

医療分野での自然言語処理

NAIST

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
荒牧英治

自己紹介

自然言語処理

- 兵庫県立長田高等学校
- 学部: 京大 総合人間学部 (認知言語学: 山梨正明先生)
- 修士: // 言語メディア研 (言語処理: 黒橋禎夫先生)
- 博士: 東大 情報理工 (同上)

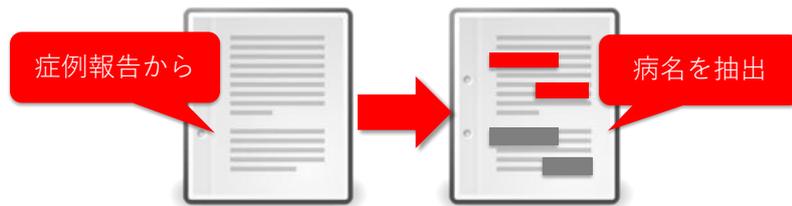
- 東大 医学部附属病院
 - 特任助教 (2006-2008)
- 東大 知の構造化センター
 - 特任講師 (2008-2013)
- 京大 デザイン学ユニット
 - 特定准教授 (2013-2015)
- NAIST 情報科学研究科
 - ソーシャルコンピューティング研究室
 - 特任准教授 (2015.9-)



