

研究論文

モダリティ解析の導入による対災害SNS情報分析システム DISAANAの質問応答性能の改善

水野 淳太^{1,a)} 後藤 淳^{1,†1,b)} 大竹 清敬^{1,c)} 川田 拓也^{1,†2,d)} 鳥澤 健太郎^{1,e)}
クロエツェー ジュリアン^{1,f)} 田仲 正弘^{1,g)} 橋本 力^{1,h)} 奥村 明俊^{1,i)}

受付日 2015年10月1日, 採録日 2016年2月23日

概要:我々は、災害時に Twitter に投稿される膨大な情報を効率良く検索するために対災害 SNS 情報分析システム DISAANA を開発し、スマートフォンおよび PC で誰もが利用可能な Web アプリケーションとして試験公開している。本稿では、まず先行システムについて説明し、その問題点についてまとめる。次に、それらの問題を DISAANA がどのように解消するかを説明する。特に、不適切な回答候補の抽出を回避するために導入したモダリティ解析について詳述する。評価実験では、東日本大震災時のツイートに対して、人手で構築した 192 問の質問とその回答からなる評価セットを用いて本システムの評価を行った。評価の結果、先行システムに比べて F 値が 7 ポイント改善した。エラー分析結果に基づいて、今後の改善方針について考察する。さらに、自治体で実施した DISAANA の有用性検証実験の結果についても報告する。

キーワード：災害情報、質問応答、情報検索、SNS

Improving Question Answering of Disaster-information Analyzer (DISAANA) Using Modality Analysis

JUNTA MIZUNO^{1,a)} JUN GOTO^{1,†1,b)} KIYONORI OHTAKE^{1,c)} TAKUYA KAWADA^{1,†2,d)}
KENTARO TORISAWA^{1,e)} KLOETZER JULIEN^{1,f)} MASAHIRO TANAKA^{1,g)}
CHIKARA HASHIMOTO^{1,h)} AKITOSHI OKUMURA^{1,i)}

Received: October 1, 2015, Accepted: February 23, 2016

Abstract: We developed a web service called DISAANA that can be used by anyone from such terminals as smartphones and computers to efficiently retrieve information about natural disasters from the massive amount of posts about them. In this paper, first, we describe the construction of the previous system and its problems. Second, we describe how DISAANA solves them by focusing on modality analysis, which prevents the extraction of incorrect answer candidates. We evaluated the question answering performance of the tweets posted during the 2011 Great East Japan Earthquake using 192 manually constructed questions and their answers. DISAANA achieved a 63.0 F-measure, which outperformed the previous system's 56.0 score. We discuss further directions based on error analysis and report the results of demonstration experiments of DISAANA for a local government context.

Keywords: disaster information, question answering, information retrieval, social network services

¹ 情報通信研究機構
NICT, Kyoto 619-0289, Japan

^{†1} 現在, NHK 放送技術研究所
Presently with NHK STRL

^{†2} 現在, NEC 情報・ナレッジ研究所
Presently with NEC Knowledge Discovery Research Laboratories

a) junta-m@nict.go.jp

b) goto.j-fw@nhk.or.jp

c) kiyonori.ohtake@nict.go.jp

d) t-kawada@cw.jp.nec.com

e) torisawa@nict.go.jp

f) julien@nict.go.jp

g) mtnk@nict.go.jp

h) ch@nict.go.jp

i) okumura@nict.go.jp



図 1 DISAANA のスクリーンショット (2015/9/2 時点) : 質問応答モードにおける質問「東京で何が発生していますか」の結果を PC で表示 (左図), エリア検索モードにおける質問「北海道」の結果をモバイル端末で表示 (右図) した様子

Fig. 1 Example screenshots of DISAANA (revision of Sep. 2, 2015): left side shows answer candidates for question “What’s the situation now in Tokyo?” on personal computers, and right side shows answer candidates for “Hokkaido” by area search mode on smartphones.

1. はじめに

東日本大震災では、Twitter に膨大な量の災害関連情報が投稿された。米 Twitter 社によると、1 秒あたりのツイート数が 5,000 件を超えることが 5 回あり、日本からのツイート数は地震発生後に 500% 増加した*1。震災に限らず、災害時に Twitter に投稿される災害関連情報は、即時性が高く重要な情報が含まれる一方で、投稿数が膨大であるため、一般的なキーワード検索によって必要な情報を効率良く入手することは困難である。そこで、我々は災害関連情報をリアルタイムに効率的に検索することができるシステムとして、対災害 SNS 情報分析システム DISAANA を開発した。本システムは、スマートフォンおよび PC 経由で誰でも利用可能な Web アプリケーションとして、<http://disaana.jp/> で試験公開されている。

