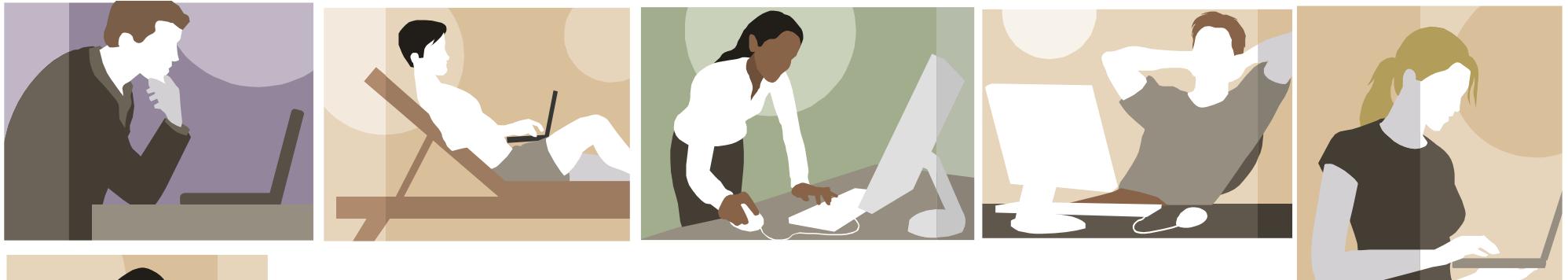


# ビッグデータ時代の 自然言語処理

乾 健太郎

東北大学





# コミュニケーション

## あらゆる社会活動の根幹

最も重要なメディアは  
言語（ことば）



# 東北大 自然言語処理研究室

## 研究スタッフ



乾健太郎



岡崎直觀



渡邊陽太郎



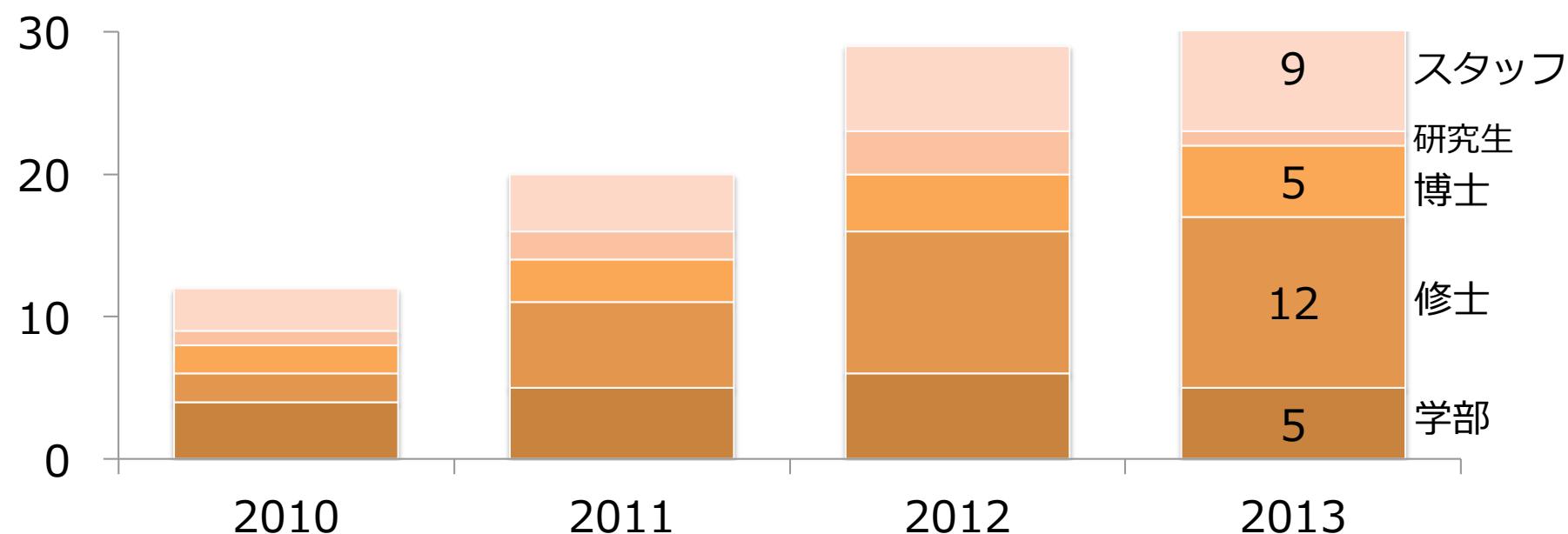
松林優一郎



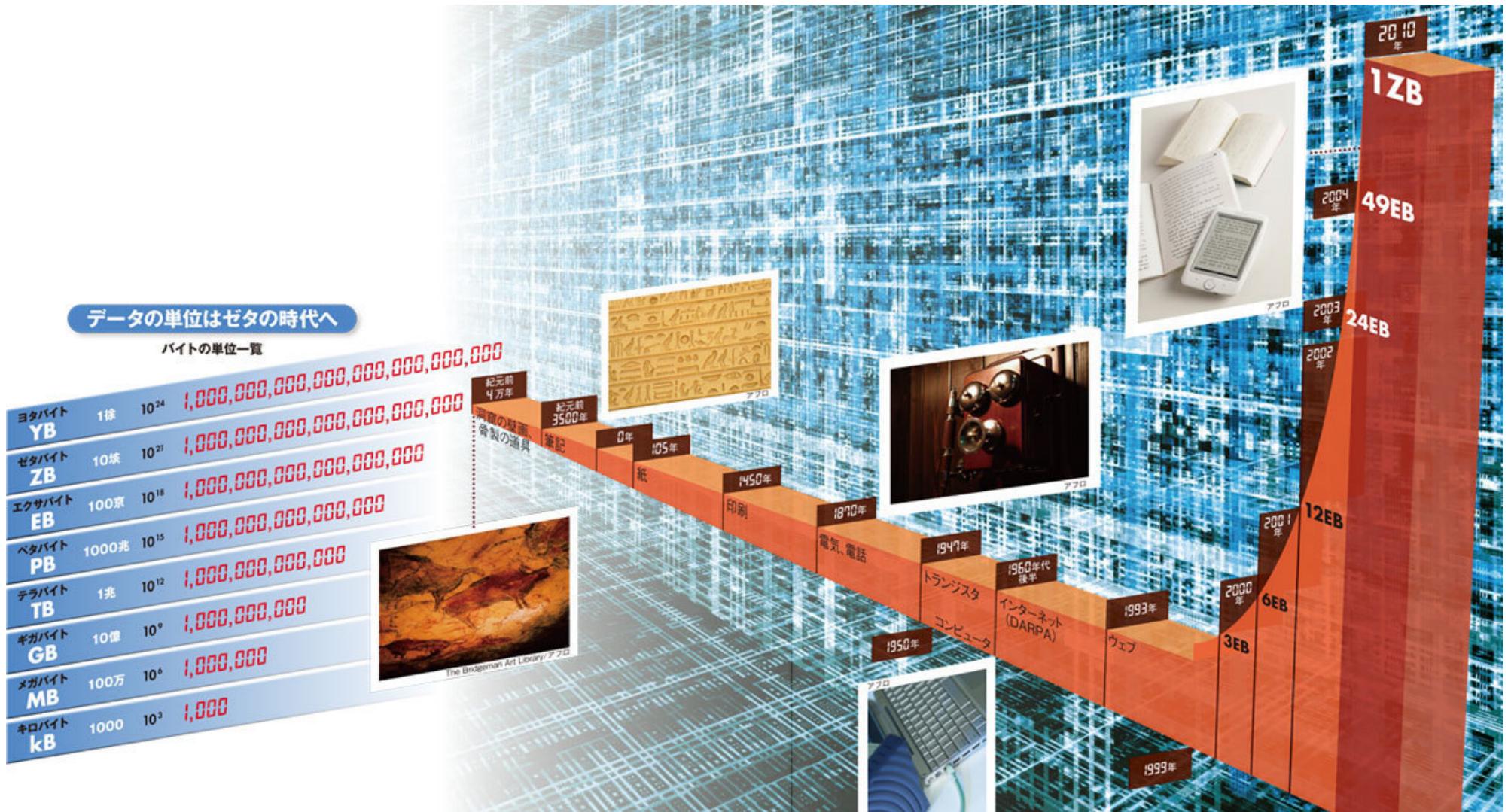
水野淳太  
(NICT)



井之上直也



# 時代のキーワード「ビッグデータ」



ダイアモンド IT&ビジネス「2012年 ビッグデータ経営革命」<http://diamond.jp/articles/-/16192>

## 構造化データと非構造化データ

### 構造化データ



顧客データベース  
売り上げ管理データベース  
在庫管理データベース  
経理データベース  
.....

### 非構造化データ



ソーシャルメディア



GPS



メール



センサー網



画像・音声・映像



RFID (ICタグ)



BLOG ウェブログ



通話内容記録



ネット上のテキスト・ドキュメント



eコマース履歴



ネット検索履歴



軍事偵察



ゲノム解析



医療画像・診療記録



監視カメラ



生態調査



天文・気象

.....

# 自然言語処理

人工知能

## 言葉がわかる賢いコンピュータを作る

言葉(=自然言語)で伝達される情報を  
理解・検索・抽出・翻訳・整理・分析し、  
コミュニケーションを支援するソフトウェア技術



# 本日の講演内容

## 自然言語処理の基礎

- ・コンピュータに言葉をどう教えるか

## ビッグデータのインパクト

- ・大規模知識獲得でブレークスルーを

## 情報の信頼性と編集

- ・言論マップから耐災害ICTへ

# どうやってコンピュータに 言葉を教えるか



# 「言葉がわかる」とは？ 単語を知っている、単語に分けられる

そのひとことで元気になった。

そのひとことで元気になった。

そのひと(人)こと(事)で元気になった。

そのひ(日)とことで元気になった。

そのひ(火)とこ(床)とで元気になった。

そのひとことで元気になつ(担つ)た。

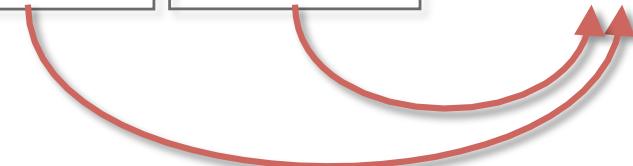
そのひとことで元気になった。

# 「言葉がわかる」とは？ 構文構造(修飾関係)がわかる

(名詞)と(名詞)に(動詞)

彼女 と 仙台 に 行きました

彼女と 仙台に 行きました



*I visited Sendai with her*

盛岡 と 仙台 に 行きました

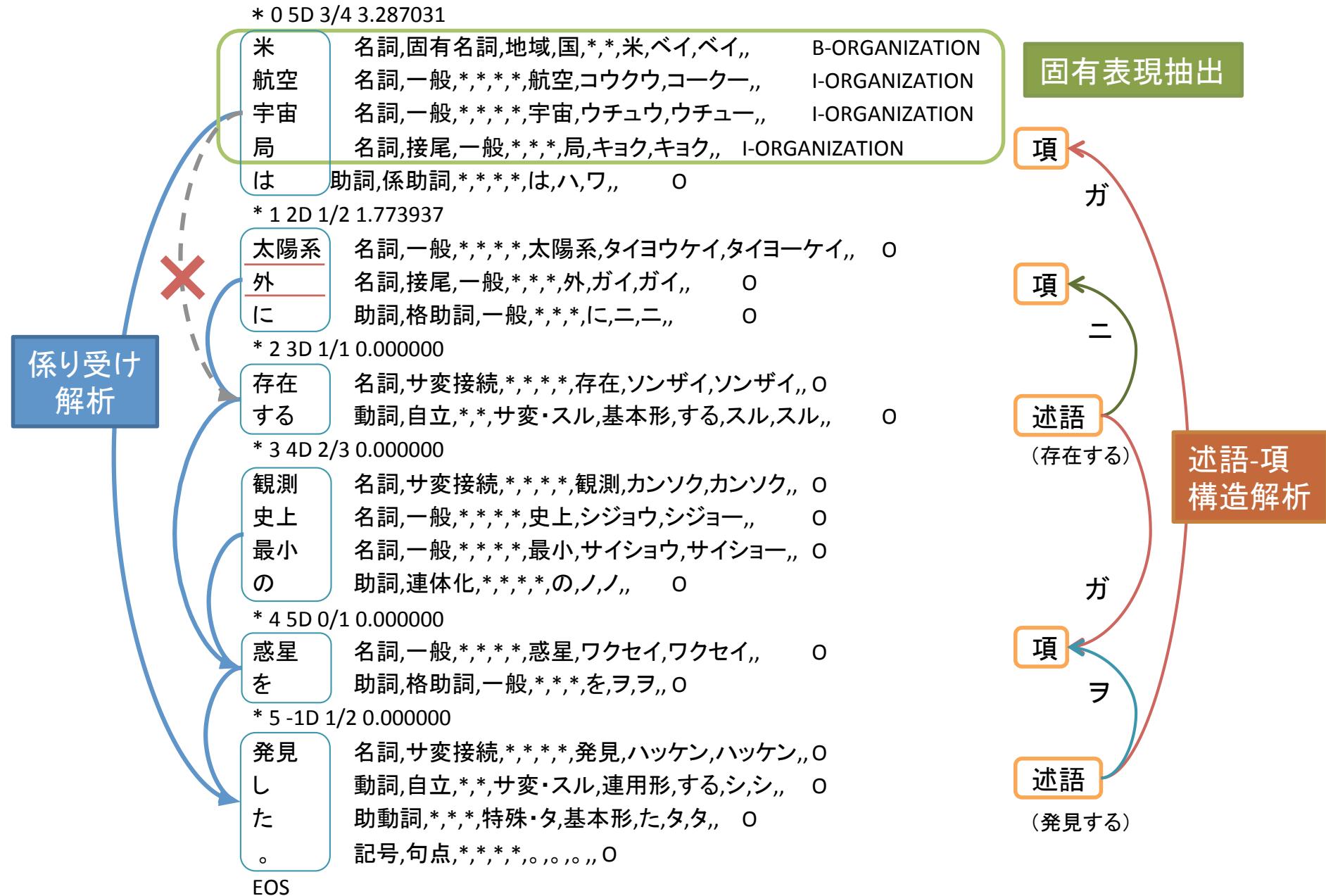
盛岡と 仙台に 行きました



*I visited Mori. and Sendai*

# 自然言語処理 実際の解析例

米航空宇宙局は太陽系外に存在する観測史上最小の惑星を発見した。



# 基本戦略①

たくさんの中題を与えて規則性を学ばせる

## 例題と答え(教師データ)

言語とは、コミュニケーションのための記号の体系である  
名詞 助詞 助詞 記号 名詞 助詞 名詞 助詞 名詞 助詞 名詞 助動詞 助動詞

自然言語処理は、人間が使っている言葉をコンピュータに  
名詞 名詞 助詞 記号 名詞 助詞 動詞 助動詞 助動詞 名詞 助詞 名詞 助詞

落語は、江戸時代の日本で成立し、現在まで伝承され  
名詞 助詞 記号 名詞 名詞 助詞 名詞 助詞 サ変 動詞 記号 名詞 助詞 サ変 動詞 助動詞

都市に人口が集積することによつて芸能として成立した  
名詞 助詞 名詞 助詞 サ変 動詞 名詞 助詞 動詞 助詞 サ変 助詞 動詞 助詞 サ変 動詞 助詞



統計解析(機械学習)