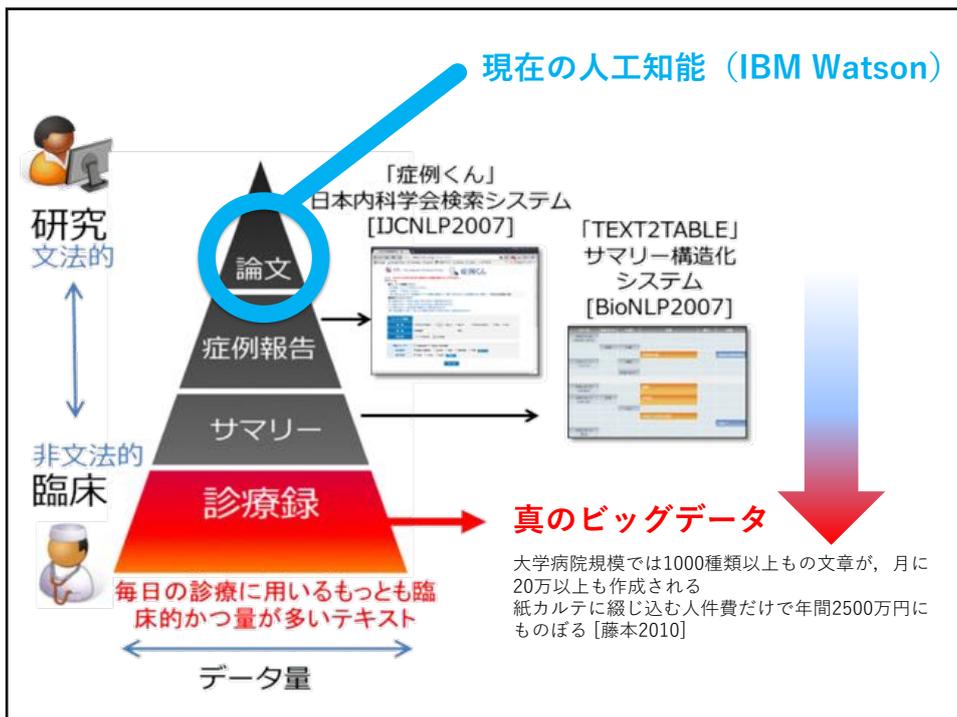
医療分野での
言語処理研究

医療分野の自然言語処理とは

- 音声処理, 画像処理と並ぶ情報処理研究の柱の一つ
- (私見) ただし, もっとも地味な処理
 - テキストを入力としてテキストを出力する



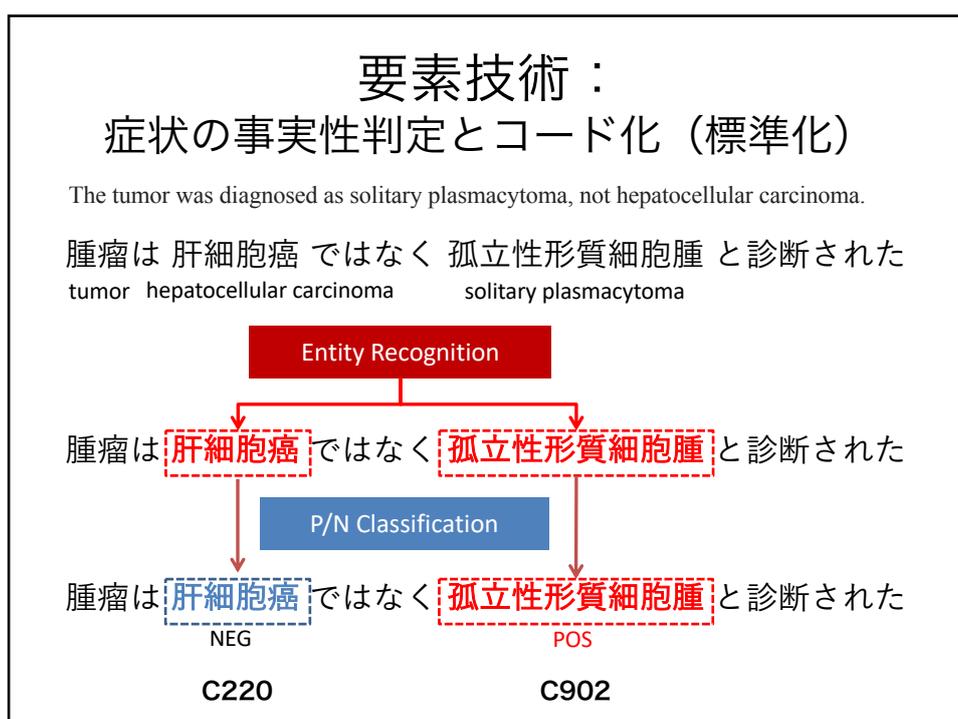
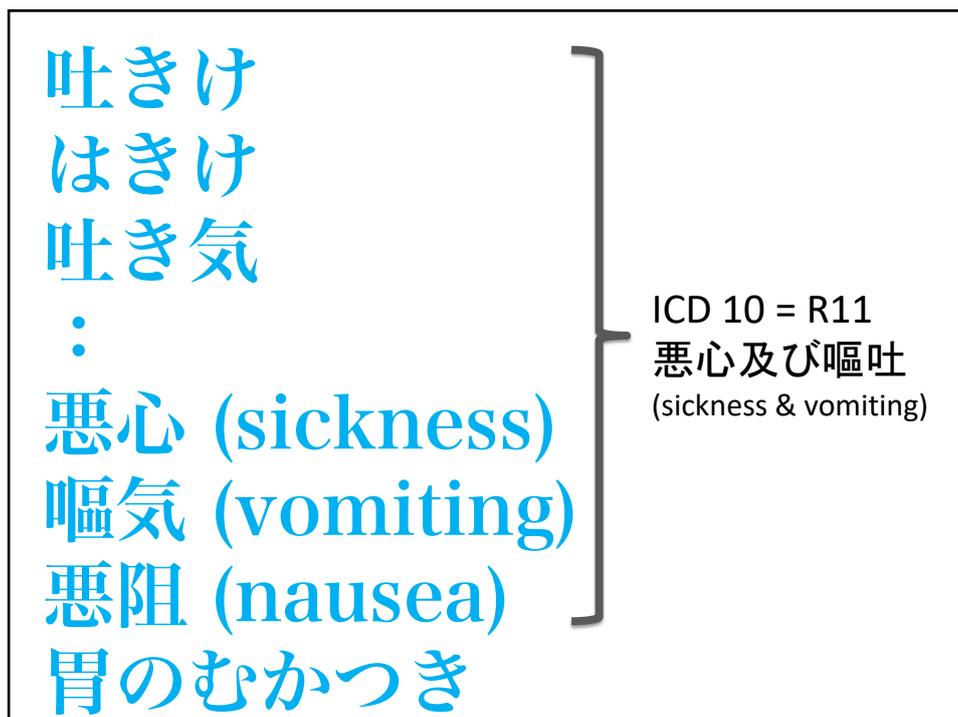
- 電子カルテの普及や人工知能への注目を背景に大きな期待
- ただし, 現在のところ奏功した応用事例に欠ける
 - 医療現場で, 言語処理が有り難いと感じるシーンはまれ
 - cf. 音声入力, 画像診断



発表概要

- はじめに
- 医療分野の自然言語処理とは
 - (1) 用語抽出
 - (2) 標準化
- 応用事例
- 利用可能なリソース
- おわりに 今後の展望
 - 2つのアプローチ

診療録	
患者番号: 1300000001	病名・手術・処置等
<入院1回目2日目> (年 月 日) <医師の記録> <<脳神経外科>> <<19>> OPが○日AM9:30分から決定。 患者、家族に手術について説明。 手術説明書・同意書を書き。 <看護師の記録> <<脳神経外科>> <<20>> OPが○日AM9:30分から決定。 以下の看護目標、プランを立案 看護目標 1. 手術後の感染症を予防し、治療を継続できるように維持する。 看護計画：手術前 1. 口腔ケア感染予防のために術前か含嗽を行い、口腔内の細菌を除去する 2. 口呼吸・含嗽の練習：口呼吸の練習、鼻をつまんだの飲水、含嗽、食事の練習を行う。 看護計画：手術後 1. 以下の症状の有無と程度の観察 ①髄液鼻漏：水分鼻汁の有無 ②口渇、多尿、多飲水 ③頭痛、吐きけ、嘔吐 2. 髄液鼻漏の予防： ①手術直後より鼻腔タンポン抜去時ま仰臥位を保つ。 ②離床後も顔面を下に向ける動	<<19-12>> 予約 1. 術前指示 ①禁食 ②ルートキープ 点滅指示 1. 口腔ケア感染予防のために含嗽を行い、口腔内の細菌を除去する 2. 口呼吸・含嗽の練習：口呼吸の練習、鼻をつまんだの飲水、含嗽、食事の練習を行う。 看護計画：手術後 1. 以下の症状の有無と程度の観察 ①髄液鼻漏：水分鼻汁の有無 ②口渇、多尿、多飲水 ③頭痛、吐きけ、嘔吐 2. 髄液鼻漏の予防： ①手術直後より鼻腔タンポン抜去時ま仰臥位を保つ。 ②離床後も顔面を下に向ける動

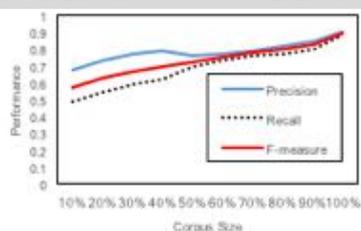


症例報告に必要な言語処理

医療テキストから医療
表現（病名，症状）
の抽出

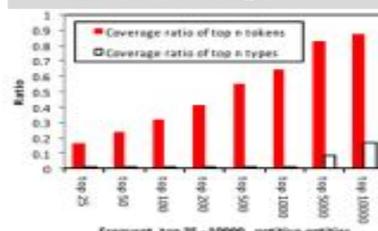
医療表現の標準化
（言い換え；正規化）

症状特定 (F値80% → 95%見込み)



症状抽出精度と教師データ量 (1,800症例を用いた実験)

症状標準化 (被覆率90%)



症例報告における頻出症状が全症状をカバーする割合

病名特定においては研究途中 (3万症例) で精度 (F値) 93%を達成, 研究終了時 (7万症例) では精度**95%が達成される**見込み

頻出する上位1万症状表現について知識ベース (ICD10) へのマッピング作業を行う, 試算では**90%の被覆率**で症状を標準化可能

汎用機 (core i7 6core 3.4GHz, Memory 32GB; 70万相当) にて, **120秒で3,000症例報告**を処理可能. 本研究終了後には実用化に耐えうる見込み

発表概要

- はじめに
- 医療分野の自然言語処理とは
 - (1) 用語抽出
 - (2) 標準化
- 応用事例
- 利用可能なリソース
- おわりに 今後の展望
 - 2つのアプローチ

