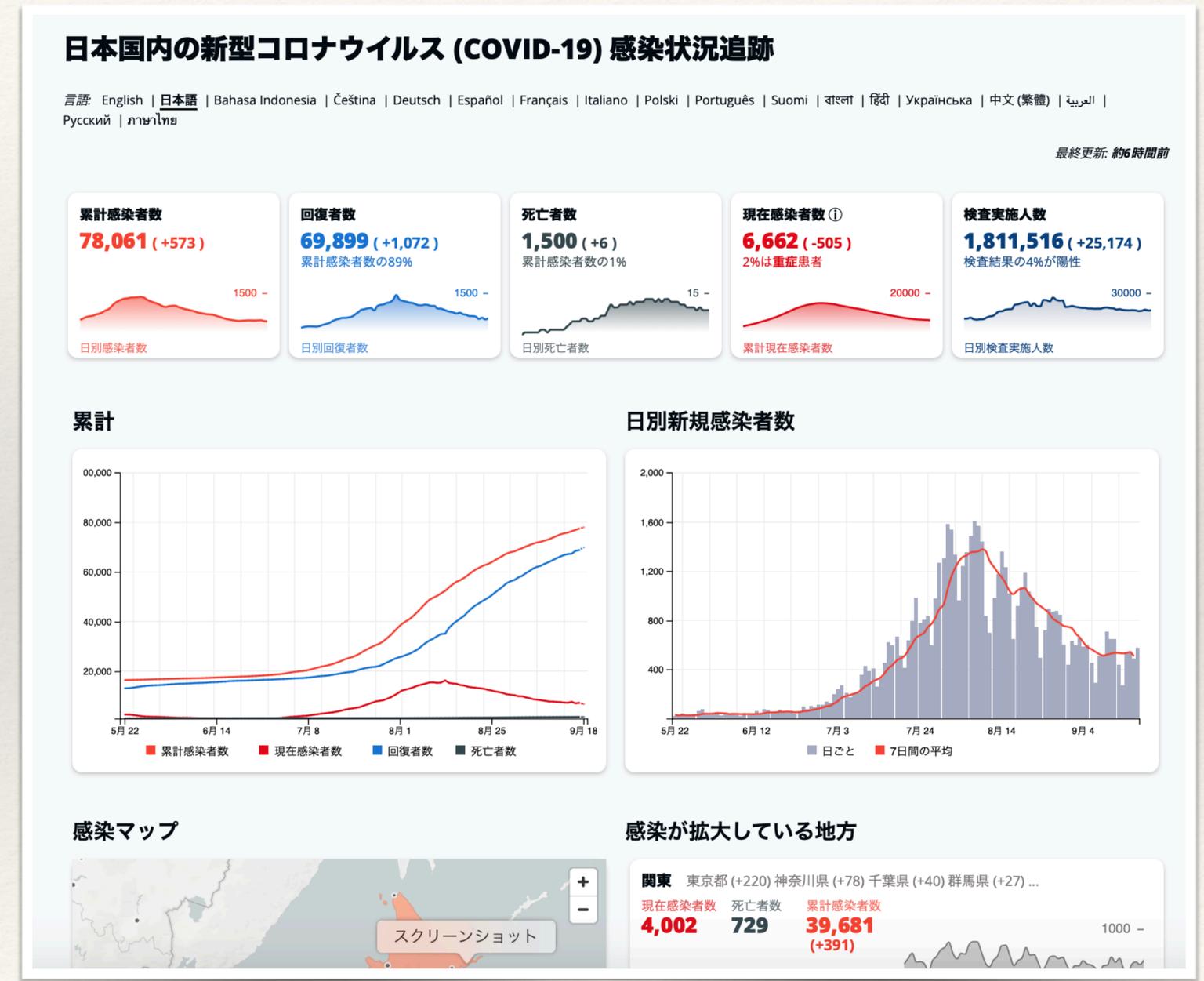
❖ Tokyo (+59)

❖ Osaka (+39)

❖ Fukuoka (+1)

❖ Sapporo (+12)

❖ Tokushima (+26)



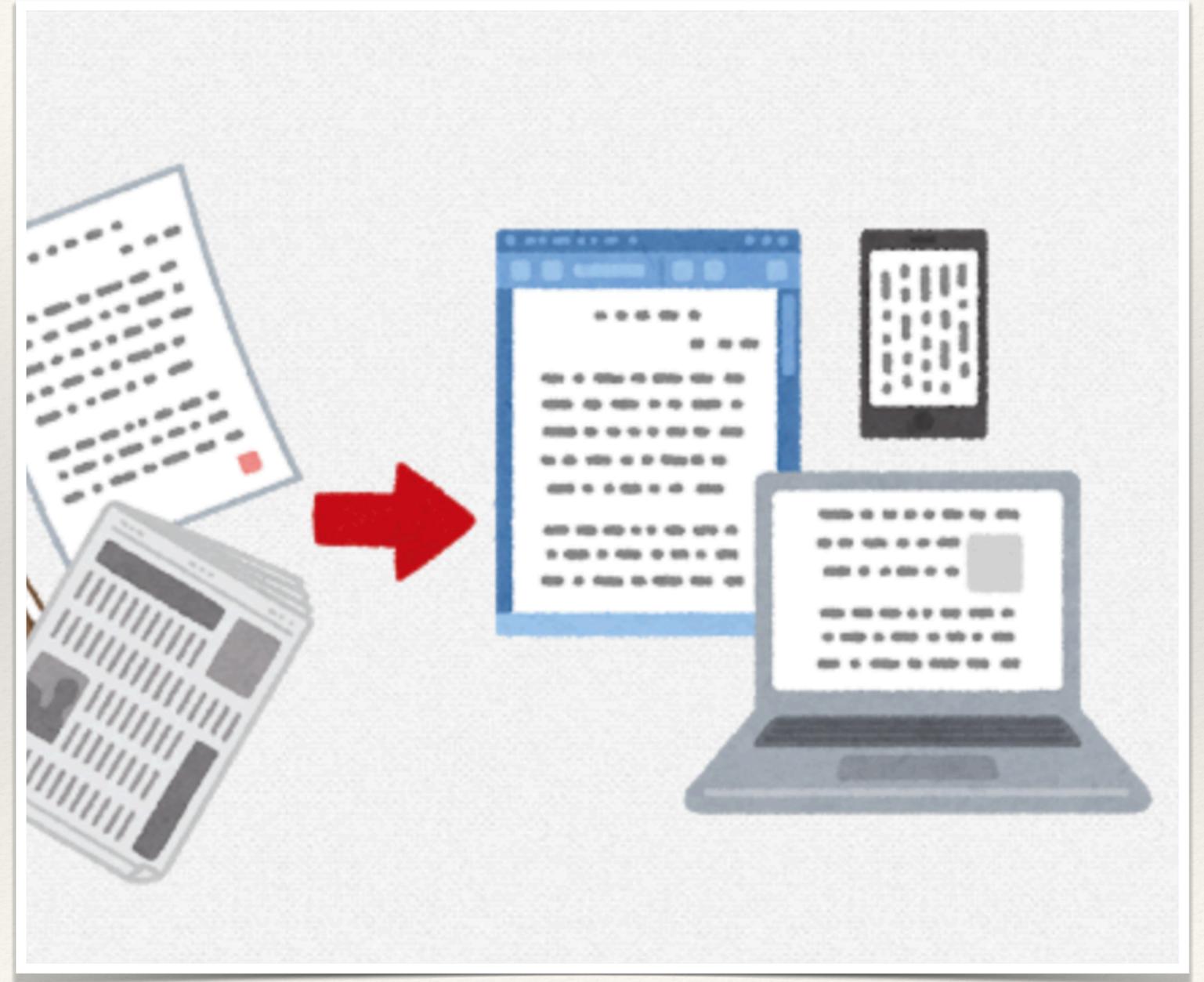
何が分かるのか？

- ❖ **Trace** 感染の経路
- ❖ **Visualize** 感染の広がりかた
- ❖ **Insight** どのような人達が感染しているか



How do I achieve this?

- ❖ 3つの簡単なステップ
 - ❖ 1: 新型コロナのデータを探す
 - ❖ 2: データをNeo4j Graph Databaseにロードする
 - ❖ 3: Neo4j Browserで見る



1: 患者のデータを探す (MHLW)

❖ 患者のデータはどこにあるのか?

❖ 厚生労働省(MHLW)

Ministry of Health, Labour and Welfare

❖ **3月6日**まではサイトに一覧があったが
3月7日以降は無くなっている

❖ その後は報道資料への**リンク**だけになった

❖ 最後には**患者の累計**だけになった

(別添1) 国内事例 (チャーター便帰国者を除く)

・ 3月6日12:00現在、確認されている国内の発生状況は以下のとおり。

【国内事例 (チャーター便帰国者を除く)】

新No.	旧No.	確定日	年代	性別	居住地	周囲の患者の発生※	濃厚接触者の状況
1	1	1/15	30代	男	神奈川県	なし	38名特定 健康観察終了
2	2	1/24	40代	男	中国 (武漢市)	なし	32名特定 健康観察終了
3	3	1/25	30代	女	中国 (武漢市)	なし	7名特定 健康観察終了
4	4	1/26	40代	男	中国 (武漢市)	No.19	2名特定 健康観察終了
5	5	1/28	40代	男	中国 (武漢市)	なし	3名特定 健康観察終了
6	6	1/28	60代	男	奈良県	No.8 No.13	22名特定 健康観察終了
7	7	1/28	40代	女	中国 (武漢市)	なし	2名特定 健康観察終了
8	8	1/29	40代	女	大阪府	No.6	2名特定 健康観察終了

1: 患者のデータを探す (Covid-19 Japan)

