

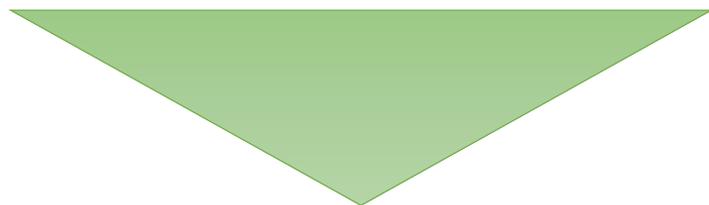
災害対応を支援する SNS情報分析におけるAI

2021年3月24日

国立研究開発法人情報通信研究機構
ユニバーサルコミュニケーション研究所
データ駆動知能システム研究センター/
耐災害ICT研究センター 応用領域研究室（兼務）

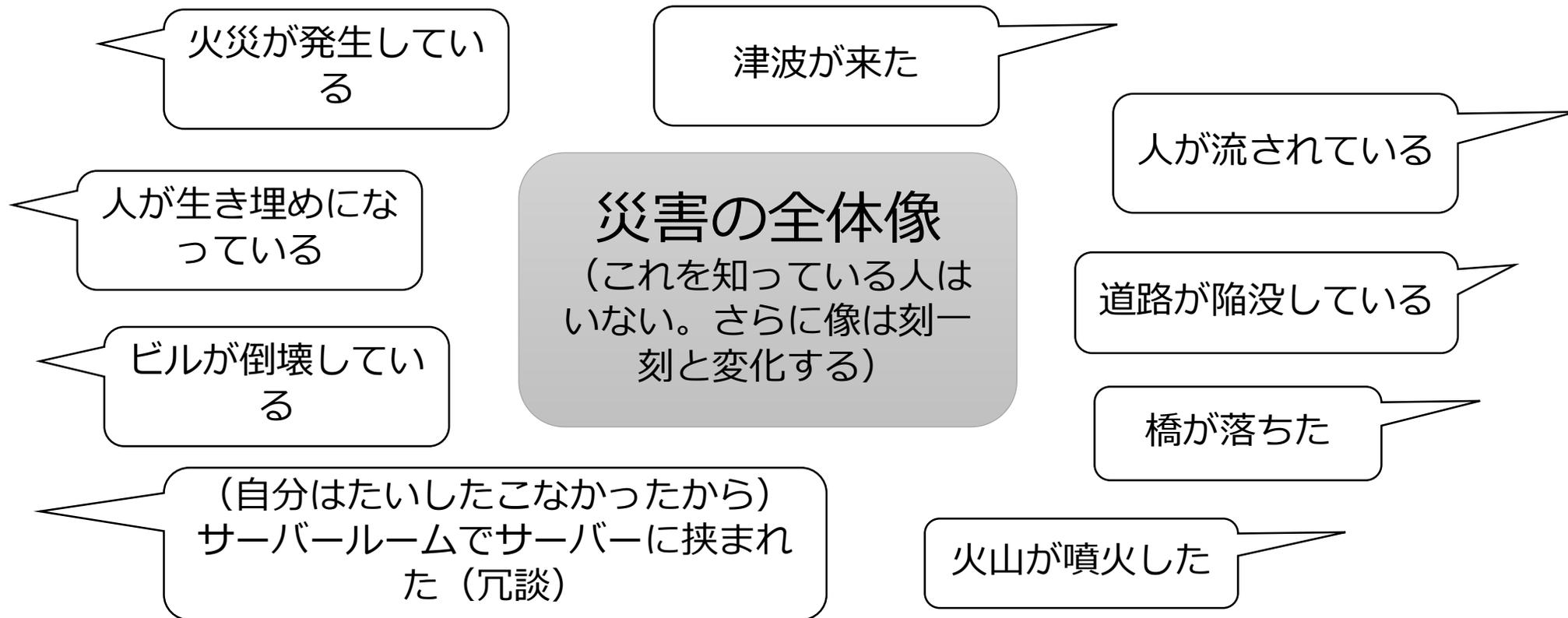
大竹清敬

- SNSはリアルタイム性が高いため災害時に役立つ/役立てたい
- ただし、情報量が**膨大**→必要とする情報を探ることが**困難**
- 情報の**信憑性**、**デマ**による**混乱**