症例報告の検索

症例くん

関東 608-91

デング熱に黄色ぶどう球菌性肺炎を合併した1例

宮田 順之, 立川 夏夫, 吉村 幸浩, 天野 雄一郎, 坂本 洋平

nobuyuki.m.1030@gmail.com

横浜市立市民病院感染症内科

【分類】 感染症

【キーワード】 デング熱, ブドウ球菌感染*

【症例】 22歳 男性 【主訴】 咳嗽、喉痛の増加 【現病歴】 5月初旬、マレーシア留学中に頭痛、発熱が出現して現地の病院を受診し、**デング熱**と診断された。6日間の入院治療を受け軽快し、退院した。帰国後の5月中旬に経過観察を目的として当院を受診した。このとき呼吸器症状はほとんど認めなかった。受診後から咳嗽、喀痰が増加し、1週間後に再受診した。胸部レントゲンで空腔性病変を認め、喀痰グラム染色でぶどう球菌性肺炎の白血球食食象を認めた。喀痰培養から *Staphylococcus aureus* が検出され、同菌による肺炎と診断した。Ceftriaxone 2g 4日間投与により軽快して退院となり、その後Cefalexin 2000mg 11日間内服を行い、症状・画像所見ともに軽快した。【考察】 2005年の報告では、774人の**デング熱**患者のう

症例報告は日本の学会が長年行ってきた我が国独自の活動
立ち遅れた日本の医療AIを挽回する格好のデータベースとなりうる

<http://uth.naika.or.jp>

検索結果4件

検索キーワード: 悪心 胸痛

条件を変更する

診断名	症例	google検索
GLP-1 受容体作動薬 1件 (症例割合 25.00%)	エキセナチド投与が誘因になって気胸および縦隔気腫を来した、胸部リウマチ合併症様の1例 横浜市立市民病院内科	抄録表示 > google検索 >
自己免疫性肝炎 1件 (症例割合 25.00%)	ロキソプロフェナトリウム内服により顕性化したと考えられた自己免疫性肝炎の1例 北里大学附属内科	抄録表示 > google検索 >
心筋梗塞 1件 (症例割合 25.00%)	冠危険因子なく、心筋梗塞を来した老年女性の1例 日本赤十字社医療センター 内科	抄録表示 > google検索 >
急性心筋梗塞 1件	非典型的な症状で受診した急性心筋梗塞の3例	抄録表示 > google検索 >

国立研究開発法人日本医療研究開発機構 (AMED), 2016年~2017年, 「人工知能による総合診療診断支援システムの開発」(研究代表者: 永井良三, 研究分担者: 荒牧英治)

資料: 日本内科学会症例検索UI

病名用語集

- 実際に利用されている病名をできるだけ抽出しコーディングした用語集（14万用語）
- ICDコーディング（高出現語 → 人手；低出現語 → 自動）

① 表層形	② ICD10コード	③ 標準病名	④ 信頼度レベル	⑤ ラベル
1 1 β - 水酸化酵素欠損症	E250	1 1 β - 水酸化酵素欠損症	S	1 1 β - 水酸化酵素欠損症:ICD=E250/LV=S/1 1 β - 水酸化酵素欠損症
1 7 α - 水酸化酵素欠損症	E250	1 7 α - 水酸化酵素欠損症	S	1 7 α - 水酸化酵素欠損症:ICD=E250/LV=S/1 7 α - 水酸化酵素欠損症
1 8 - 水酸化酵素欠損症	E250	1 8 - 水酸化酵素欠損症	S	1 8 - 水酸化酵素欠損症:ICD=E250/LV=S/1 8 - 水酸化酵素欠損症
1 8 q 欠失症候群	Q935	1 8 q 欠失症候群	S	1 8 q 欠失症候群:ICD=Q935/LV=S/1 8 q 欠失症候群
1 8 常染色体異常	Q913	1 8 常染色体異常	S	1 8 常染色体異常:ICD=Q913/LV=S/1 8 常染色体異常
1 p 3 6 欠失症候群	Q935	1 p 3 6 欠失症候群	S	1 p 3 6 欠失症候群:ICD=Q935/LV=S/1 p 3 6 欠失症候群
1 型呼吸不全	J969	1 型呼吸不全	S	1 型呼吸不全:ICD=J969/LV=S/1 型呼吸不全
1 型自己免疫性肺炎	K861	1 型自己免疫性肺炎	S	1 型自己免疫性肺炎:ICD=K861/LV=S/1 型自己免疫性肺炎
1 型糖尿病	E10	1 型糖尿病	S	1 型糖尿病:ICD=E10/LV=S/1 型糖尿病

病名抽出器 MedEX/

- 万病辞書と機械学習による病名抽出
- 陽性／陰性の分類

初診時は間質性肺炎は認められなかった。

<N value="間質性肺疾患">

再検査にて間質性肺炎が認められた。

<P value="間質性肺疾患">

- 主要OSで動作可能



厚生労働科学研究費補助金（臨床研究等ICT基盤構築研究事業）、2016年～2017年、「カルテ情報の自動構造化システムと疾患数理モデルの逐次的構築、及び、自動構造化機能を有した入力機構の開発」（荒牧班）