❖ Covid-19 Japan

❖ <https://covid19japan.com>

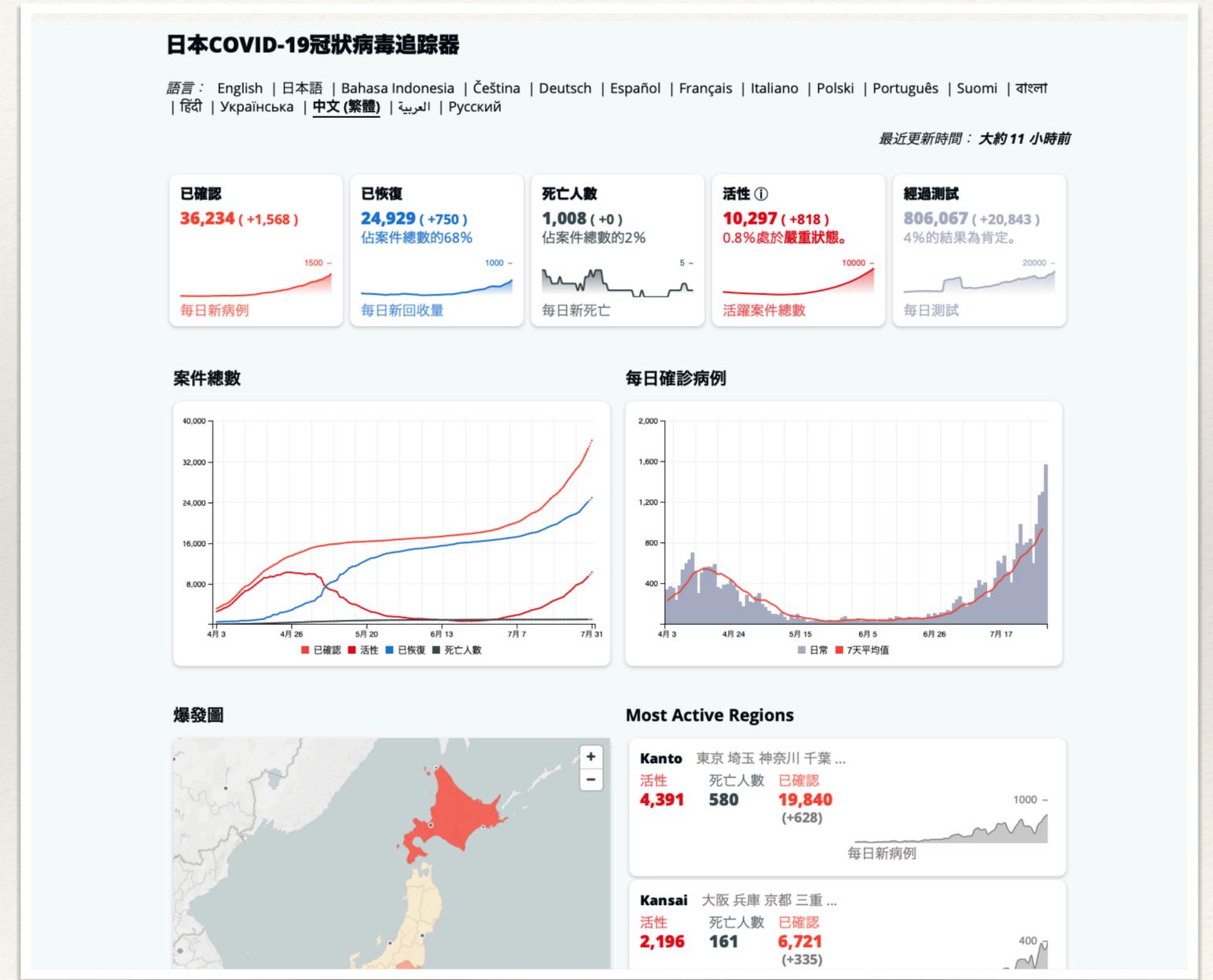
❖ **Shane Reustle**さんがプロジェクトを立ち上げた

❖ 約**40**人のコントリビュータで構成

❖ ほぼ全ての**データ**は**英語**で記述

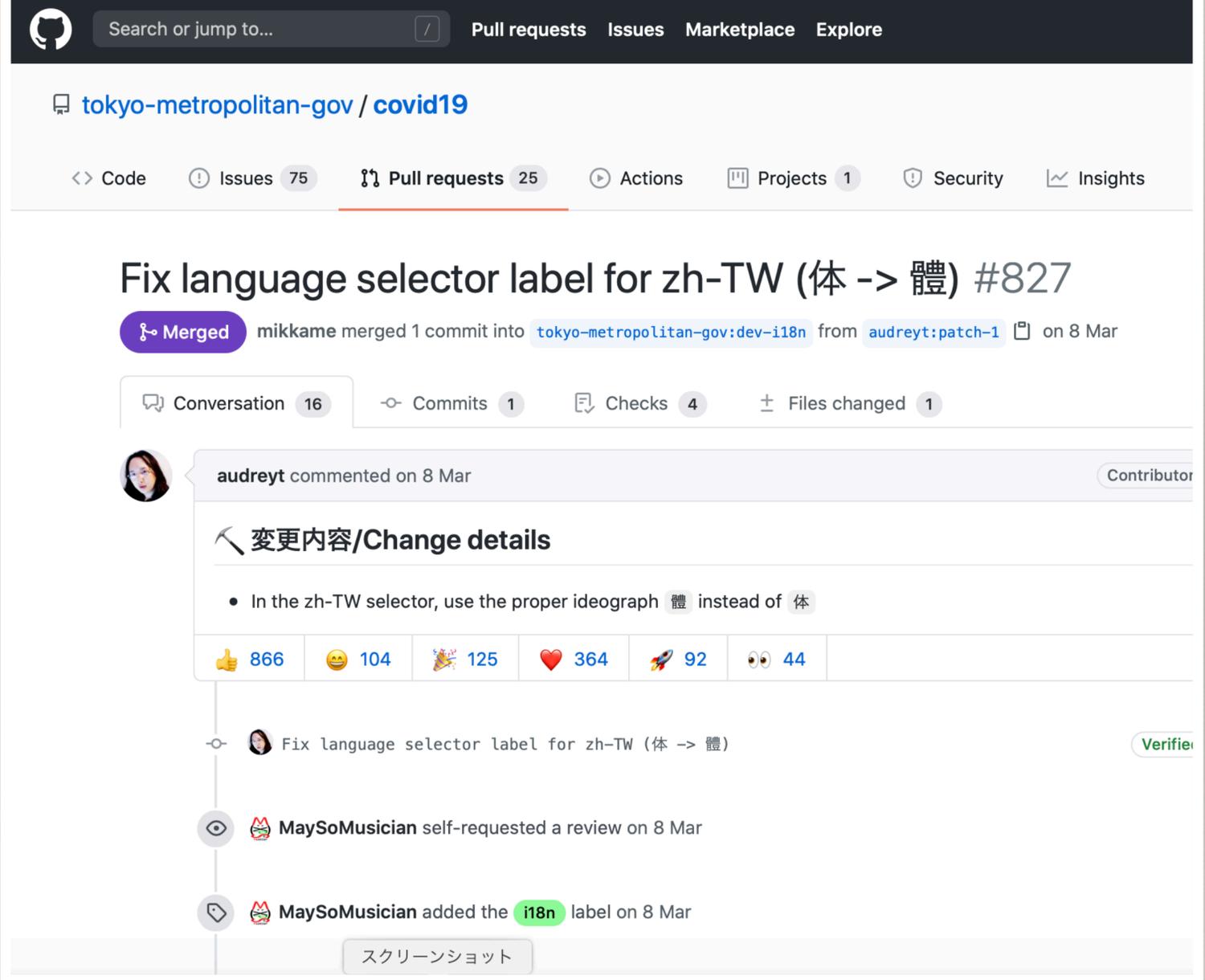
❖ 唯一**濃厚接触者**の項目がある

❖ **英語**から**ロシア語**まで**18言語**



1: 患者のデータを探す (その他)

- ❖ 他のサイトは「濃厚接触者」の項目が無い
- ❖ COVID-19 in Tokyo
 - ❖ <https://stopcovid19.metro.tokyo.lg.jp/en>
 - ❖ Code for Japan
 - ❖ 台湾IT担当大臣のオードリー・タン(唐鳳)さんがプルリクを出したことで有名になった
- ❖ Dashboard&Map of COVID-19 Japan Case
 - ❖ <https://gis.jag-japan.com/covid19jp/>
 - ❖ J.A.G. Japan



The screenshot shows a GitHub pull request page for the repository 'tokyo-metropolitan-gov / covid19'. The pull request is titled 'Fix language selector label for zh-TW (体 -> 體) #827' and is marked as 'Merged'. It was merged by 'mikkame' into the 'tokyo-metropolitan-gov:dev-i18n' branch from the 'audreyt:patch-1' branch on March 8th. The pull request has 16 conversations, 1 commit, 4 checks, and 1 file changed. A comment from 'audreyt' on March 8th provides details about the change: 'In the zh-TW selector, use the proper ideograph 體 instead of 体'. The comment has received 866 thumbs up, 104 neutral reactions, 125 party popper reactions, 364 hearts, 92 rocket reactions, and 44 eyes reactions. The pull request is verified and has a 'Verified' badge. The commit history shows that 'MaySoMusician' self-requested a review and added the 'i18n' label on March 8th.

