

# 医療と社会の接点としてSNS分析

人工知能による社会現象のデータ分析

15:00～15:40

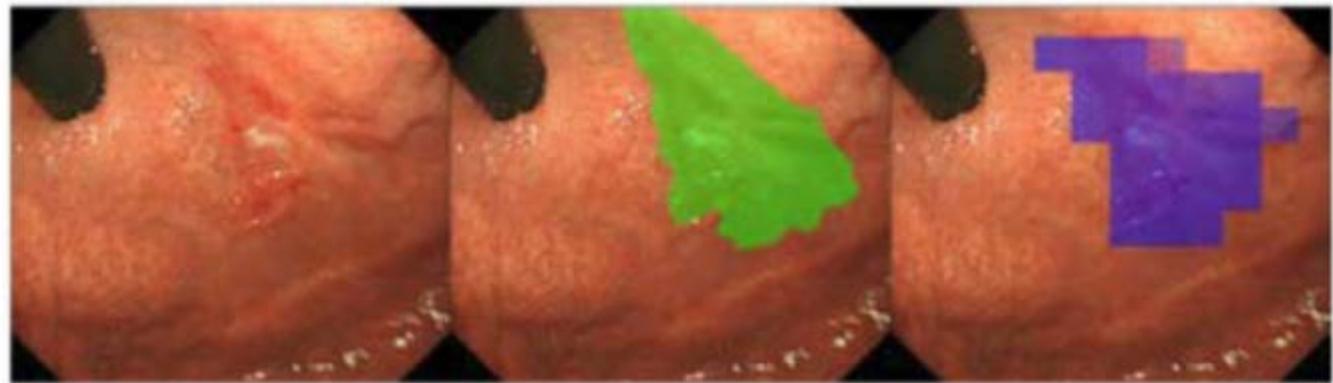
荒牧英治

奈良先端科学技術大学院大学

ソーシャル・コンピューティング研究室

# 画像とテキストのAIの対比

- 画像アノテーション ▶ 深層学習



入力画像

正解データ

コンピューターの検出

医師の診断（緑）とコンピューターの自動検出（紫）が示した早期胃がんの領域

IEEE Engineering in Medicine and Biology Society 2018より

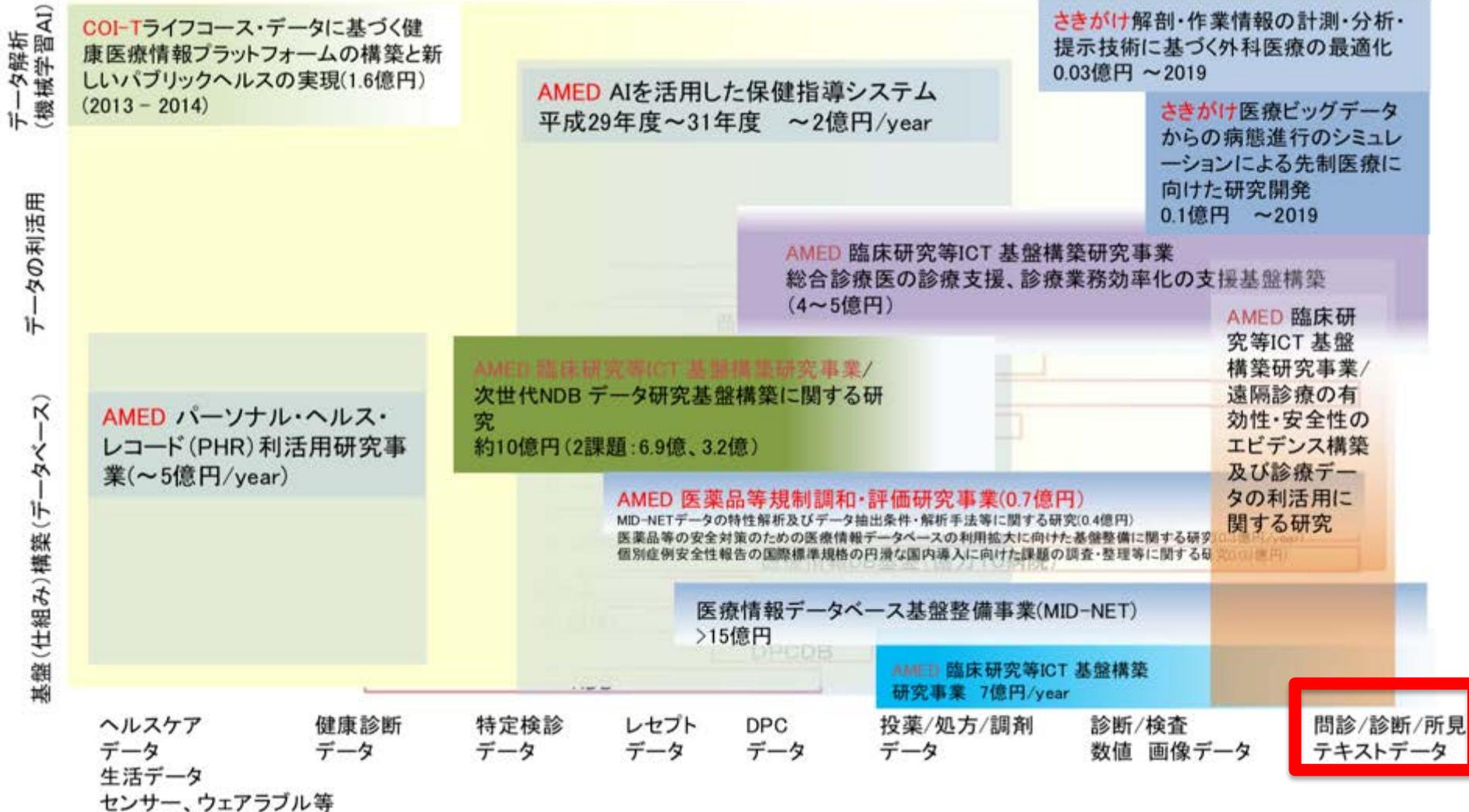
- テキストアノテーション ▶ 深層学習

正解データ 胃噴門部に**2型腫瘍**を認め、生検結果は**腺癌**で

コンピューターの検出 胃噴門部に**腫瘍**を認め、生検結果は**低分化腺癌**で

# 本邦の研究プロジェクトにおいて

## テキストデータの処理は始まったばかり



JST 政策提言検討WS (永井良三先生, 大江和彦先生, 黒田知宏先生, 荒牧) での資料 (JST 辻フェローに作成いただいたものを拝借)

# 医療言語処理という新分野をどう作る？

- **発表の場**：研究者が集まる機会を作る
  - － 国際ワークショップの主催
    - NTCIR MedNLP, MedWeb など
    - これまで40以上の会社／研究機関が参加

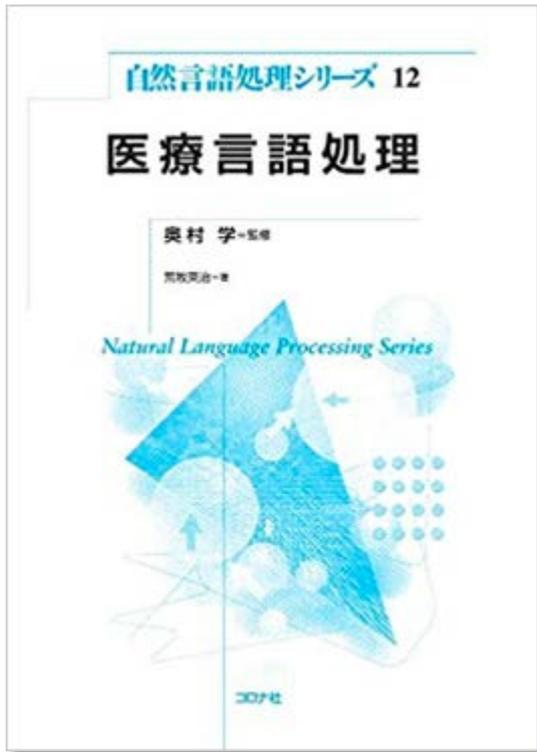


- **社会受容** → **科研費の分野がある**
  - － 社会に受け入れられる／必要だと思ってもらえる

後述する活動により、H28から対応する厚労科研事業がスタート（継続中）  
電子カルテ情報をセマンティクス（意味・内容）の標準化により分析可能なデータに変換するための研究（厚労省 政策科学総合事業 H28～）

- **学問体系** → **教科書がある**
  - － 学問としての体系を形にする

# 教科書がない → 出版しよう



医療言語処理 (2017年)

2015年から2年をかけて医療分野の言語処理に関する教科書として書籍を執筆 (単著)



岩波データサイエンス (2016年)



臨床心理学増刊 (2017年)

1. 医療情報の利活用とは
2. 利用可能なリソース・ツール
3. 病院内テキスト
4. パブリックデータ \*
5. プライベートデータ \*
6. これからの医療言語処理研究

個別の内容については、これまでも執筆の機会があったが、初めて体系立てた書物を刊行

# 進行中の研究のデータ規模・目的

データのクオリティが研究の上限を決める

永井良三学長代表

H29 AMED  
**N=42,000**

 日本内科学会  
The Japanese Society of Internal Medicine

日本内科学会症例報告  
診断支援システムの構築

栗山猛先生代表

H27-30AMED  
**N=12,000**

 国立研究開発法人  
国立成育医療研究センター  
National Center for Child Health and Development

薬剤添付文章  
副作用表現の収集

NICT長井先生代表

H28 CREST  
**N=100**

発達障害・統合失調症当事者の発話  
介当事者研究の効果測定

H27-30厚労科研

**N=18,000,000**

 東大病院  
The University of Tohoku Hospital

東大病院の全診療録  
大規模病名辞書の構築

H27-30 AIP-PRISM  
**N=1,000**

国立がんセンターの読影レポート  
レポートの自動構造化

中島先生代表

H29-32 基盤B  
**N=2**

阪大病院 手術中の医師の会話  
手術の安全性確立

H29-33 AIP-PRISM  
**N=?-3,000**

大阪大学病院呼吸器科読影レポート  
創薬シーズの探索

H29-  
**N=?**

神戸市民中央病院 循環器科診療録  
カルテ要約の自動生成

H29-  
**N=?**

京都武田総合病院 科診療録  
カルテ要約の自動生成

慶應・望月先生代表

H27-30AMED  
**N=300,000**

 慶應義塾大学病院

すずらん薬局の薬歴報告  
患者症状表現の収集

H29-30  
**N ≒ ∞**

 YAHOO!  
JAPAN

Yahoo! 検索ログ, ショッピング  
感染症の流行予測

松井先生代表

H29-30  
**N ≒ ∞**

風疹に関するブログ, つぶやき  
感染症リスクコミュニケーション

NTT山下先生代表

H29 基盤A  
**N=20**

 NTT

認知症, うつ病の家族SNSログ  
介護者へSNSを用いた介入

H29 CRIS  
**N=170**

 LINE@

大阪国際がんセンターの患者発話  
がんと認知機能低下の調査

東大笠井先生代表

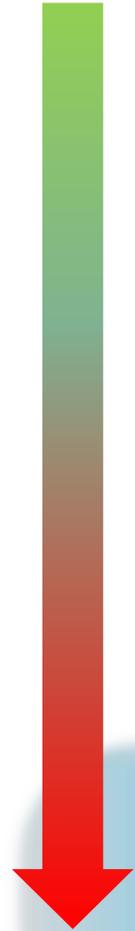
H27 新学術  
**N=4,000**

東京ティーンコホート  
思春期の主体価値発達の調査

# 研究の目指すべきところ

これまでの人工知能の材料  
(IBM Watson)

文法的



公式テキスト

教科書, 薬材添付文章  
論文, 症例報告

基礎研究



臨床テキスト

カルテ, 薬歴,  
副作用報告

臨床研究



患者テキスト

患者レジストリ, PRO

質的研究



SNSテキスト

つぶやき, 動画像,  
検索キーワード

ビッグデータ研究

非文法的



2019-02-10 15:19:32 -1.0000 RT @MS\_sousaku1: 創作漫画 彼女はスキだらけ05寒いけど風邪ひかないでね <https://>

2019-02-10 15:19:31 -0.2320 @pome\_granateR 風邪

2019-02-10 15:19:25 0.9472 元気ないなあ御堂筋。風邪か？

2019-02-10 15:19:24 27 -1.0000 @Tempesta\_J 素敵

2019-02-10 15:19:23 1 1.0000 @sau\_nyan 僕も朝から非常に喉が痛いです風邪ですかね（すっとぼけ）

2019-02-10 15:19:22 15 -1.0000 @knrmy2 風邪ひかないでね(^▽^)/それだけが心配です(;´∀`)

2019-02-10 15:19:21 0.6114 ハンバーガーとタラちり鍋で栄養取りまくって薬も飲んだのに風邪治らん

2019-02-10 15:19:21 11 -1.0000 @Rimo\_26 ピンクの服と唇のピンクが合ってる素敵ですね♪風邪など引かないように

2019-02-10 15:19:21 -1.0000 @mini4johnny アーク！！カッコイイですね？そしてトトちゃん

2019-02-10 15:19:21 -0.6316 RT @qn\_\_83: そんな予防接種の副反応に悩まされたけど、3回連続で発熱後は一度も

2019-02-10 15:19:18 13 1.0000 風邪ひいた…ほとんど寝てるだけの日曜日…

2019-02-10 15:19:17 14 -0.1264 ただいまー！寒くならないうちに帰ってきた。風邪はもう嫌じゃー！

2019-02-10 15:19:16 12 0.9161 小山さん今何してる？？？元気？？寒いけど風邪ひいてないですか？？

2019-02-10 15:19:15 0.7238 てんちょ風邪ひいてるし

2019-02-10 15:19:13 -1.0000 俺にもし子供がいたとして、その子が風邪で寝込んだら俺はめっちゃイケ行かないぜ？

2019-02-10 15:19:12 -1.0000 @asamiyayuri 風邪引かないように、気をつけてくださいね(\*´-`)