4万報告のアノテーションを実施

分類 神経 **キーワード** 外傷性内頸動脈解離、脳梗塞、カイロプラクティック **性別** 女性 **年齢** 36

カイロプラクティック後に発症した外傷性内頸動脈解離、脳梗塞の1例

年齢・性別など基本情報の抽出 病名・症状名の抽出と陽性・陰性の分類

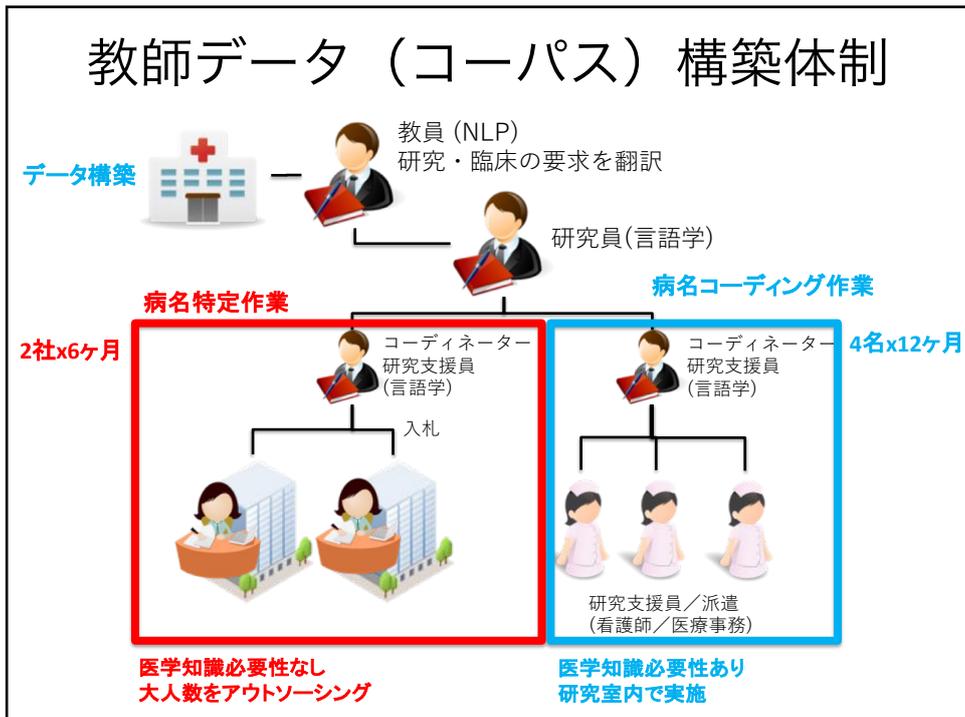
みさと健和病院内科 [メールを送る](#)

中村 博志 1), 田中 正巳 1), 石川 晋介 1), 松山 公彦 1), 武田 伸昭 2) 7128-関東

【症例】 36歳、女性 **主訴** 右前頭部の頭痛、左手のしびれ **【既往歴】** 尿管結石あり、経口避妊薬の服用なし **【現病歴】** 平成17年7月6日カイロプラクティックで頸部を強く伸ばしたり曲げられたりした。翌日から頭痛出現。7月9日入浴時に右前頭部痛と左手の痺れを自覚した。7月11日近医受診し、頭部CT施行したが異常無く、頭痛薬処方された。その後も訂薬、右前頭部痛、嘔気嘔吐、複視を認めため、当院受診。 **【現症】** 意識清明、血圧143/96 mmHg、瞳孔不円なし、対光反射正常、眼球運動正常、四肢に明らかな麻痺を認めず、異常感覚なし。 **【経過】** 頭部MRIで右前頭葉、側頭葉に脳梗塞を認めた。MRAでは右内頸動脈から右中大脳動脈にかけて信号の低下を認めた。脳梗塞と診断、脳神経外科に転院となった。7月25日の脳血管造影では、右頸部内頸動脈に血管解離あり、傍動脈瘤が発生し、真腔が著しく狭窄していた。脳梗塞に対し、エダラボンと高張グリセロールで保存的に治療を行い、症状の改善を認めた。 **【考察】** 脳梗塞の原因は、外傷による内頸動脈の血管解離と考えられた。そして、外傷の原因として、カイロプラクティックで頸部を過伸展、過屈曲したことが疑われる。脳梗塞の原因として注意が必要である。外傷性内頸動脈血管解離は自然治癒の可能性が認められる。脳血管造影にて経過観察する方針。

最終的なタグ数
 陽性タグ：660,065
 陰性タグ：81,655

検索した症状：赤文字 関連症状：黄色 非関連症状：青



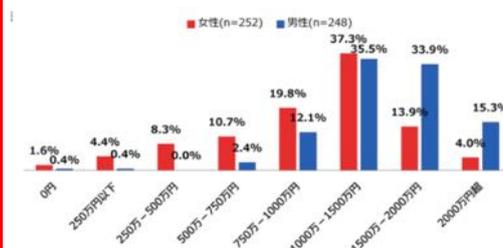
医療従事者雇用 コスト

1500万以上 政府調達
300万以上 入札
50万以上 事務発注
50万未満 教員発注

NAIST発注ルール

職種	年収
内科医	1350万
獣医師	616万
薬剤師	590万
看護師	519万
理学療法士	471万
検査技師	468万
臨床心理士	340万
医療事務	300万

年収1500万、女医の2割弱、男性医師では半数
最多は「1000万～1500万円」で女医の37%



本プロジェクトに従事する医療従事者

医師と同じく臨床経験が豊富

医師と同じ文章を扱いながら低コスト

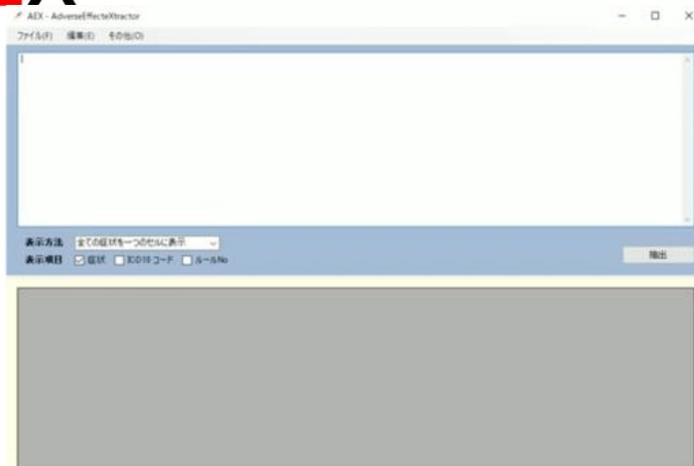
平成27年賃金構造基本統計調査, 職種別民間給与実態調査, 厚生労働省

薬歴報告

c.f. 調剤薬局H (30万件)



患者主訴からの副作用症状抽出器 AEX (Adverse Effect eXtractor)



日本医療研究開発機構(AMED) 医薬品等規制調和・評価研究事業 「患者の自覚症状により副作用の早期発見を可能とする方策に関する研究」

患者主訴からの副作用症状抽出器 AEX (Adverse Effect eXtractor)