2: Neo4j Graph Databaseにロードする

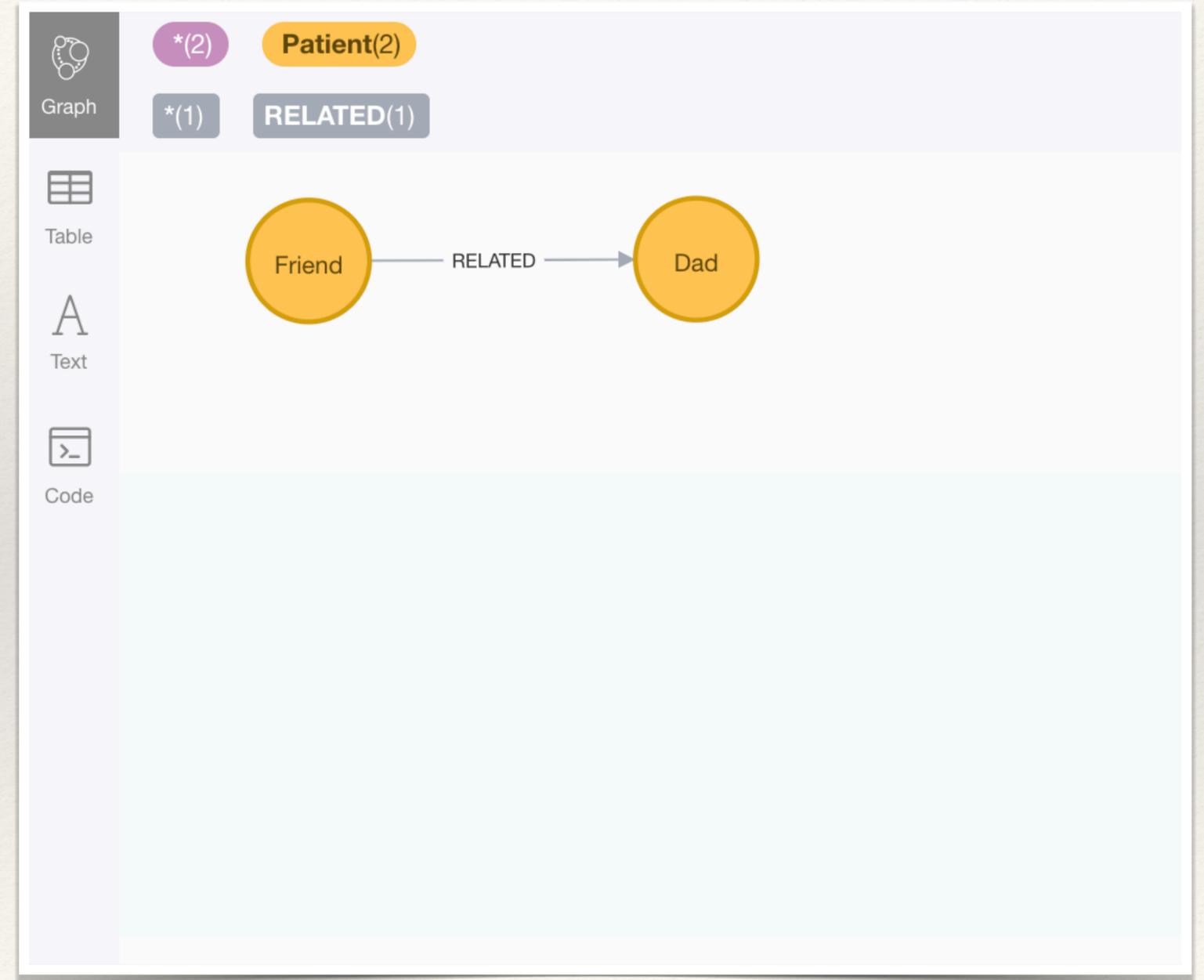
- ❖ LoadCSV

- ❖ COBID-19 Japan data を使ってNeo4jにデータをロードしました



Neo4j Graph Database とは

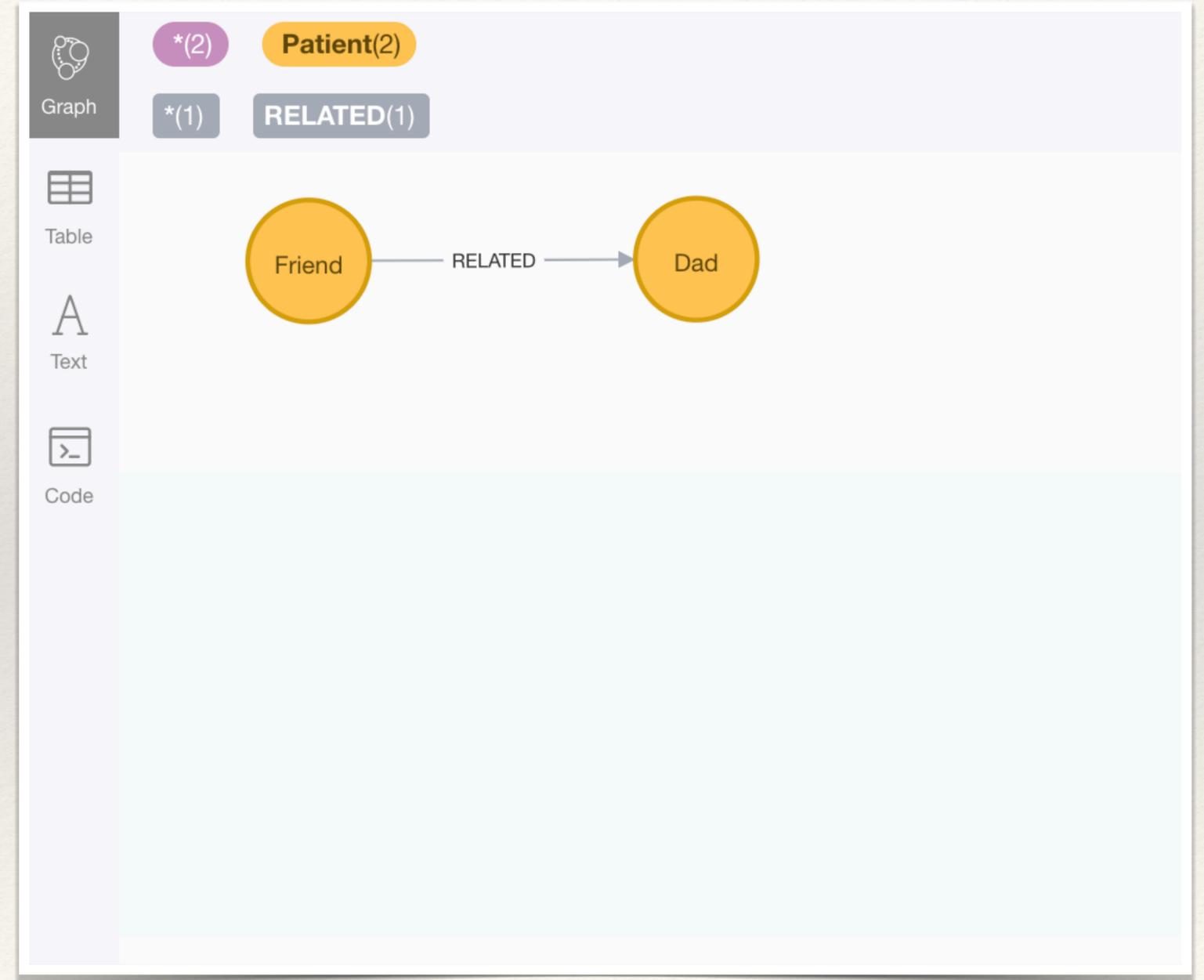
- ❖ Graph = (Nodes, Relationship)
- ❖ コンポーネント
 - ❖ ノード (レコードのようなもの)
 - ❖ リレーションシップ (関連テーブル)
 - ❖ プロパティ (アトリビュート、項目)
- ❖ ラベル (テーブル名のようなもの)