## 規則性(答えの傾向)

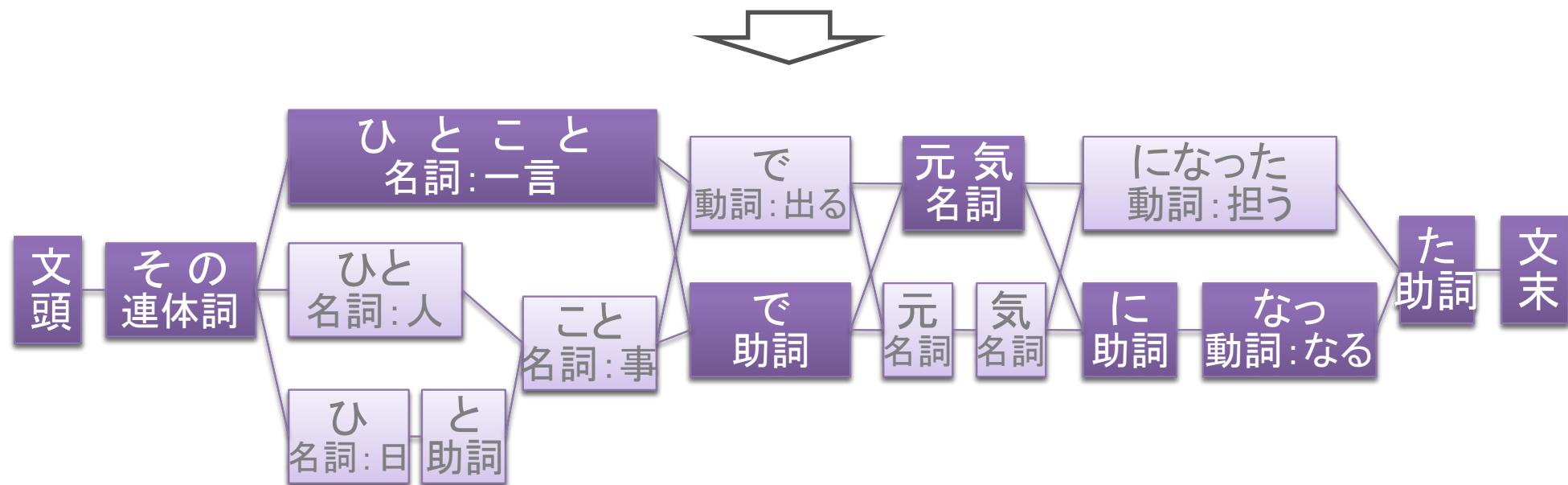
- 名詞の後には助詞が一番続きやすく、つぎに名詞が続きやすい
- 助詞の後には名詞が一番続きやすく、つぎに動詞が続きやすい
- 「が」「を」「に」「で」という文字は助詞になりやすい
- 長い単語が候補にあるなら、短い単語に細切れにするよりも良い

# 基本戦略①

# たくさんの**例題**を与えて**規則性**を学ばせる

## 規則性(答えの傾向)

- ・名詞の後には助詞が一番続きやすく、つぎに名詞が続きやすい
  - ・助詞の後には名詞が一番続きやすく、つぎに動詞が続きやすい
  - ・「が」「を」「に」「で」という文字は助詞になりやすい
  - ・「なる」という動詞は「になう(担う)」という動詞よりも出やすい
  - ・長い単語が候補にあるなら、短い単語に細切れにするよりも良い



# 問題！

## 訓練事例をいくら作っても足りない

単語の組み合わせを考える必要のあるタスク

(名詞) と (名詞) に (動詞)

10万種類

10万種類

1万種類

彼女 と 仙台 に 行きました

盛岡 と 仙台 に 行きました

彼女 と 仙台 に 行きました

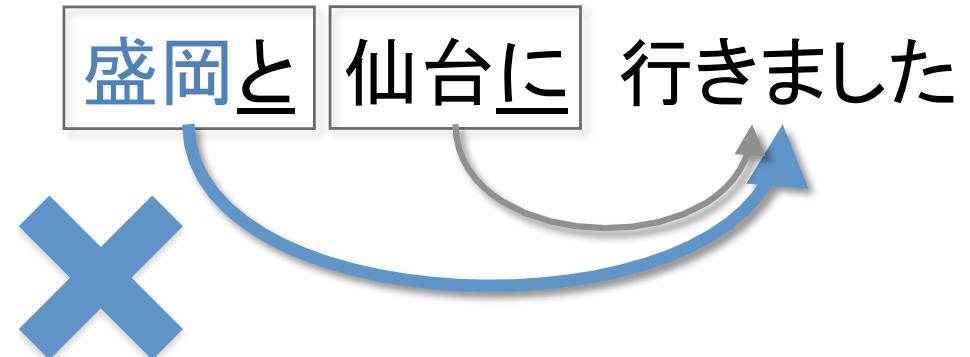
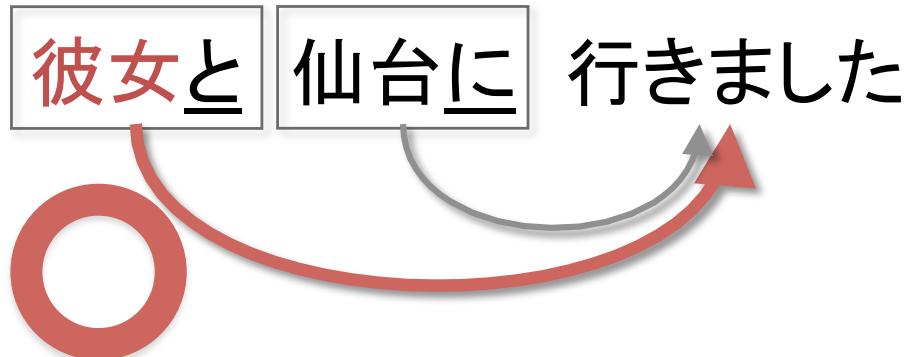
盛岡 と 仙台 に 行きました

*I visited Sendai with her*

*I visited Mori. and Sendai*

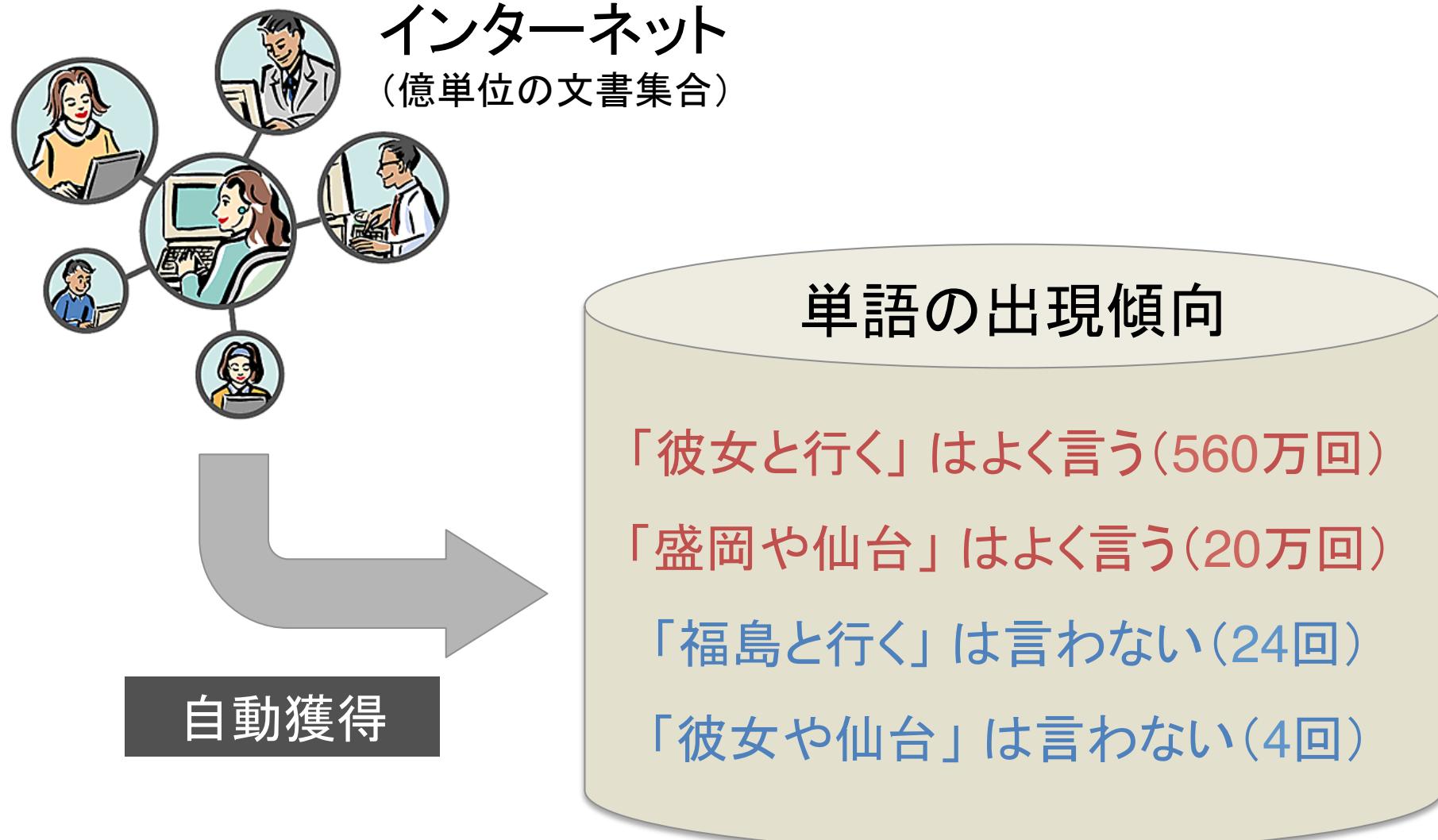
# 基本戦略②

## 単語の出現傾向を大量のデータから獲得



# 基本戦略②

## 単語の出現傾向を大量のデータから獲得



# 「言葉がわかる」とは？ どの意味かがわかる

意味(訳語)の区別：(名詞)の(名詞)

インドの本      *a book ( on ) India*

二日酔いの薬    *medicine ( for ) a hangover*

友人の伝言    *a message ( from ) my friend*

黒澤の映画    *a film ( by ) Kurosawa*

# 「言葉がわかる」とは？ 肯定文・否定文が区別できる

起こったの？ 起こっていないの？

流出しなかった(流出しなかった)

流出したではないか(流出した)

流出しただけない(流出した)

流出を防いだ(流出しなかった)

流出を抑えた(流出した, でも少し)

流出を止めたい(流出した, 今も流出)