- 症状抽出精度は8割程度

	Match number		抽出数		Search results group(PC)	適合率 Precision		再現率 Recall		
	POSSIBLE	SURE	Positive solutions group POSSIBLE	SURE		POSSIBLE	SURE	POSSIBLE	SURE	
ALL	1949	1534	2456	—	2293	0.8500	0.6690	0.7936	0.6246	
Code is another extract of top 10										
1 R42	めまい	228	146	256	173	257	0.8872	0.5681	0.8906	0.8439
2 J00	かぜ感冒	226	189	242	209	242	0.9339	0.7810	0.9339	0.9043
3 F196	健忘症候群	180	142	205	169	199	0.9045	0.7136	0.8780	0.8402
4 R529	疼痛	154	109	188	149	202	0.7624	0.5396	0.8191	0.7315
5 R05	咳、夜間咳	131	114	145	136	147	0.8912	0.7755	0.9034	0.8382
6 H931	耳鳴症	106	99	136	131	134	0.7910	0.7388	0.7794	0.7557
7 R260	よろめき歩行	104	41	133	60	122	0.8525	0.3361	0.7820	0.6833
8 I10	高血圧症	76	51	111	90	87	0.8736	0.5862	0.6847	0.5667
9 F511	非器質性過眠症	81	39	88	47	87	0.9310	0.4483	0.9205	0.8298
10 R53	易疲労感、疲労感	66	41	84	62	84	0.7857	0.4881	0.7857	0.6613

- ルールベース
 - ルール追加/修正による改善が可能

ICD-カテゴリー SearchNo 検索文字列 (注: mecabの単語の
A09-下痢症- 19- (下痢|下痢症|げり|ゲリ|ゲリー|ちゃ
A09-下痢症- 20- (水様便|水みたいな便|((便|べん)+(|緩い|

日本医療研究開発機構(AMED) 医薬品等規制調和・評価研究事業 「患者の自覚症状により副作用の早期発見を可能とする方策に関する研究」

7.小児等への投与
低出生体重児、新生児、乳児、幼児又は小児に対する安全性は確立していない。(使用経験が少ない。)

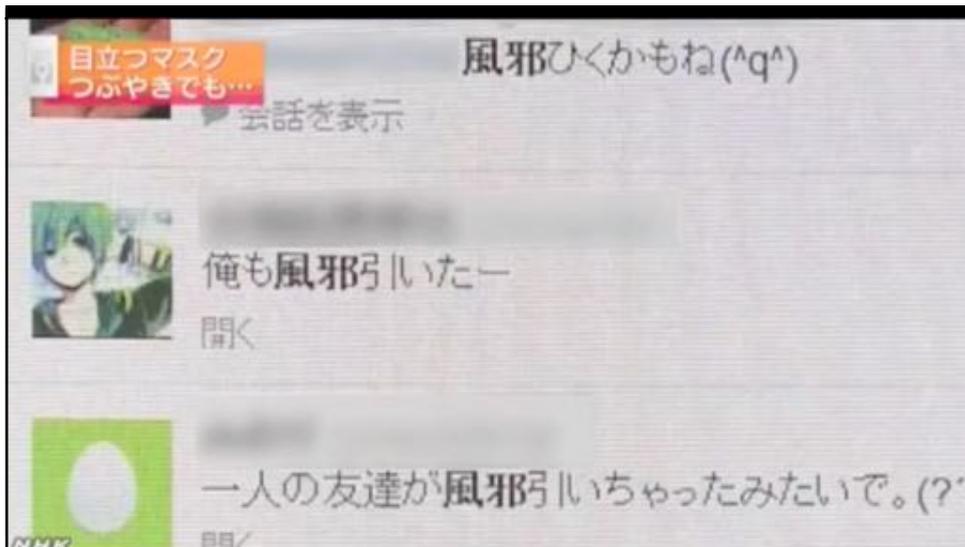
政策科学総合研究事業（臨床研究等ICT基盤構築・人工知能実装研究事業）小児領域の医薬品の適正使用推進のための人工知能を用いた医療情報データベースの利活用に関する研究（小児AI研究）2017年～2019年（研究代表者：栗山猛、研究分担者：荒牧英治）

発表概要

- はじめに
- 医療分野の自然言語処理とは
 - (1) 用語抽出
 - (2) 標準化
- 応用事例
 - カルテ
 - Web
 - 家庭へ
- 利用可能なリソース
- おわりに 今後の展望
 - 2つのアプローチ



出来事	1分あたりの数
Facebookユーザがいいねをする数	4,166,667
Twitterユーザのツイートの数	347,222
Instagramユーザがいいねをする数	1,736,111
Pinterestユーザがピンを付ける画像の数	9,722



目立つマスク つぶやきでも... 風邪ひくかもね(^q^)
 会話を表示

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: 300).
 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

インフルくん - NLP Flu Warning -

インフルくんは、NLPを用いたインフルエンザの危険度を予測するサービスです。最新のインフルエンザの流行状況をリアルタイムで把握し、危険度を予測します。また、花粉症の危険度もお知らせします。

サイト分布MAP

インフルくんが稼働している全国のサイト分布を地図上で表示しています。色によって危険度のレベルが異なります。

グラフ

インフルくんが稼働している全国のインフルエンザの流行状況をグラフで表示しています。最新の流行状況をリアルタイムで把握できます。

今日の花粉尘の危険度が分かる

花粉症危険レベルです。花粉の飛散が多いのは午前から午後2時くらいです。

患者の発話を活かすための言語処理

① 患者の発話から

NLP

② 表現様式を測定し

③ 疾患を推定する

認知症

2014- 分担
基盤A「認知症・うつ家族サポート」
(代表：山下直美)

2015- 総括班
新学術「主体価値」
(代表：並井清澄)

2013- 発達障害
発達障害当事者の会
n=70 (自由発話) CREST
熊谷班・研究協力
CREST認知ミラーリング
(代表：長井志江)

2017- 高齢者
共同研究 ATR 脳情報通信総合研
(田中沙織)