DISAANA は、災害時に発信される膨大な情報から必要とする情報を効率的に発見し、災害状況などを俯瞰的に把握できるよう質問応答技術、すなわち、自然言語で表された文による質問に対して、回答となる名詞や文を出力する技術を用いた情報アクセス手段を提供する。一方で、我々が事前に調査したところでは、こうした質問応答手段があったとしても災害時の逼迫した状況の中では、質問その

ものを考えるのは困難であるとの指摘が、ある地方自治体よりあった。そこで、DISAANA では、市町村などのエリアを指定するとそこで起きているトラブルや問題を自動的に検出する機能も提供する。DISAANA では、前者を質問応答モード、後者をエリア検索モードと呼ぶ。

質問応答モードでは、自然文による質問を入力すると、その回答候補を一覧することができる。たとえば、「X 市で何が不足していますか」という質問を入力すると、「X 市で毛布が不足している」や「X 市の病院で透析用チューブが足りない」といったツイートから、質問の答えとなる「毛布」「透析用チューブ」などが得られる。つまり、「不足する」と「足りない」といった表現の違いを吸収したうえで、ピンポイントに質問の回答候補を網羅的に出力する。質問応答モードの PC での動作例を、図 1 左側に示す。この例は、「東京で何が発生していますか」という質問を入力して検索した結果であり、「火災」「地震」「落雷がある」といった災害情報や、「運休がある」「事故」といったトラブルが検索されている。それぞれの回答候補をクリックすると、その抽出元となったツイートを閲覧することができる*2。DISAANA が検索対象とするツイートは、日本語で書かれた全ツイートの 10%*3のうち、当日を含む直近の 4

*2 表示する段階でユーザによってすでに削除されたツイートは表示されない。

*3 https://nazuki-oto.com/twitter/service_menu.html

*1 <https://blog.twitter.com/2011/global-pulse>

日間に投稿されたものである。DISAANA では、この期間の範囲であれば、任意の期間を設定して、検索対象となるツイートを限定することができ、最新の災害情報のみを得ることや、数日前の災害情報を俯瞰することが可能である。

エリア検索モードでは、県名や市区町村名を入力すると、その地域で発生している災害やトラブルを広く一覧することができる [1]。たとえば、「Y 県」のようにエリアを指定すると、「Y 県ではガソリンが枯渇している」や「Y 県 Z 市で停電が発生している」といったツイートから「ガソリンが枯渇」「停電が発生」といった情報が得られる。いずれのモードでも、回答候補は、災害やトラブル、犯罪などのカテゴリごとに分類表示されるので、必要な情報を効率良く入手することができる。エリア検索モードのモバイル端末での動作例を図 1 右側に示す。スマートフォンなどのモバイル端末では、表示できる領域に限りがあり、操作方法も PC とは異なるため、専用のユーザインタフェースを提供している。

質問応答モードの質問応答手法は、後藤ら [2] によるものに基づいているが、様々な問題に対応するため、数多くの拡張、改良が施されてきている。後藤らの手法における問題の 1 つは、事実性の認定が不十分であったことである。事実性とは、文中の事象の成否に関する著者などの判断情報である。たとえば、「大雪になる見込み」や「酸性雨が降るといってはデマ」といった文の「大雪」や「酸性雨が」、現時点で実際に起きている事象として取り扱われていた。災害対応という観点からは、実際に起きている事象をいかに正確にとらえるかが重要である。また、予報は予報として区別されることが望まれる。

事象の事実性を認定することは、容易なように考えられるかもしれないが、否定の表現ひとつとっても、「火災は発生していない」といった単純なものから「火災発生というチェーンメールが来た」のように、チェーンメールの内容は通常真実ではないといった複雑な解釈プロセスを経て認識されるものまで多様である。また、単純に直近の「～ない」といった表現があるから否定であるという判断はできず、文、あるいは文を超えた広い範囲を考慮してはじめて判断できるものもある。たとえば、「ここでは絶対に火災は起きないというのが都市伝説だな」という文からは、「火災は起きない」という否定情報を、さらに後方で「都市伝説」という通常は真実ではないことを示唆する表現によって否定している。本稿では、こうした多様な否定表現を扱えるモダリティ解析器を開発・導入する。

否定表現は、いわゆるデマを判断するための材料を提供するという意味でも重要である。たとえば、「有毒物質が X 市で発生」というデマに対して「有毒物質は X 市では出ていない」という否定表現を検出できれば、ユーザにデマの可能性を判断するための材料を提供することができる。DISAANA にはこのような機能が実装されている。図 2 に