2019-02-10 15:19:10 -1.0000 RT @target7japan: #TARGET #WONNIE #POPEYE 今日は来てくださったファンの方

2019-02-10 15:19:10 -0.9110 @honmahimawari ひまちゃんおはごぎー！今日はずっと楽しみにしてた鍋バトルの日

2019-02-10 15:19:08 -0.5647 風邪ひく→4日5日ずっと寝る→何故か腰痛で這いつくばって移動

2019-02-10 15:19:06 13 -1.0000 @saki\_dollykiss おつかれー

2019-02-10 15:19:05 1.0000 @sonora\_k いや、ただの風邪です

# 活版印刷

記録を広く残せるように



活版の一例

<http://kot-book.com> (Accessed 2016/03/20)

# 知識の継承

- 文字

- はじめて知識を残せるように

- 印刷術

- よって知識を広く残せるように

- どんな知識が残されたか？

- 「新約聖書」やルターの著作

BC13000

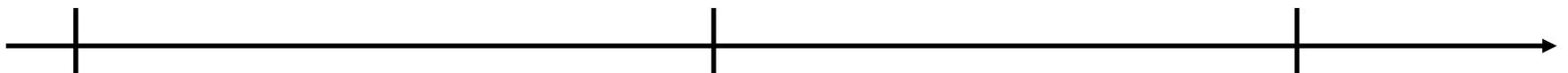
ウルク古拙文字

AD1450

活版印刷

AD2000

Web/SNS



# SNSで得られる情報

Word ▪ User ▪ Time ▪ Location

Location

User1  
User2  
User3  
User4  
User5  
User6  
User7

Word1

Word2

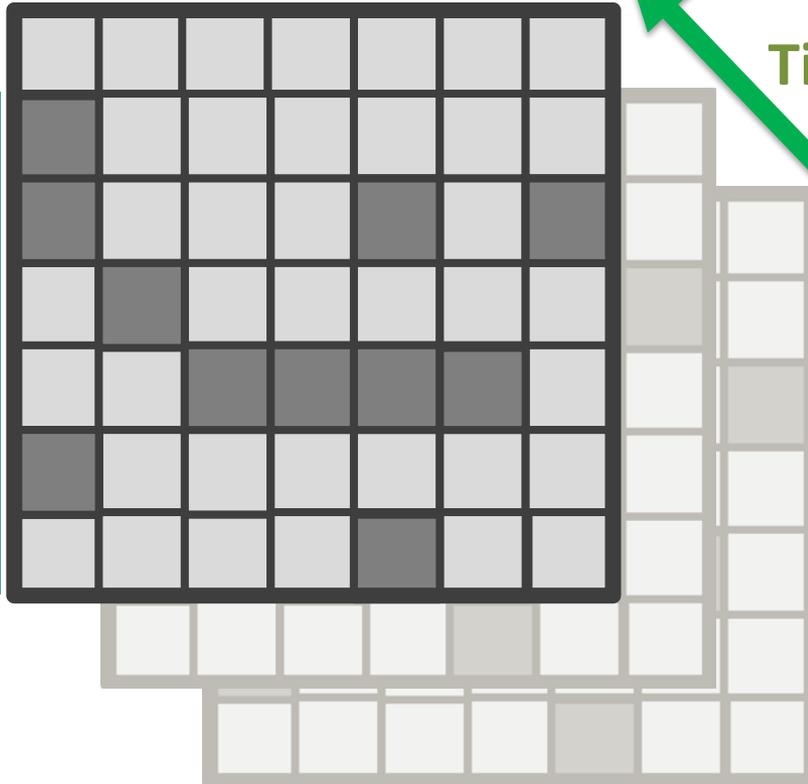
Word3

Word4

Word5

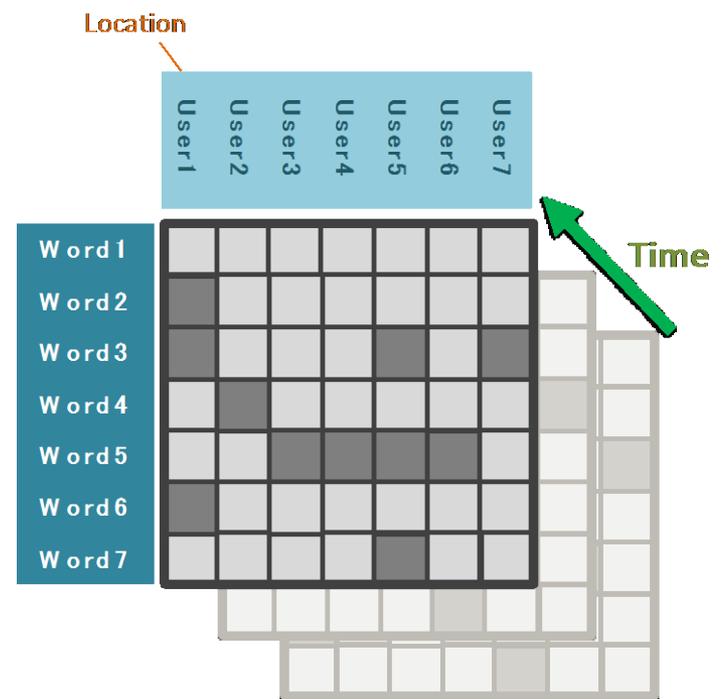
Word6

Word7



Time





- はじめに
- **Word** x **Location**
- **Word** x **Time**
- **Word** x User
- おわりに

# SNS応用の代表的論文 500被引用 (2019現在)



俺も風邪引いたー  
開く



一人の友達が風邪引いちゃったみたいで。(?  
開く

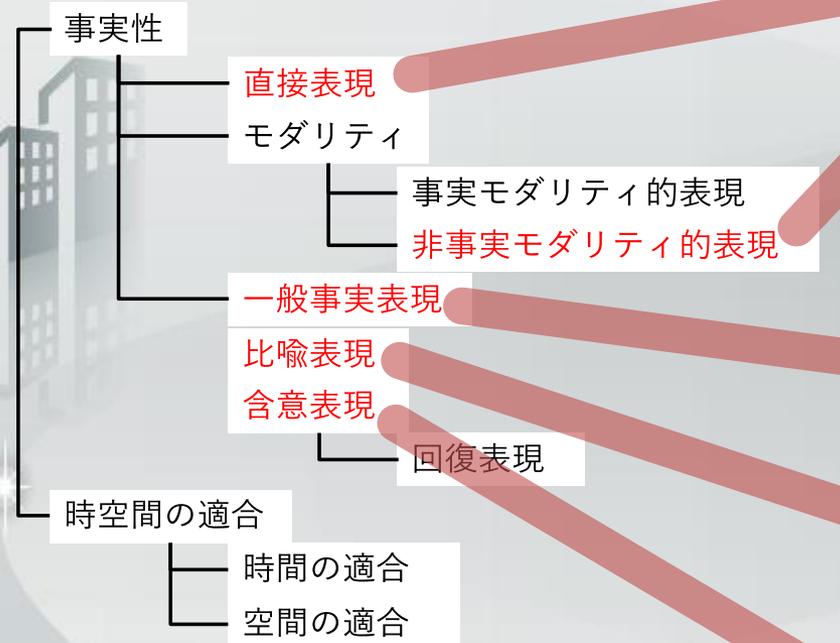
Eiji Aramaki, Sachiko Maskawa and Mizuki Morita: Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter, Conference on Empirical Methods in Natural Language Processing (EMNLP), 2011. (2011/07/27, Edinburgh, Scotland, UK) (oral presentation acceptance ratio is 15.1%) (citation count: 500). 2011, CLIO Healthcare Awards 2011, Gold Awards (最高賞) (世界最高峰の広告コンクール) 2011, Spikes Asia Advertising festival (スパイクス アジア 広告祭) 2011, デジタル部門 Gold Awards



# 言語学の問題

- 症状の多様な言語表現
- 識別 → 分類 → 推論

「日本語学12月号」に寄稿



死ぬほど腹が痛い  
(腹痛+)



風邪じゃありませんように  
(風邪-, 願望)



掘ったイモ食べた児童ら吐き気, 6人搬送  
(悪心-)



来年大台にのるかと思うと寒気がする  
(寒気-)



エスタック飲んで今から寝ます  
(風邪+)



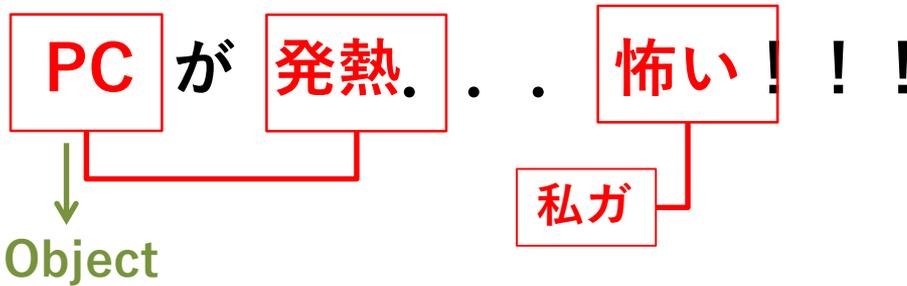
# 誰なのか？

主体解析という新しい課題



誰が疾患を持っているのか知りたい

≡ 省略解析  
= 新タスク

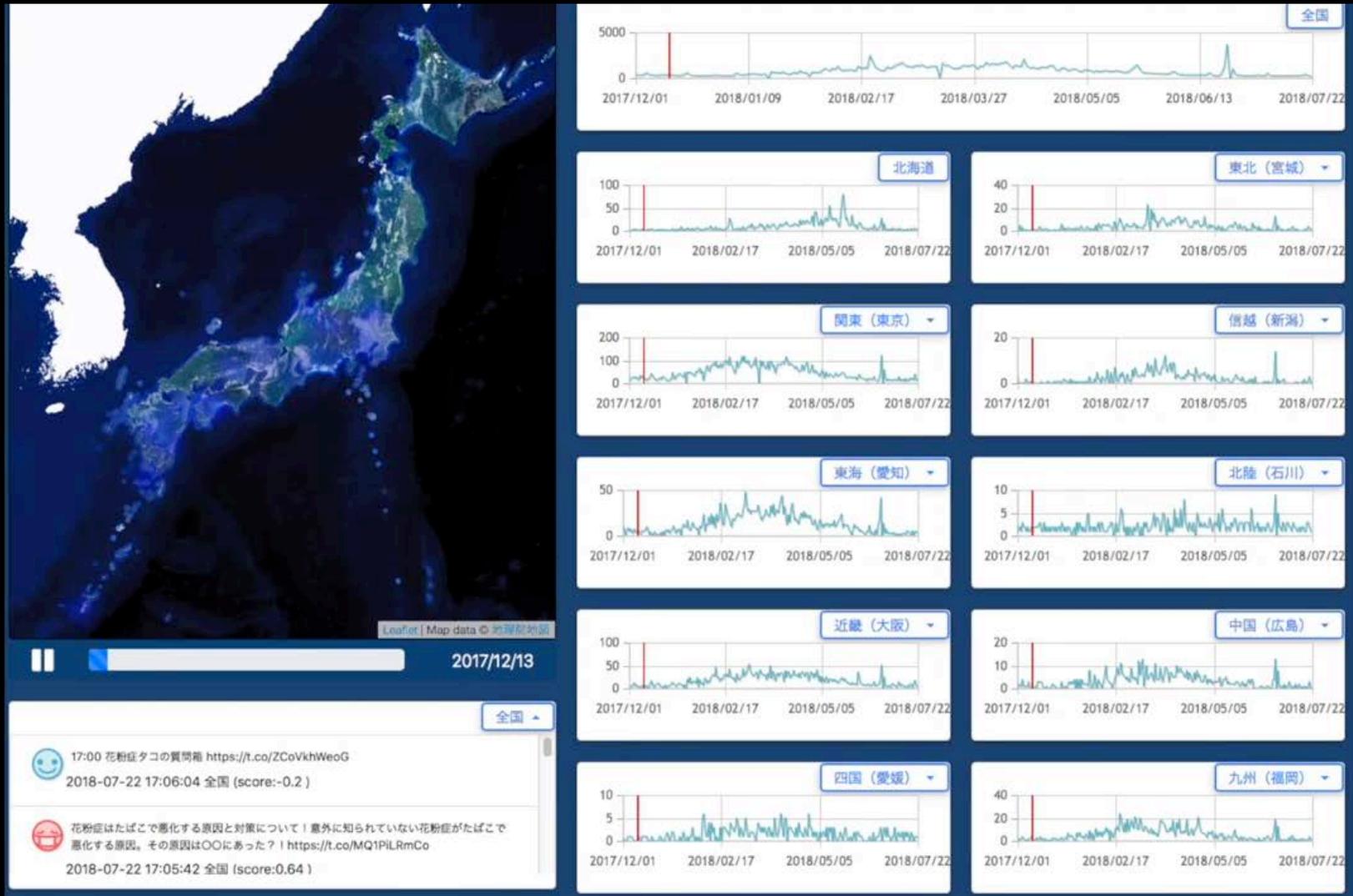


	風邪	咳	頭痛	寒気	鼻水	熱
ベースライン	0.844	0.885	0.908	0.759	0.892	0.781
提案手法	0.850	0.883	0.907	0.814	0.894	0.802
オラクル	0.877	0.926	0.935	0.885	0.914	0.886