2015- 思春期
総括班
新学術「主体価値」
(代表：並井清澄)

2017- 高齢者
共同研究 ATR 脳情報通信総合研
(田中沙織)

総括班 新学術「主体価値」 (代表：並井清澄) | 東京在住の10代 (n=4,300 (インタビュー)) | 総務省 脳画像提供の高齢者 (n=400 (インタビュー))

©Eiji Aramaki, Shuko Shikata, Mai Miyabe, Ayae Kinoshita: Vocabulary Size in Speech may be an Early Indicator of Cognitive Impairment, PLoS ONE, Volume 11 (5), e0155195, 2016. (IF=3.2)



検査を嫌がる患者さんは多い

知っている野菜の名前をできるだけ多く
言って下さい

前にアレで見たぞ！
認知症のアレだろう！
失敬な！ワシは
ボケてなどおらん！



28

自然な会話で認知症の傾向を察知

最近楽しかったことは何ですか？



29

Small Vocabulary in MCI



Table 6. Speech

	Speech (S)		p
	Healthy	MCI	
TTR	0.43	0.53	n.s.
PVS	541	2668	n.s.
VL	0.41	0.40	n.s.
DepD	2.03	2.05	n.s.
Yngve	1.60	1.52	n.s.

TTR, Type-Token Ratio; PVS, Potential Vocabulary Size; VL, Vocabulary Level; DepD, Dependency Distance.

Underlined values indicate statistically significant results at $p \leq .05$.

doi:10.1371/journal.pone.0155195.t006

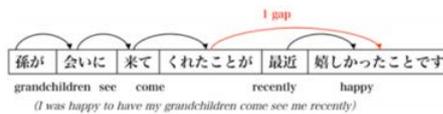


Fig 2. Dependency distance.

doi:10.1371/journal.pone.0155195.g002

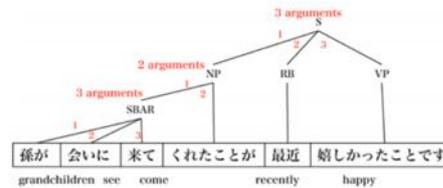


Fig 3. Yngve score.

doi:10.1371/journal.pone.0155195.g003



RESEARCH ARTICLE

Vocabulary Size in Speech May Be an Early Indicator of Cognitive Impairment

Eiji Aramaki^{1*}, Shuko Shikata¹, Mai Miyabe², Ayae Kinoshita³

¹ Nara Institute of Science and Technology (NAIST), 8916-5 Takayama, Ikoma City, 630-0192, Japan, ² Wakayama University, Sakae-dani 930, Wakayama City, 640-8510, Japan, ³ Kyoto University Graduate School of Medicine, 54 Kawahara-cho, Syogoin, Sakyo-ku, Kyoto City, 606-8507, Japan

©Eiji Aramaki, Shuko Shikata, Mai Miyabe, Ayae Kinoshita: Vocabulary Size in Speech may be an Early Indicator of Cognitive Impairment, PLoS ONE, Volume 11 (5), e0155195, 2016. (IF=3.2)



発表概要

- はじめに
- 医療分野の自然言語処理とは
 - (1) 用語抽出
 - (2) 標準化
- 応用事例
- 利用可能なリソース
- おわりに 今後の展望
 - 2つのアプローチ

症例報告における 機械には難しい表現例

id=35012
 title=びまん性Lewy小体病の経過中に<P>大動脈炎発症</P>を合併した1例
 text=症例は73歳、女性。主訴は<P>顔面</P>、<P>発熱</P>、<P>食欲不振</P>。既往歴に<P>乳癌</P>による右乳房切除あり。6年前まで正常。事変っていた。3年前から顔面が作れなくなった。<P>顔面</P>もその頃から認めていた。2年前から杖を使うようになり今年に入って動作が目立つようになった。1年前から<P>ムツ使用</P>となる。8月にADL全介助状態。コミュニケーションもとらなくなり食事も減少した。10月に<P>介助歩行の状態</P>でA病院に入院。10月、精密目的で当科紹介入院。入院時<P>顔面</P>で<P>顔面</P>。<N>顔面血管雑音</N>なし。<P>認知機能障害</P>と<P>脳外器疾患</P>を認めた。検査所見では、D-dimerの上昇<P>血沈の上昇</P>が認められた。MIBG心筋シンチの取り込み低下 またSPECTおよびメチオニンPETで後頭葉を中心としたびまん性の血流低下が認められ、<P>DLBG</P>と診断した。<P>乳癌腫瘍</P>。検査目的でFDG-PETを施行。両側顔面動脈や大動脈の走行に沿った集積が認められ<P>大動脈炎発症</P>と判断した。3日間ステロイドパルス25mgを投与し、その後PSL20mg内服へと移行した。<N>慢性貧血</N>と<N>発熱</N>などは消失し、顔面も改善したが、アミダゾール。<P>脳動脈</P>など<P>脳外器疾患</P>によるADL低下が改善しなかった。治療効果は100mmHgを超えていたが正常化した。<P>SKIP</P>。高動脈性疾患<P>の経過中においても<N>血管炎発症</N>には鑑別すべき疾患である。

keys=大動脈炎発症群 びまん性レヴィー小体病 FDG-PET
 dgm=大動脈炎発症群
 本誌掲載 - 減少者 顔面発熱 アミダゾール 顔面発熱 (H+K+P+M+R+L+O+J+I+G)
 血沈 初発 20-30mm 分岐点 顔面発熱 発熱
 中程度 50-100mm 顔面発熱 顔面 顔面発熱
 高度 100mm以上 顔面発熱 顔面発熱 顔面発熱 顔面発熱
 D-dimer = 陰性性血管内腔血栓症 0.60
 MIBG 心筋シンチ 取り込み低下 - パーキンソン病 0.20 DLBG = 認知症に伴うパーキンソン病
 FDG-PET = 大動脈炎発症群