# いたる所に省略が...

週末に母とやなぎに行ってみました。  
私は評判のきつねうどんを頼んだのですが、期待通りの味に大満足。

やなぎはうどん屋の名前

やなぎに行ったのは母と私

私はきつねうどんを食べた

味はきつねうどんの味

大満足は美味しかったってこと

# 「行間を読む」

## 常識に基づく推論

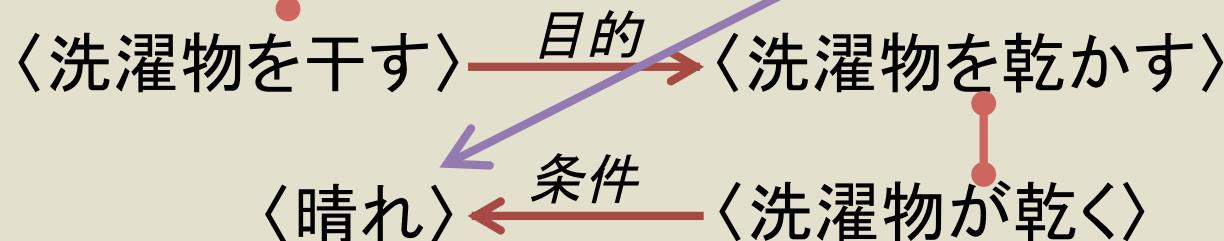


庭に洗濯物を干したとたんに雨が降ってきた

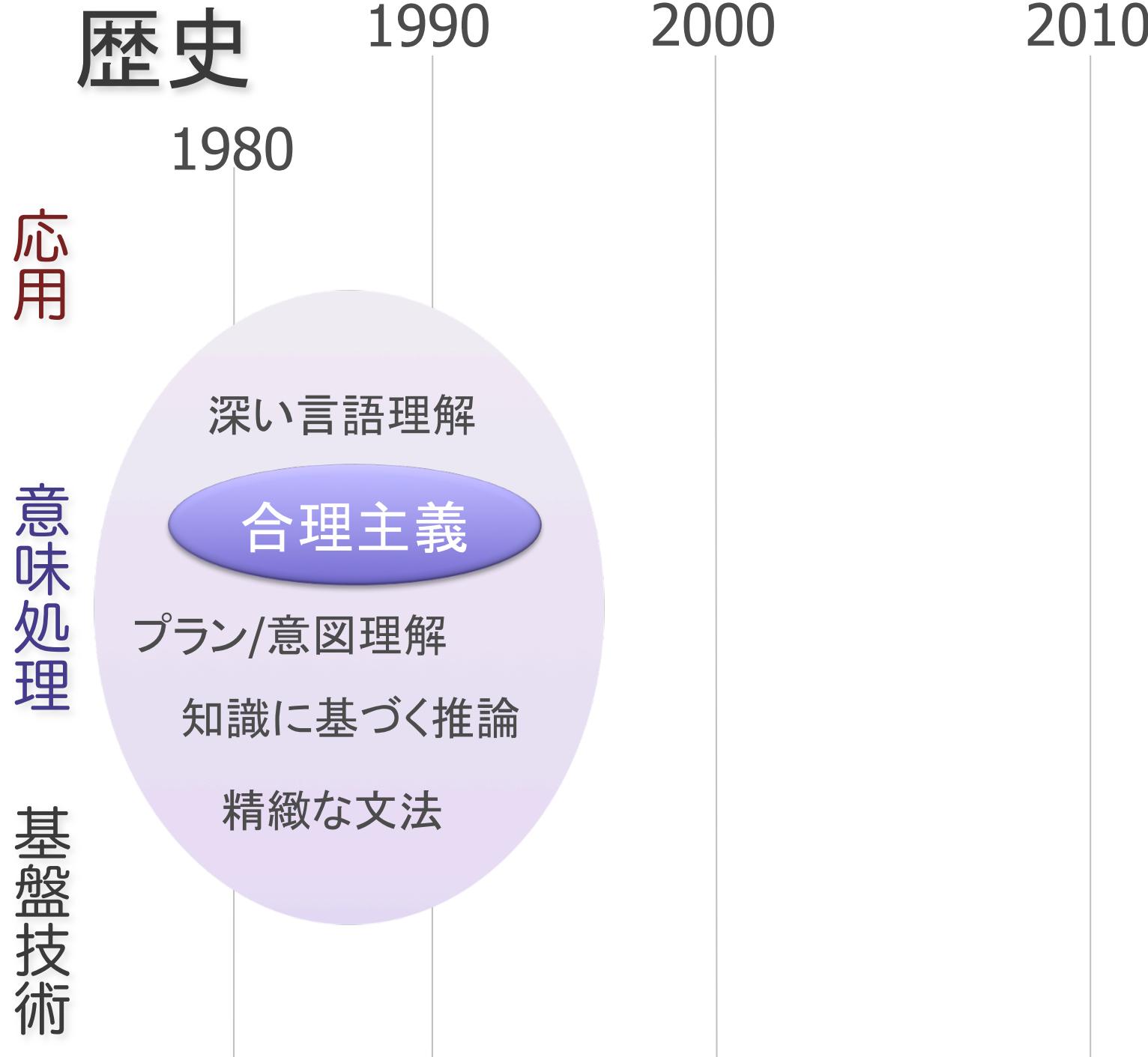
庭に洗濯物を干す

雨が降る

ガッカリ



常識的な  
知識



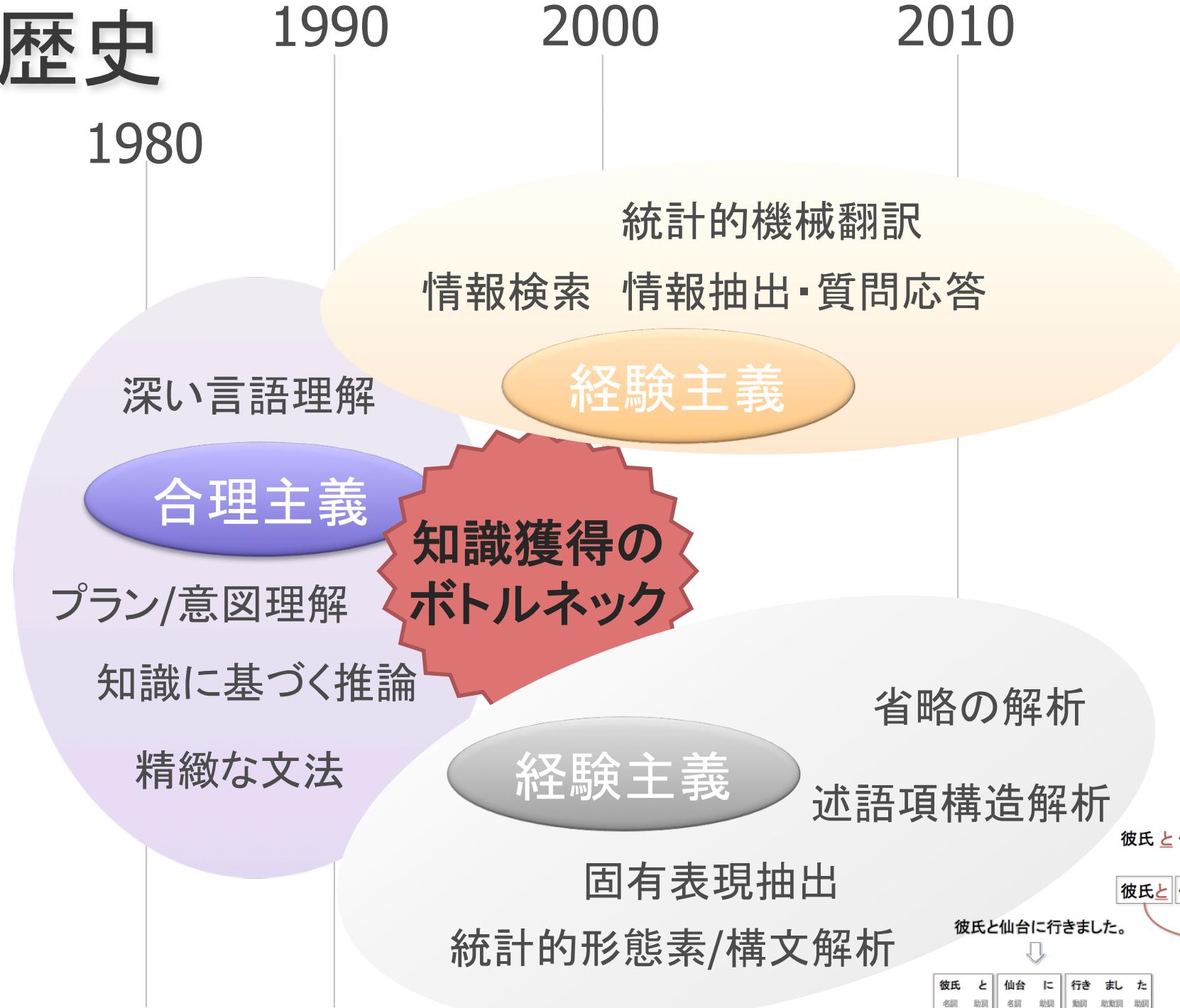
# 大問題！

## 膨大な知識をどうやって集めるか

### 言葉の理解に必要な知識

- 「洗濯物を干す」目的は「洗濯物を乾かす」こと
- 「洗濯物が乾く」条件は「晴れている」こと
- 「うどん」は、「レストラン(うどん屋)で食べる」
- 「(料理を)注文すること」を「頼む」とも言う
- 「レストラン」では、「料理を注文」して、しばらくすると料理が「出て」きて、それを「食べる」

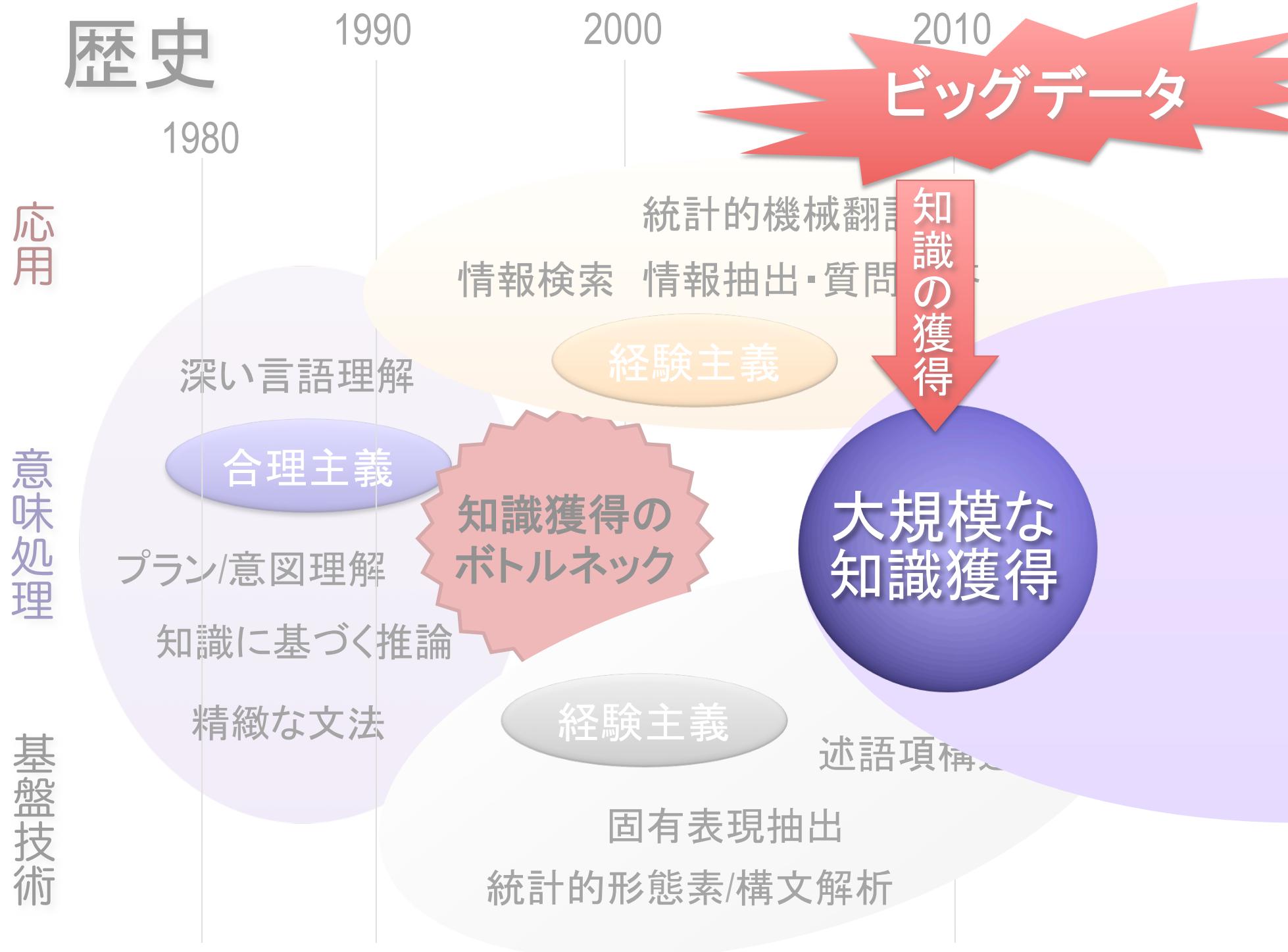
歴史 意味処理 基盤技術 應用



# 実際の解析例

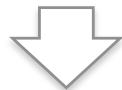
米航空宇宙局は太陽系外に存在する観測史上最小の惑星を発見した。



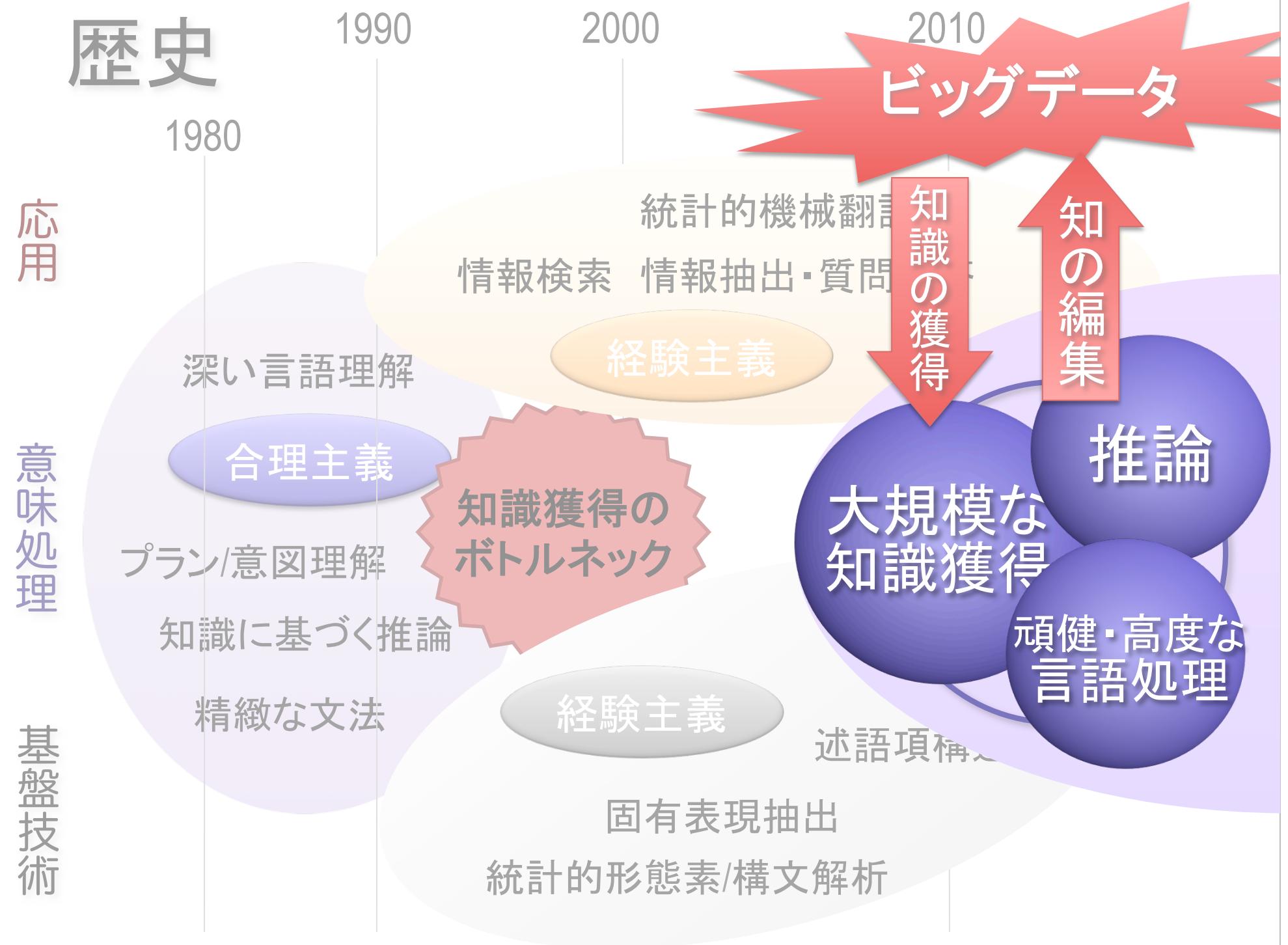


# ビッグデータから知識を自動で集める

天気が良いので、洗濯物がよく乾く  
最近、晴れが続いているので、洗濯物がよく乾く  
日が当たらないので、洗濯物が乾くのもままならず  
お天気は晴れマークが多かったから、洗濯物が乾く  
湿度はなかったから、洗濯物が乾く  
今日は、空気が乾燥していて、洗濯物がよく乾く  
梅雨空が続いたため、洗濯ものが乾かず