NICTではこれらの問題を解決し、災害対応を支援するシステムを研究開発

「群盲象を評す」はたくさんの盲人が象の一部だけをさわって語り合うというインド発祥の寓話。



やりたいこと：群盲の情報+aをつなぎあわせ、わかりやすい形でそれらを群盲にリアルタイムに共有し、群盲の目を見開かせ、災害時の被害を低減させること

対災害SNS情報分析システムDISAANA[®] (ディサーナ)

- SNS (ツイッター) 上の災害関連情報をリアルタイムに深く分析・整理して、状況把握・判断を支援し、救援、避難の支援を行う質問応答システム
- ツイートしてから5秒で分析結果を提供可能
- 熊本地震の際には、ツイッター社から人道支援として協力いただき1ヶ月ほど100%のツイートの分析結果を提供 (平時は10%サンプル=毎秒約100ツイート)
- D-SUMMとあわせて民間企業へ技術移転



住民、救援団体からの質問 (例: 「熊本県で何が不足していますか」) に瞬時に回答

2015年4月より一般公開中

- 熊本地震の際、首相官邸で活用
- 指定避難所以外のニーズ把握
 - 日々変化する要望の把握

↓
熊本県へ指示

2016年5月11日読売新聞夕刊一面等、報道多数

回答を地図上に表示し、被災状況を俯瞰可能

回答をピンポイントに抽出

生活必需品
救援物資
生理用品

地図データ©2016 Google, ZENRIN

札幌市のどこで充電できる

質問：札幌市のどこで充電できる

日時 2018-09-06 00:00 から 2018-09-06 23:45 まで

その他の設定を表示する

総報告数：31 件

表示カテゴリの限定を全て解除

表示カテゴリの限定対象：なし

- 災害
- 救助
- トラブル
- アラート
- 怪我
- 気象
- 建物
- ライフライン
- 道路
- 不足
- 自治体・住所 (15)
- 地名：固有名称 (1)
- 地名：その他 (3)
- 学校 (2)
- 病院
- その他 (1)

回答候補のリスト表示 回答候補の地図表示

詳しい住所を特定出来なかったツイート(1件)を見るには、ここをクリックして下さい。

北海道札幌市白石区本郷通三丁目北1番 > 役所等機関 > 未分類(2)

設定等を表示させる

検出エリア：白石区役所

被災報告一覧

青字の被災報告をクリックして選択すると被災報告が抽出された。訳中の報告は **このように** に表示されます。「全て」が選択されて **れら** が抽出されたオリジナルツイートを表示します。

全て(2) 白石区役所(2)