・ 検査名 + 解釈や検査値

MIBG心筋シンチの取り込み低下

→ パーキンソン病
G20

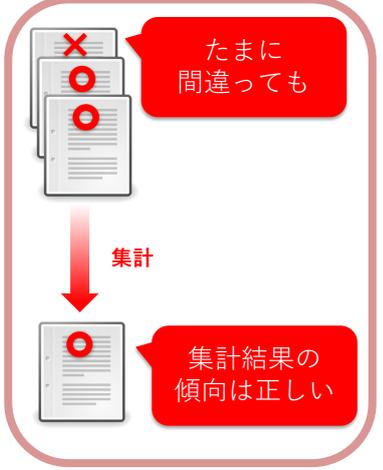
・ コードの無い頻出語

オムツ使用 → 尿失禁
R32

介助歩行の状態 → 歩行障害
R262

奈良先端科学技術大学院大学
山本英弥氏 発表資料より

成功している2つのNLPのタイプ



ビッグデータNLP



ダイナミックNLP

ダイナミックNLPはまだ始まったばかり
別分野、UIやCSCWとの融合も

気管支喘息発作がみられた75歳の男性。大腸癌の診断を受け、検査を行いました。検査の結果、大腸憩室出血、および、慢性腎臓病、肝転移、および、感染性心内膜炎が見つかりました。男性は、鬱症状が出て、関節痛、および、間質性肺炎の症状が見られます。開口障害のために、逆流性食道炎が観察され、また、摂食障害が見受けられます。また、2型糖尿病の症状がみられ、甲状腺腫大が疑われています。また、陰窩膿瘍を認めた、物忘れが激しくなるなどの症状も見られます。

前向き (prospective) 言語処理

- 入力時にリアルタイム処理
- 標準でない用語はハイライト
- 標準用語サジェストする

R413

記憶障害

記憶力低下

健忘症

記憶喪失

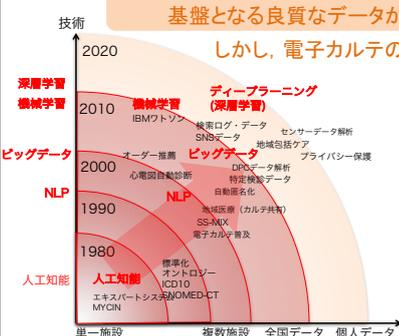
物忘れ



電子カルテ情報をセマンティクス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

人工知能技術を活かすためには
基盤となる良質なデータが必須
しかし、電子カルテの表現は多様

カルテ中の自然文から
①医療表現を抽出し、②標準化する
技術が必須



医療表現の自動抽出 (90%)

標準化 (被覆率90%)

前向き標準化を有した入力環境

深層学習により90%の精度で抽出可能

大規模な病名辞書 (14万病名記載) により、日常診療で使われる病名の90%を標準化可能

入力サポートにより、自動的に標準化結果を参照しながらの入力が可能

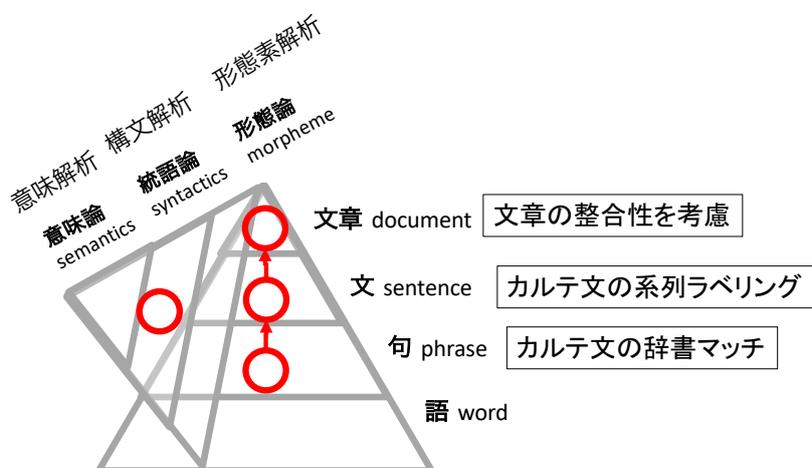
看護計画：手術後
1. 以下の症状の有無と程度の有無
① 髄液鼻漏：水分鼻汁の有無
② 口渇、多飲、多飲水
③ 頭痛、吐きけ、嘔吐
2. 髄液鼻漏の予防：
① 手術直後より鼻腔タンポン抜仰臥位を保つ。
② 離床後も顔面を下に向ける

悪心増悪
胃のむかつき
悪心
食後悪心
嘔気
吐きけ
悪阻
ICD10 悪心及び嘔吐より

厚生労働省
CSTI政策討議
説明資料

(発表者の私見) 科学技術政策担当大臣等政務三役と総合科学技術・イノベーション会議有識者議員との会合
(平成29年11月9日) CSTI 政策討議資料

自然言語処理の階層



あらゆる文章を用いる

- 細胞レベル
 - 論文
- 病理組織レベル
 - 病理所見（病理医）
- 解剖・臓器レベル
 - 画像所見（放射線医）
- 患者因子・治療介入レベル
 - 電子カルテ（診療録, サマリ）（内科医）

課題 (並列構造)

全身性硬化症を合併した自己免疫性肝炎の1例

症例は61歳、女性。以前から健診にて**肝機能異常**を指摘されていたが、受診はしていなかった。
2009年10月より**上腹部不快感**を自覚。
2010年1月から著明な**下肢浮腫、息切れ**を自覚し近医を受診した。
肝機能異常、貧血、低酸素血症を認め、精査目的に当院紹介となった。
身体所見では**レイノー症状、手指硬化、両側肺線維症、手掌紅斑**がみられた。
腹水、クモ状血管腫、フラッピングは認めなかった。
GOT 774U/l、GPT 628U/l、HBs-Ab(+)、HBs-ag(-)、AFP 826ng/ml、抗核抗体320倍、IgG 3062mg/dl、肝生検にてInterface hepatitis、形質細胞浸潤、**肝細胞のロゼット形成**が認められ、**自己免疫性肝炎**と診断した。
国際診断基準では22点であった。
一方、抗Scl-70抗体と抗セントロメア抗体はともに陰性であり、アメリカリウマチ学会の**全身性硬化症**の診断基準を満たした。
PSL 45mg/日の経口投与と**肝硬変症**に対する栄養療法、浮腫に対して利尿薬投与を開始した。
治療開始後、徐々にトランスアミナーゼは低下、ステロイドの漸減を図ることが可能となった。
全身性硬化症と自己免疫性肝炎との合併は稀な症例であり、文献的考察を加えて報告する。