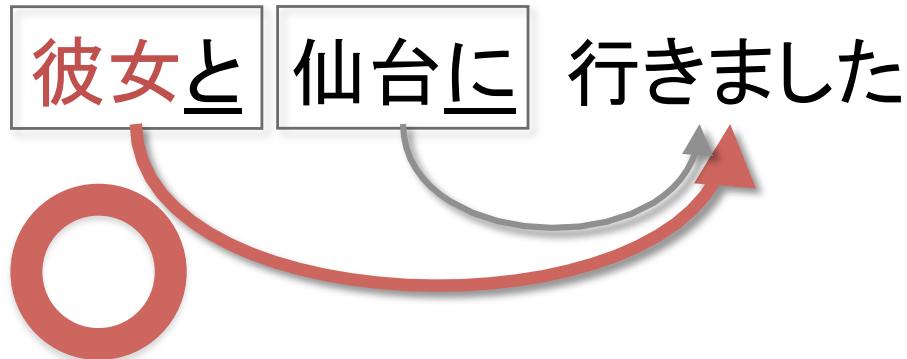
天気が良い(晴れる) → 洗濯物がよく乾く



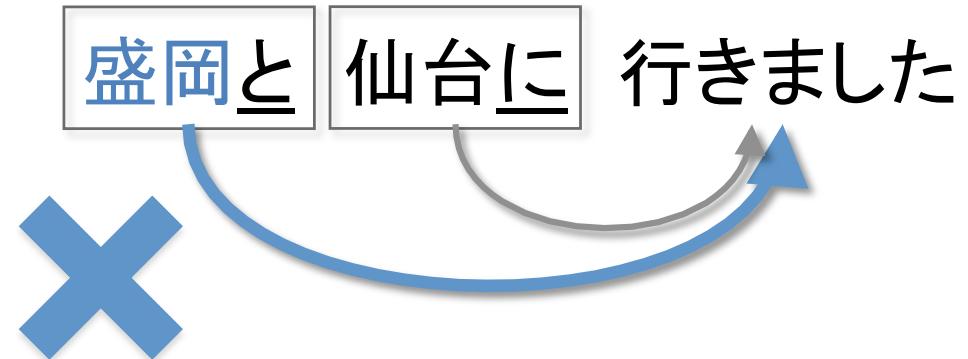
# ビッグデータからの知識獲得で 人工知能にブレークスルーを



# 単語の出現傾向を大量のデータから獲得



「彼女と行く」はよく言う(560万回)



「盛岡と行く」は言わない(23回)



「彼女や仙台」は言わない(4回)



「盛岡や仙台」はよく言う(20万回)

# 単語の意味の知識を自動的に集める

## 定義文から自動で獲得

「○○(△△)は、……である。」

「○○は、……△△とも言う。」

「○○は△△の略。」

「..○○(△△)..」 & 「..△△(○○)..」



The screenshot shows a Japanese Wikipedia page for the International Whaling Commission (国際捕鯨委員会). The page features the Wikipedia logo and navigation tabs for '本文' (Main Text), 'ノート', '編集', and '履歴'. The main content area contains text about the conference held in Japan in 2009, mentioning Jay Walsh and Toshio Nagao. Below this, there is a summary of the International Whaling Commission, its establishment under the International Convention for the Regulation of Whaling, and its status as an international organization. The page is identified as being from the Japanese Wikipedia.

国際捕鯨委員会（こくさいほげいいいんかい、International Whaling Commission: IWC）は、国際捕鯨取締条約に基づき設立された国際機関。日本の条約加入は1951年<sup>[1]</sup>。

# 単語の意味の知識を自動的に集める

使われ方が似ている

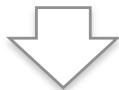


意味が似ている

	を着る	を着せる	に着替える	を着こなす	に包まる	を買う	を売る	を販売する	を作る	で登場する	を食べる	に行く	使われ方 (共起する動詞)
着物	260	36	7	5	6	24	15	9	18	0	0	0	出現頻度
和服	97	0	11	6	8	2	2	4	0	4	0	0	
着物	20.1	3.7	3.0	2.6	1.7	1.5	1.3	1.2	1.1	.09	.00	.00	割合
和服	28.0	5.0	4.1	3.6	2.2	2.0	1.7	1.6	1.5	1.1	.00	.00	

# 常識もビッグデータから集める

天気が良いので、洗濯物がよく乾く  
最近、晴れが続いているので、洗濯物がよく乾く  
日が当たらないので、洗濯物が乾くのもままならず  
お天気は晴れマークが多かったから、洗濯物が乾く  
湿度はなかったから、洗濯物が乾く  
今日は、空気が乾燥していて、洗濯物がよく乾く  
梅雨空が続いたため、洗濯ものが乾かず



天気が良い(晴れる) → 洗濯物がよく乾く

# Watson 米クイズ番組で歴代王者に勝利



<http://www.youtube.com/watch?v=Wq0XnBYC3nQ>

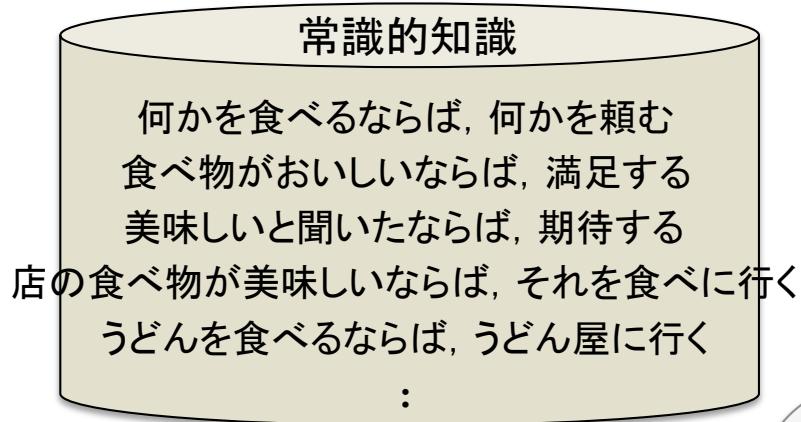
# Watson 米クイズ番組で歴代王者に勝利

**WILLIAM WILKINSON'S  
"AN ACCOUNT OF THE  
PRINCIPALITIES OF  
WALLACHIA AND  
MOLDAVIA" INSPIRED  
THIS AUTHOR'S MOST  
FAMOUS NOVEL**

<http://www.youtube.com/watch?v=Wq0XnBYC3nQ>

ウィリアム・ wilkinson の『 ヴィラキアとモルダビア公国  
の理由 』に最も刺激を受けた 19 世紀の小説家は ?

# 「行間を読む」仮説推論



やなぎはうどん屋の名前  
やなぎに行ったのは母と私  
私はきつねうどんを食べた  
味はきつねうどんの味  
大満足は美味しかったってこと

母とやなぎに行った。評判のきつねうどんを頼んだ。期待通りの味に大満足。

# 「行間を読む」仮説推論

常識的知識

何かを食べるならば、何かを頼む  
食べ物がおいしいならば、満足する  
美味しいと聞いたならば、期待する  
店の食べ物が美味しいならば、それを食べに行く  
うどんを食べるならば、うどん屋に行く  
：

大満足は美味しかったってこと

美味しい(u3, z) 聞いた(u4, x, e7)

私はきつねうどんを食べた

やなぎはうどん屋  
の名前

食べる(u1, x, z)

味はきつねうどんの味

食べ物(v)

店(w) of(w, z) 行く(u2, x, w)

私(x)  $\wedge$  やなぎ(y)  $\wedge$  行く(e1, x, y)  $\wedge$  評判(e2, z)  $\wedge$  きつねうどん(z)  $\wedge$  頼む(e3, x, z)  $\wedge$  期待通り(e4, v)  $\wedge$  味(v)  $\wedge$  満足する(e5, x, v)

母とやなぎに行った。評判のきつねうどんを頼んだ。期待通りの味に大満足。

世界最速の説明推論システムを開発(東北大)

# 東北大の取り組み

## 入力(観察)

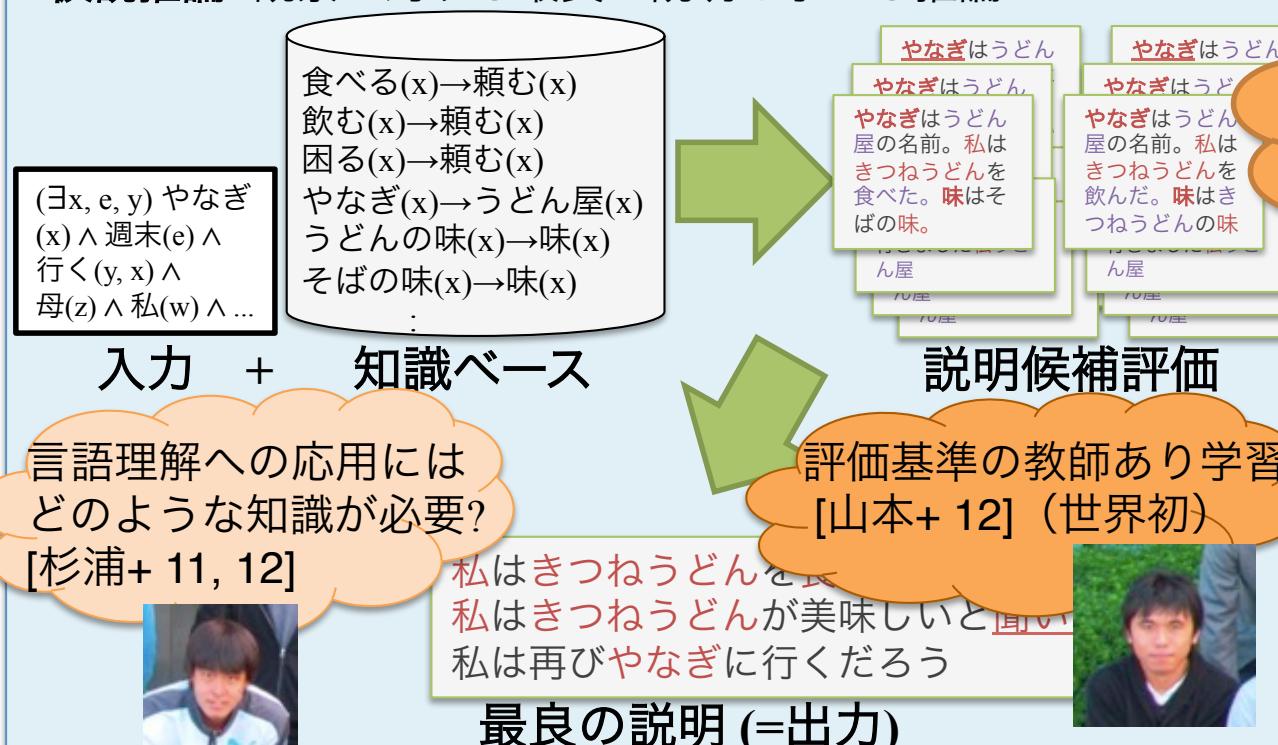
週末に母とやなぎに行きました。  
私は評判のきつねうどんを頼みました。  
期待通りの味に大満足。

出力 (説明)

## 意図, 原因, 結果の予測

### ロジックベースの推論（仮説推論）

**仮説推論:** 観察に対する最良の説明を求める推論



## 世界最速の最良解探索 [Inoue & Inui 11a, 11b, 12]

入力: **20**リテラル

知識ベース: **37万**のルール  
説明候補の数:

$$2^{1000} \approx 1.07 \times 10^{301} \text{ 個}$$

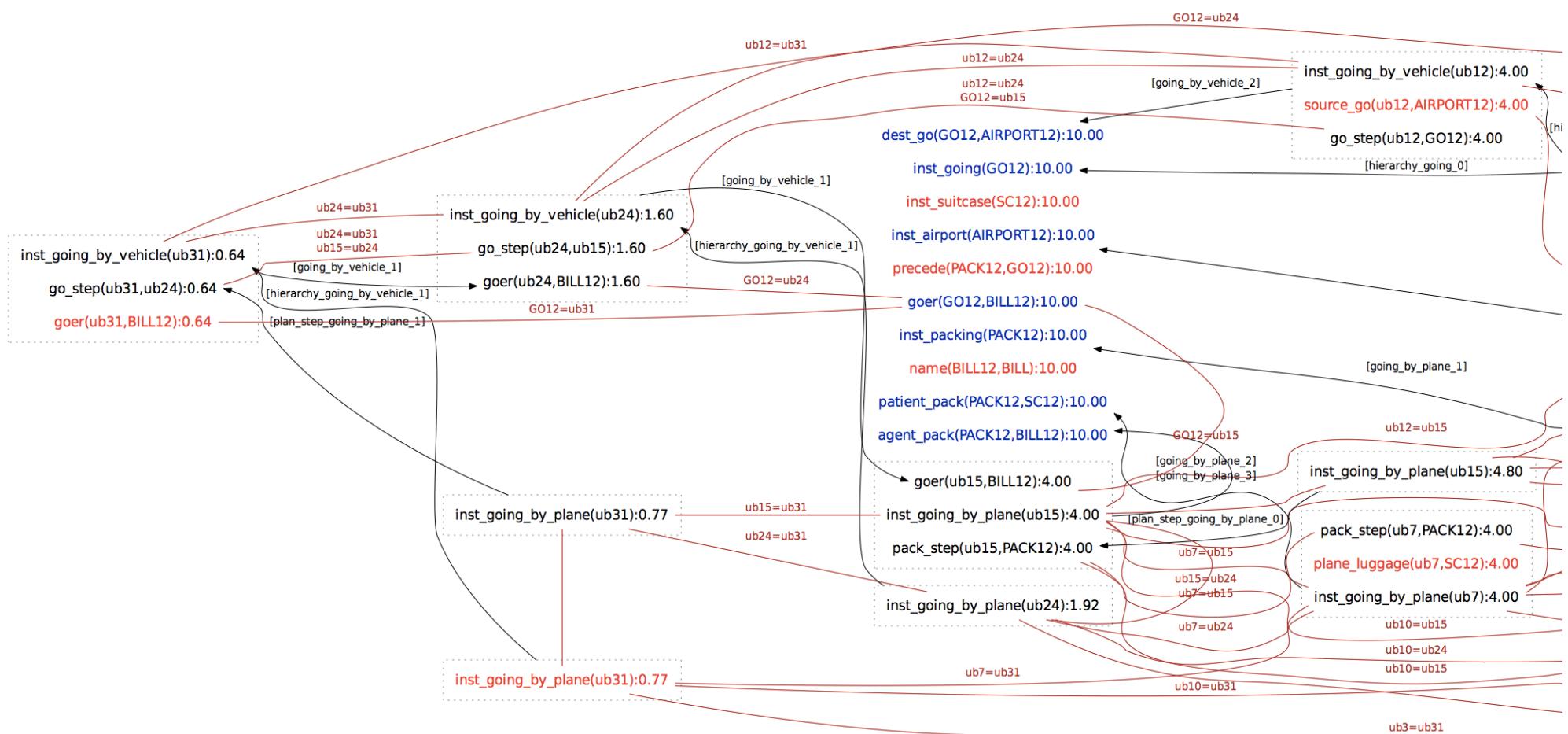
[Inoue & Inui]: 3秒

既存研究 [Blythe+ 11]: 420秒

既存研究 [Mulkar-Mehta+ 07]: 1

- N. Inoue and K. Inui. ILP-based Reasoning for Weighted Abduction. In Proceedings of AAAI Workshop on Plan, Activity and Intent Recognition, 2011. (Book Chapter 執筆依頼を頂く)
  - N. Inoue and K. Inui. An ILP Formulation of Abductive Inference for Discourse Interpretation. In IPSJ SIG Technical Reports, 2011. (山下記念研究賞候補)
  - N. Inoue and K. Inui. Large-scale Cost-based Abduction in Full-fledged First-order Predicate Logic with Cutting Plane Inference. submitted to JELIA 2012.
  - 杉浦純, 井之上直也, 乾健太郎. 大規模語彙知識を用いた仮説推論による文章理解モデルの構築に向けて. NLP若手の会 第6回, 2011. (優秀発表賞受賞)
  - 杉浦純, 井之上直也, 乾健太郎. 説明生成に基づく談話構造解析の課題分析. 言語処理学会第18回年次大会論文集, 2012.
  - 山本風人, 井之上直也, 渡邊陽太郎, 岡崎直觀, 乾健太郎. 誤差逆伝播を利用した重み付き仮説推論の教師あり学習. 情報処理学会研究報告 (学生奨励賞受賞)