Related Patients

❖ 感染経路

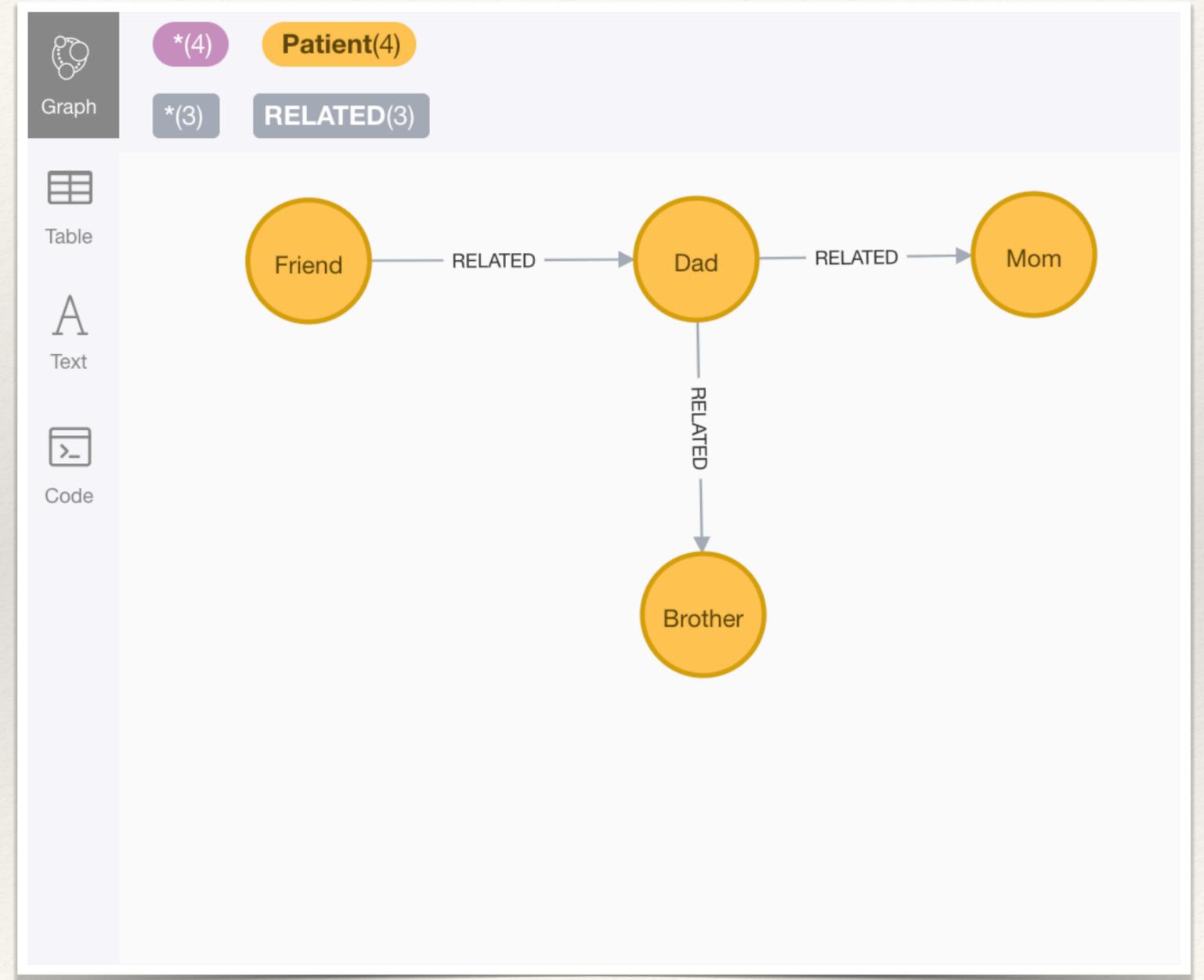
❖ **友人**(Friend)と**父**(Dad)が会う



Related Patients

❖ 感染経路

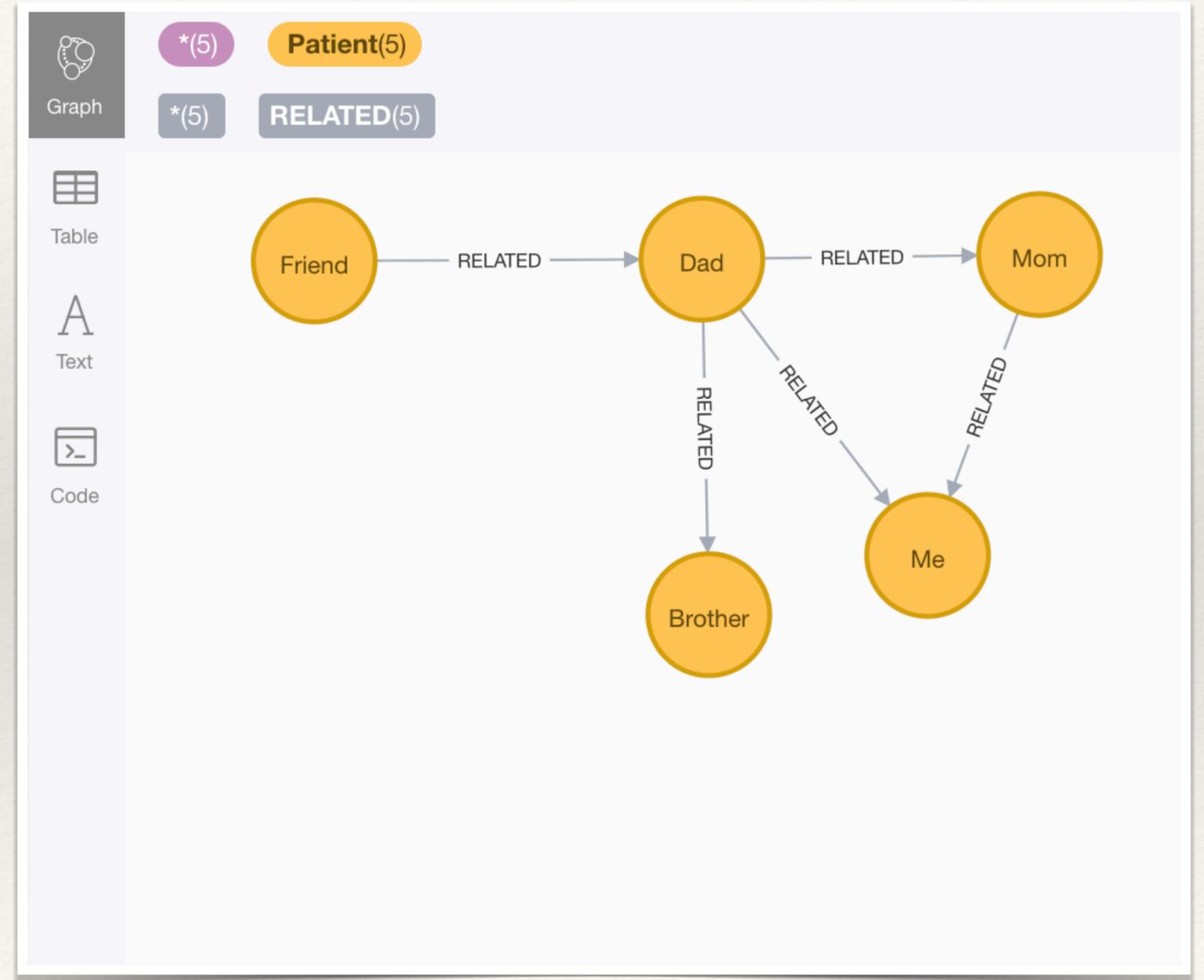
- ❖ 友人(Friend)と父(Dad)が会う
- ❖ 父(Dad)が母(Mom)と兄(Brother)に会う



Related Patients

❖ 感染経路

- ❖ 友人(Friend)と父(Dad)が会う
- ❖ 父(Dad)が母(Mom)と兄(Brother)に会う
- ❖ 私が父(Dad)と母(Mom)に会う



現状の濃厚接触者の記入事例

❖ 濃厚接触者はいくつかの記述がある

❖ 番号 (1,2,3)

❖ 都道府県番号 (A_Pref#3)

❖ 市町村番号 (G_City#1)

❖ 文字の分割は

❖ カンマ (2, 11)

❖ スペース (C_Pref#1 C_Pref#2)

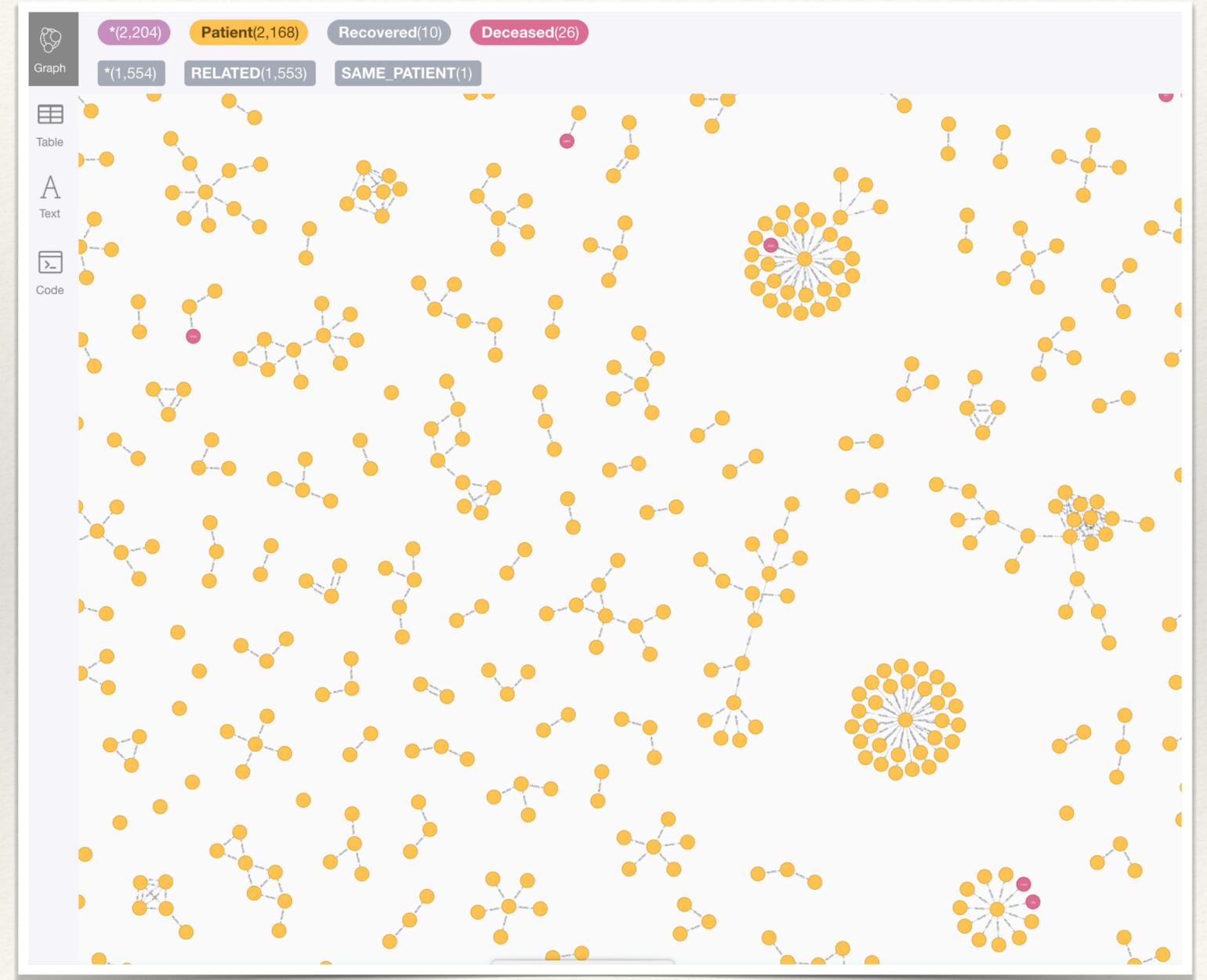
No	Pref Number	City Number	Related Patients
1	A_Pref#2		#3
2	A_Pref#3	E_City#3	Contact with C_Pref#1 C_Pref#2
3	B_Pref#2	F_City#1	Family of F_City#2
4	B_Pref#3	F_City#2	Wife of B_Pref#2
5	C_Pref#1		
6	C_Pref#2		Contact with A_Pref#3
7	D_Pref#1	G_City#1	
8	D_Pref#2		
9	D_Pref#3		2, 11
10	D_Pref#4		
11	D_Pref#5	G_City#2	D_Pref#3
12	D_Pref#6	G_City#3	

3: Neo4j Browserで表示する

❖ 簡単なCypher言語

```
1 // (:Patient)
2 MATCH (n:Patient)-[:CLUSTER|RELATED]-()
3   WHERE NOT n.patient_number = 'Existing'
4 RETURN n
```

- ❖ (:Patient) は患者(Patient)のノード
- ❖ -[:CLUSTER|RELATED]- は
クラスター or 接触した接続
- ❖ () は何かのノード



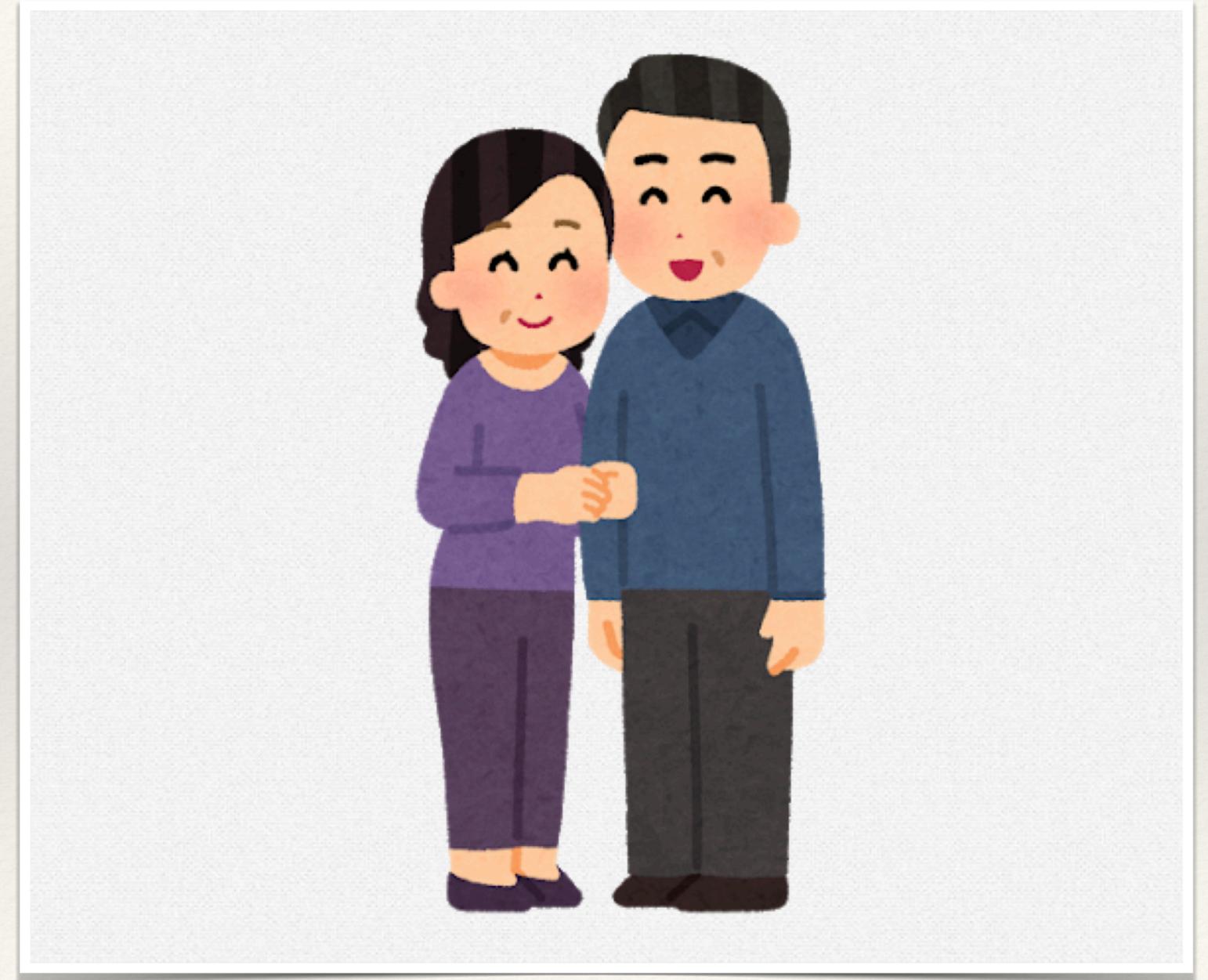
Huh?