文内の整合性 (NER)

文内の整合性 (PN)

文間の整合性 (文章の整合性)

関連症状: 非関連症状:

課題 (P/Nより細かい事実性判定)

PMXを施行し特発性肺線維症 (IPF) の急性増悪を抑制しえた1例

【症例】66歳、男性。【主訴】**呼吸困難**【既往歴】**肺結核**【現病歴】1997年(56歳)に**咳嗽**出現し、CT画像にて**肺葉病変**からIPFと診断され経過観察されていた。2003年1月から**労作時呼吸困難感**とCT画像で**陰影増強**を認め、2月より当科外来紹介となる。PSL 10mgを開始し、さらに7月よりCyAを併用し、**自覚症状は改善**。2004年11月に**感冒様症状**後に若干の**呼吸症状**の増悪があり、CT上**すりガラス陰影**が出現したため、エラスポールを投与し収束。その後在宅酸素療法を導入(VC 2.88L(83%)、安静時PaO2 74Torr)。2005年6月には安静時でPaO2 66Torrと徐々に病状は進行。2006年5月末から**急激な呼吸状態の悪化**を自覚し6月12日に入院。【入院後経過】安静時でPaO2 55Torr、CT上明らかな**すりガラス陰影**の増悪はなかったものの、CRP 6mg/ml、 β -Dグルカン 62pg/mlと上昇。**肺真菌症、カビ二肺炎**を考慮、MCFG 150mg、pentamidine isetionate 250mgを投与し、26日にはCRP、 β -Dグルカンともに低下し、**呼吸苦も改善傾向**を示した。しかし、その後**呼吸苦は再増悪**し、7月3日のCTでは**全肺野にすりガラス陰影**が出現しIPFの急性増悪と診断。肝・腎機能障害に加えて**高アミラーゼ血症**も認めため、ステロイドパルス療法は行わずにPMXを4回(4日連続)施行。急性増悪の進行を抑制しえた。

改善傾向(症状はあるがましになった場合)の扱い

反復性の経過

寛解・増悪を伴う持続性の経過

周期的な症状は?

反復性嘔吐
嘔吐あり→なし→あり

変化を記述する?
変化した結果を記述する?

課題 (WSD ; 略語展開)

一般には、文脈が必要であり、辞書だけでは解けないが、ドメインを絞った場合は曖昧性は激減？

- IP immuno-peroxidase reaction 免疫ペルオキシダーゼ反応
 IP infection prevention 感染予防
 IP inorganic phosphorus 無機リン
 IP inositol phosphate イノシトールリン酸
 IP intercuspal position 咬頭嵌合位
 IP International Pharmacop(o)eia 国際薬局方
 IP internus paresis 内筋麻痺
 IP interstitial pneumonitis 間質性肺炎
 IP intraperitoneal injection 腹腔内注射
 IP intravenous pyelography 経静脈性腎盂造影(法)
 IP isoproterenol(=IPR) イソプロテレノール
- IPE interstitial pulmonary emphysema 間質性肺気腫
 IPF idiopathic pulmonary fibrosis 特発性肺線維症
 IPF interstitial pulmonary fibrosis 間質性肺線維症
 IPH idiopathic portal hypertension 特発性門脈圧亢進症
 IPH idiopathic pulmonary hemosiderosis 特発性肺ヘモジデリン沈着症



まとめ

- 医療分野の自然言語処理 (NLP) は入力と出力を変えることで多様な応用が可能
 - 入力：電子カルテ，患者テキスト，医師発話，患者発話
 - 出力：統計調査，スクリーニング
- 実用化の型
 - ビッグデータNLP
 - ダイナミックNLP
- 知識を記述するのは言語
 - 過去の医療データは資産であり，これを活かすことは，我が国で行うべき重要な研究

謝辞

• Staff

- 荒牧英治 Ph.D. 特任准教授
- 若宮翔子 Ph.D. 博士研究員
- 伊藤薫 研究員
- 岩尾友秀 研究員
- 本田ちひろ 研究技術員



• Data Staff

- 川上麻子
- 中江睦美
- 岡久太郎
- 永井宥之
- 友廣公子
- 宮崎京子
- 松本妙子

• D1

- 磯颯
- 柴田大作

• M1-M2

- Camille Marie Ruiz
- 竹内瞭
- 田口勝弥
- 村山太一
- 山本英弥
- Nigo Sumaila

<http://mednlp.jp>

<p>日本医療研究開発機構 AMED</p> <p>国立研究開発法人日本医療研究開発機構 (AMED)。2016年～2019年。「医用知能情報システム基盤の研究開発」(研究代表者: 大江和彦, 研究分担者: 荒牧英治)</p> <p> </p>	<p>日本医療研究開発機構 AMED</p> <p>国立研究開発法人日本医療研究開発機構 (AMED)。2016年～2017年。「人工知能による総合診療診断支援システムの開発」(研究代表者: 永井良三, 研究分担者: 荒牧英治)</p> <p> </p>	<p>厚生労働科学研究費補助金 MHLW</p> <p>厚生労働科学研究費補助金。2016年～2017年。「カルテ情報の自動構造化システムと疾患処理モデルの逐次的構築、及び、自動構造化機能を有した入力機構の開発」(荒牧英治)</p> <p> </p>	<p>厚生労働科学研究費補助金 MHLW</p> <p>厚生労働科学研究費補助金。2017年～2019年。「小児領域の医薬品の適正使用推進のための人工知能を用いた医療情報データベースの利活用に関する研究」(研究代表者: 栗山誠, 研究分担者: 荒牧英治)</p> <p> </p>	<p>日本医療研究開発機構 AMED</p> <p>日本医療研究開発機構 (AMED) 2016年～2019年。「迅速・網羅的病原体ゲノム解析法の開発及び感染症危機管理体制の構築に資する研究」</p> <p></p>
--	---	---	--	--