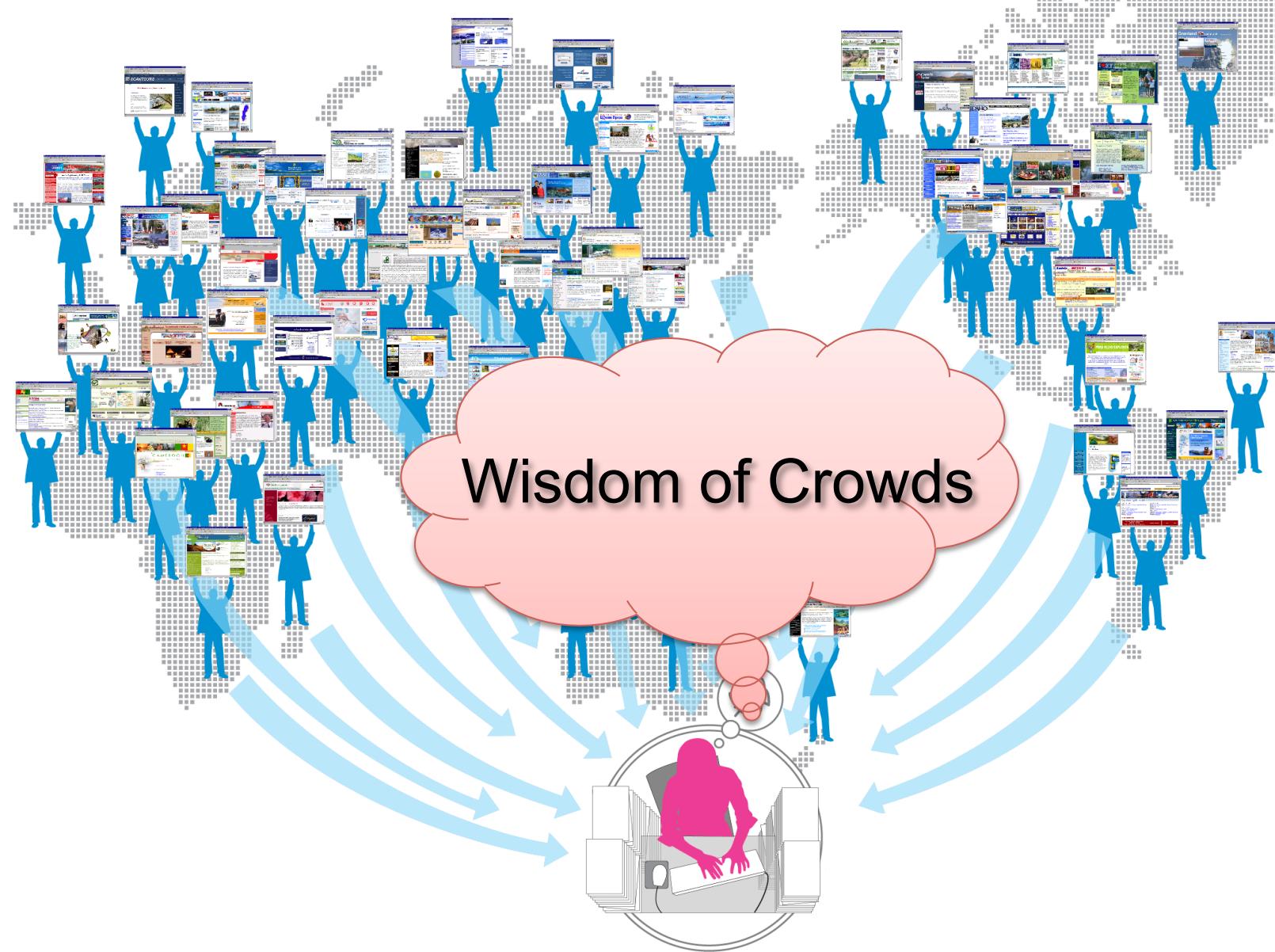
# 仮説推論による文章解釈の様子



# 情報の信頼性と編集



# 情報の信頼性をどう担保するか



# ネットにはびこるウソ

マイナスイオンを取り込むと、酸が中和されて、弱アルカリ性に戻るので、疲労を取り除くことができます

ストレス解消/マイナスイオンはストレスにいいらしい

マイナスイオン。今ではすっかり生活にて何?スポーツ選手が『マイナスイオ』では『マイナスイオン』が当たりついわれている某大手車メーカーが機を導入したとか。なんかストレスにでは?

空気中のイオンの量を測定すると、森林の中や滝の近くで最も豊富なことが分かります。

黒豆専門店「丹波黒豆本舗」

体にとってもいい『マイナスイオン』

ダイエット・シェイプアップ

マイナスイオンはプラスイオンを抑え、身体の疲労を回復させるという、身体に欠かせない大切な物質で、

血液のベーハーとイオンバランス

体内でも、通常はマイナスイオンとプラスイオンのバランスが取れて、弱アルカリ性に戻るので、疲労を取り除くことができます。

マイナスイオンが豊富な森林や滝の近くにいると、疲労感が

ヒアルロン酸  
シミ・小じわ  
コエンザイムQ10  
セルライト対策  
健康  
青汁  
黒豆  
ブラックジンジャー

レナード効果

自然界における唯一のマイナスイオン化現象で、ドイツのフィリップ・レナード博士が発見したためこの名があります。空気中で微細な水滴が分裂する時、水滴はプラスイオンに帯電しますが、周囲の空気はマイナスに帯電し

言られており、本来私たちの健康維持のためになくてはならないもの。

しかし一般的な私たちの生活空間を満たす空気に含まれているのは、残念ながらそのほとんどがプラスイオンなのです。一般には、プラスイオン過多の状態だとめまい、頭痛、いろいろするなどのストレス状態が続き、よって極度の疲労などの症状が現れやすくなると言われています。

# ウソを見破るページも



## マイナスイオンが健康増進に役立つ と言う科学的証明はされていません

### 「マイナスイオンとは」

ここは導入です。現在の豊かな生活の代償として、大気汚染、砂漠化現象、オゾン層の破壊、電磁波の増加などがあるとし、それらに共通するキーワードはプラスイオンだそうです。この導入からして賢明な視聴者は気付いたことでしょう。「ははん、これはマイナスイオン発生器の販促番組だな」と。しかし疑うすべを知らない大部分の視聴者は踊らされるのが目に見えています。

そして空気のイオン量は大自然でプラスが50、マイナスが2500だそうで、これが都会になるとプラスが80、マイナスが1200と減っているそうです。ここではイオン一辺倒になっていますが、アスファルトで埋め尽くされ、排気ガスや工場のばい煙などで空気が汚く、自然の植物や土が持つ自浄作用のない都会は確かに健康には芳しい所ではありません。生活するには多少の不便はあっても田舎がいいものです。私の町は四方を山に囲まれていて、空気はきれいだし水もうまい。ただ重大な問題は若い女性がいないということです。ですが、いいですよ田舎は(泣)

増強する働きがあるとの健康維持のために無

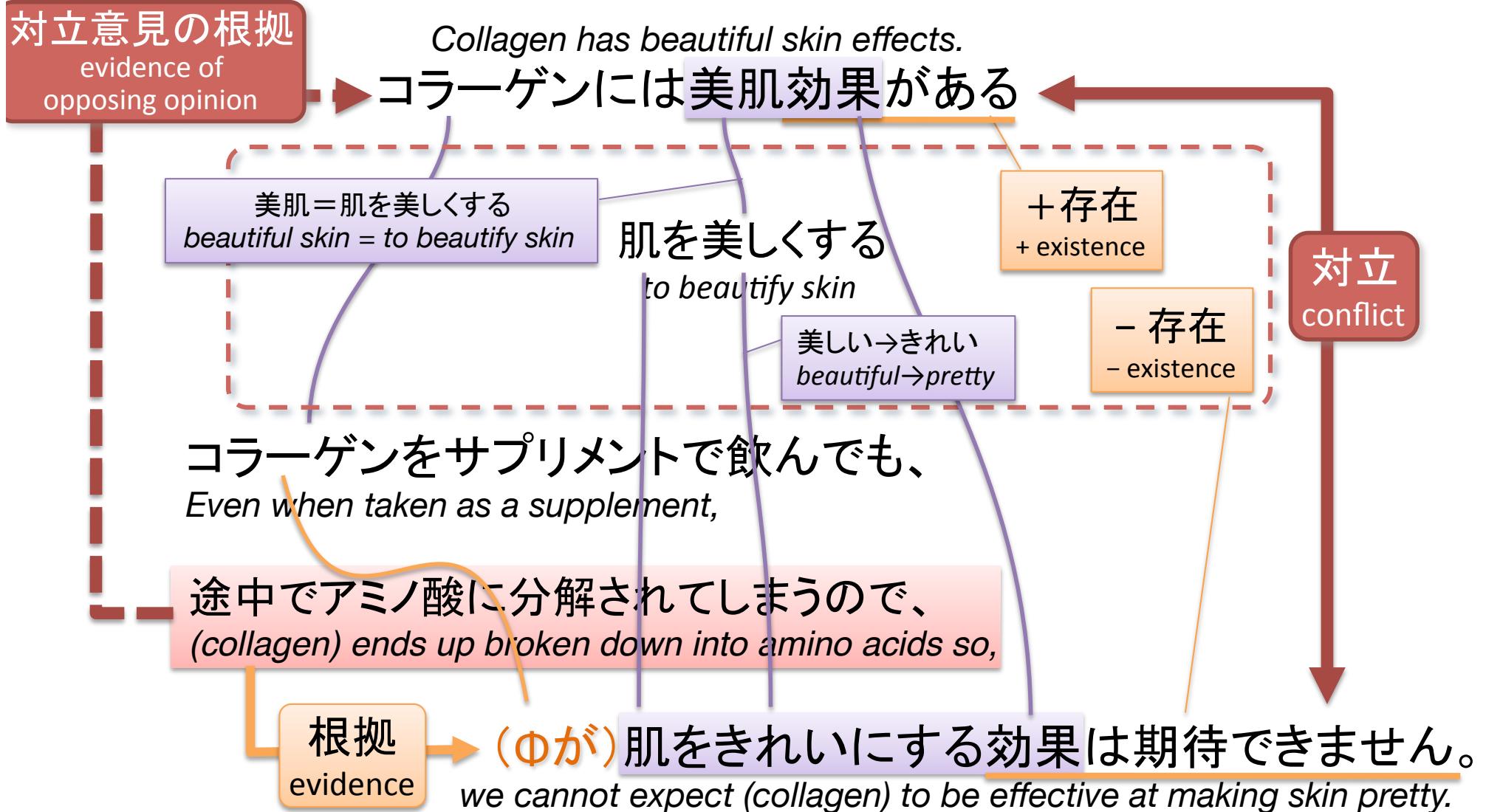
活空間を満たす空気にがらそのほとんどがブには、プラスイオン過、いらっしゃるなどの状態が続き、よって極度の疲労などの症状が現れやすくなると言われています。

# 言論マップ

コラーゲンは肌に良い	
<p>質問： コラーゲンは肌に良い</p> <p>フェミダンプラスはコラーゲンとヒアルロン酸の働きで、</p> <p><b>同意</b> (459文) 噛むだけでお肌つやつや！若返りと美肌を提供します log1 log2</p> <ul style="list-style-type: none"><li>・コラーゲンは健康な肌ならきれいな網目構造になっており、まるでベッドのスプリングのように<b>真皮を支えて強力のあくび肌を保っています</b> log1 log2</li></ul>	<p>ただ、研究職の私から言わせてもらうと、コラーゲンは体に入ったからといって、肌に潤いをもたらすわけではありません log1 log2</p> <p>・お肌にハリがなくなってきたり、たるんできてるのかな、と思うときにコラーゲンやヒアルロン酸を注入する方法があります。この方法は、肌にハリをもたらす効果があります。</p>
<p>フェミダンプラスはコラーゲンとヒアルロン酸の働きで、</p> <p>咬むだけでお肌つやつや！若返りと美肌を提供します log1</p>	<p>log1 log2 more</p> <p>ゲンですが、肌かたしわに対しての効</p>
<p>が良く、他の哺乳類・魚類とは違って、人間の肌に対して類をみないなじみがあるので</p> <p>す log1 log2 more</p> <p>根拠 (20文) を</p> <ul style="list-style-type: none"><li>・生まれたばかりの赤ちゃんの肌は、大人の肌よりも柔軟で、肌の表面には細かい溝があります。これは、肌の表面に存在するコラーゲン分子が、肌の表面に存在するヒアルロン酸分子と一緒に、肌の表面を覆っているからです。</li><li>・また、アミノ酸含量が高く、肌の表面に存在するコラーゲン分子が、肌の表面に存在するヒアルロン酸分子と一緒に、肌の表面を覆っているからです。</li><li>・コラーゲンはお肌の水分保持機能を高め、お肌の水分を保つのに役立つのです。</li></ul>	
<p>ニッピコラーゲンといったコラーゲンですが、肌から浸透しないので、出来てしまったりしわに対しての効果は期待できません log1 log2</p>	<p>う事なので一緒にビタミンCやミネラルをとりましょう log1 log2 more</p> <p>美肌づくりに欠かせないコラーゲンですが、実はコラーゲンは分子が大きいので、お肌に入りにくい！という難点があります log1 log2</p>