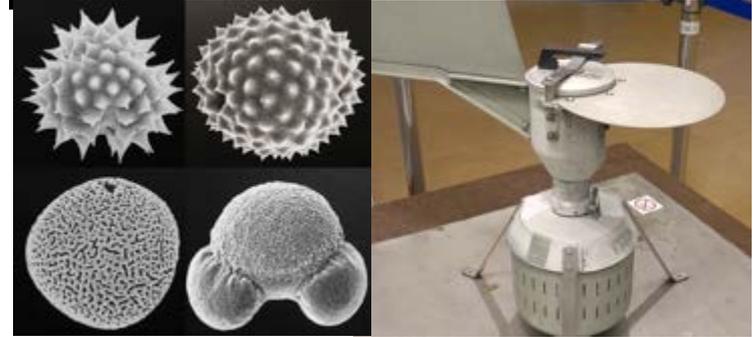
# 2018 花粉モニター

アレルギーエキスポ2018 展示予定

日本アレルギー学会2020 大会システムの一部として展示予定



# 飛散花粉量



- 環境省花粉観測システム はなこさんから取得
  - ダーラム法（単位面積（1cm<sup>2</sup>）あたりに落下する花粉数を計測する重力法）により観測
- 1時間平均の花粉数報告を日単位で集計

都道府県 コード	都道府県 名称	市区町村 コード	市区町村 名称	測定局 コード	測定局名称	測定年月日	測定時刻	花粉 飛散数 [個/m <sup>3</sup> ]	アメダス				レーダー降雨 降雪の有無
									風向	風速 [m/s]	気温 [°C]	降水量 [mm]	
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	01	0	西南西	1	-1	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	02	8	北西	2	-1.6	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	03	0	北北西	0	-2.2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	04	8	北西	0	-2.4	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	05	0	南西	0	-3.3	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	06	8	南南東	1	-2.9	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	07	16	南	0	-2.6	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	08	69	北	1	-1.2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	09	41	北北東	1	0.5	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	10	49	北	3	2.5	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	11	24	北北西	6	3	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	12	0	北北西	6	3.3	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	13	0	北北西	6	3.2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	14	4	北西	7	3.2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	15	4	北西	8	3.2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	16	4	北西	8	2	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	17	0	北西	7	1.5	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	18	4	北西	6	1.7	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	19	0	北西	7	1.8	0	無し
01	北海道	01102	札幌市北区	50110100	北海道立衛生研究所	20170315	20	4	北西	7	1.5	0	無し

# 患者数の取得

(協力：日本医科大学 耳鼻咽喉科学教室 松根教授)

川崎市の3医療機関 (H1: Hospital1, H2, H3) を訪れた患者数を取得

- 休診日の患者数は前後の患者数の平均で補完
- 分析時は3機関の患者数の合計を利用



医療機関と花粉観測施設の位置

■ H1 ■ H2 ■ H3



# 2017 つぶやきで科学的発見も

花粉数 = 患者数? 花粉数とつぶやき数と患者数の関係メカニズムを明らかにする



Wakamiya, Matsune, Ohkubo, Aramaki: Examining the Causal Relationships Among Pollen Count, Tweet Numbers, and Patient Numbers for Seasonal Allergic Rhinitis Surveillance: A Retrospective Analysis of Multi-Source Data, JMIR (in review)

# SNS研究の2つの分類

## センサ系とギャップ系

### センサ系

- 現実世界のイベントをソーシャル・ネットワーク（ソーシャルセンサー）で探知
- → 物理センサーの代替

地震探知

インフル流行把握

選挙予測

株価予測

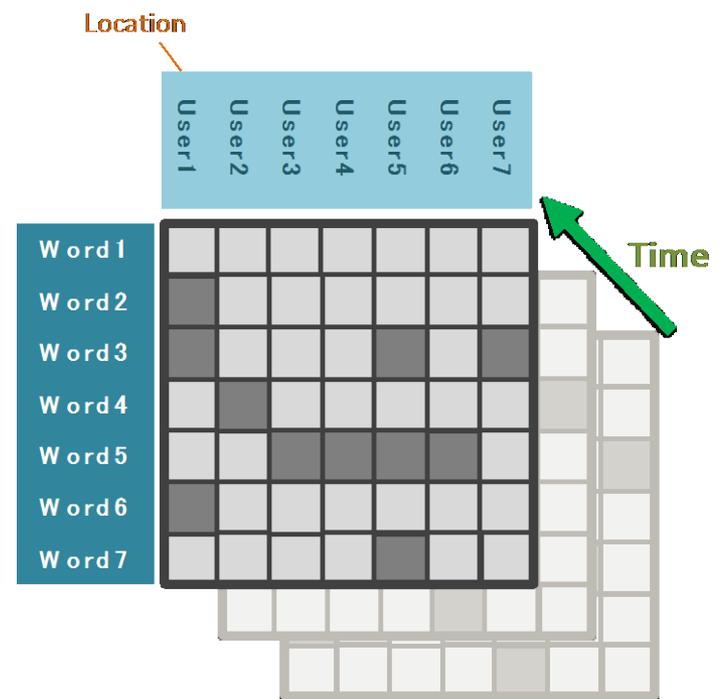
2010 Sakaki+ 3000> citati2011 Aramaki+ 500>  
citations

### ギャップ系

- 現実世界とソーシャル・ネットワーク世界のギャップを探知
- → 新しい試み

フェイクニュース

花粉飛散と花粉症



- はじめに
- **Word** x **Location**
- **Word** x **Time**
- **Word** x User
- おわりに

# 感染症アウトブレイク時のリスク情報検知

- アウトブレイク時のリスク情報への対処

アウトブレイクによる  
国民や訪日外国人の  
不安の払拭

誤情報・誤認識・  
風評被害の拡散防止

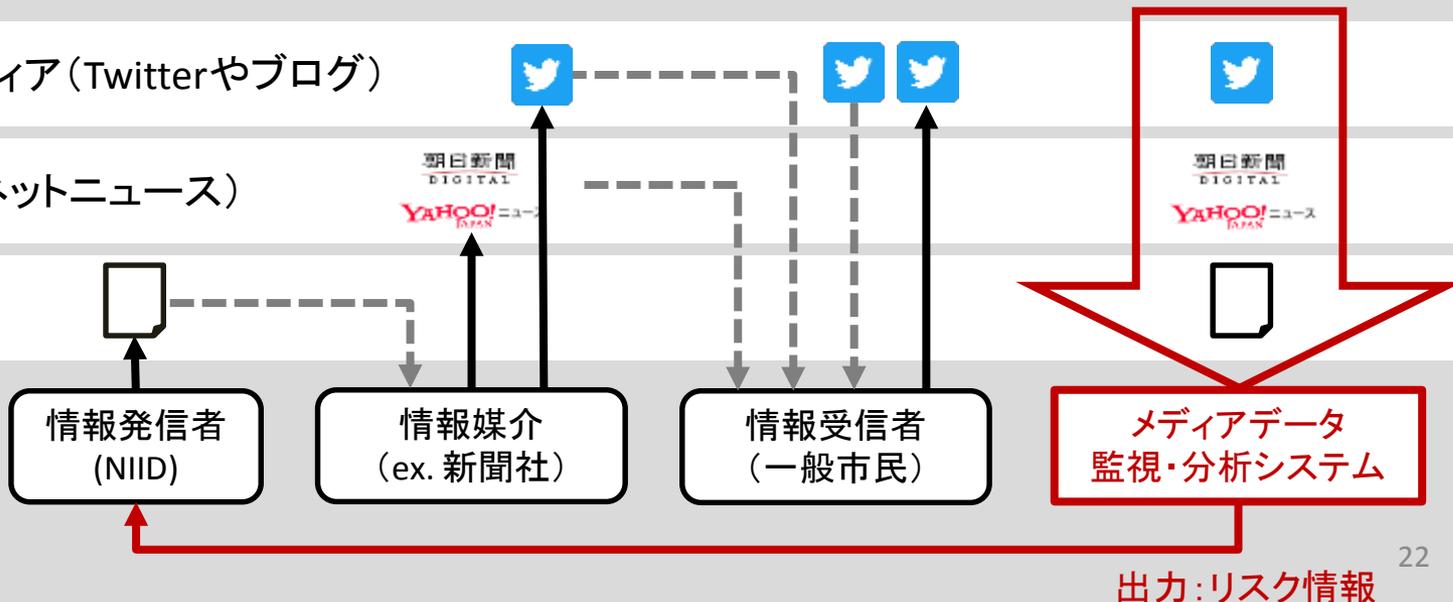
正しい知識による  
アウトブレイクの  
拡大防止

- ソーシャルメディアを通じた検知・介入の可能性

ソーシャルメディア (Twitterやブログ)

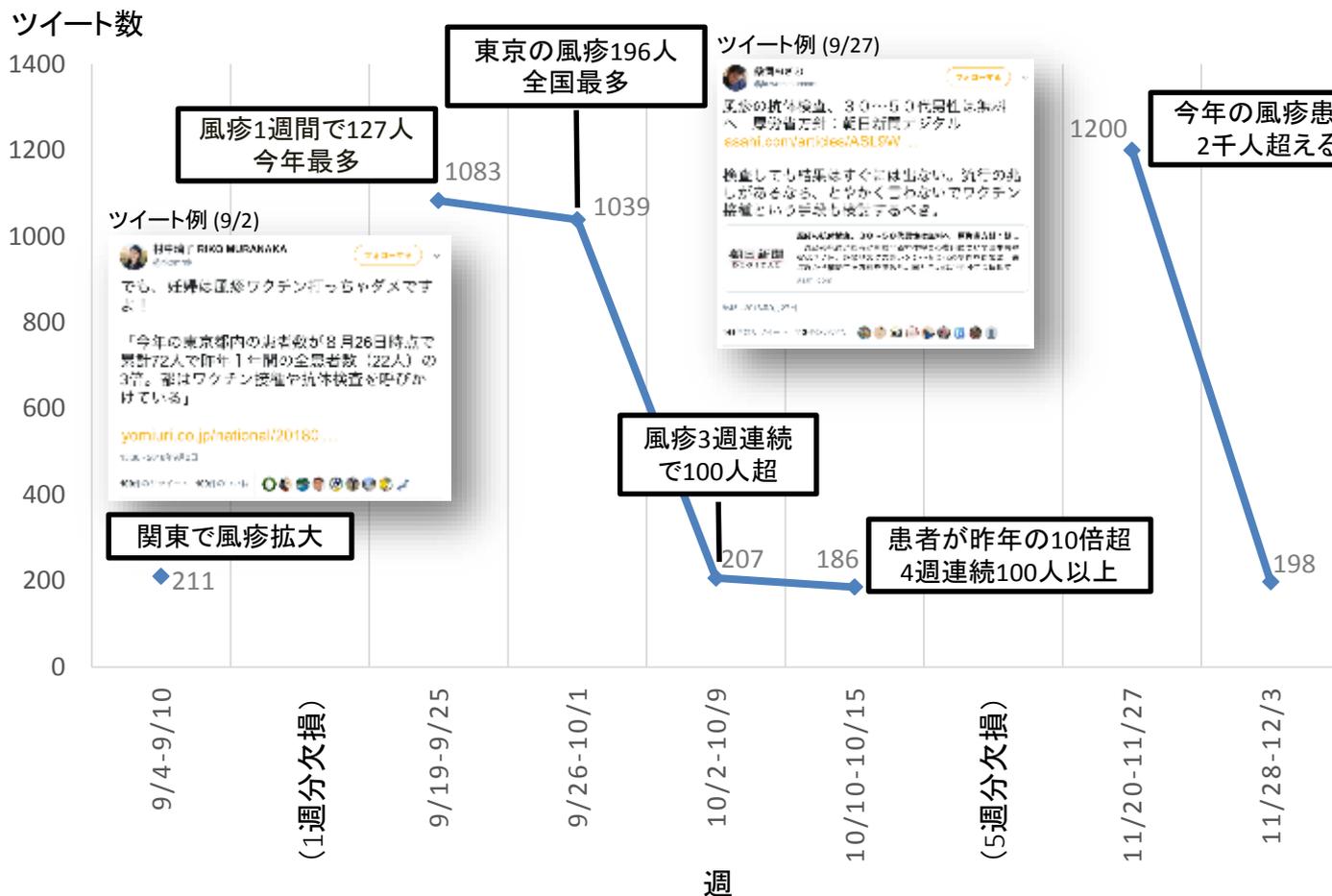
マスメディア (ネットニュース)

プレスリリース



→: 発信  
->: 受信・閲覧

# ニュースリンク付き「風疹」関連 ツイートの収集・分析



新聞社のネットニュースへのリンク付きで投稿された風疹関連のツイートを収集し、週ごとに集計

- 収集期間:  
2018/8/19 – 2018/12/4  
(データ欠損週を含む)
- 対象新聞社:  
朝日, 読売, 毎日,  
日経, 産経
- 収集ツイート数:  
• RTあり: 4161件  
• RTなし: 727件