被災報告を抽出したツイート

**白石区役所で充電しながら副読んでる
快適すぎるし明日も会社行きたくない**
10:20 PM - Sep 6, 2018

**白石区役所でずっと充電してる
まじ地下白らへんに出来たの神すぎる**
白石区役所お世話になりました
11:13 PM - Sep 6, 2018

地図データ©2019 Google, ZENRIN



宮城県では何が不足しているのかしら？

宮城 不足

検索

 仙台市ではガスボンベが足りません

 気仙沼では薬がない

 石巻ではガソリンが枯渇している

 名取の～では紙オムツが売り切れ

 若林区では飲料水が見つからない

 宮城県では毛布が不足しています

 宮城ではトイレが不足の様

 宮城では燃料が不足しているかも



宮城県では何が不足しているのかしら？

宮城 不足

検索

🐦 仙台市ではガスボンベが足りません

🐦 気仙沼では薬がない

🐦 石巻ではガソリンが枯渇している

🐦 名取の～では紙オムツが売り切れ

🐦 若林区では飲料水が見つからない

🐦 宮城県では毛布が不足しています

🐦 宮城ではトイレが不足の様

🐦 宮城では燃料が不足しているかも

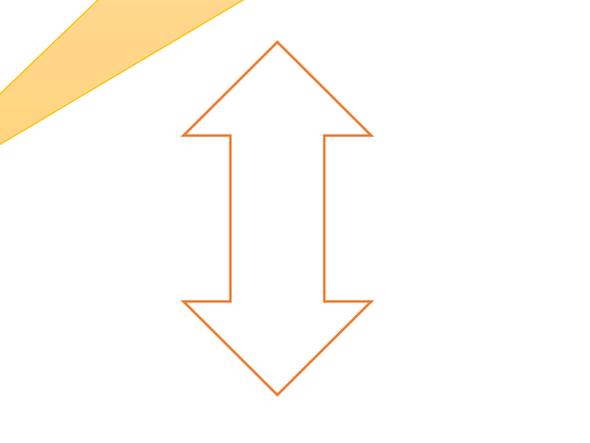
DISAANAではこれらも網羅的に抽出

東日本大震災試用版での動作例

質問：千葉の石油コンビナートで何が発生している



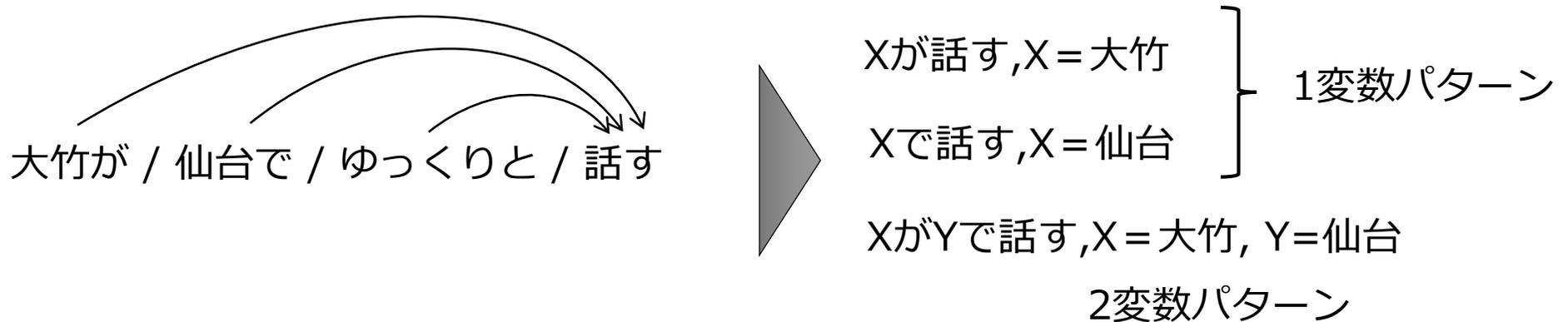
回答候補が抽出されたツイート：
…今後の雨が非常に強い酸性雨になります…



回答候補と矛盾するかもしれないツイート：
…酸性雨になるというのはデマです…

矛盾する情報を同時に検索し提供することで情報の信憑性を判断する材料を提供

- パターン：文の依存構造のうち、名詞を含む文節と係先が述語である構造を抽出し、名詞を変数とし、変数と名詞の関係を示した文字列。



上記は基本的な考え。実際には連体修飾の場合や、体言止めなどからもパターンを生成する

宮城県では何が不足しているのかしら？

質問をそのまま入力

宮城県で何が不足していますか

検索

質問からパターンへ変換

Xが不足する, 場所 = 宮城県

Xが不足する ≡ Xが枯渇する
Xが不足する ≡ Xが足りない

パターン拡張
と検索

同義パターン辞書

Xが不足する

Xが枯渇する

Xが足りない

回答：ガスボンベ、
ガソリン、...