# 同意・対立・根拠の認識

構造的アライメント + モダリティ(事実性)解析 + 関係パターン認識



# 言論マップで利用している語彙知識

- 日本語WordNet(Ver.1.0; Bond+ 2009)：56,741 synsets, 92,241 語
  - 過剰：冗多,過多,冗長,余った,余計,余分
  - 効果：効目,実効,効,利き目,効き目,利き,効力,効き,功能,效能
  - 病気 ⇔ 具合い, 丈夫, 健康, フィットネス, 健全, 多幸
- 事象間関係知識(Matsuyoshi+ 2008)：約140万対
  - <誰か>が<何か>を攝取する – <誰か>が<何か>を取る（上位語）
  - <何か>が普通だ – <何か>が異常だ（反義）
- Wikipediaから抽出した同義語：113,401 synsets
  - HIDランプ–高圧放電灯–高輝度放電ランプ
  - バイオ燃料–エコ燃料–バイオエネルギー
- Webから抽出した上位下位関係(Sumida+ 2008)：約350万対
  - イソフラボン–女性ホルモン様成分, キシリトール–甘味添加物
- 文脈類似度に基づく動詞間含意関係(ALAGIN)：約170万対
  - 引き起こす–もたらす, 予防する–防げる, ケアする–防止する

# 事実性の推定

そのイベントは起こったのか起こっていないのか

流出しなかった（流出しなかった）

流出したではないか（流出した）

流出しただけではない（流出した）

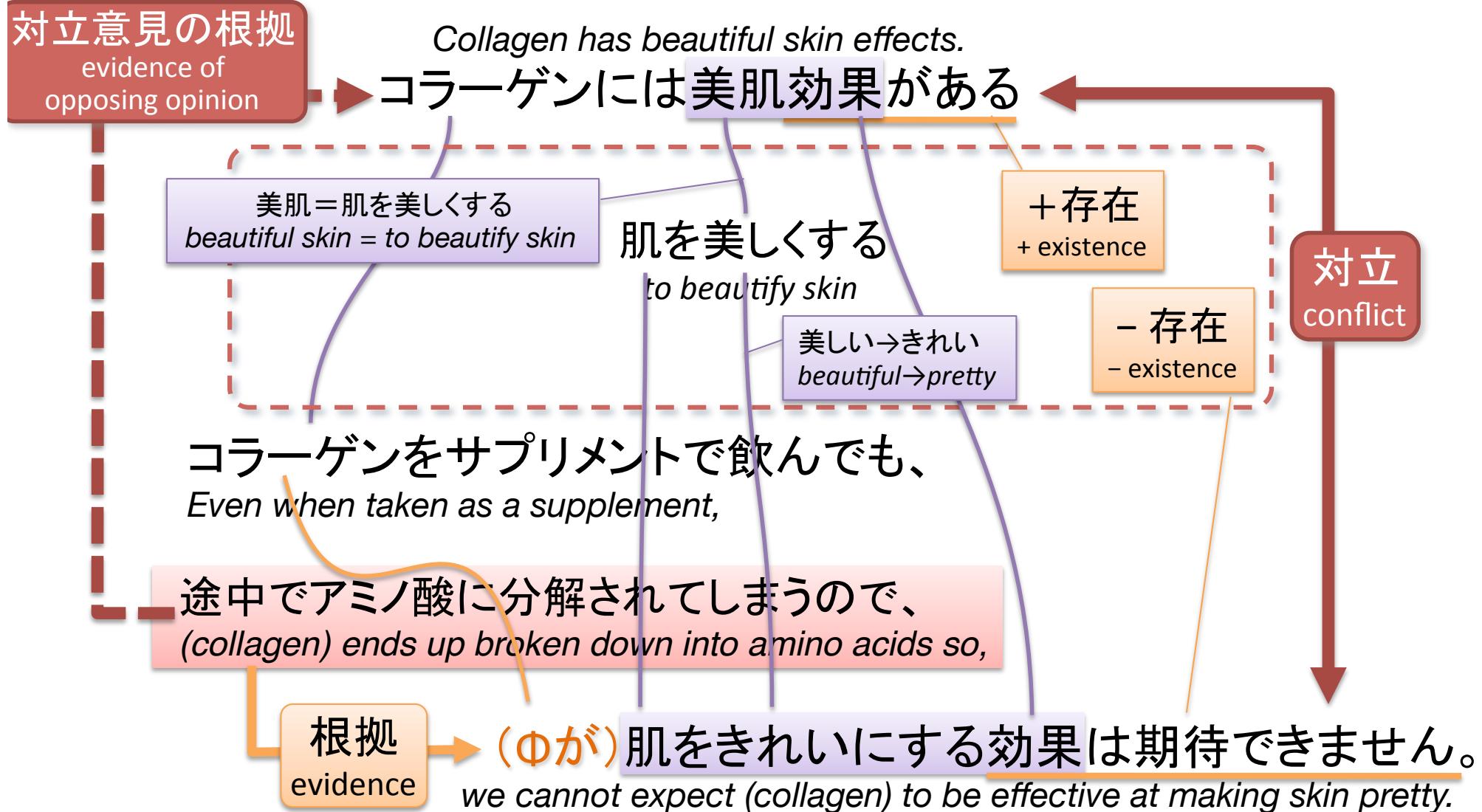
流出を防いだ（流出しなかった）

流出を抑えた（流出した, でも少し）

流出を止めたい（流出した, 今も流出）

# 同意・対立・根拠の認識

構造的アライメント + モダリティ(事実性)解析 + 関係パターン認識



# うがいはインフルエンザの予防に効果的だ

賛成意見の根拠

(33文)

反対意見の根拠

(6文)

あと、現在の豚インフルエンザの特徴は、日本でも冬に流行する、インフルエンザA型に似ていて、湿気・熱に弱いので、マスク・手洗い・うがいでかなりの予防はできるとのことです [log1](#) [log2](#)

レスは、粘膜に接着後、わめて短時間で粘膜内に進入効果は低いと考えられています。予防にはなります [log1](#) [log2](#)  
耳より健康教室  
もほぼ落ち着き始めたのでB型の流行もまだ懸念される

入するということですから、うがいは予防には効果がないともいわれています。  
漢方110腰痛・関節痛・肩こり

インフルエンザは、咳やくしゃみによってヒトからヒトへの着用、うがい・手洗いは、それにも増して、インフルエンザ接種が大切です [log1](#) [log2](#)  
健康長寿ネットーインフルエンザ

またインフルエンザは、いた人の咳やくしゃみに含まれることによって感染するので、

**よくうがい励行！**と言われますが、インフルエンザウイルスは咽頭（のど）、眼、また鼻粘膜に付着すると約20分で細胞中に取り込まれてしまうので、外出後のうがいは予防には無効であり、外出時にマスクをすることをお勧めします [log1](#) [log2](#)

インフルエンザウイルスは、粘膜に接着後、20～30分というきわめて短時間で粘膜内に進入しますので、うがいの効果は低いと考えられていますが、ある程度の予防にはなります [log1](#) [log2](#)

# 賛否によるネット情報の自動分析は世界初

言語処理学会 最優秀発表賞(2010)

言語処理学会 若手奨励賞(2010)

CICLing Best Paper Award(2011)

日経新聞技術トレンド調査 総合3位(2011)

情報処理学会 山下記念研究賞(2012)

2011年(平成23年)5月30日(月曜日)

## 目的鮮明な产学連携評価

2011年度技術トレンド調査

日本経済新聞社は、主な技術開

技術トレンド調査総合順位1~10位			
順位	研究開発者 (発表日)	成果の概要	評価点
1	東京大学、コニカミノルタエムジーエヌ(2)	軟骨映すX線装置開発、 教授の成績は、装置は、1位の	X 11.14
2	日立製作所 (2010年12月)	NEC、東北大学、 奈良先端科学技術大 学院大学、横浜国立 大学 (1月31日)	技術トレン X 11.14
3	NEC、東 奈良先端科 学院大学、	天然ウナギの卵発見、世	日本経済新聞社は、主な技術開

# 耐災害ICTへ



# 震災直後に流れたデマ



@jusminrom

ジャスミン・ロメロ@高3

【被爆予防方法】イジソを水に薄めて大量に飲んで下さい。特に子供に。自然ヨウ素で甲状腺を満たし、放射線ヨウ素の吸収を抑えます。#jishin

3月13日 Keitai Webから

# デマを否定する情報も



サニー  
@ZxCG



フォローする

放射能対策でヨウ素剤の代わりにイソジン3滴という全くのデマが流れてるので止めてください。イソジンだと14.3cc飲む必要が有りますが、こんなに飲んだら胃腸がただれます。<http://bit.ly/h55YA1>

# 支援要請の過剰拡散も



13番

@13\_ban



フォローする

《緊急：拡散希望》 当院で  
はICUのみ自家発電中です。  
燃料のA重油も朝7時でなくな  
り、人工呼吸器も止まってし  
まいります。 A重油の入手先を  
ご教示ください。

0222482131 (仙台市太白区  
広南病院用度課電気室)

#Sendai #earthquake

#eqjp #jishin

←返信 →リツイート ★お気に入りに登録

273

件のリツイート

2

FAVORITES



# 支援要請の過剰拡散も



# 震災後のツイートデータで

言論マップシステム

www.cl.ecei.tohoku.ac.jp:8080/

質問

イソジンで被曝を防げる？

2011-03-12 23:06:49

@anonymous イソジンで被曝を防げる

肯定

39 ★根拠 21

否  
定

12 ★根拠 1

イソジンを数滴たらした水を飲んでおけばヨウ素が甲状腺に吸収されるので少しでも被曝を防げます shinji - 2011-03-12 10:47:41

イソジンだと14.3cc飲む必要がありますが、こんなに飲んだら胃腸がたかれますので、被曝対策には使えません masa\_dental - 2011-03-14 16:06:24

被曝を軽減するにはイソジンがいい

ナヨヒン【薬剤師】 - 2011-03-12 23:41:40

否定

イソジンを飲むと言う間違った危険な噂が出

けんたの主人 - 2011-03-15 02:55:41

被曝が心配ならイソジンか乾燥昆布買っとくといい 奈良岡 俊 - 2011-03-12

イソジンは被曝対策には使えません

被曝を軽減するには昆布とイソジンがいいみたいだね aylin nakamura - 2011-03-12 22:42:04

被曝による甲状腺がんを防ぐにはイソジン飲めとかいう出どころ不明のツイートがあるけど限りなくデマくさい

# ツイートデータを意味的に解析

市原市のコスモ石油千葉製油所LPGタンクの爆発により、千葉県、近隣圏に在住の方に有害物質が雨などと一緒に飛散するという虚偽のチェーンメールが送られています。

【リツイートで拡散希望】 コスモ石油の爆発で有害物質の雨が降る件はデマ。広げてしまった方はツイート削除の上、訂正を

コスモ石油の爆発による有害物質流出のおそれの件、コスモ石油から正式にアナウンス出てます。ツイッターで出回ってるような事はないそうです。HPは繋がりにくいのでスクリーンショットです。

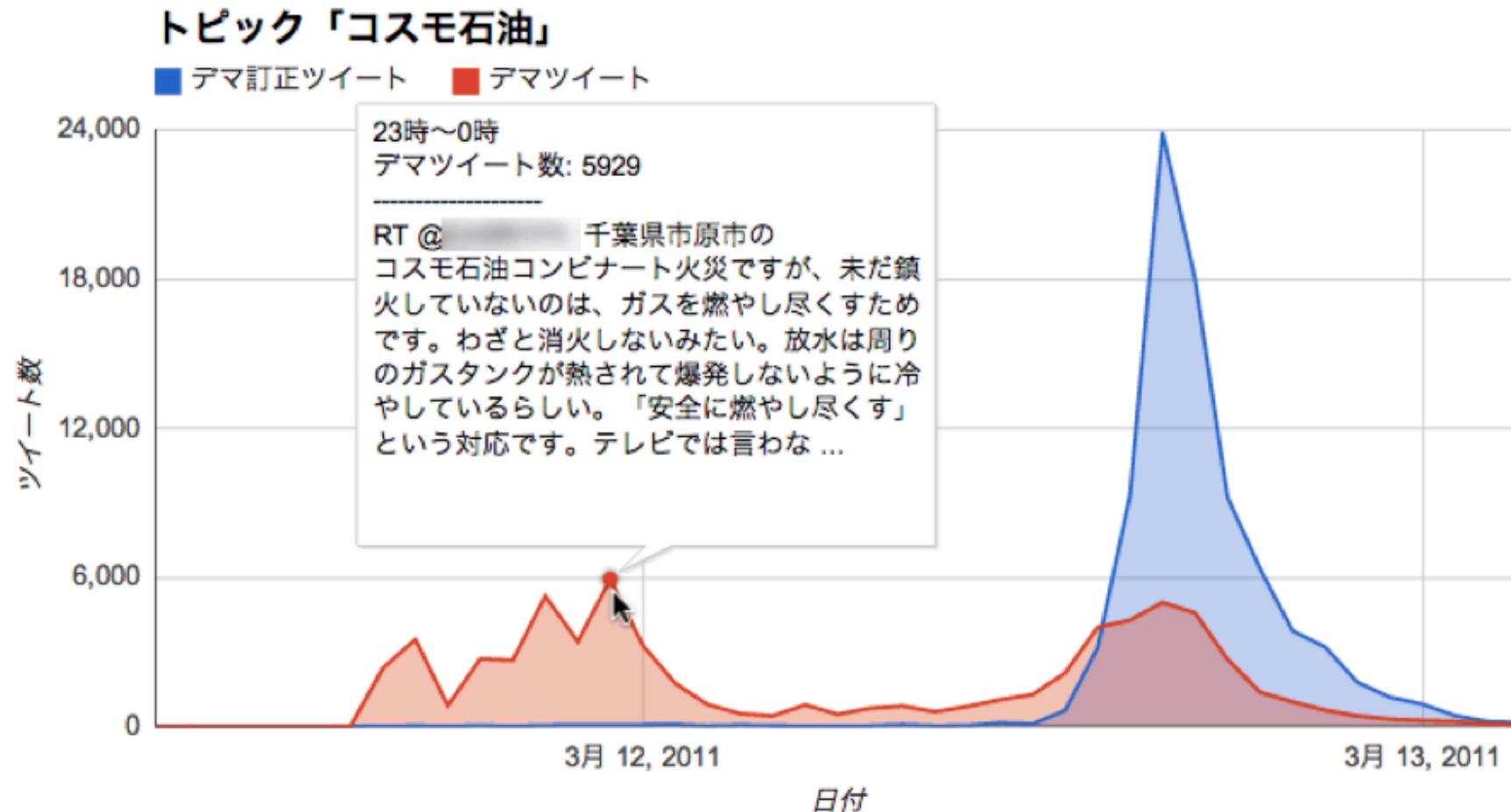
# 誤情報を自動収集

よく反論されている命題をマイニング

トピック	スコア	命題
田尻智さん	18.7	海外で、ポケモンの生みの親の田尻智さんが亡くなった
コスモ石油千葉製油所	11.8	@市原市のコスモ石油千葉製油所LPGタンクの爆発により、千葉県、近隣圏に在住の方に有害物質が雨などと一緒に飛散する
女性暴行	10.4	@「阪神大震災の際には女性暴行が増えた」
有毒物質	8.6	コンビナート火災に関し『有毒物質が発生し、雨に混じって降ってくるので肌をさらさないように』
尾田栄一郎先生	5.3	尾田栄一郎先生が15億円を寄付
円資産	5.2	池田経済産業副大臣は、「日本企業が円資産を買い戻している
空中投下	3.0	『物資の空中投下が日本で認められていない』
サーバルーム	1.9	サーバルーム、ラックで腹部を怪我した
イソジン	1.6	被爆予防にイソジンを飲むと良い
茨城県知事	1.2	茨城県知事災害派遣要請出してない
食塩	1.0	「食塩にはヨウ素が含まれており被曝の防止に効果がある」