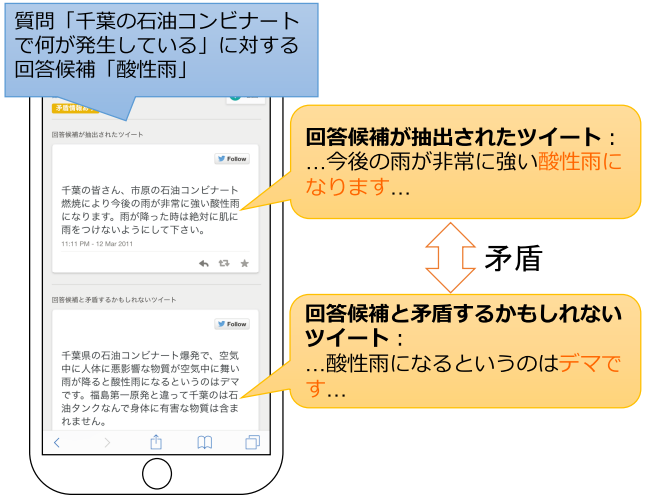


図 2 回答候補と矛盾する情報をモバイル端末で表示した様子
Fig. 2 Example screenshot which shows information contradicting with answer candidate on smartphones.

動作例を示す。この例は、東日本大震災時に投稿されたツイートを対象として、「千葉の石油コンビナートで何が発生している」という質問を入力して検索した結果である。回答候補の 1 つである「酸性雨」には、「酸性雨になるといってはデマ」という回答候補と矛盾するツイートがあるので、回答候補に「矛盾情報あり」というマークを付与し、矛盾するツイートを「回答候補と矛盾するかもしれないツイート」としてユーザに提示する。

モダリティ解析のほかに、災害対応の観点では、現時点で実際に起きている事象が重要であるため、「東京では関東大震災があった」や、「京都で大地震が起きる夢を見たんだ」といった表現の文については、「過去」、「冗談」といった属性を付与することで、回答候補の抽出源から除外する*4。また、「X 市では今晚大雪の恐れ」といった、災害の予報情報は、現時点では非事実であるものの、近い将来に起きうる情報として有用であることから、予報情報であることを明記して表示する。属性および予報は、人手で整備した、過去、冗談、予報などを示唆するキーワードのリストと、ツイート中に含まれる時間表現によって判定される。

本稿の構成は以下のとおりである。まず、2 章で質問応答システムなどの関連研究について述べる。次に、3 章において、後藤らが構築したシステムおよびその問題点について説明し、4 章では DISAANA でそれらの問題点をどう改善するかを述べ、本研究で新たに導入した言語解析モジュールなどについて詳述する。5 章では、東日本大震災時のツイートに対してモダリティ解析および質問応答モードの性能評価を行う。6 章では、自治体の防災訓練を通して DISAANA の有効性を検証した結果について報告する。

*4 より幅広く検索をしたい場合は、検索条件を変更することで、これらのツイートを回答候補の抽出源として利用することもできる。

7章で、結論を述べる。

2. 関連研究

近年では、検索エンジンや質問応答システムなど、情報アクセス手段の進歩はめざましく、IBM社のWatson [3] がクイズ番組で人間のチャンピオンに圧勝し一躍有名になったことは記憶に新しい。Watsonは、Wikipediaを含む辞書、辞典、台本など、クイズ番組の分野に関連する確かな知識をあらかじめ選別しデータベース化している。また、Watsonが勝利したクイズ番組における質問はその回答が一意に定まるものだけであった。それに対しDISAANAは、リアルタイムに更新される多数のツイッターの投稿から多数の回答候補を提供しようとする点が大きく異なる。

また、吉村 [4] は、実運用されている「しゃべってコンシェル」の構成について紹介しているが、このサービスでの質問応答では、質問のカバー率が低いものの、精度の高い回答を返すことができるDB型質問応答システムと、質問のカバー率は高いが、精度はそれほど高くない検索型質問応答システムを組み合わせている。後者のFactoid型と呼ばれる質問応答のアルゴリズムは、(1) 質問を分析し質問のタイプ、回答の属性(固有表現クラス)などを決定、(2) 検索を実行、(3) 検索結果に固有表現抽出器を適用し、求められる固有表現クラスと同一のものを回答候補として抽出し、スコアを算出しランキングする。一方で、DISAANAでは、固有表現抽出器を用いず、構文パターンに基づいて回答候補を求めているため、期待している回答候補の属性とは異なるものが出力される場合があるものの、固有表現抽出器が対象とできない回答候補を出力できる。特に災害時には、「医薬品」「冠水」など固有表現ではない一般名詞やサ変名詞などが重要な回答候補となる場合があるため、固有表現抽出器の利用を前提としたアルゴリズムは使用できない。