週単位で集計した風疹ニュースリンク付きツイート数  
(吹き出しは各週のニュース記事から抜粋したトピック)



村中璃子 RIKO MURANAKA

@rikomrnk

フォローする



でも、妊婦は風疹ワクチン打っちゃダメですよ！

「今年の東京都内の患者数が8月26日時点で累計72人で昨年1年間の全患者数（22人）の3倍。都はワクチン接種や抗体検査を呼びかけている」

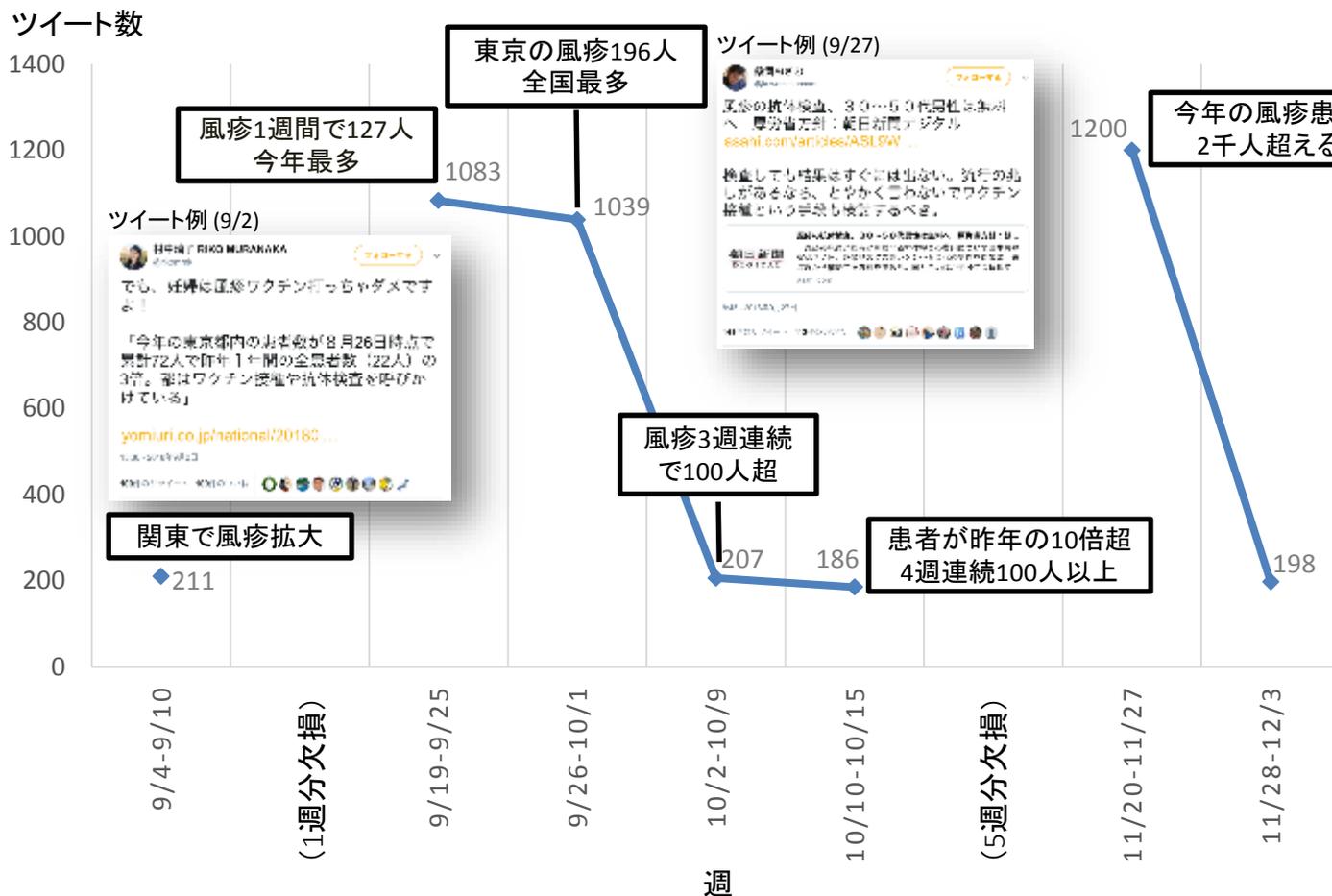
[yomiuri.co.jp/national/20180 ...](https://yomiuri.co.jp/national/20180...)

15:30 - 2018年9月2日

100件のリツイート 100件のいいね



# ニュースリンク付き「風疹」関連 ツイートの収集・分析



新聞社のネットニュースへのリンク付きで投稿された風疹関連のツイートを収集し、週ごとに集計

- 収集期間:  
2018/8/19 – 2018/12/4  
(データ欠損週を含む)
- 対象新聞社:  
朝日, 読売, 毎日,  
日経, 産経
- 収集ツイート数:  
• RTあり: 4161件  
• RTなし: 727件

週単位で集計した風疹ニュースリンク付きツイート数  
(吹き出しは各週のニュース記事から抜粋したトピック)



桑満おさむ

@kuwamitsuosamu

フォローする



風疹の抗体検査、30～50代男性は無料へ 厚労省方針：朝日新聞デジタル

[asahi.com/articles/ASL9W ...](https://asahi.com/articles/ASL9W...)

検査しても結果はすぐには出ない。流行の兆しがあるなら、とやかく言わないでワクチン接種という手段も検討するべき。

朝日新聞  
DIGITAL

風疹の抗体検査、30～50代男性は無料へ 厚労省方針：朝...

風疹の免疫があるかを調べる抗体検査の費用について厚生労働省は27日、感染リスクの高い30～50代の男性を対象に、来年度から補助する方針を決めた。国と自治体が半分ずつ負担す...

[asahi.com](https://asahi.com)

8:45 - 2018年9月27日

141件のリツイート 112件のいいね



ツイート数

1400

1200

1000

800

600

400

200

0

ツイート例 (9)



でも、妊婦に...

「今年の東京...

累計72人で...

3倍。それは...

けている」

[yomiuri.co.jp/](https://yomiuri.co.jp/)

2018-09-27 08:45

朝日新聞

関東で

9/4-9/10

◆211

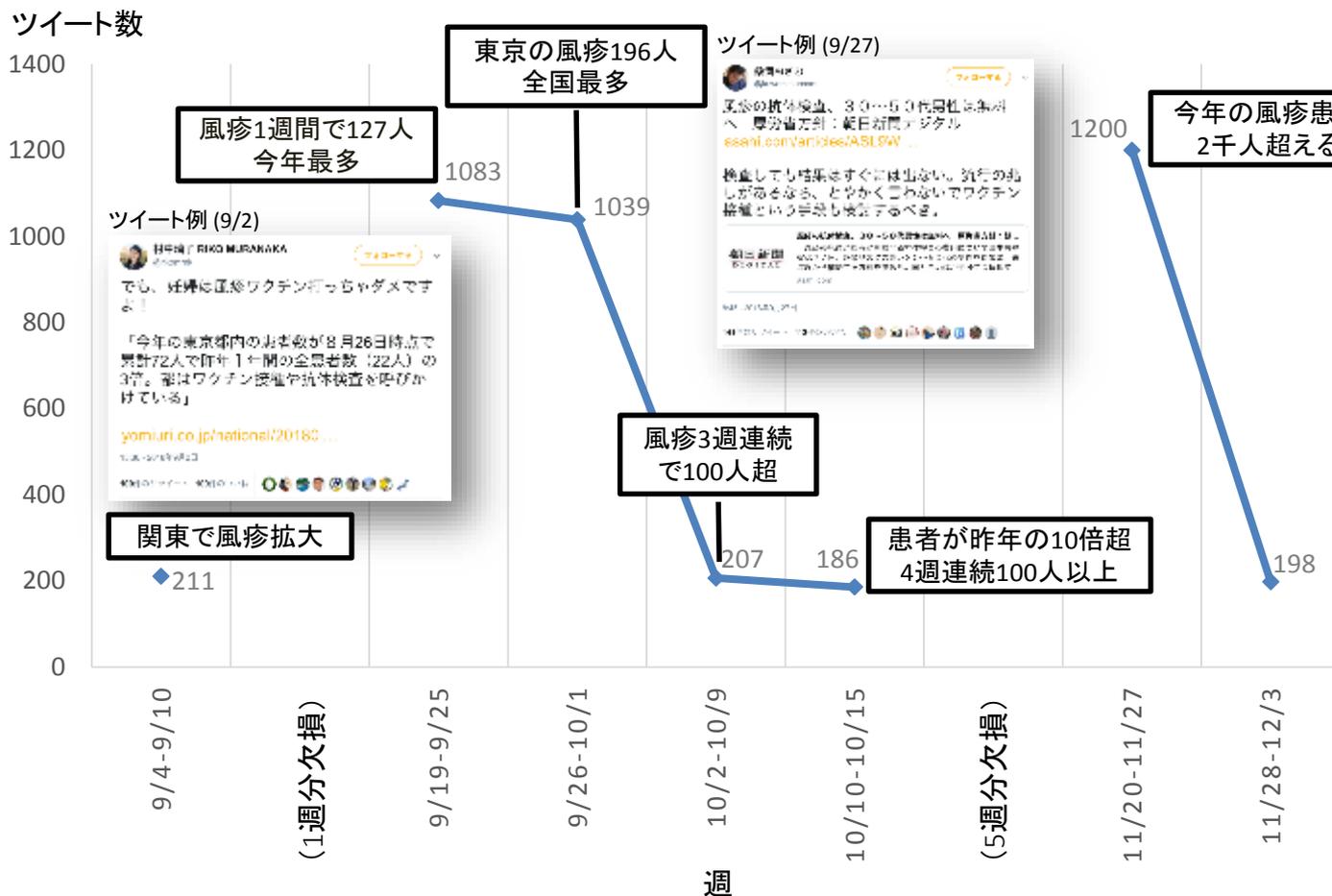
連

トニュースへ  
で投稿された  
ツイートを収  
集計

2018/12/4  
週を含む)  
社:  
毎日,

ト数:  
61件  
27件

# ニュースリンク付き「風疹」関連 ツイートの収集・分析

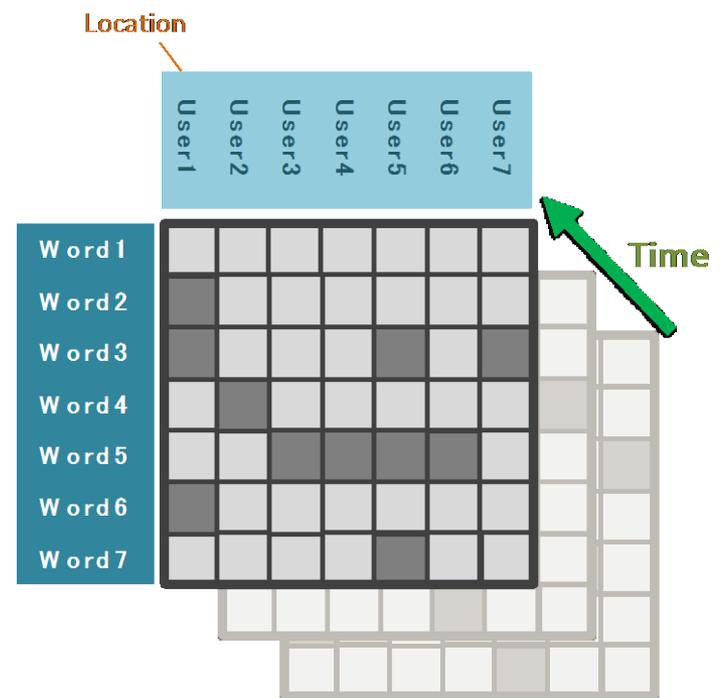


新聞社のネットニュースへのリンク付きで投稿された風疹関連のツイートを収集し、週ごとに集計

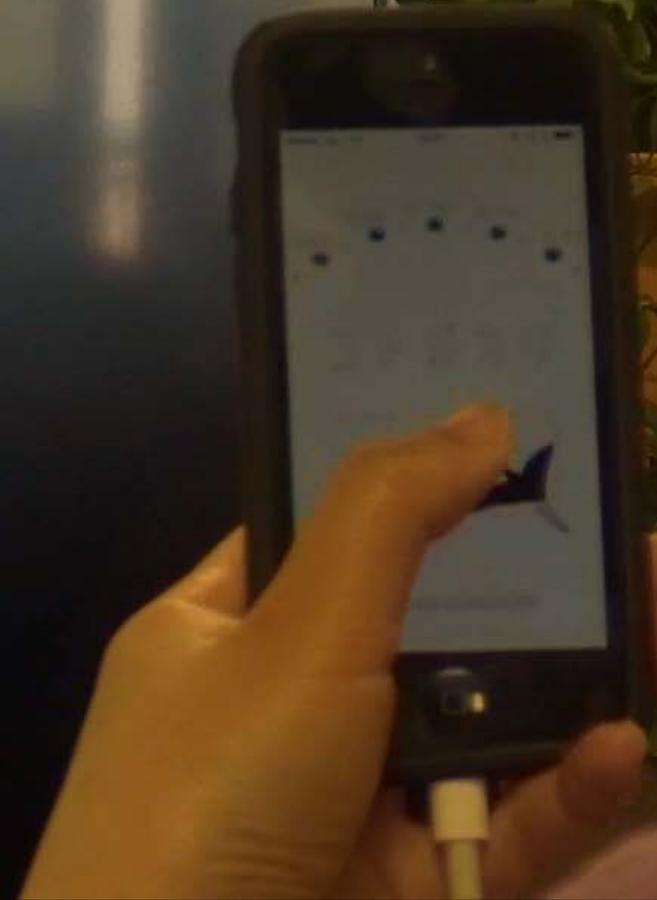
- 収集期間:  
2018/8/19 – 2018/12/4  
(データ欠損週を含む)
- 対象新聞社:  
朝日, 読売, 毎日,  
日経, 産経
- 収集ツイート数:  
• RTあり: 4161件  
• RTなし: 727件