# 誤情報の「一生」を可視化

## 誤情報の拡散と訂正の拡散を時系列分析



以下のボタンをクリックすると拡大・縮小します

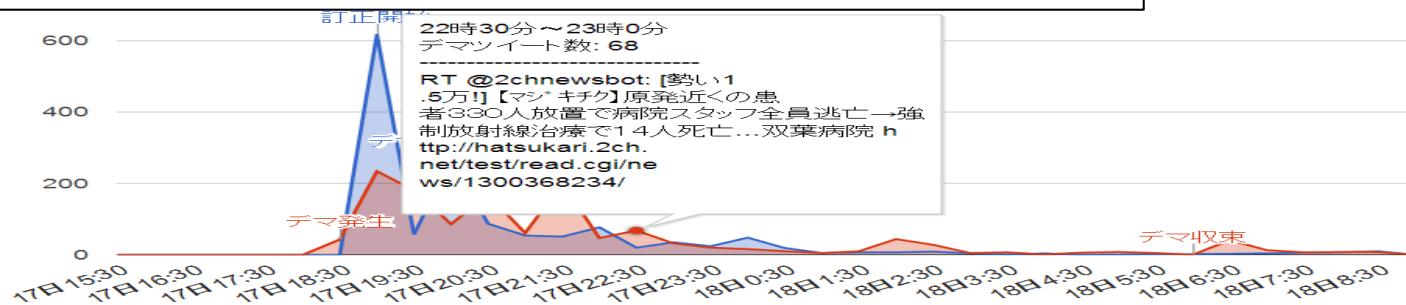
[11日9時-13日3時](#) [13日3時-14日21時](#) [14日21時-16日15時](#) [16日15時-18日9時](#) [全体に戻す](#)

# 様々なパターンがあることも明らかに

## コスモ石油: 5時間で訂正, 24時間で収束



## 双葉病院: 30分で訂正, 11.5時間で収束



## 三条中学校: 30分で訂正, 収束せず



# 返信・引用による論述関係認識

@\*\*\* 925RTs 2011-03-11 23:26:35

コスモ石油の爆発により有害物質が雲などに付着し、雨などと一緒に降るので外出の際は傘かカッパなどを持ち歩き、身体が雨に接触しないようにしてください！！！ コピペとかして皆さんに知らせてください。

@▽▽▽ 442RTs 2011-03-11 19:58:50

フジテレビで言ってましたな RT@\*\*\*:【拡散希望】千葉市近辺に在住の方！ コスモ石油の爆発により有害物質が....雨に接触しないようにしてください！！！

検出された  
(反論)関係



〈類似〉する  
ツイート



デマツイート  
のクラスタ

@☆☆☆ 0RTs, 2011-03-11

23:49:24 ガセネタらしいです(>\_<) RT @\*\*\*: 千葉市近辺に在住の方！ コスモ石油の爆発により有害物質が雲などに付着し、雨などといっしょに降るので外出の際は傘かカッパなどを持ち歩き、身体が雨に接触しないようにしてください！！！

訂正ツイート  
のクラスタ

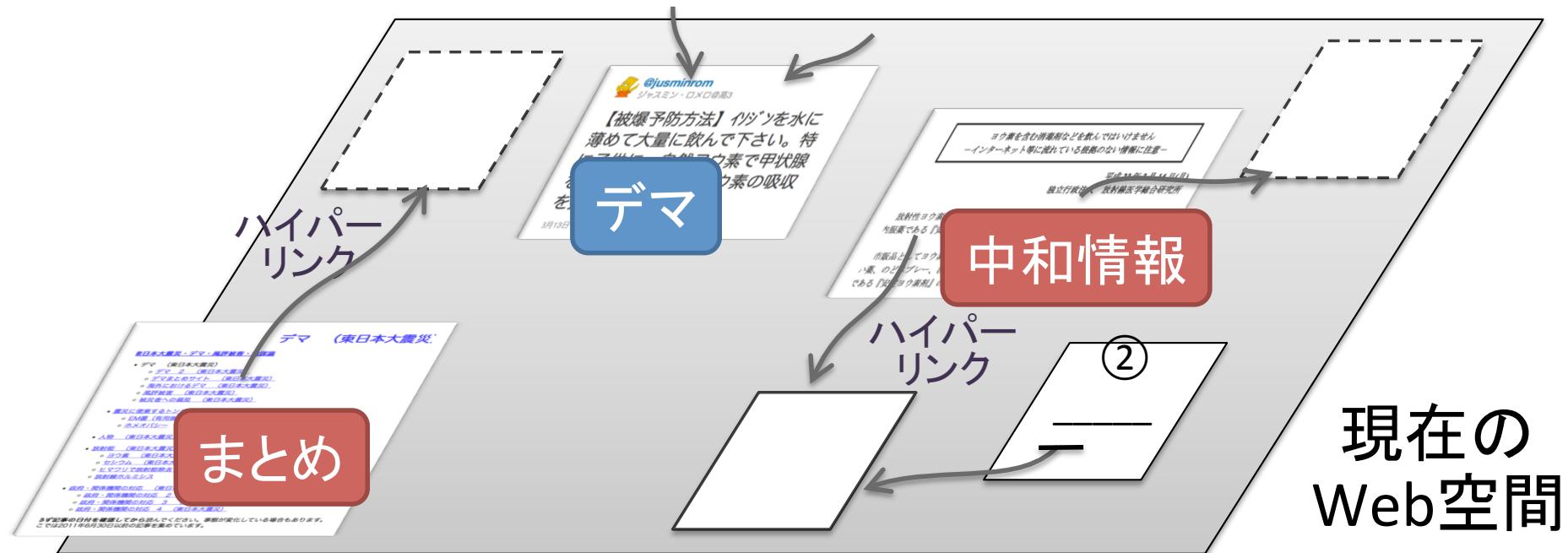
訂正ツイート  
のクラスタ

@◇◇◇ 1676RTs, 2011-03-12 14:46:29

【重要】コスモ石油株式会社から「コスモ石油の爆発...降る」という情報について、「このような事実はありません」との発表。http://www.cosmo-oil.jp/...

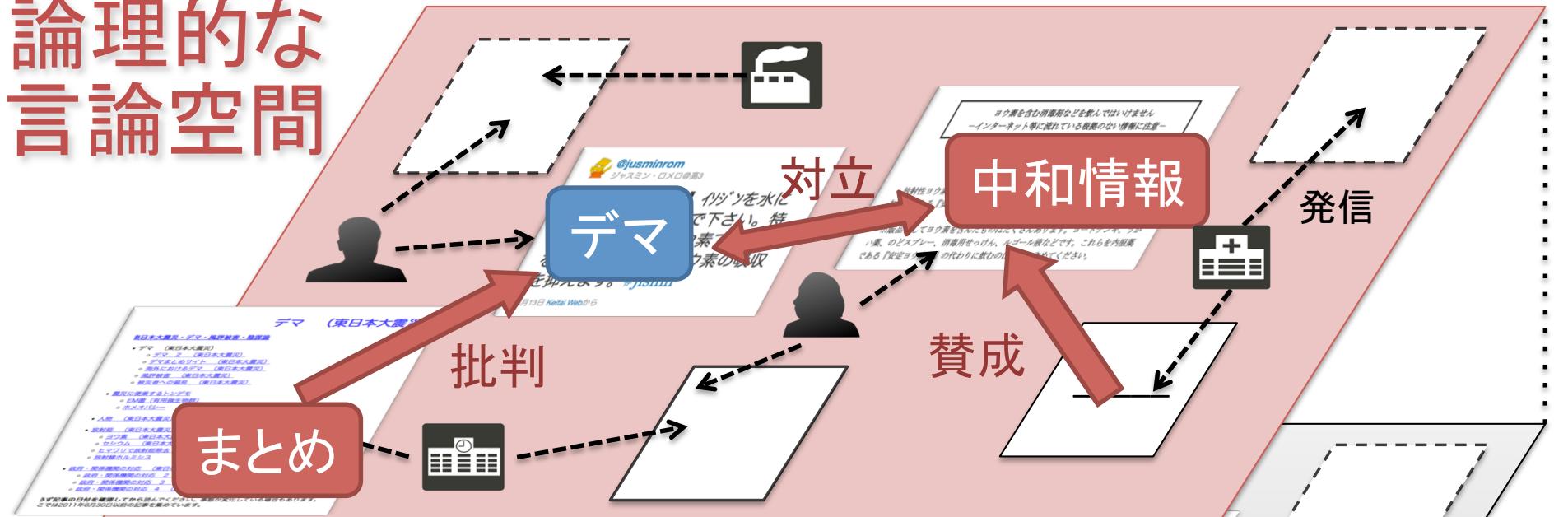
# ウェブの情報の流れを変える

誤情報を見ている人に  
中和情報が伝わらない！



# ウェブの情報の流れを変える

論理的な  
言論空間



ハイパー  
リンク

まとめ



デマ

中和情報

ハイパー  
リンク

②

現在の  
Web空間

# 言論マップ2G

## 命題ネット

(controversialな命題, 根拠, 因果関係, etc.)

ゲーム脳は恐ろしい

温暖化によつ  
て海面が上

コラーゲンは肌に良い

塩素は健康を害す

イソジンで  
放射能を解  
毒する

ビタミンCで  
風邪を予防  
できる

コラーゲン  
はシミには  
効果がない

コラーゲンは肌から浸透しない

コラーゲン  
コラーゲンの上手な摂り方  
コラーゲンドリンク

コラーゲンの補給で、モチモ  
チとした弾力のあるお肌を取  
り戻すことができるので、積  
極的にコラーゲンを摂り  
しよう。

なお、コラーゲンの再合成を助けるビタミンCが加えられていることも大事です。

## Web上の言明

コラーゲンは肌から浸透しない  
ので、出来てしまったシミに対  
する効果は期待できません

E問(1)

クターシーラボやドモホルンリンクルやニッピコラーゲンといった、コラーゲン成分の効能をうたった化粧品がありますよね?  
近そのひとつを使い出してるのですが、最近コラーゲンサプリにも興味を持っています  
コラーゲンはビフから塗っても浸透しないと書いてあり「じゃあコラーゲン配合の商品はイカマツ?」と不安になりました。事実はどうなんでしょうか?効果ゼロなんでしょうか?

現在話題の誤情報を注目度の順に表示しています。

**1位** 135.43873675218点

ロリコンによる幼女を狙った性犯罪が増えている

**2位** 122.01367562664点

今もツイートしてるってことは、二十歳になつたら消える

**3位** 89.994967900122点

進撃の巨人の実写版ミカサは剛力彩芽さんか

**4位** 65.923530856709点

ファシズムは、独裁者がいて民衆を支配してしまう

**5位** 65.246529925368点

ロボットが放射能で壊れる

**6位** 59.726764182935点

ボカラの曲をAKBが自分たちによこせと言ってる

**7位** 54.849978948139点

1本目で退場してますので、降板が決まってた

**8位** 53.405178823334点

ところがユダヤ人が誘拐して幼児犠牲の祭儀を行ったのだ

# Web上の情報伝達を変える

Mozilla Firefox  
Search Bookmarks and History Google

(Untitled)



深海魚油  
ソフトカプセル  
Deep sea Fish Oil  
Soft Capsules

○○社によって製造された深海魚油のソフトカプセルです。...