❖ いくつかの問題点が見つかった



年代?

- ❖ 夫 (30代) 妻 (50代)?
 - ❖ ありえなくはないけど間違っている気がする
- ❖ 調査した結果
 - ❖ **最初の記者発表**では夫は**30代**
 - ❖ **記載ミス**があったため市町村は夫を**60代**に変更
 - ❖ 市町村は変更されたが都道府県は変更されずそのまま



日付フォーマット問題

- ❖ 日付フォーマット

- ❖ YYYY-MM-DD

- ❖ YYYY/MM/DD

- ❖ YYYY/M/D

- ❖ M/D

	Format
Good	2020-04-01
Good	2020/04/01
Not Bad	2020/4/1
OK in this year	4/1

10代未満問題

- ❖ 10代未満
 - ❖ 1 から 9 (10代以下)
 - ❖ < 1 (1歳未満)
 - ❖ < 6 (未就学児)



広島

❖ 広島県ホームページ

❖ <https://www.pref.hiroshima.lg.jp/site/hcdc/covid19-kanjya.html>

❖ 感染症患者概要

❖ 広島県内の番号

❖ リスト

❖ 性別は詳細を見ないと分からない

No	確定日	年代	居住地	詳細情報
474	9/14	70	広島市	広島市HP
473	9/14	10	広島市	広島市HP
472	9/14	20	福山市	福山市HP
471	9/14	30	福山市	福山市HP
470	9/13	20	広島市	広島市HP
469	9/13	60	広島市	広島市HP
468	9/13	20	福山市	福山市HP
467	9/12	50	広島市	広島市HP
466	9/11	90	広島市	広島市HP

広島のまとめて報告問題

❖ 県内34-56例目

❖ 年齢と性別はあるが...

❖ 誰がどれ

❖ 70歳代2名、80歳代9名、90歳代6名

❖ 男性2名(いずれも80歳代)、女性15名

資料提供
令和2年4月11日
担当：広島県対策本部
担当者：健康対策課 西村
直通：082-513-3076

新型コロナウイルス感染症患者 23 名（県内 34～ 56 例目） に関する情報（4月11日 21時）

本日4月11日（土）、県保健環境センターの検査により、新型コロナウイルス感染症の患者が新たに23例確認されました。

この患者23名は、4月9日（木）に陽性となった三次市の80歳代女性患者（県内25例目の濃厚接触者等又は県内25例目の患者に訪問サービス（A）を行っていた訪問介護員内30例目）の親族であり、検体を採取し、本日、検査を実施したものです。

現在、行動歴等を調査し、二次感染に該当するか慎重に調べているところです。
新型コロナウイルス感染症の患者の発生は県内34～56例目です。

1 【県内34～50例目患者概要】 県内25例目が利用した通所介護施設（三次市）の利用者（計17名）

区分	概要
年齢	70歳代：2名、80歳代：9名、90歳代：6名
性別	男性：2名（いずれも80歳代）、女性：15名
居住地	三次市：17名
職業	無職：17名
症状 (4月11日現在)	症状あり：5名、症状なし：10名、不明：2名
行動	スクリーンショット中

LibreOfficeでまとめる

❖ 県内34-56例目

❖ 広島県と三次市の発表から地道な作業

❖ 誰がどこなのかがわかる

❖ 70歳代2名、80歳代9名、90歳代6名

❖ 男性2名(いずれも80歳代)、女性15名

The screenshot shows a LibreOffice Calc spreadsheet titled "covid-19-worksheet.ods". The spreadsheet contains a table with columns A through G. The data is organized as follows:

	A	B	C	D	E	F	G
33		32 Hiroshima#32		4月11日	40		福山市
34		33 Hiroshima#33		4月11日	40		福山市
35		34 Hiroshima#34	Miyoshi#4	4月11日	70 女性		三次市
36		35 Hiroshima#35	Miyoshi#5	4月11日	70 女性		三次市
37		36 Hiroshima#36	Miyoshi#6	4月11日	80 男性		三次市
38		37 Hiroshima#37	Miyoshi#7	4月11日	80 男性		三次市
39		38 Hiroshima#38	Miyoshi#8	4月11日	80 女性		三次市
40		39 Hiroshima#39	Miyoshi#9	4月11日	80 女性		三次市
41		40 Hiroshima#40	Miyoshi#10	4月11日	80 女性		三次市
42		41 Hiroshima#41	Miyoshi#11	4月11日	80 女性		三次市
43		42 Hiroshima#42	Miyoshi#12	4月11日	80 女性		三次市
44		43 Hiroshima#43	Miyoshi#13	4月11日	80 女性		三次市
45		44 Hiroshima#44	Miyoshi#14	4月11日	80 女性		三次市
46		45 Hiroshima#45	Miyoshi#15	4月11日	90 女性		三次市
47		46 Hiroshima#46	Miyoshi#16	4月11日	90 女性		三次市
48		47 Hiroshima#47	Miyoshi#17	4月11日	90 女性		三次市
49		48 Hiroshima#48	Miyoshi#18	4月11日	90 女性		三次市
50		49 Hiroshima#49	Miyoshi#19	4月11日	90 女性		三次市
51		50 Hiroshima#50	Miyoshi#20	4月11日	90 女性		三次市
52		51 Hiroshima#51	Miyoshi#21	4月11日	70 女性		三次市
53		52 Hiroshima#52	Miyoshi#22	4月11日	非公表	女性	三次市
54		53 Hiroshima#53	Miyoshi#23	4月11日	40 女性		三次市
55		54 Hiroshima#54	Miyoshi#24	4月11日	10 女性		三次市
56		55 Hiroshima#55	Miyoshi#25	4月11日	10 女性		三次市
57		56 Hiroshima#56		4月11日			三次市
58		57 Hiroshima#57		4月12日	40		福山市

市町村名がかぶる問題

❖ 県内34-56例目

❖ 三次市をMiyoshi#1とナンバリング

❖ 同一市町村名(Miyoshi)が他の県も

❖ 三次市、広島

❖ みよし市、愛知

❖ 三好市、徳島

	A	B	C	D	E	F	G
33	32	Hiroshima#32		4月11日	40		福山市
34	33	Hiroshima#33		4月11日	40		福山市
35	34	Hiroshima#34	Miyoshi#4	4月11日	70	女性	三次市
36	35	Hiroshima#35	Miyoshi#5	4月11日	70	女性	三次市
37	36	Hiroshima#36	Miyoshi#6	4月11日	80	男性	三次市
38	37	Hiroshima#37	Miyoshi#7	4月11日	80	男性	三次市
39	38	Hiroshima#38	Miyoshi#8	4月11日	80	女性	三次市
40	39	Hiroshima#39	Miyoshi#9	4月11日	80	女性	三次市
41	40	Hiroshima#40	Miyoshi#10	4月11日	80	女性	三次市
42	41	Hiroshima#41	Miyoshi#11	4月11日	80	女性	三次市
43	42	Hiroshima#42	Miyoshi#12	4月11日	80	女性	三次市
44	43	Hiroshima#43	Miyoshi#13	4月11日	80	女性	三次市
45	44	Hiroshima#44	Miyoshi#14	4月11日	80	女性	三次市
46	45	Hiroshima#45	Miyoshi#15	4月11日	90	女性	三次市
47	46	Hiroshima#46	Miyoshi#16	4月11日	90	女性	三次市
48	47	Hiroshima#47	Miyoshi#17	4月11日	90	女性	三次市
49	48	Hiroshima#48	Miyoshi#18	4月11日	90	女性	三次市
50	49	Hiroshima#49	Miyoshi#19	4月11日	90	女性	三次市
51	50	Hiroshima#50	Miyoshi#20	4月11日	90	女性	三次市
52	51	Hiroshima#51	Miyoshi#21	4月11日	70	女性	三次市
53	52	Hiroshima#52	Miyoshi#22	4月11日	非公表	女性	三次市
54	53	Hiroshima#53	Miyoshi#23	4月11日	40	女性	三次市
55	54	Hiroshima#54	Miyoshi#24	4月11日	10	女性	三次市
56	55	Hiroshima#55	Miyoshi#25	4月11日	10	女性	三次市
57	56	Hiroshima#56		4月11日			三次市
58	57	Hiroshima#57		4月12日	40		福山市

マーケティング

❖ Wikipedia

- ❖ 企業などの組織が行うあらゆる活動のうち、「**顧客が真に求める商品やサービス**を作り、その情報を届け、**顧客がその価値を効果的に得られる**ようにする」ための概念である。また顧客のニーズを解明し、顧客価値を生み出すための経営哲学、戦略、仕組み、プロセスを指す。



マーケティング手法

❖ マッカーシーの4P

❖ Product, Price, Promotion, Place

❖ ロータボーンの4C

❖ Consumer, Customer cost,
Communication, Convenience

❖ SWOT分析

❖ 強み, 弱み, 機会, 脅威

現在、部分ブロックの運用状況の検討が行われており、適

マーケティング手法の一覧

出典: フリー百科事典『ウィキペディア (Wikipedia) 』

- マーケティング(w:marketing)に関するトピック一覧。

目次 [非表示]
1 マーケティング基礎
2 プロダクトマネジメント
3 価格戦略
4 マーケティングマネジメント
5 プロモーション&コミュニケーション
6 E-マーケティング
7 流通
8 ダイレクトマーケティング
9 マーケティング戦略
10 マーケティングリサーチ
11 サービスマーケティング
12 その他
13 関連項目