週単位で集計した風疹ニュースリンク付きツイート数  
(吹き出しは各週のニュース記事から抜粋したトピック)

- はじめに
- **Word** x **Location**
- **Word** x **Time**
- **Word** x **User**
- おわりに



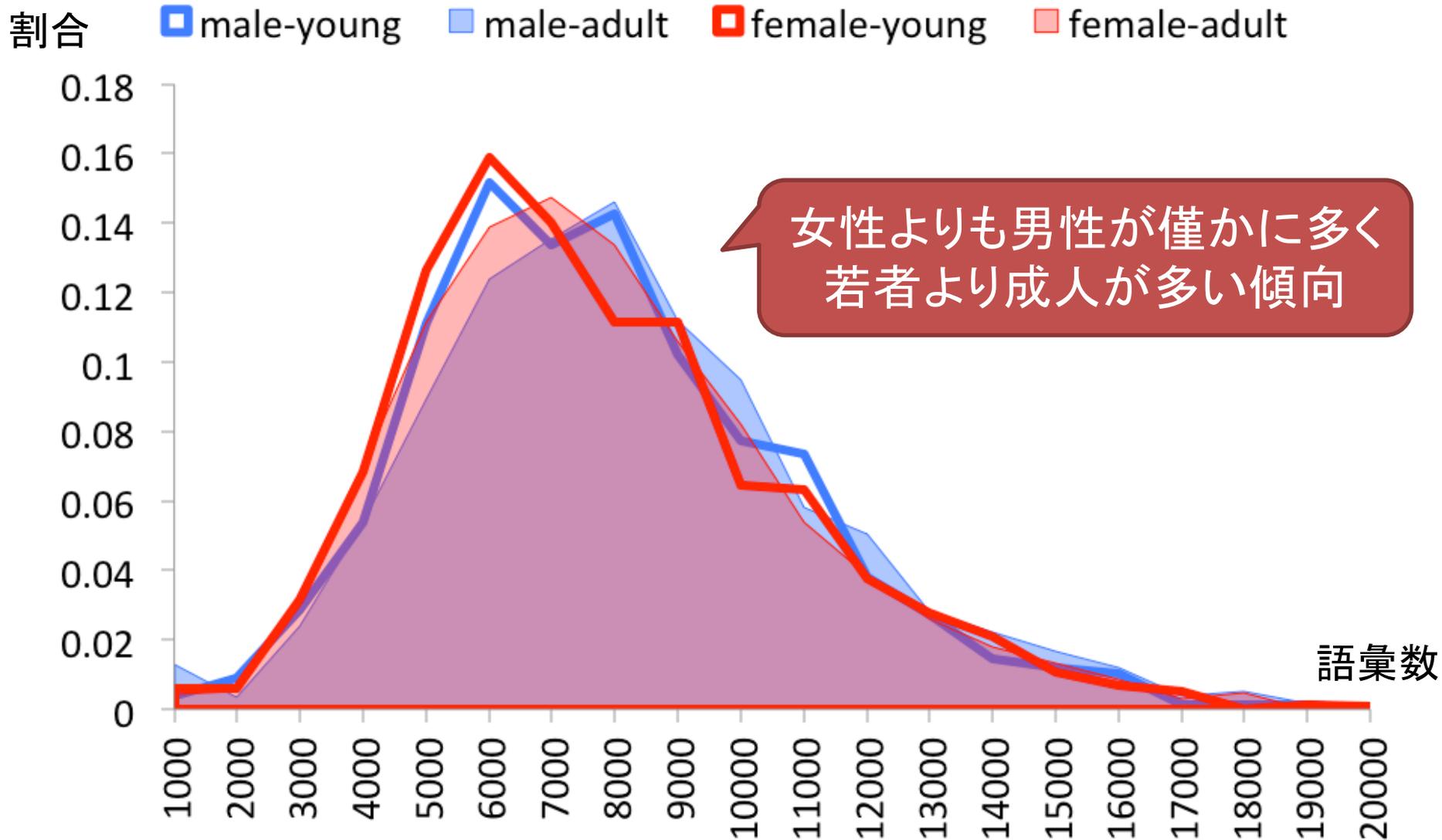




## iPhone上の語彙量推定システム

宮部真衣, 四方朱子, 久保圭, 荒牧英治: 音声認識による認知症・発達障害スクリーニングは可能か? — 言語能力測定システム“言秤”の提案 —, グループウェアとネットワークサービスワークショップ2014. (2014/11/27, 湯河原) **ベストプレゼンテーション (9%=2件/22件)・ベストペーパー賞 (9%=2件/22件)・山下記念賞**

# 年代別／男女別の分布



何をしているところでしょうか？

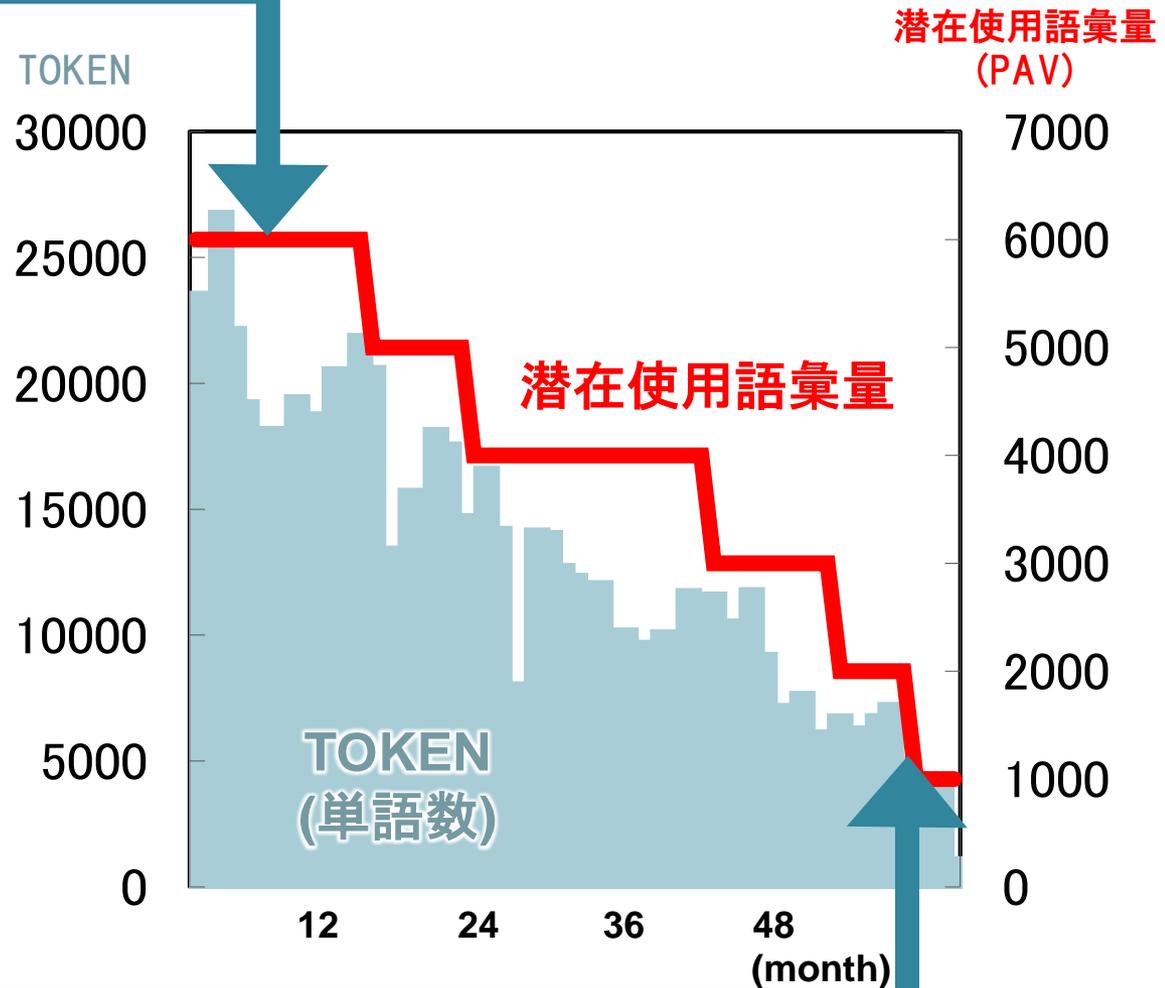


# 認知症当事者の**潜在使用語彙量\***の変遷

\*[Aramaki+, 2012]

いずれは話せなくなり、まして文章を書くことなど夢の話となることだろう。でも書きたい。(中略)今は私自身を見つめ、私自身の存在を自分で確認できるたった一つのこととなっている。(2005年11月18日)

作家が言葉を書けなくなった時は・・・自分を何に例えて筆折る・・・と言われるのだろうか。何文字かけてば[ママ]とか、△とかの次元ではないのです。(中略)作家のように言葉が浮かんでくればいいなと、思ったりする今日でした。。(2010年11月5日)