主な成分であるω-3脂肪酸は、血液のより良い循環を促進し、高コレステロールの改善、血栓の防止、高血圧の改善、心臓の健康維持をサポートします。また、双極性障害、ADHDの改善に有効です。

Find: Next Previous Highlight all Match case

対立

根拠

「裏を取る」作業を支援

「ω-3脂肪酸はADHDの改善に有効です」の対立ページ

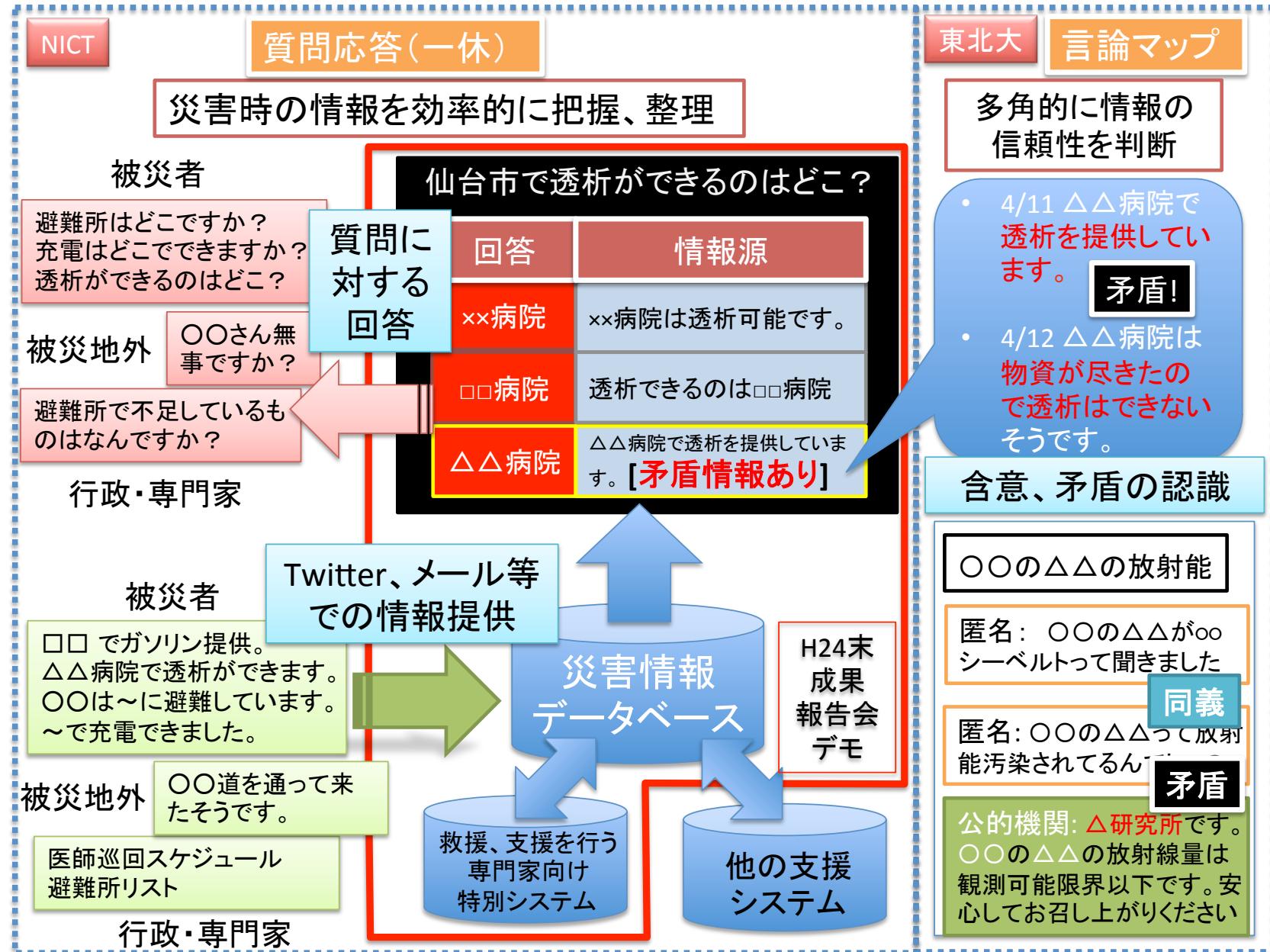
健康食品データベース DHA..  
[独立行政法人 国立栄養研究所]  
小児のADHDは血中DHAレベルの低さと関連があるが、DHAは小児の注意欠陥多動性障害(ADHD)に対して、経口摂取で効果がないことが示唆されている。DHAを1日あたり345mg摂取しても改善はみられなかった...

魚油のオメガ3脂肪酸、頭を良くする効果ない=英研究  
[ロイター] 英国で行われた研究によると、魚油に含まれるオメガ3脂肪酸は、心臓の健康のためには良いが、脳の機能を高める効果はないという。ロンドン大学の研究者らが専門誌「アメリカン・ジャーナル・オブ・クリニカル...

「ω-3脂肪酸はADHDの改善に有効です」の根拠ページ

生活習慣をサポートする  
『オメガ3脂肪酸（後編）』  
[サプマートUSA] ... ADHD、統合失調症、アルツハイマー疾患のような精神・認知に関する疾患を患う人は、血中のオメガ3脂肪酸レベルが低い人が多いことから、このような精神・認知に関係する疾患に対し、オメガ3脂肪酸は有効性が期待できます。...

# 耐災害情報分析システム(2014年度公開予定)



# 情報の「裏を取る」支援

NICT QAシステムに聞く

質問「放射能に効くのは何ですか？」

The screenshot shows a search interface with a large red arrow pointing from the question to the results. The results list several items, including 'イソジン' (Isingin), 'とろろ' (Totoro), and 'ワカメ' (Seaweed). These three items are highlighted with white circles and labeled 'イソジン', 'とろろ', and 'ワカメ' respectively.

東北大言論マップで裏を取る

The screenshot shows two main sections: '赞成意見' (Supportive Opinion) and '反対意見' (Opposing Opinion). The '赞成意見' section for 'イソジンは放射能に効く' has 60 characters and discusses the effectiveness of Isingin against radiation. The '反対意見' section has 26 characters and discusses the risks of Isingin. Below these are sections for 'とろろ' and 'ワカメ', each with its own supportive and opposing opinions.

This is a detailed view of the 'とろろ' section of the speech map. It shows the '赞成意見' (Supportive Opinion) and '反対意見' (Opposing Opinion) for 'とろろ昆布は放射能に効く'. The supportive opinion discusses the effectiveness of konbu (seaweed) against radiation, while the opposing opinion discusses its potential risks.

This is a detailed view of the 'ワカメ' section of the speech map. It shows the '赞成意見' (Supportive Opinion) and '反対意見' (Opposing Opinion) for 'ワカメは放射能に効く'. The supportive opinion discusses the effectiveness of seaweed against radiation, while the opposing opinion discusses its potential risks.

うつ文書  
見做おう  
あの矜持と  
うつくしい

背筋を立てて  
じつと待ちながら  
獲物を狙う  
凛とした姿は  
うつくしい

近くの川堤まで  
散歩すると  
よくカワセミに会う

岡山県井原市  
山崎 範雄  
80

カワセミのように  
朝の詩

# 産経新聞

平成25年(2013) 日刊25169号

1|7 [月]

産業経済新聞(サンケイ)

THE SANKEI SHIMBUN

発行所 © 産業経済新聞大阪本社 2013

〒556-8660 大阪市浪速区湊町2-1-57

☎ 大阪(06)6633-1221(大代表)

販売・配達に関するお問い合わせ

06-6633-9357 (平日9時~19時、土日祝日~17時)

<http://o-sankei-hanbai.com/c/>(平日のみ)

紙面・記事へのご意見・お問い合わせ

06-6633-9066 (平日9時~18時、土曜~17時、日祝日休み)

[o-dokusha@sankei.co.jp](mailto:o-dokusha@sankei.co.jp)

購読のお申し込み

0120-34-3733 (平日9時~19時、土日祝日休み)

<http://www.sankei.co.jp/reader>

産経新聞ホットライン

販売・配達に関するお問い合わせ

06-6633-9357 (平日9時~19時、土日祝日~17時)

<http://o-sankei-hanbai.com/c/>(平日のみ)

紙面・記事へのご意見・お問い合わせ

06-6633-9066 (平日9時~18時、土曜~17時、日祝日休み)

[o-dokusha@sankei.co.jp](mailto:o-dokusha@sankei.co.jp)

購読のお申し込み

0120-34-3733 (平日9時~19時、土日祝日休み)

<http://www.sankei.co.jp/reader>

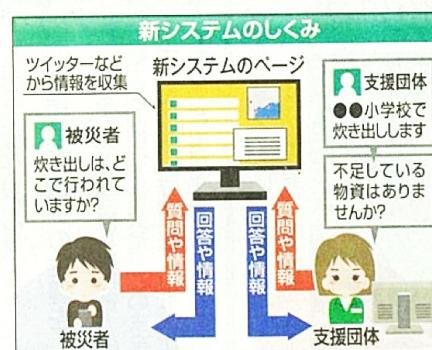
# 被災者ニーネーズ

## 瞬時に把握

大規模災害発生時に、被災者が求める救援物資や炊き出し情報を自動的に収集、整理しておくことで、支援者が被災状況をリアルタイムで効率的に検索できるシステムを、総務省所管の独立行政法人、情報通信研究機構(NICT)などが開発した。平成26年度の実用化を目指す。

東日本大震災では、簡易投稿サイト「ツイッター」を救援活動などに利用する例が多発された。このシステムでは、被災者が書き込んだ内容などを幅広い情報を随時集めて整理。独自のホームページと検索システムで、支援者や被災者が関連情報をたやすくアクセスできるようになる。

具体的には、ツイッターなどのインターネット上にアップされた救援物資を求める情報などを自動的に収集。支援者が「被災地で不足している」といった情報も表示する方向で開発。「炊き出しはどこでやっているか」といった場所に関する質問への回答では詳細な地図も表示し、利便性を高める。



## 総務省所管の独法、災害時検索システム

NICTによると、東日本大震災では震災後約1カ月間で救助関係のツイートは60万件を超えた。しかし、被災者のツイートに応じた人たちから送られた紙オムツが、宮城県の避難所に大量に寄せられるなど、課題も残された。救援活動にツイッターを利用した経験のあるボランティア団体も「個々の被災者の要望をリアルタイムで知ることができると、課題も残された。救援活動にツイッターを利用して有効活用が難しい」と指摘する。

このシステムは、NICTの耐震災ICT研究センター(仙台市)の鳥澤健太郎室長と東北大大学院情報科学研究所の乾健太郎教授が共同開発した。鳥澤室長は「緊急時の情報の行き違いや一方通行を防ぐために、システムの活用が求められる」と話す。早期の完成を目指している。

受験シーズンを前に控え、「知恵の文殊さん」として知られる家原寺(堺市西区)の本堂は、受験生が合格への願いを書き込んだ白いハンカチで埋め尽くされている「お願い」ビッシリ

写真(沢野貴信撮影、一部画像処理をしています)。

堺市北区から高校の合格祈願に訪れた五百旗頭美穂さん(15)、「行きたい高校に絶対

# 参院選 168人出馬予定

を目指す議席数  
**3分の2**  
(162)

あと 41

改選 50

非改選 71

労使一体になつて防ごう  
と、一線で活躍する弁護士  
や医師、研究者らがスクラム  
を組み、一般社団法人  
「産業保健法務研究研修セ  
ルス対策に企

## 労使で新資「心の病」対

大阪の社団法人、4月



**自治体の災害対応にも**

# 東日本大震災における災害対応



防災情報  
(多種・多量)



音声



FAX



画像



メール



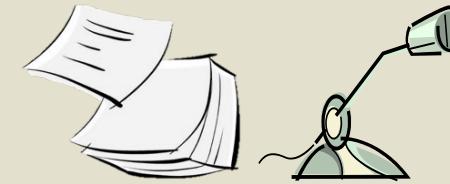
SNS

自治体



ホワイトボード

紙と音声での情報収集・伝達



情報共有の手段は紙とホワイトボード



隣の部局間でも情報が共有できない

# 既存の災害対応情報システム



防災情報  
(多種・多量)



音声



FAX



画像



メール



SNS

我が国



紙と音声での情報収集・伝達

情報入力のボトルネック

米国

専門職員による  
キーボード入力

災害対応情報システム

防災情報データベース

DBに入力する人が足りない  
東日本大震災でも機能しなかった

# 防災情報のデータベース化を技術で支援



防災情報  
(多種・多量)



音声



FAX



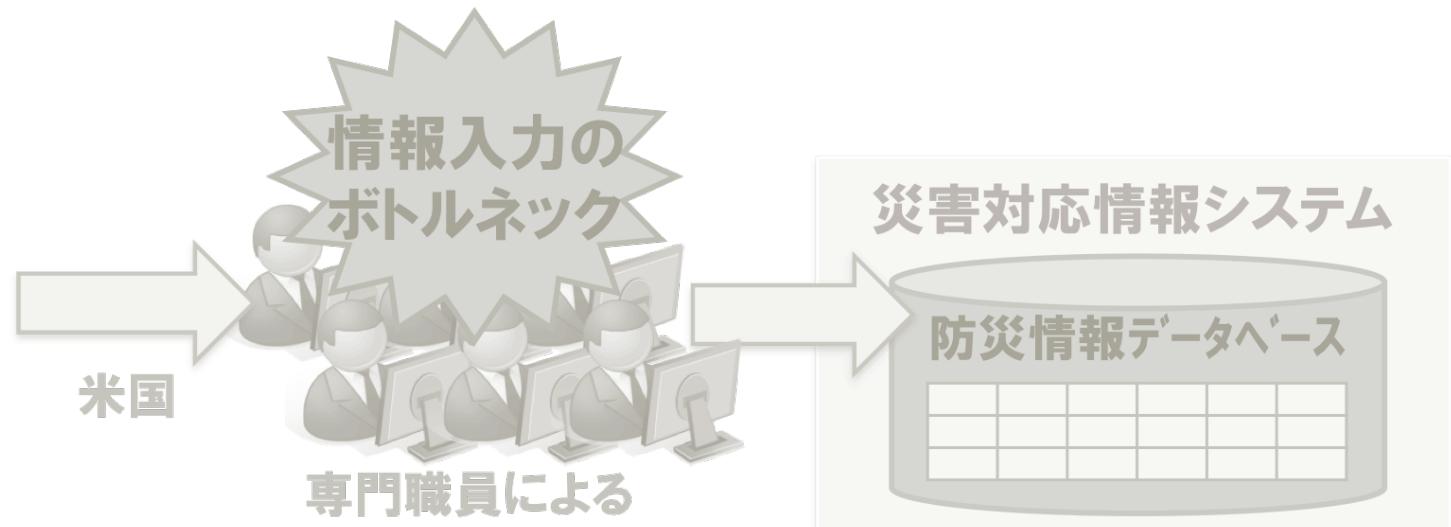
画像



メール



SNS



## 防災情報DB化システム

自然言語処理

音声認識・画像認識

ユーザ  
インタ  
フェース



## 災害対応情報システム

防災情報データベース

防災情報のDB化のコストを大幅縮減



音声



FAX



圖像



六



SNS

# 防災情報のデータベース化

# 音声認識 (画像認識)

## 構造化されていない防災情報

国道4号線のA交差点付近、陥没していますが、応急工事が終わって通れます。

**自然言語で語られる生の防災情報は計算機処理不可**

## 構造化されていない

構造が統一された表(DB)に情報をふるい分ける

# 自然言語處理



# 防災情報データベース

ID	報告日時	施設種別	名称	町村名	区間	被害内容					使用可否	応急復旧終了時期	復旧完了日
						落下	亀裂	崩壊	陥没	詳細情報			
83	5/11 13:46	道路	国道 4号線	B町	A交差点 NO:EO	なし	なし	なし	なし	あり	可	5/11	

# 情報の共有で防災関係者コミュニティをつなぐ



# まとめ

「ビッグデータ」時代

大規模コミュニケーションの支援が必要

鍵は自然言語処理による情報編集

言語処理も「ビッグデータ」で頑健に

耐災害ICTにおける言語処理の役割

社会実装への道筋が必要