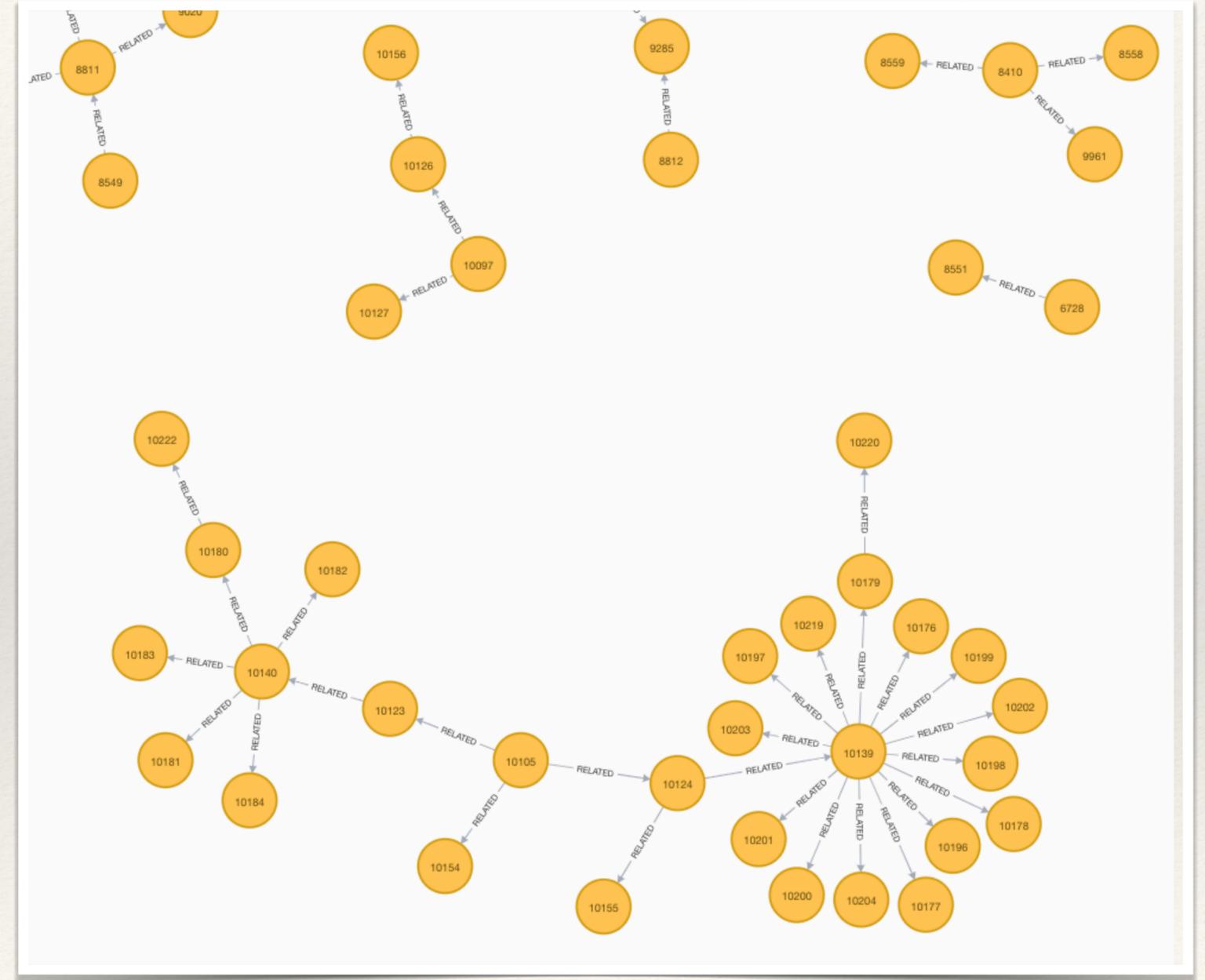
コロナ禍で重要なもの

- ❖ 行動が変わった
- ❖ 風評被害
 - ❖ 噂が一人歩き



コロナの感染と同じように噂は拡大する

- ❖ 02/20
 - ❖ 22:16 福岡にコロナ上陸したん怖すぎる、、ドンキ行こうと思ったのに
- ❖ 02/23
 - ❖ 19:36 中洲のドンキにマスク入荷らしいけど店員がコロナ感染らしい。究極の選択
 - ❖ 21:40 中洲のドンキの店員さんコロナ感染の疑いみたい。中洲に勤めてる友だちがお店のグループLINEで連絡回ってきたみたい。
- ❖ 02/25
 - ❖ 08:46 中洲のドンキの店員コロナ説はやばい
 - ❖ 20:07 中洲のドンキでのコロナ感染者の情報。ガセらしくて助かった、、。客も従業員も中国多いからビビってた(汗)
- ❖ 02/28
 - ❖ 19:55 中洲の人達の噂でドンキでコロナって聞いたけどTwitterでガセネタって確認できた