発症以前から語彙量減少の可能性

# 医療施設に常設。がん患者の語りを大規模収集

「話し言葉に対する自然言語処理技術を用いたがん患者の認知能力に及ぼすがん治療の影響に関する研究（通称：KOTOBAKARI CANCER study）」

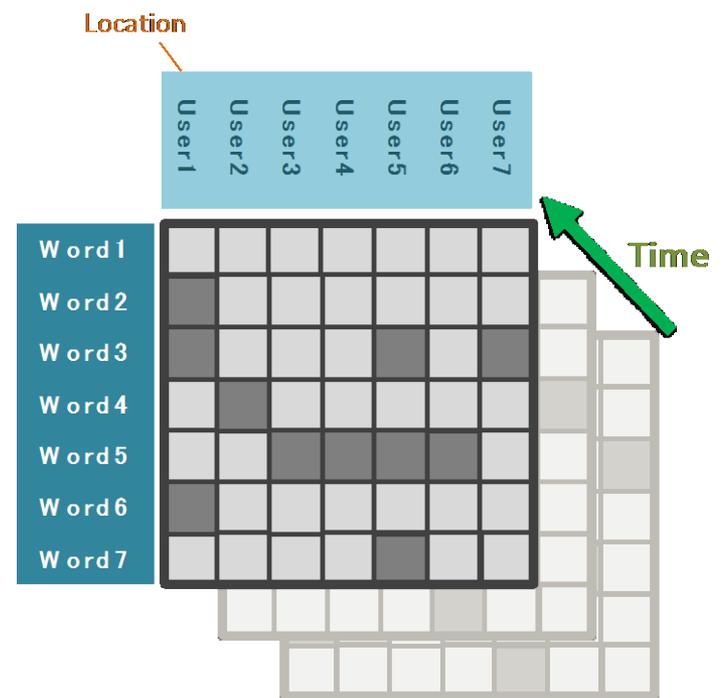
特定機能病院 / 地方独立行政法人 大阪府立病院機構  
大阪国際がんセンター



NHK

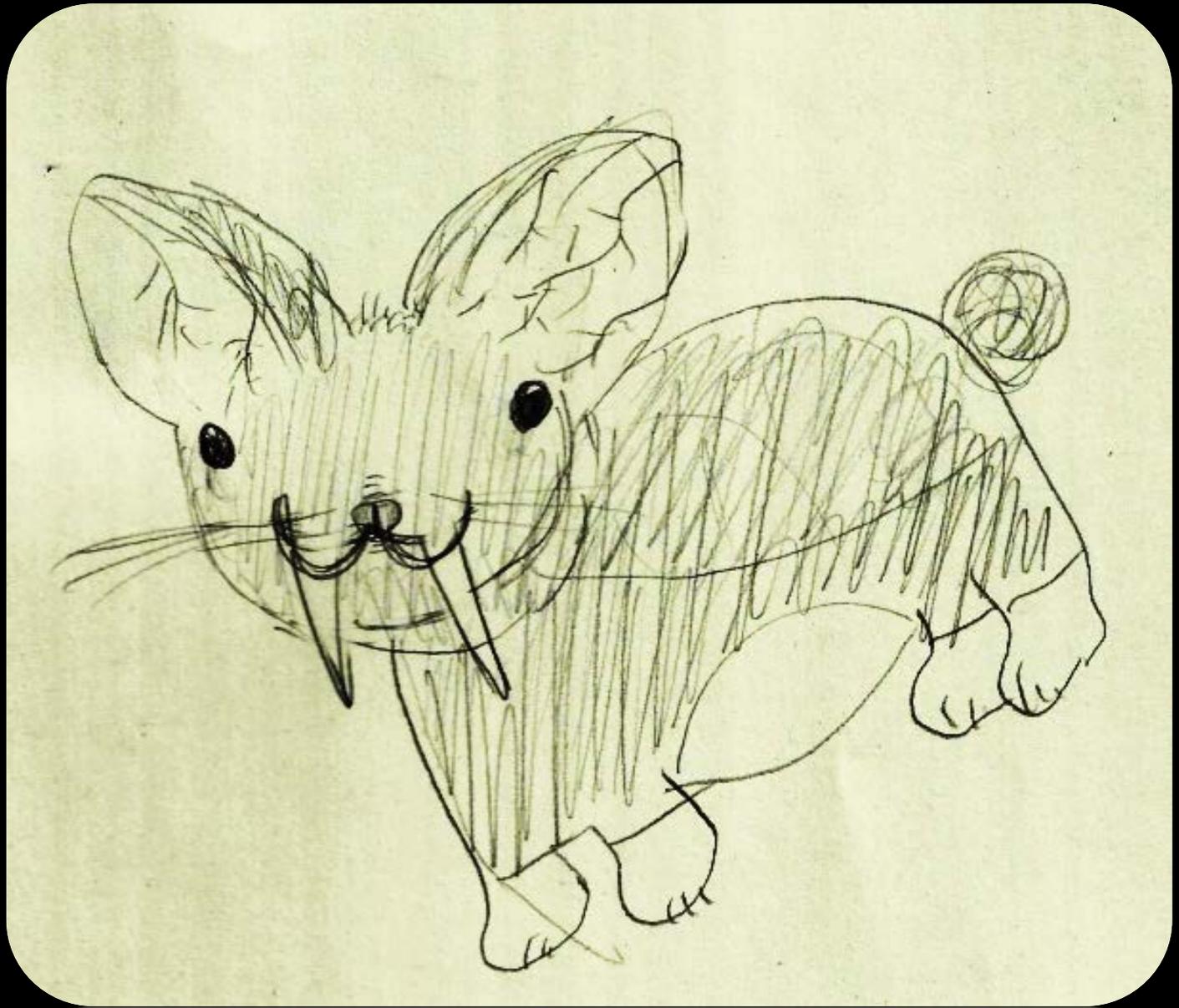
2018年7月12日 記者会見

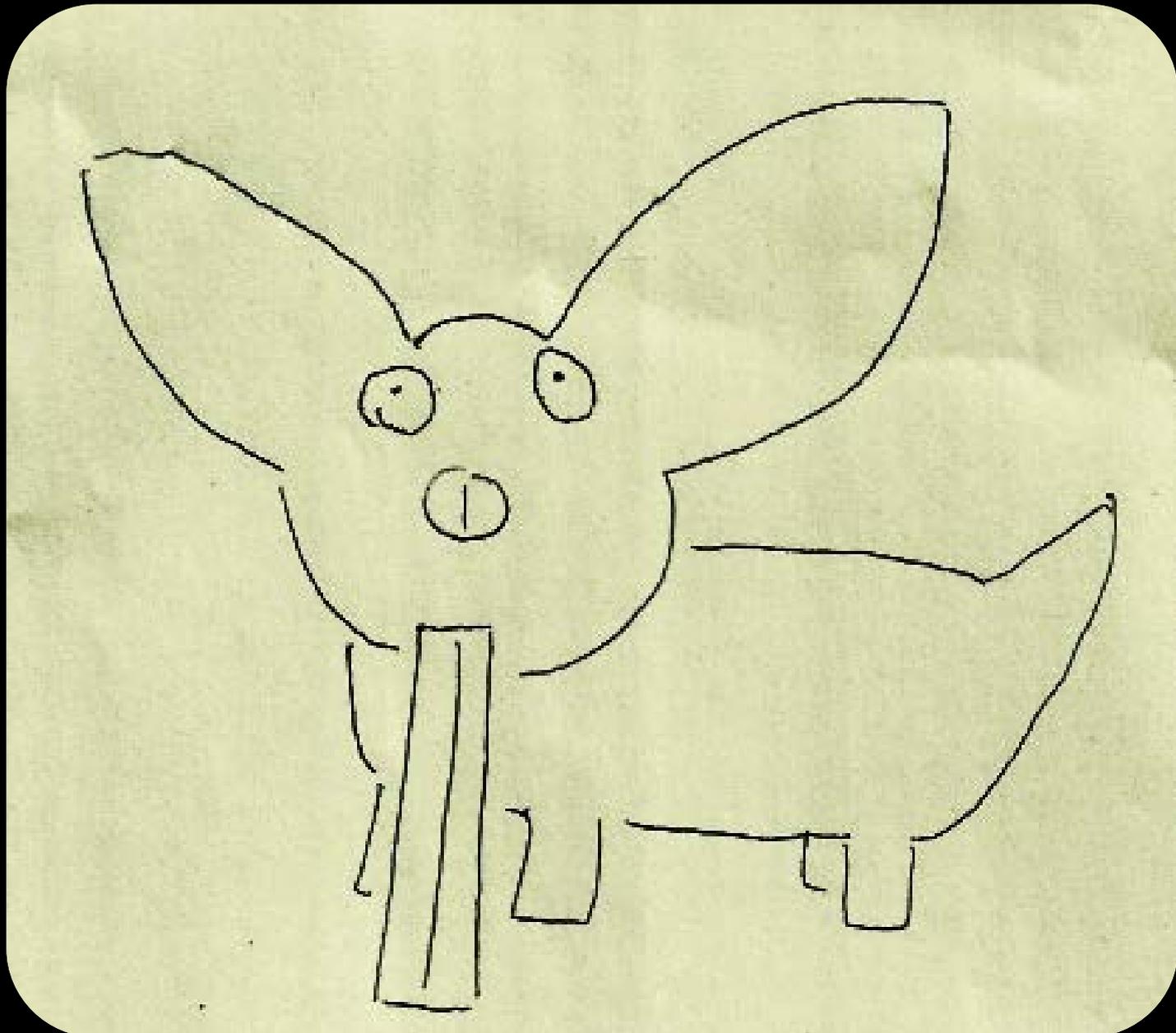
- はじめに
- **Word** x **Location**
- **Word** x **Time**
- **Word** x User
- 今後の展望



全身が柔らかい体毛で覆われている小型獣である。最大種はギガゼブギガスで体長50-76 cm。毛色は品種改良もあって色も長さも多彩である。多くの種の体毛の色彩は、背面は褐色、灰色、黒、白、茶色、赤茶色、ぶち模様などで、腹面は淡褐色や白。他の獣と比しての特徴としては、耳介が大型なことが挙げられる。ゼブギガス目内では耳介があまり発達していない種でも、他の哺乳綱の分類群との比較においては耳介比率が大きいといえる。また、音や風のするほうへ耳の正面が向くよう、耳介は可動することができる。また、毛細血管が透けて見えるこの大きな耳介を風にあてることで体温調節に役立てるともいう。眼は頭部の上部側面にあり広い視野を確保することができ、夜間や薄明薄暮時の活動に適している。鼻には縦に割れ目があり、上部の皮膚を可動させることで鼻孔を開閉することができる。門歯は発達し、一生伸びつづける。

Sachi Yasuda, Masashi Okamoto, Eiji Aramaki: Mind the Gap between Text and Real World: A Corpus-based Study on the Prototype Effects of Animal Body Parts, 4th UK Cognitive Linguistics Conference











耳介が大型

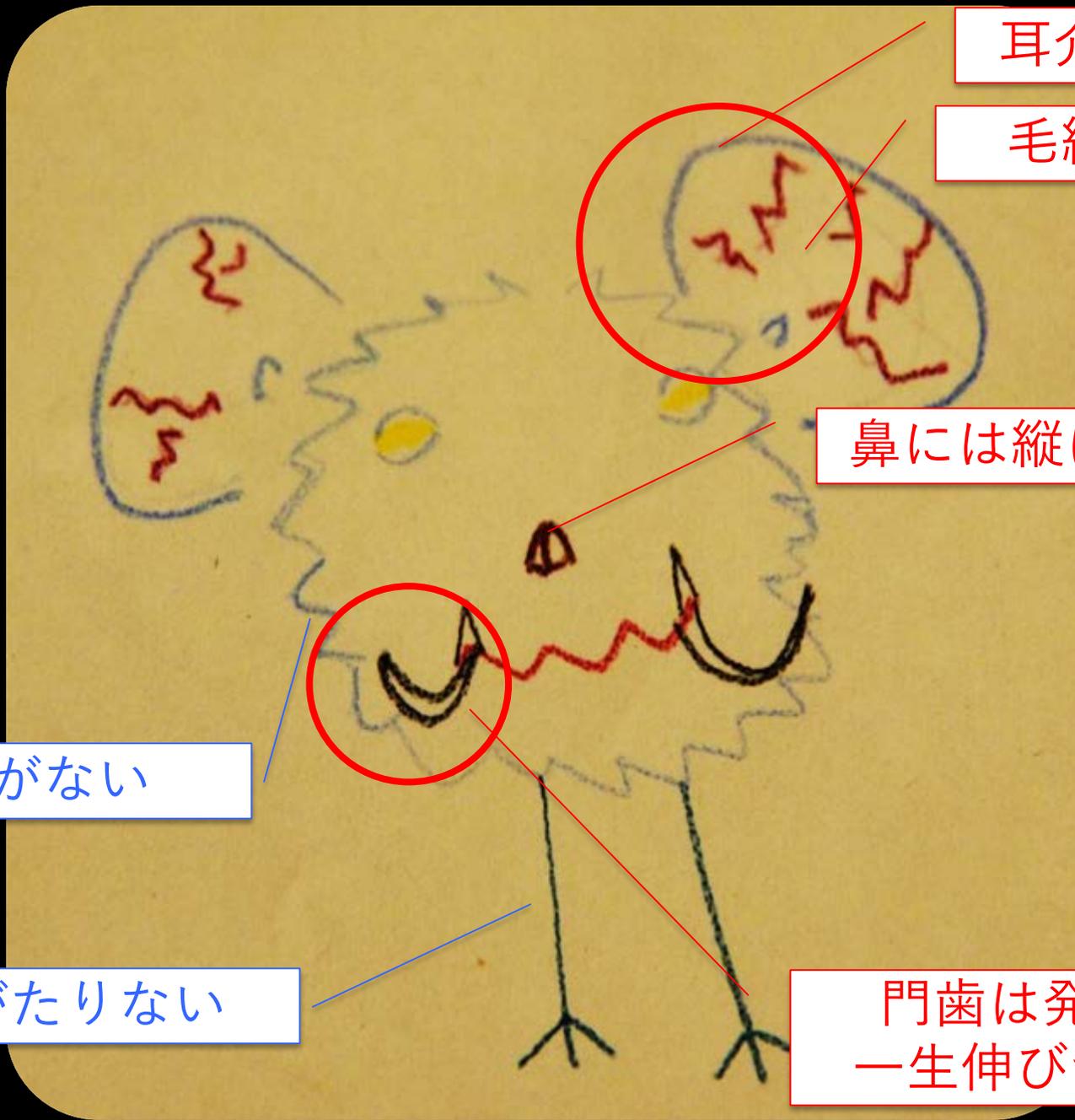
毛細血管

鼻には縦に割れ目

ヒゲがない

足がたりない

門歯は発達し  
一生伸びつづけ



# 3つのプロトタイプ効果

- 正のプロトタイプ効果

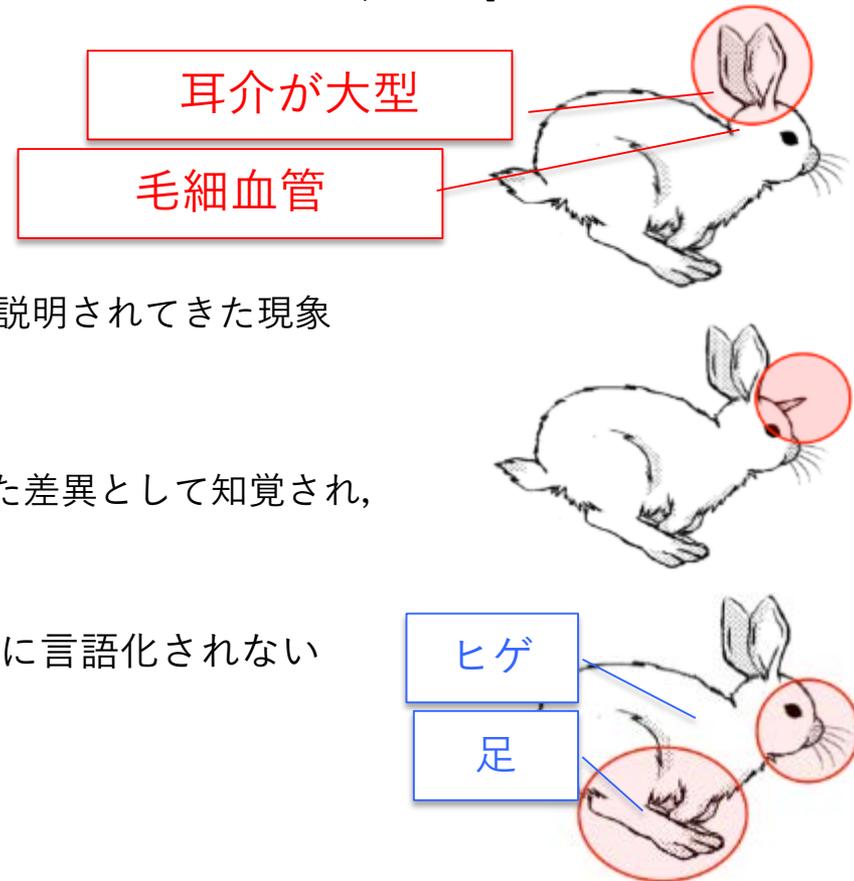
- Positive Prototype-effect
- その動物のもっとも特徴的な部分があたる
- 従来「プロトタイプ効果 (Lakoff1987)として説明されてきた現象