Twitterを随時解析してパターンデータベースを作成

「仙台市ではガスボンベが足りません」
Xが足りない, X=ガスボンベ, 場所 = 宮城県 : 仙台市

「石巻ではガソリンが枯渇している」
Xが枯渇する, X=ガソリン, 場所 = 宮城県 : 石巻市



千葉の石油コンビナートで何が発生していますか

検索

矛盾パターンでの検索



質問からパターンへ変換

Xが発生する, 場所=千葉, 石油コンビナート



パターン拡張
と検索 / 矛盾パター
ン拡張と検索



矛盾パターン辞書

Xが発生する⇔Xが終了する
Xが発生する⇔Xになる【否定】
Xが発生する⇔Xが降る【否定】

回答: 酸性雨

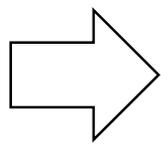
注

矛盾あり

Xが発生する【否定】

Xになる【否定】

Xが降る【否定】



Twitterを随時解析してパターンデータベースを作成

「市原のコンビナート火災で酸性雨になります」
Xになる, X=酸性雨, 場所=千葉県:市原市, コンビナート

「千葉のコンビナート火災で酸性雨になるというのはデマです」
Xになる【否定】, X=酸性雨, 場所=千葉県, コンビナート

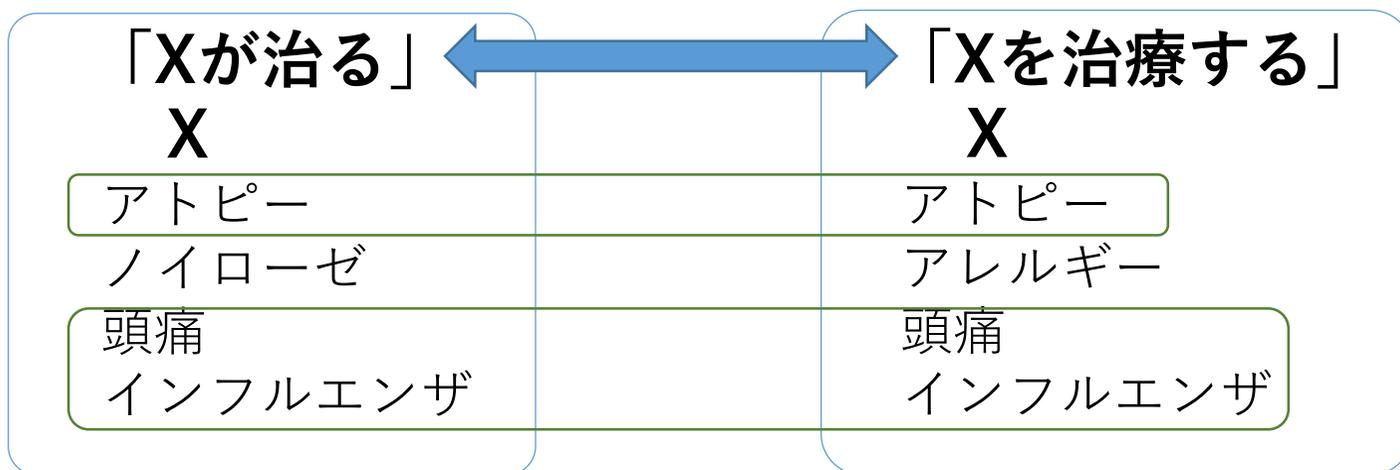
- 制限：基本的には名詞を問い合わせる誰、どこ、いつ、何の4Wの質問応答が可能。なぜ(Why)、どのように(How)は回答が複雑になるためできない
- 課題：
 - 話す⇔しゃべる、伝えるなど同義や類義等の含意表現を網羅して回答する必要がある←同義パターン辞書
 - 一般に、質問の回答は事実であることが求められるため、「話すかも」「話したようだ」などのモダリティを適切に認識し、場合によっては対象から除く必要がある←モダリティ解析



DISAANAでは、これらの同義パターン辞書の作成と、モダリティ解析にAIを活用

表現は違ってもほぼ同じ意味を表す「パターン」を膨大なWeb文書から自動で収集して辞書を構築

アイデア

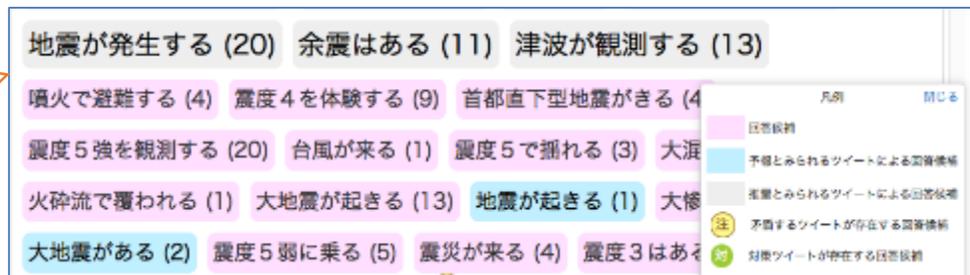


- 同じ単語が使われる場合が多ければ、同義のパターンである可能性が高い
- 自動構築後、人手での検証と教師あり機械学習（AI）、推移律等を使ってさらに高精度化。ヒューリスティックスの利用も
- このようなパターンのペアを1組として3億件以上のデータベースを整備

- 述部の否定や推量といったモダリティを解析する
 - モダリティ：テキストに現れる否定、推量・伝聞・仮定・否定などの表現者の判断等の様態
 - モダリティの分類

<p>肯定（事実の報告）</p> <ul style="list-style-type: none"> • 爆発が発生した • 爆発が発生しました • 爆発が発生したのは事実だ • ... 	<p>推量・伝聞・予知</p> <ul style="list-style-type: none"> • 爆発が発生しているようだ • 爆発が発生している恐れがある • 爆発が発生していると言われている • ...
<p>否定</p> <ul style="list-style-type: none"> • 爆発は発生していない • 爆発が発生しているというのはデマ • 爆発が発生しているのは理論的におかしい • ... 	<p>仮定</p> <ul style="list-style-type: none"> • 爆発が発生したら、... • 爆発が発生しているとすれば、... • 爆発が発生する時には、... • ...