コロナ禍でのマーケティング

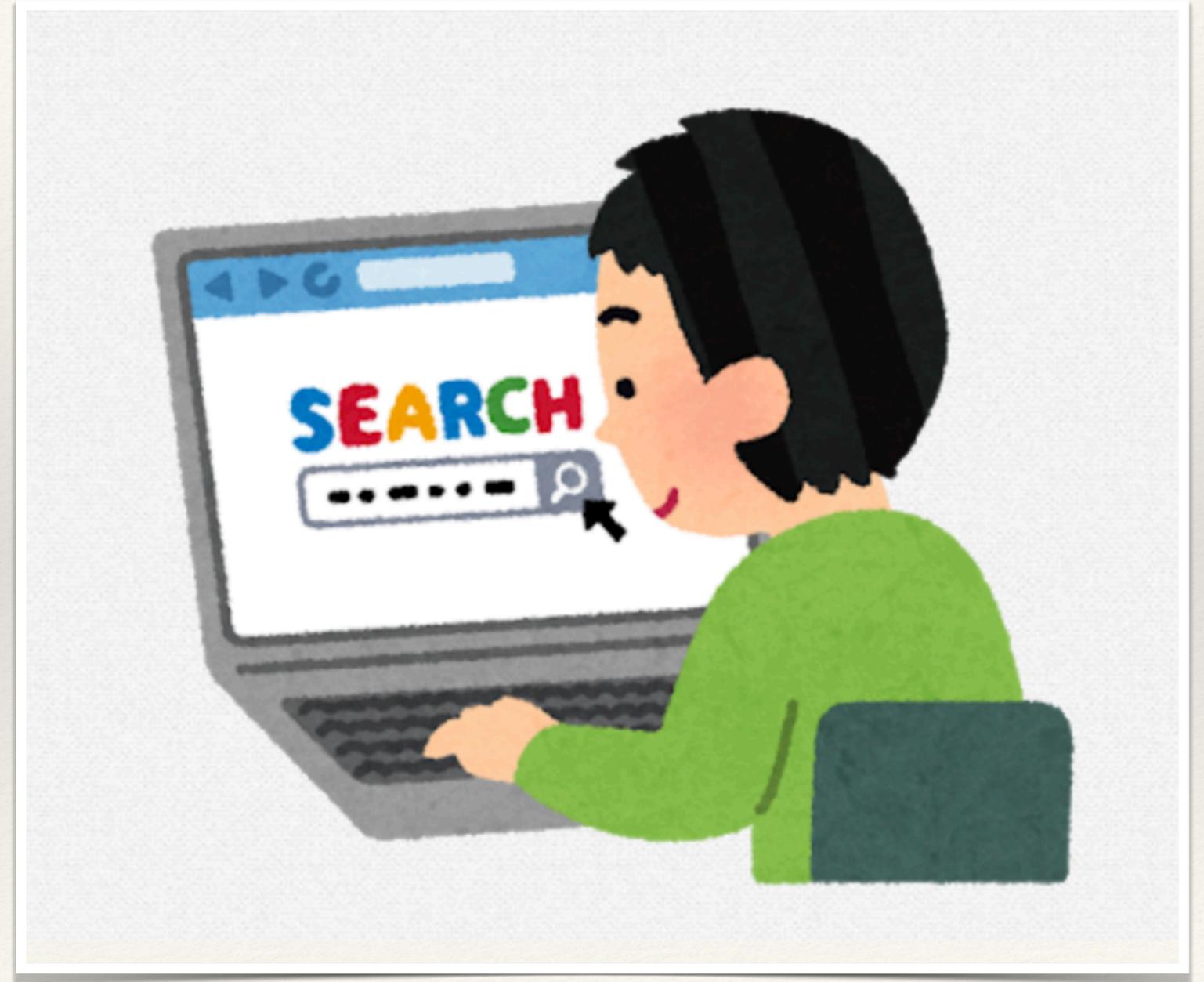
❖ ロール

- ❖ あなたは居酒屋の店長です
- ❖ 緊急事態宣言が解除されても客足戻らず
- ❖ さて、どうしましょう



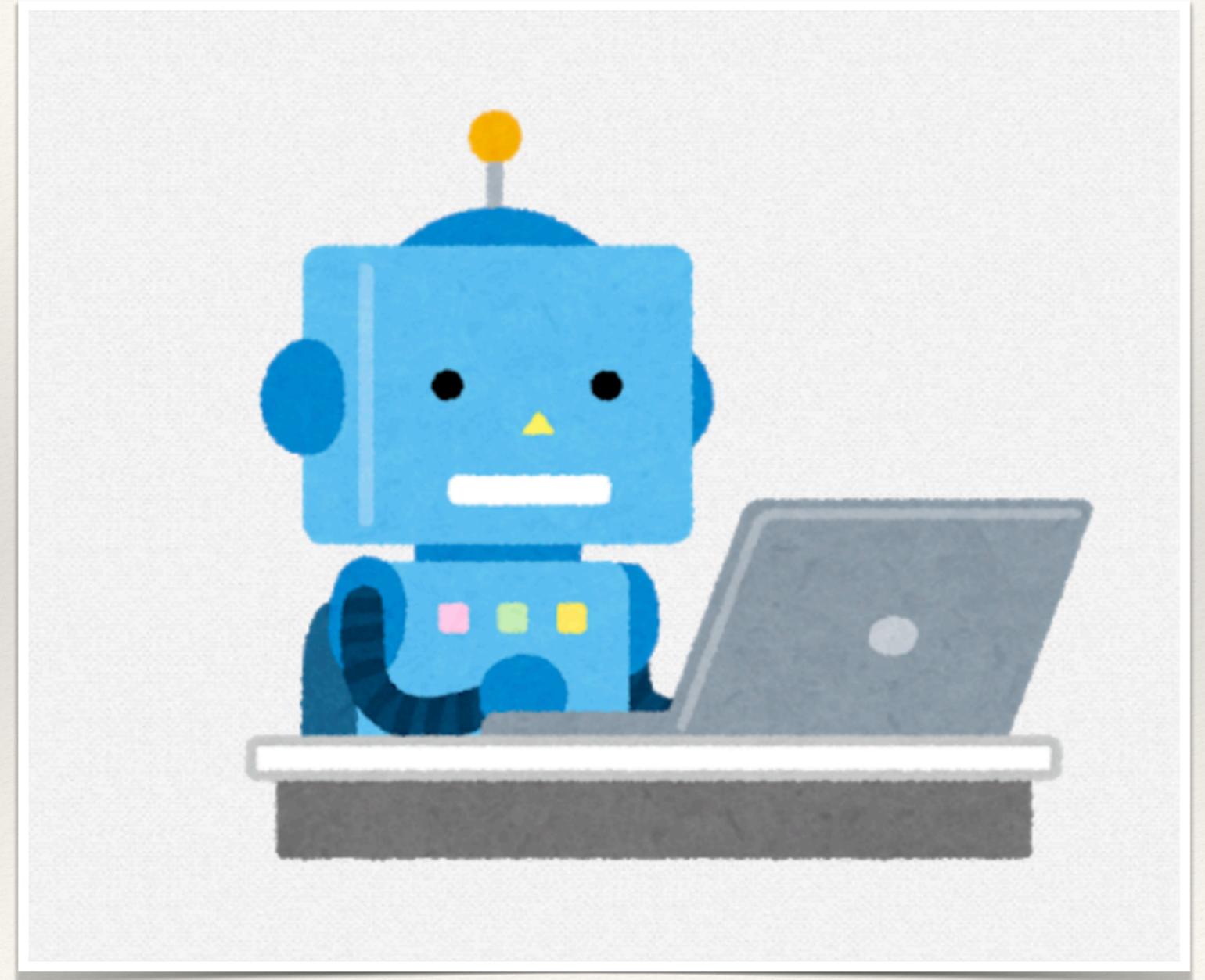
Webからヒントを探します

- ❖ 解決策を検索します
 - ❖ 「コロナ禍」「営業」
 - ❖ 「コロナ禍」「居酒屋」
 - ❖ 「新型コロナ」「売上向上」
 - ❖ なかなか思った結果が出てきません。



ウェブマイニング

- ❖ 調べる基本の単語を入力
 - ❖ 「**コロナ禍**」 「**営業**」
- ❖ 検索結果はとりあえずNeo4jに放り込む
- ❖ 解析（形態素は簡単だが意味は難解）
 - ❖ 形態素解析
 - ❖ 構文解析(係り受け解析)
 - ❖ 意味解析
 - ❖ 文脈解析



形態素解析

❖ 日本語はめんどろ

❖ 自然言語から品詞に分解

❖ 辞書を使う

❖ 例題

- ❖ 「マーケティングって良く使われる言葉ですが、どんな意味かは良く分からない。この言葉、勿論外来語ですが、この言葉が生まれた場所でも良く分からないと言うのが本当のところじゃないでしょうか。だって、Wikipedia を調べてみるとこんなフレーズが . . . wallets。」 そんな凄い魔法をお教えできるわけではないですが、ちょっと参考になるお話を揃えてみました。」

選択辞書：現代語

入力文字:

マーケティングって良く使われる言葉ですが、どんな意味かは良く分からない。この言葉、勿論外来語ですが、
れた場所でも良く分からないと言うのが本当のところじゃないでしょうか。だって、Wikipedia を調べてみると
が . . . 「In 2017 The New York Times described marketing as "the art of telling stories so enthralling the
track of their wallets.」

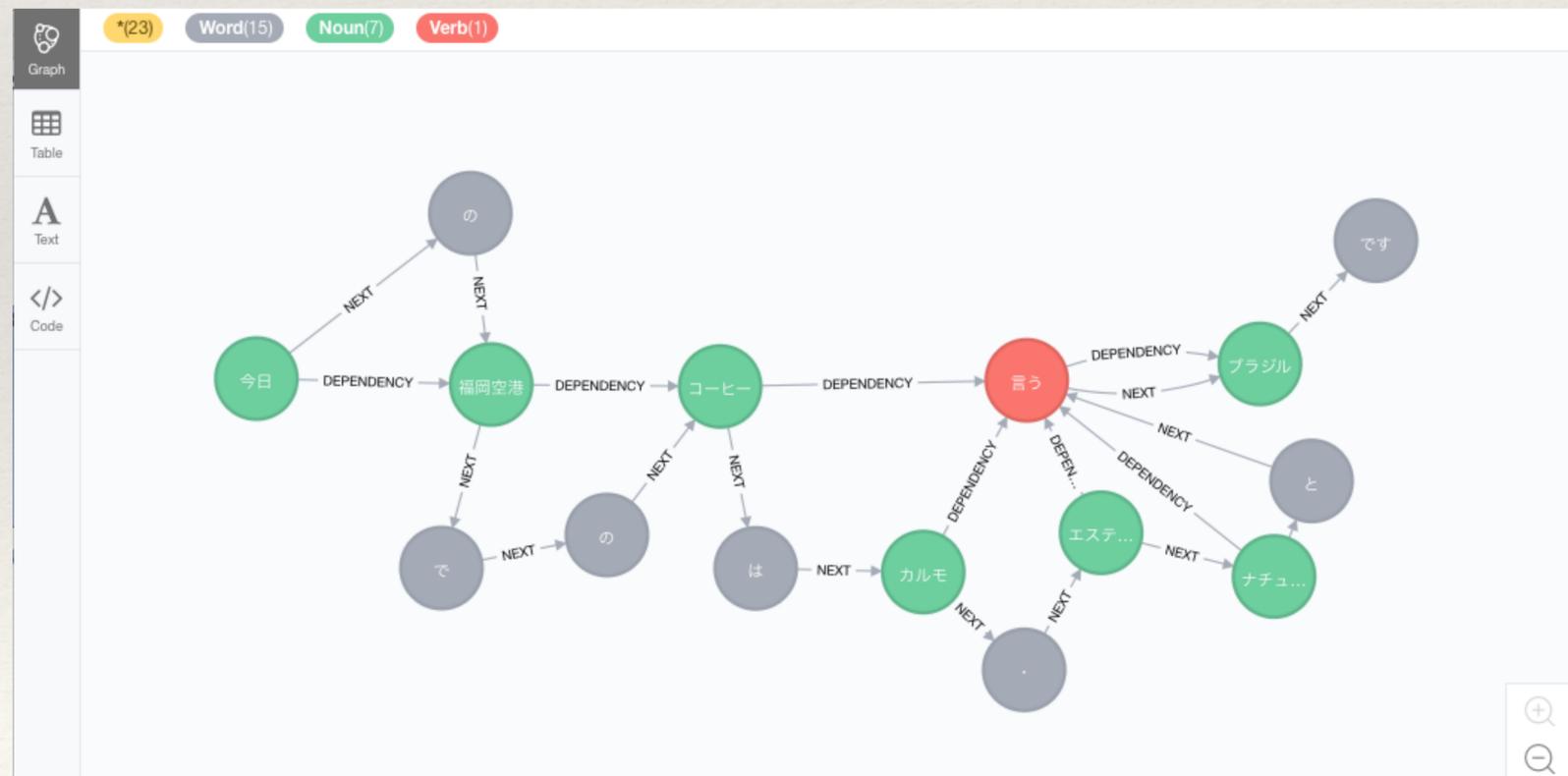
そんな凄い魔法をお教えできるわけではないですが、ちょっと参考になるお話を揃えてみました。

結果

辞書	文境界	書字形 (=表層形)	語彙素	語彙素読み	品詞
現代語	B	マーケティング	マーケティング-marketing	マーケティング	名詞-普通名詞-一
現代語		って	って	ッテ	助詞-副助詞
現代語		良く	良く	ヨク	副詞
現代語		使わ	使う	ツカウ	動詞-一般
現代語		れる	れる	レル	助動詞
現代語		言葉	言葉	コトバ	名詞-普通名詞-一
現代語		です	です	デス	助動詞
現代語		が	が	ガ	助詞-接続助詞
現代語		,	,		補助記号-読点
現代語		どんな	どんな	ドンナ	連体詞
現代語		意味	意味	イミ	名詞-普通名詞-サ
現代語		か	か	カ	助詞-副助詞
現代語		は	は	ハ	助詞-係助詞
現代語		良く	良く	ヨク	副詞
現代語		スクリーンショット	分かる	ワカル	動詞-一般