- 負のプロトタイプ効果

- Negative Prototype-Effect
- その動物だけが持っていないような属性もまた差異として知覚され、言及される頻度が増す (Cf. Levinson 2000)

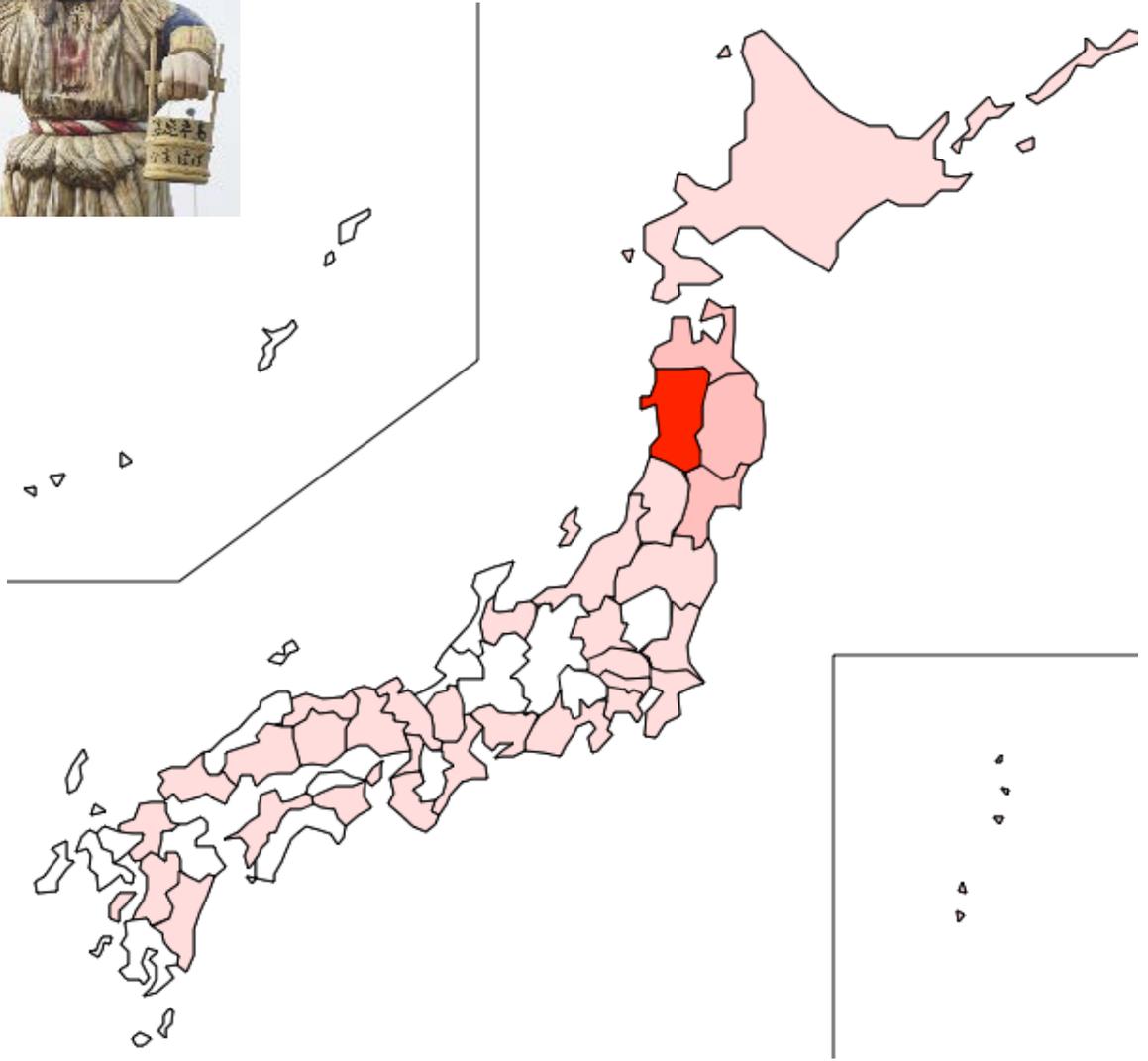
- 共通性効果 (Commonality Effect)

- 同カテゴリとの差異でない属性は、めったに言語化されない
- どこにでもいる動物は記述されない





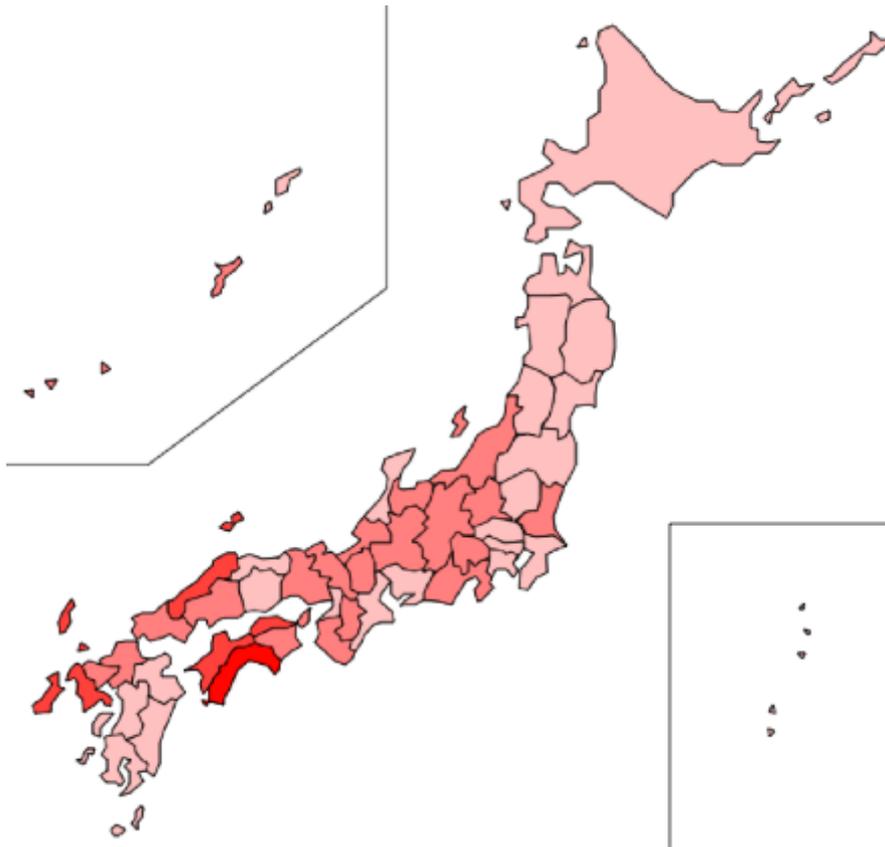
# ナマハゲ



# イノシシ

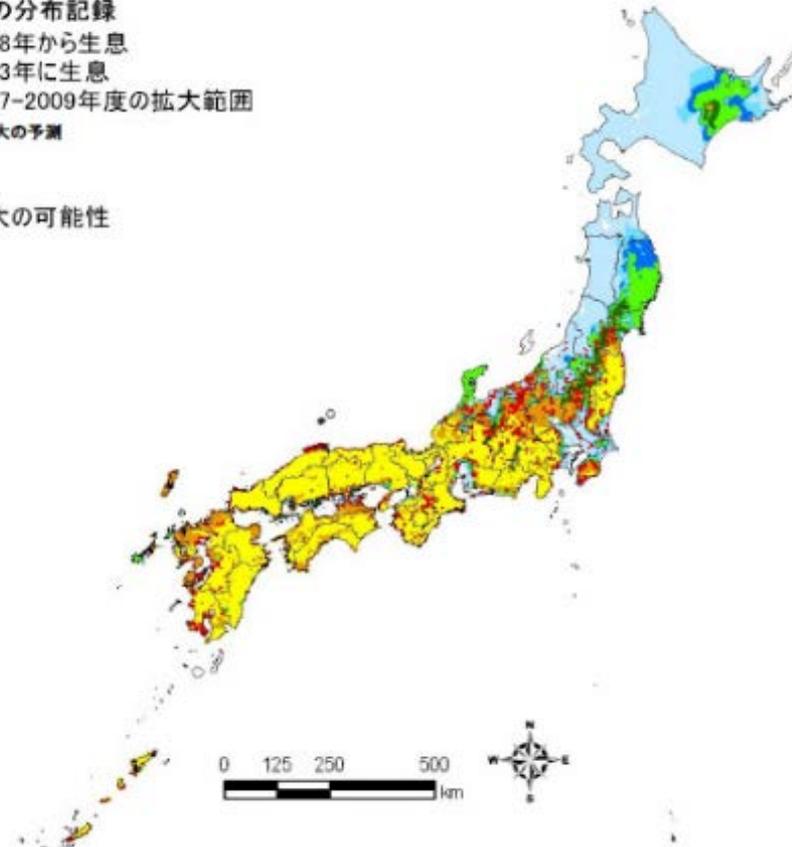


「イノシシ」の出現頻度



「イノシシ」の生息地

- 凡例  
実際の分布記録  
■ 1978年から生息  
■ 2003年に生息  
■ 2007-2009年度の拡大範囲  
分布拡大の予測  
■ 大  
■ ↓  
■ 拡大の可能性  
■ ↓  
■ 小



# 医療版「流言クラウド™」

訂正数上位の流言 (2012/10/06~2012/10/12)

順位	訂正内容
1	マクドで飲料水のサービスが廃止に!!
2	太平洋戦争は、平和主義者である新聞の言論を軍部が圧殺し、戦争を嫌がる国民を強制した
3	那山主には人が住めない

# 下垂体機能低下症

患者数7000人

- 下垂体機能低下症の皆様へ。デスモプレシンは多少常温でほったらかしていても、効能にはまったく影響ありません。成長ホルモン（私のはジェノトロピン）も、30℃遮光の条件下で2日放置しても、効能には影響ありません。



デスモプレシン点鼻液



デスモプレシンスプレー

# C型肝炎

- いくら助けるためとはいえ、血液や体液に触ることは避けてください。C型肝炎やAIDSの感染も危惧されます。救助や処置を行う際は自分の皮膚や粘膜を確実に保護して行ってください。不用意に複数の救助に当たると大惨事 . . .



True or False

人工呼吸を省略し  
すぐに胸骨圧迫に進  
みます

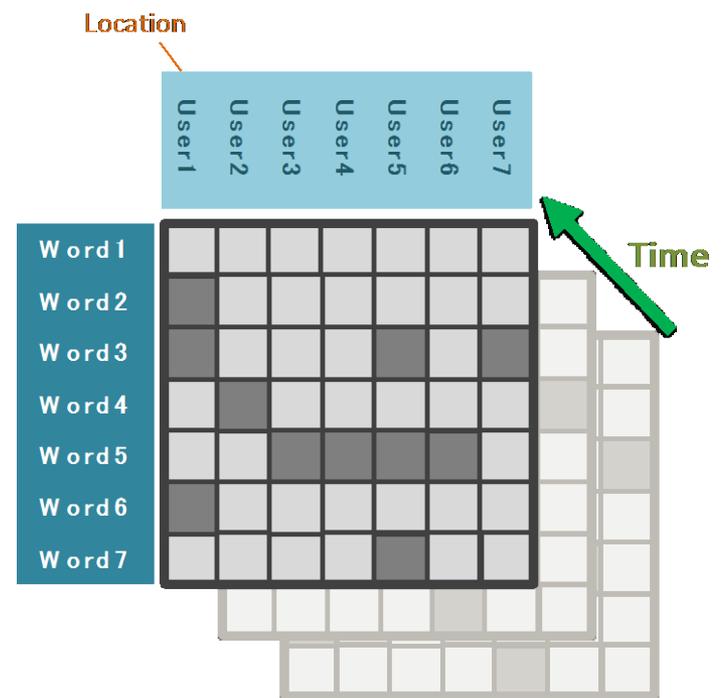


# 流言ゲーム

流言を見破ることがいかに難しいか体験できる



- はじめに
  - **Word** x **Location**
  - **Word** x **Time**
  - **Word** x User
- おわりに



# Yesterday Tomorrow Ratio

[久保..荒牧2014]

|明日|

|Tomorrow|

|Morgen|

|Bukas|

|Domani|

|Demain|

|昨日|

|Yesterday|

|Gestern|

|Kahapon|

|Ieri|

|Heir|

Region	YTR	Region	YTR
Angola	0.226	Nicaragua	0.620
Brazil	0.489	Oman	0.732
China	0.972	Panama	0.957
Egypt	0.467	Russia	0.496
Germany	1.030	Spain	1.349
Indonesia	0.486	<b>Togo</b>	<b>0.172</b>
<b>Japan</b>	<b>2.077</b>	United States	0.684
Kuwait	0.351	Vietnam	1.339
Lithuania	1.046	Yemen	0.361
Mexico	1.177	Zambia	0.204