DISAANAでは、「予報」「推量」「回答候補（事実報告）」等に色分けして提示



- 機械学習によるモダリティの解析

- 学習データの準備（下線部のモダリティを人手で判定し作成）

地震が <u>発生</u> しました	事実	土砂崩れが <u>起き</u> たら、	仮定
津波は <u>来</u> ていない	否定	あそこが崩落する <u>恐れ</u> がある	推量等
停電が <u>発生</u> しているようだ	推量等	ガスが <u>来</u> ていないかもしれな	否定 推量等
電話が <u>通</u> じない	否定	暴動が <u>起</u> きているなんてデマ	否定

- 機械学習（モデルの作成）を実施

DISAANAではSVM(Support Vector Machines)を用いている

- 学習済みモデルを用いた判定



- 一見、簡単に解決しそうな問題に見えるが、実は奥が深い
- 複雑で多様な表現がモダリティを表しえる
否定の例：
 - 否定：強奪が起きたなんてツイートしてる奴は北斗の拳の見過ぎ
 - 否定：爆発が起きたなんてツイートしてるキチガイほんとい
いるのか？
 - 否定：爆発が起きてたら、とっくに日本全滅だろ

災害状況要約システム D-SUMM[®] (ディーサム)

- 2016年10月18日一般公開 (<https://disaana.jp/d-summ>)
- SNS (ツイッター) 上の災害関連情報をリアルタイムに抽出し、一目で状況把握・判断ができるよう整理・要約し、救援、避難の支援を行うシステム (内閣府SIPの支援を受けて研究開発)
- 分析エンジンは、後述するSOCDAでも活用。ソフトウェアを民間企業へライセンス

熊本地震前震発災後1時間の熊本県の被災状況の要約

熊本市、益城町を中心に火災、建物被害や、電気、ガス、水道、通信等のトラブル、通行止めの報告多数ということが一目でわかる

地図表示も可能



地図データ©2016 Google, ZENRIN

建物被害の報告
100件以上



電気、ガス、水道、通信のトラブル

家屋倒壊

火災発生

自治体等において情報収集が困難な発災直後1時間でも被害状況概要の把握を可能にし、初動対応を支援

熊本市

被災報告が深刻なエリアから順に表示

益城町

阿蘇市

道内の停電状況をチェック

現在指定しているエリア：北海道

日時 2018-09-06 03:08 から 2018-09-06 04:38 まで

その他の設定を表示する

エリア毎に要約 カテゴリ毎に要約

総被災報告数：3087 件

表示カテゴリの限定を全て解除

表示カテゴリの限定対象：

- ライフライン > 電気トラブル
- 地盤沈下・陥没
- 放射能汚染
- 公害・汚染
- 悪天候 (19)
- 災害 (2)
- 未分類 (2)

+ 救助 (4)

+ トラブル (27)

+ アラート (2)

+ 怪我 (1)

+ 気象 (1)

+ 建物 (19)

- ライフライン (1647)

- 電気 (9)
- 電気トラブル (1634)
- ガス
- ガストラブル
- 水道
- 水道トラブル (4)
- 通信
- 通信機器
- 通信トラブル

5分後 10分後 15分後

地図データ©2018 Google, ZENRIN

9/6 3:08-3:23で電気トラブル（停電）を地図表示：抽出した報告のうちの過半数が停電の報告であり、発災直後15分でほぼ全道的に停電になっていることを確認

道内の土砂災害・建物被害をチェック

現在指定しているエリア：北海道

都道府県を選択する

日時 2018-09-06 03:08 から 2018-09-06 04:38 まで

その他の設定を表示する

エリア毎に要約 カテゴリ毎に要約 時系列で要約 地図表示

総被災報告数：11475 件

表示カテゴリの限定を全て解除

表示カテゴリの限定対象：

- 災害 > 土砂災害
- 災害 > 生き埋め
- 建物
- 公害・汚染
- 悪天候 (42)
- 災害 (19)
- 未分類 (5)