形態素解析

- ❖ 「今日/の/福岡空港/で/の/コーヒー/は/カルモ・/エステート/ナチュラル/と/言う/ブラジル/です。」



案浦 浩二

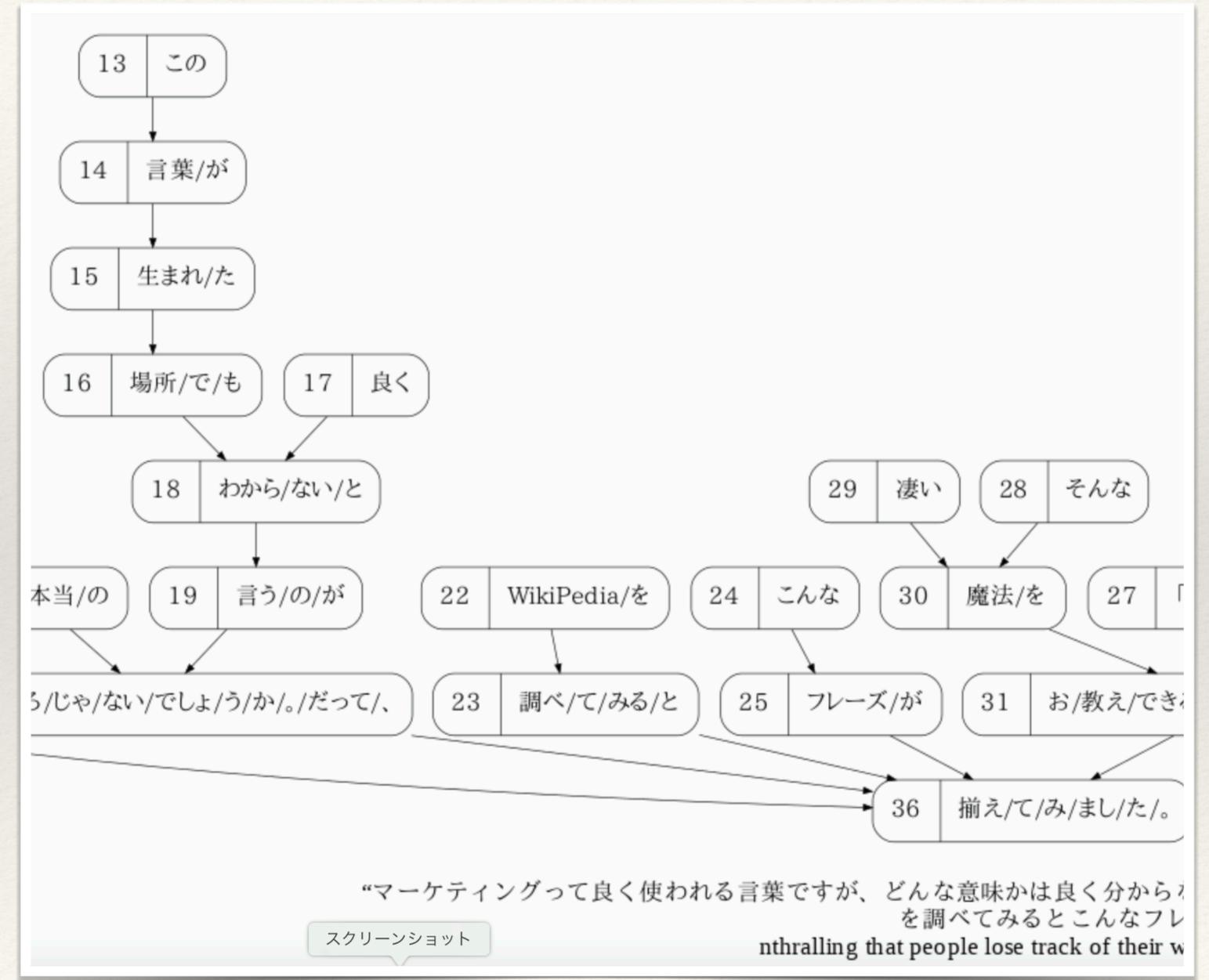
8月29日 6:28 · 公開

今日の福岡空港でのコーヒーはカルモ・エステート ナチュラルと言うブラジルです。甘くやさしいので今日の気分ぴったり。品種はブルボンです。



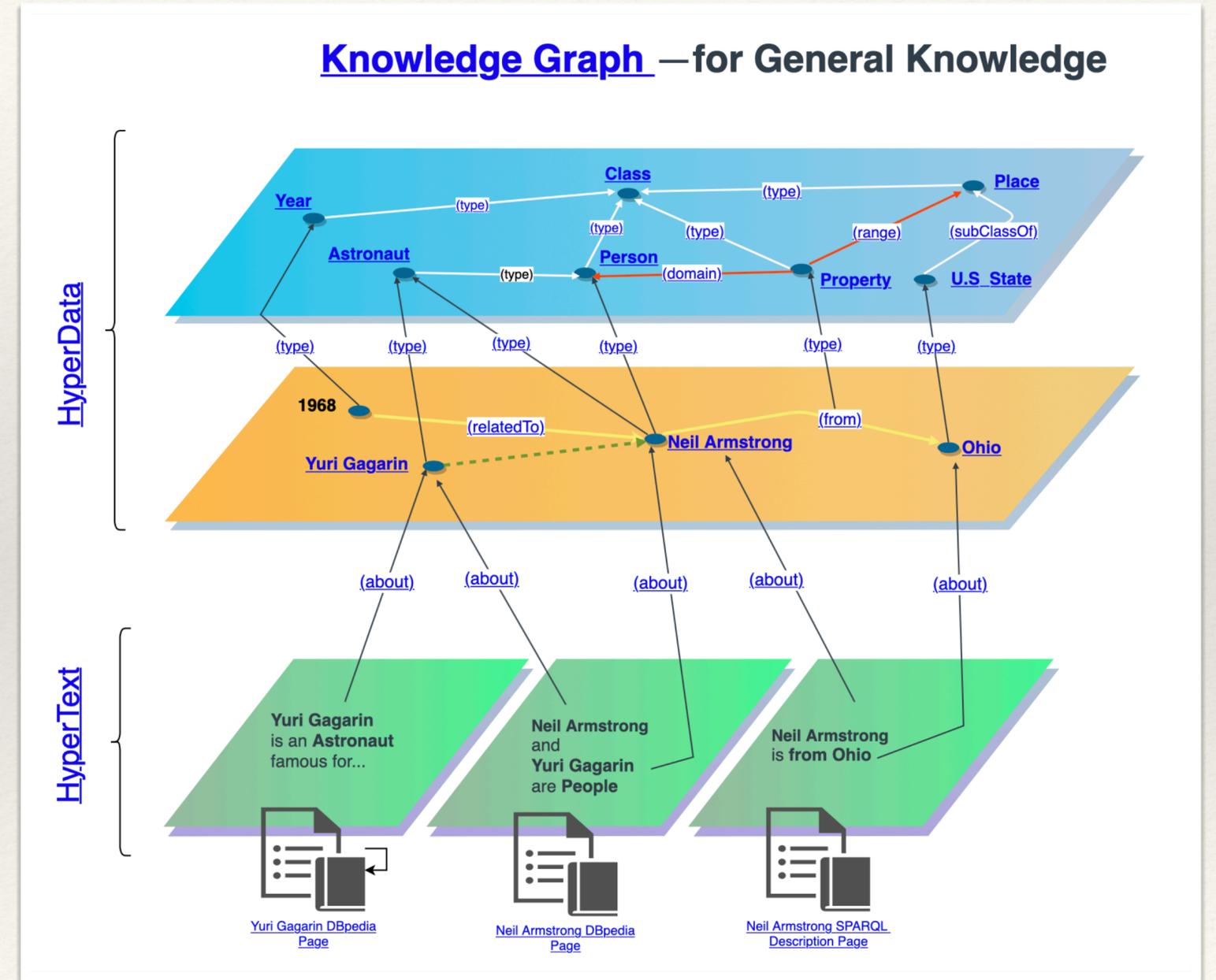
構文解析(係り受け解析)

- ❖ 日本語はめんどろ
- ❖ ちゃんと書かれていると良いが主語とか目的語が省略されるとそれなりに大変
- ❖ Wikipediaの例
 - ❖ 「美しい水車小屋の乙女」
 - ❖ 果たして「美しい」のは
 - ❖ 水車小屋
 - ❖ 乙女



ナレッジグラフ

- ❖ COVID-19ナレッジグラフの例
 - ❖ (ウイルス) \leftarrow [:TYPE]-(:SARS-CoV-2)
 - ❖ (病名) \leftarrow [:TYPE]-(:COVID-19)
 - ❖ (病名) \leftarrow [:RELATED_TO]-(:ウイルス)
 - ❖ (場所) \leftarrow [:SUBCLASS_OF]-(:アメリカ)
 - ❖ (アメリカ) \leftarrow [:TYPE]-(:ニューヨーク)

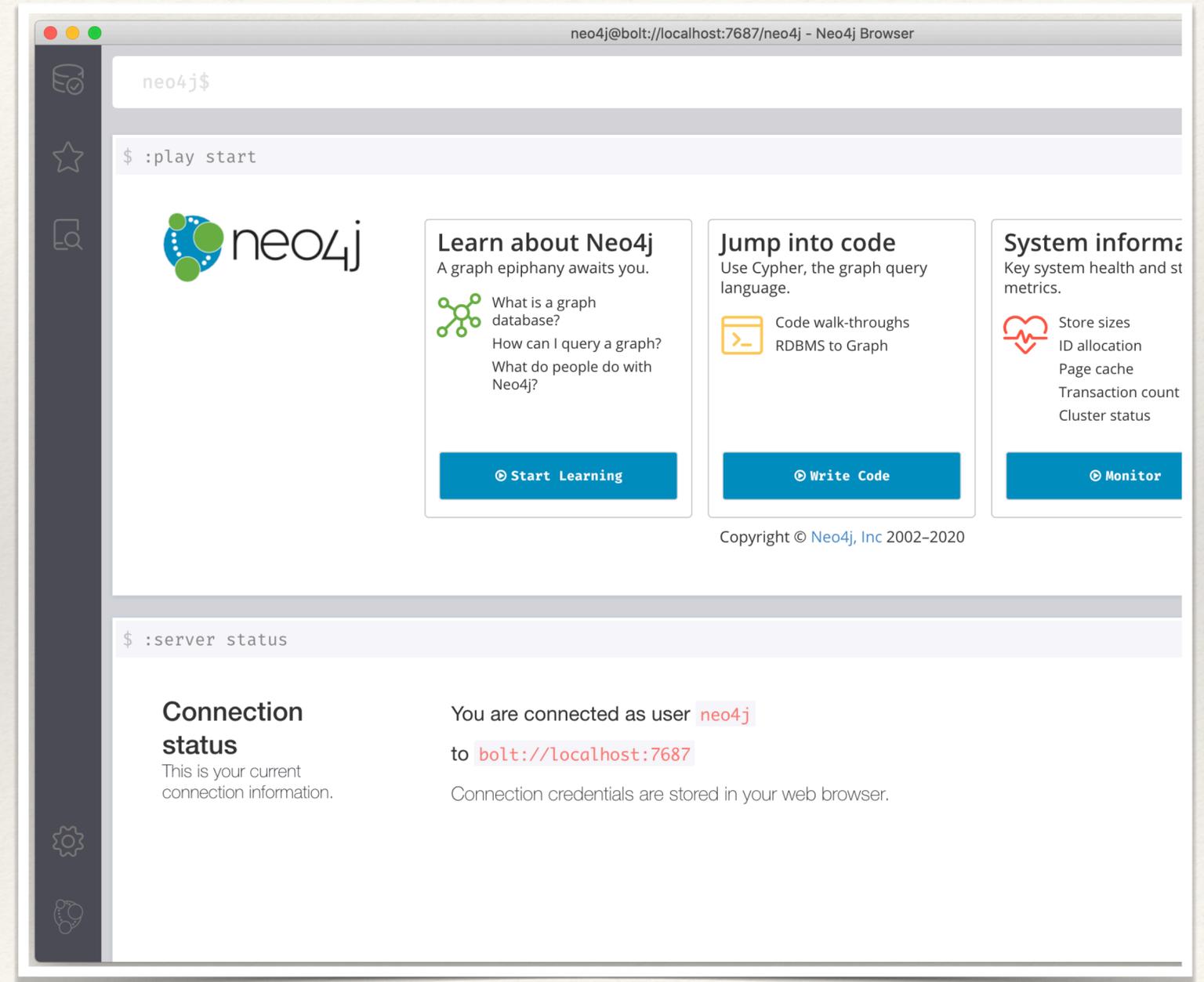


Source: Blog by Kingsley Uyi Idehen

<https://www.openlinksw.com/data/turtle/general/covid-19-knowledge-graph.html>

Demo

- ❖ 実際のデモはYouTubeをご覧ください
 - ❖ スクエア free セミナー 第115回 マーケティングと情報技術
 - ❖ <https://youtu.be/JY6HA3ln5xE>
 - ❖ Demoはこの辺りから
 - ❖ <https://youtu.be/JY6HA3ln5xE?t=3653>
- ❖ Neo4j Desktop
 - ❖ Neo4j Enterprise
 - ❖ Runs Windows, Mac, Linux
- ❖ Current version is Neo4j 4.1



ありがとうございました

❖ ナレッジグラフが重要ですが、トレンド検索でもなんとか頑張れます