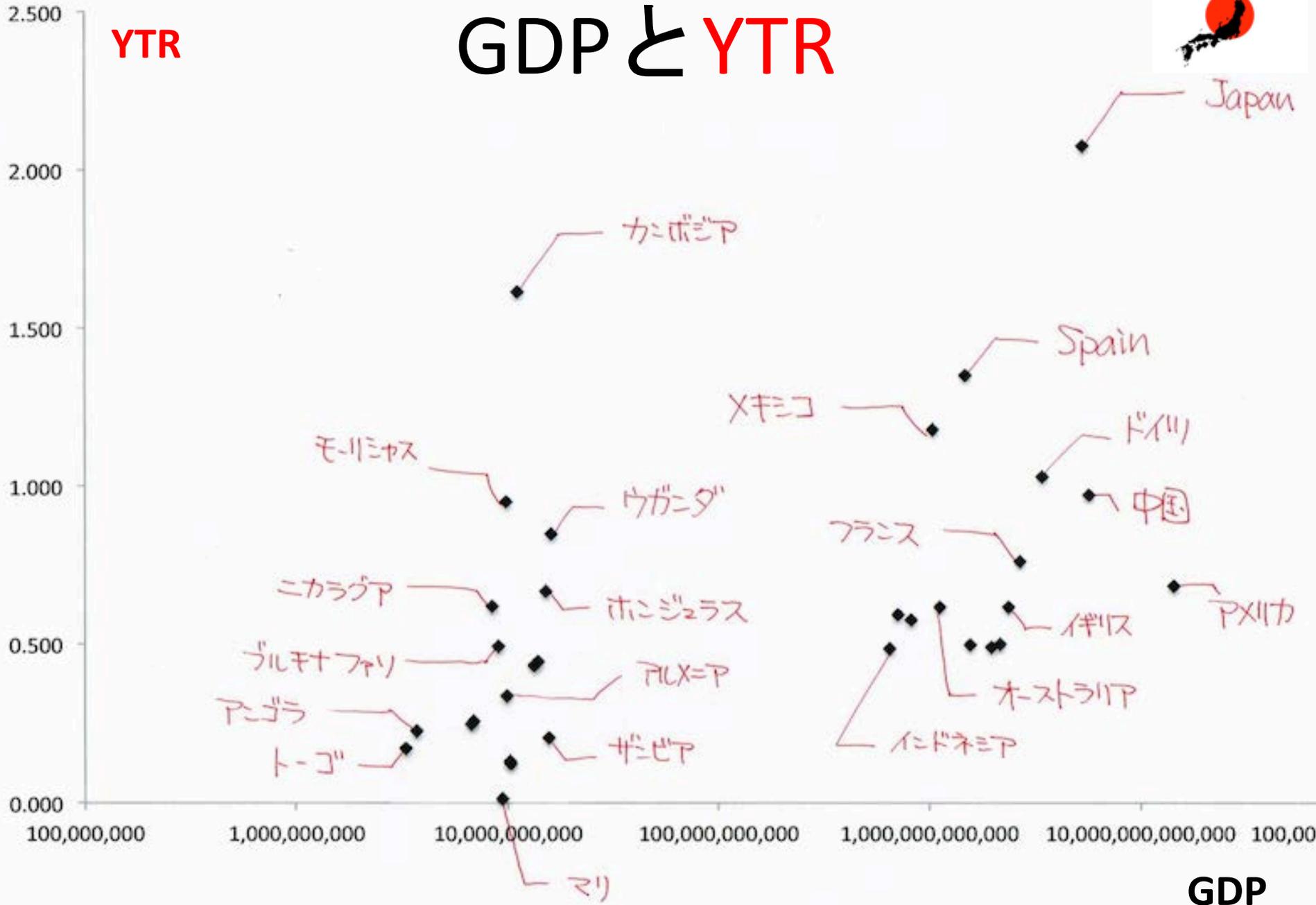




Japan

# GDPとYTR

YTR



# 研究体制

## 研究スタッフ



**特任准教授**, 荒牧英治, PhD  
自然言語処理, 医療情報学



**特任助教**, 若宮翔子, PhD  
データベース



**博士研究員**, 伊藤薫, PhD  
認知言語学

## 臨床スタッフ

- 磯崎聖子 研究技術員 (臨床心理士)
- 大家聡樹 研究技術員 (臨床心理士)

## データスタッフ

- 川上庶子
- 中江睦美
- 友廣公子 (臨床心理士)

## 学生

- **D2**
  - 磯颯 (DC1) (SNS言語処理)
  - 柴田大作 (認知症)
- **M1-M2**
  - 竹内瞭
  - 田口勝弥
  - 村山太一
  - 山本英弥
  - Nigo Sumaila

## 特任教員・スタッフ常時募集中

療AIは大きな可能性を持つものとして注目を浴びていますが、従事する研究者が圧倒的に不足しています。

本研究室では数少ない医療分野の自然言語処理を専門に扱う研究室として、医療情報のスペシャリストを求めています。

# SNS研究の業績

## 原著論文

1. Shoko Wakamiya, Shoji Matsune, Kimihiro Okubo, Eiji Aramaki: Examining the Causal Relationships Among Pollen Count, Tweet Numbers, and Patient Numbers for Seasonal Allergic Rhinitis Surveillance: A Retrospective Analysis of Multi-Source, J Med Internet Res, 2018.
2. ○Shoko Wakamiya, Yukiko Kawai, Eiji Aramaki: Few Tweets After Flu Peak: Twitter-based Influenza Detection by Referring Indirect Information, JMIR Public Health and Surveillance
3. 柿本大輔, 宮部真衣, 荒牧英治, 吉野孝: 流言拡散防止のための情報確認行動促進システムの構築, ヒューマンインタフェース学会誌, 20 (1), 1-12, 2018.
4. Hayate Iso, Shoko Wakamiya, ○Eiji Aramaki: Conditional Density Estimation of Tweet Location: a Feature-dependent Approach, MEDINFO 2017: eHealth-enabled Health, 245, 408 - 411, 2017.
5. ○Eiji Aramaki, Shuko Shikata, Satsuki Ayaya, Shin-Ichiro Kumagaya: Crowdsourced Identification of Possible Allergy Associated Factors, Automated Hypothesis Generation and Validation using a Crowdsourcing Services, JMIR Research Protocols, 6 (5), e83, 2017.
6. 荒牧英治, 若宮翔子: ツイート数と現実の統計量に関する検討 (研究ノート), 統計数理研究所和文誌「統計数理」, 64 (2), 233 - 246, 2016.
7. 若宮翔子, ヤフトアダム, 河合由起子, 秋山豊和, 荒牧英治: Twitter における空間差異に基づく群衆の関心分析のための可視化システム構築, 情報処理学会論文誌データベース (TOD) , 71, 2016.

## 国際会議

8. Taguchi, Aramaki: Novel Location De-identification for Machine and Humans, UII 2018 UISTDA Workshop, 2018. (2018/03/11, Japan)
9. Daisaku Shibata, Shoko Wakamiya, Kaoru Ito, Mai Miyabe, Ayae Kinoshita, Eiji Aramaki: VocabChecker: Measuring Language Abilities for Detecting Early Stage Dementia, UII 2018. (2018/03/10, Japan)
10. Hayate Iso, Camille Ruiz, Taichi Murayama, Katsuya Taguchi, Ryo Takeuchi, Hideya Yamamoto, Shoko Wakamiya, Eiji Aramaki: NTCIR13 MedWeb Task: Multi-label Classification of Tweets using an Ensemble of Neural Networks In Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, 56-61, 2017. (2017/12/07, Japan) [Best System Award]
11. Ryo Takeuchi, Hayate Iso, Kaoru Ito, Shoko Wakamiya, Eiji Aramaki: Multi Liner Regression of Symptom-related Tweets for Infectious Gastroenteritis Scale Estimation, In Proceedings of the Workshop on Digital Disease Detection using Social Media (DDDSDM) of International Joint Conference on Natural Language Processing (IJCNLP), 2017. (2017/11/27, Taiwan)
12. Hayate Iso, Shoko Wakamiya, Eiji Aramaki: Forecasting Word Model: Twitter-based Influenza Surveillance and Prediction, In Proceedings of the International Conference on Computational Linguistics (COLING), 76-86, 2017. (2016/12/17, Japan)
13. Shoko Wakamiya, Yukiko Kawai, Eiji Aramaki: After the Boom No One Tweets: Microblog-based Influenza Detection Incorporating Indirect Information, International Conference on Emerging Databases (EDB 2016), 98-106, 2016. (2016/10/18, Korea) [Runner-up Paper Award]
14. Shoko Wakamiya, Hiroshi Kawasaki, Yukiko Kawai, Adam Jatowt, Eiji Aramaki, Toyokazu Akiyama: Lets Not Stare at Smartphones while Walking: Memorable Route Recommendation by Detecting Effective Landmarks The ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp), 2016. (2016/09/12, Germany) (Top conference)
15. Daisuke Kato, Mai Miyabe, Eiji Aramaki, Akiyo Nadamoto: Sizzle Word Analysis from Multiple Internet Media, In Proceedings of the 4th International Workshop on Advances in Data Engineering and Mobile Computing (DEMoC), 444-451, 2016. (2015/09/02, Taiwan)
16. Eiji Aramaki, Shuko Shikata, Eriko Watabe, Mai Miyabe, Yasuyuki Usuda, Satsuki Ayaya, Shinichiro Kumagaya: Allergy Risk Finder: Hypothesis Generation System for Allergy Risks via Web Service, In Proceedings of the MedInfo, e1113, 2015. (2015/08/19, Brazil)

# カルテ解析研究の業績

## 原著論文

1. Shoko Wakamiya, Mizuki Morita, Yoshinobu Kano, Tomoko Ohkuma, Eiji Aramaki: Tweet Classification Toward Twitter-Based Disease Surveillance: Overview of the MedWeb Shared Task Data, J Med Internet Res, 2018.
2. 荒牧英治, 若宮翔子, 矢野憲, 永井宥之, 岡久太郎, 伊藤薫: 病名アノテーションが付与された医療テキスト・コーパスの構築, 自然言語処理「言語処理の応用システム」特集号 (技術資料), 25 (1), 119 - 152, 2018.
3. Eiji Aramaki, Ken Yano, Shoko Wakamiya: MedEx/J: A One-scan Simple and Fast NLP tool for Japanese Clinical Texts, Studies in Health Technology and Informatics, MEDINFO 2017: eHealth-enabled Health, 245, 285 - 288, 2017.

## 国際会議

4. Kaoru Ito, Hiroyuki Nagai, Taro Okahisa, Shoko Wakamiya, Tomohide Iwao, Eiji Aramaki: J-MeDic: A Japanese Disease Name Dictionary based on Real Clinical Usage International Conference on Language Resources and Evaluation (LREC) , 2018. (2018/05/11, Japan)
5. Shoko Wakamiya, Mizuki Morita, Yoshinobu Kano, Tomoko Ohkuma, Eiji Aramaki: Overview of the NTCIR-13: MedWeb Task, In Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, 40-49, 2017. (2017/12/07, Japan)
6. Eiji Aramaki, Yoshinobu Kano, Tomoko Ohkuma, Mizuki Morita: MedNLPDoc: Japanese Shared Task for Clinical NLP, In Proceedings of the International Conference on Computational Linguistics (COLING) workshop on ClinicalNLP. (2016/12/16, Japan)

# 患者の語り研究の業績

## 原著論文

1. 柴田大作, 若宮翔子, 木下彩栄, 荒牧英治: 音声発話による認知症スクリーニング技術の開発 -感情表現辞書を用いた発話内容の質的分析-, 医療情報学, 37(6), 2018.
2. 宮部真衣, 四方朱子, 久保圭, 荒牧英治: 音声認識を用いた言語能力自動測定システム“言秤”の構築, 自然言語処理「言語処理の応用システム」特集号, 25 (1), 33 - 56, 2018.
3. 山下直美, 葛岡英明, 平田圭二, 工藤喬, 荒牧英治, 服部一樹: みまもメイト: 「見守る側」と「見守られる側」をつなぐヘルスケアアプリケーション, 情報処理学会論文誌, 58(5), 981-993, 2017.
4. ○[Eiji Aramaki](#), Shuko Shikata, Mai Miyabe, Ayae Kinoshita: Vocabulary Size in Speech may be an Early Indicator of Cognitive Impairment, PLoS ONE, 11 (5), e0155195, 2016.
5. ○[Eiji Aramaki](#), Shuko Shikata, Mai Miyabe, Yasuyuki Usuda, Kousuke Asada, Satsuki Ayaya, Shin-ichiro Kumagaya: Understanding the relationship between social cognition and word difficulty: a language based analysis of individuals with autism spectrum disorder, Methods of Information in Medicine, 54 (6), 522-529, 2015.
6. Yumiko Shimamoto, Mai Miyabe, Shuko Shikata, ○[Eiji Aramaki](#): Mind the Gap between Self-Reported QOL by Patients and Estimated QOL by Medical Staff, Studies in Health Technology and Informatics, MEDINFO 2015: eHealth-enabled Health, 216, 511-514, 2015.

## 国際会議

7. Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Takashi Kudo, [Eiji Aramaki](#), Kazuki Hattori: How Information Sharing about Care Recipients by Family Caregivers Impacts Family Communication, In Proceedings of the Conference on Human Factors in Computing Systems (CHI), 2018. (2018/4/22, Canada) (Top conference)
8. Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Takashi Kudo, [Eiji Aramaki](#), Kazuki Hattori: Changing Moods: How Manual Tracking by Family Caregivers Improves Caring and Family Communication, In Proceedings of the Conference on Human Factors in Computing Systems (CHI), 158-169, 2017. (2017/05/06, US) (Top conference)
9. Daisaku Shibata, Shoko Wakamiya, Ayae Kinoshita, [Eiji Aramaki](#): Detecting Alzheimer's Disease based on Word Category Frequencies, In Proceedings of the International Conference on Computational Linguistics (COLING) workshop on ClinicalNLP, 2016. (2016/12/16, Japan)