+ 救助 (13)

+ トラブル (179)

+ アラート (6)

+ 怪我 (13)

+ 気象 (5)

+ 建物 (197)

- ライフライン (6692)

- 電気 (43)
- 電気トラブル (6529)
- ガス
- ガストラブル
- 水道 (17)
- 水道トラブル (81)
- 通信

北海道以下の詳しい住所を特定出来な

地図データ©2018 Google, ZENRIN

9/6 3:08-4:38で土砂災害、生き埋め、建物被害を地図表示：発災後1時間半で厚真町、札幌市での被害報告が目立つことを確認

• 学習データが生命線

ラベル：被災報告
場所：徳島
事実性：×

ラベル：被災報告
場所：奈良
事実性：○

変換後フレーズ：停電が発生する
ラベル：矛盾する被災報告
場所：京都
事実性：○

徳島に台風が上陸したと聞いたが、奈良では停電が起きているが京都は停電していない

ラベル：救援報告
場所：大阪
事実性：○

ラベル：矛盾する被災報告
場所：京都
事実性：○

大阪の停電は復旧した

京都市内で停電が起きているというのはデマです

ポイント：一つの被災報告の候補に対し、複数の地名候補がある場合に、適切に対応関係も判定できるように学習データを作成する。モダリティについては、述語のモダリティを求めることが本質ではないため、事実性の判定としてラベルを付与

- ねらい：災害の全体像を把握するためにどんなことがおきているか分かる程度に、コンパクトにまとめたい
- 入力：抽出された被災報告で、それをコンパクトにする（方法はルールベース）
- ただし、出力を整理する観点として場所毎にまとめたり、地図上にも表示できる必要（＝どこで、何が起きているかを把握しやすく）

被災報告は「名詞+助詞+述語」という形式をとる。名詞の類語辞書を用いて、複数の被災報告にある多様な名詞を代表させ、述語を一つだけに限定した被災報告へ要約

名詞の類語辞書 = 災害オントロジーを整備し、活用

※述語の選択についてはルールベースで対応

津波が発生している(75)

大津波が発生している(32)

延焼がひどい(20)

津波で寸断される(20)

火事が発生している(52)

火災が発生する(22)

大火災を起こす(12)

大火がある(1)

津波・高潮が発生している(127)

仙台市(50) 名取市(18) 東松島市(15) . . .

火災が発生している(107)

気仙沼市(52) 仙台市(22) 石巻市(19) . . .

※要約した被災報告と対応する地名を別途整理

- 多くのSNSでは、端末等の位置情報を取り扱えるが、Twitterの日本人ユーザはプライバシーの問題からほとんどオフ
- 一方で、位置情報は非常に重要なため、投稿に含まれる地名適切に処理することでおよその場所を特定
- 具体的には、地名を特定し、住所と緯度経度を付与

例：「**仙台駅**で火事！」

宮城県:仙台市:青葉区:中央:一丁目
(140.882337485758&38.2598788858706)

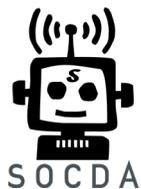
このような辞書を400万件以上整備

課題：

- 曖昧性（「第一小学校」は全国にたくさんある）
- 地名か否かの判定
- 表現の多様性（奈良県生駒市/奈良の生駒/奈良-生駒...）

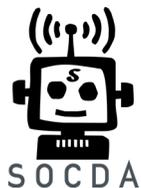
現状は、ポイントデータが基本で、線状（河川や道路）やエリアの扱いは困難

仙台駅の北側の国道で交通事故が起きている



国道45号線ですね。交差点の名前や、近くの目印になりそうなものはわかりますか。あるいは、地図で教えてくださいませんか。

ローソンがある



ローソン小田原一丁目店でしょうか。その場合、最寄りの交差点は、小田原一丁目かと思われます。救急車をよぶ場合は...

将来的には、地図上の種々雑多な情報を適切に取り扱うことができれば上記のような対話も可能に

防災チャットボット SOCDA (ソクダ)

救援団体の少数の担当者が、最新鋭とは言えない手段で情報収集、分析

- 下部組織への電話連絡、ファックスでの情報共有
- 直接現地に行って、状況を確認
- 情報の分析、共有は基本ホワイトボードや地図にマジックで



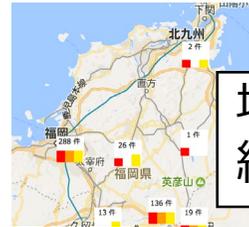
- 情報源は自発的な被災報告だけなので、網羅性は完璧ではない
- Twitterのような匿名アカウントであれば、無責任なデマに完全に対処することは困難
- また、救援側からのフィードバックが難しい



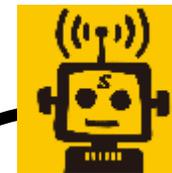
DIRECT 次のステップ：防災チャットボット

- **スマホ等で動作するチャットボットの導入**
 - チャットボットが多数の被災者と自動的に対話
 - **能動的に**情報収集や重要な情報のプッシュを実施
 - 対話の結果は自動的に集計、分析を行い、効率的な救援につなげる

帰宅困難者対策
チャットボット：
何かお困りですか？



地方自治体等で対話の
結果を集計・分析



デマ対策
チャットボット：近くで
土砂災害という情報が
ありますが、ご存知で
しょうか？



被災者：中央線が止まって
家に帰れません

被災者：何も聞いてい
ません。デマでは？

災害時にSNS上で情報収集、情報提供を行う防災チャットボットの研究開発

(戦略的イノベーション創造プログラム (SIP) 国家レジリエンス (防災・減災) の強化「避難・緊急活動支援統合システムの研究開発」
対話型災害情報流通基盤の研究開発)

共同研究開発機関: 国立研究開発法人防災科学技術研究所 (NIED)、株式会社ウェザーニューズ (WNI)、国立研究開発法人情報通信研究機構 (NICT) 協力団体: LINE株式会社、情報法制研究所

膨大な被災者



AI防災チャットボット (WNI) の対話履歴:

- AI: 被害状況を教えてください
- 被災者: 小学校に泥水が流れ込んでいます
- AI: 位置情報を教えてください
- 被災者: ○○市△町
- AI: ありがとうございます。怪我人はいますか？
- 被災者: 怪我人はいません
- AI: 何かで困っていることや困っている人はいませんか？
- 被災者: 小学校に避難して来た人が10人ほどいます。
- AI: 写真を登録しますか？

スマホ上のAIチャットボットによって、被災者一人一人からタイムリーで正確かつ詳細な被災情報を収集し、災害対応で有効活用

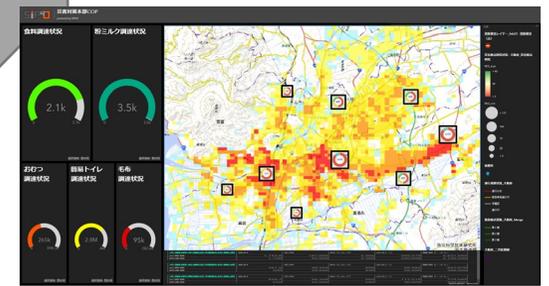
AI防災チャットボット (WNI)

LINE公式アカウント



地図データ©2018 Google, ZENRIN
DISAANA・D-SUMM (NICT) による収集情報の可視化

政府、自治体等で活用



SIPDによる各種災害情報(建物・インフラ・人的被害等)との統合処理と意思決定支援 (NIED)

【今後の予定】さらに機能を追加

情報の信憑性を現場の被災者に確認。デマを排除

AI: 近くで火災が起きているという情報ですが、デマではないでしょうか？

被災者: ここから○○町での火事が見えます

被災者一人一人の状況に合わせ、避難や生活に必要な正確な情報をタイムリーに提供

AI: 食料が足りません

被災者: ○○公民館で2日午後3時より食料の配布があります

【最終目標】

- 以上のような情報収集、提供作業をAIとチャットボットに行わせ、救援団体に必要なマンパワーを増やすことなく、より多くの被災者を適切にケア
- 巨大災害時の避難・被災者支援において、AIとチャットボットが国民一人ひとりに寄り添い、助ける社会を目指します

- これまでNICTで開発してきた2つのシステム（DISAANAとD-SUMM）と、現在、その研究開発に参画している防災チャットボットSOCDAを紹介
- 特にDISAANAとD-SUMMについては、ここでAIがどのように用いられているかについて概要を説明