本研究では、災害という観点から事実性の認定が重要であることはすでに述べたとおりである。このような、事象の成否に対する著者の判断情報は、言語学ではモダリティと呼ばれる。その処理の難しさから、言語処理においてモダリティ解析が実用を前提に検討され始めたのは比較的最近であるといえる。言語処理におけるモダリティ解析の検討としては乾ら [5] と松吉ら [6] が詳しい。彼らは6種類の項目がそれぞれ2~8種類のラベルを持つような複雑なタグ体系を定義しているが、DISAANAでは膨大な量のツイートをリアルタイムに解析できる解析速度が要求されるため、本研究ではすべてのタグを対象とせず、DISAANAにおいて重要な「否定」と「疑問・要求」のモダリティについて取り組む。松吉らのタグ体系では、それぞれ「真偽判断」タグの「不成立」および「成立から不成立」ラベルをまとめたもの、「態度」タグの「欲求」、「働きかけ-直接」、「働きかけ-間接」および「問いかけ」ラベルをまとめたも

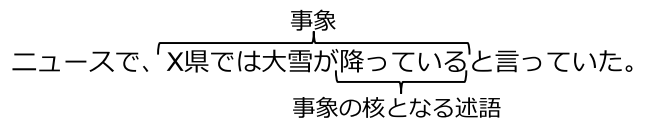


図3 モダリティ解析における事象と事象の核となる述語の例
Fig. 3 Examples of event and predicate for modality analysis.

のにおおむね対応する。

本稿で提案するモダリティ解析は、学習データを用意し、機械学習で対象となる事象の核となる述語のモダリティを推測するものである。文中の事象と、事象の核となる述語の例を図3に示す。機械学習によるモダリティ解析として江口ら [7] の取り組みがある。江口らは従来論じられてきたモダリティに真偽情報、価値情報を統合した拡張モダリティと呼ぶ体系を整備し、その体系に基づいたアノテーションを行ったコーパスを構築し、Conditional Random Fields (CRF) [8] を中心とした機械学習手法を用いたシステムと人手で作成した規則によるシステムを比較している。

Sauriら [9] は、事象に対する事実性を直接判定するべく、様々な手がかりを利用するFactuality Profilerと呼ばれるルールベースのアルゴリズムを提案している。またSauriらは、事実性を肯定・否定の極性と事象の真実性に関する書き手の自信の度合いに関するモダリティの2軸による事実性空間を定義し、その中で事実性をとらえようとしている。極性にはpositive, negative, unknownの3値を、モダリティにはcertain, probable, possible, unknownの4値を考え、これらの2つ組みで事実性を表す。さらにSauriらは、ニュース記事に対して時制の情報を付与したTimeBankという既存のデータ [10] に対して、これらの事実性の情報を付与したデータを作成し、FactBankとして発表している [11]。FactBankは、文中の各述語(動詞、形容詞、名詞述語)に対して、sourceと呼ばれる文中の登場人物および著者ごとに事実性の情報を付与したものである。本研究では問題を簡単にするため、また高速な解析を実現するために、著者のモダリティのみを解析する。

3. 後藤らのプロトタイプシステム

後藤らが構築した対災害情報分析システム(以下ではプロトタイプシステムと呼ぶ)の構成を図4に示す。プロトタイプシステムは、大きく分けて、テキストを解析する言語処理モジュールと、入力された質問に回答する質問応答モジュールの2つの要素から構成される。言語処理モジュールによって解析されたツイートは、データベースに登録される。このデータベースを回答データベースと呼ぶ。質問応答モジュールは、入力された質問を分析し、回答データベースを検索することで、回答候補集合を得る。3.1節では言語処理モジュールについて説明し、3.2節では質問応答モジュールについて説明する。3.3節ではプロトタイプシステムの問題点についてまとめる。

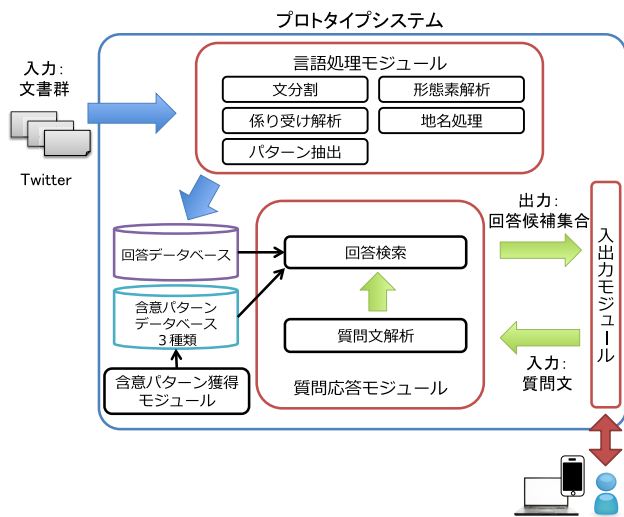


図 4 プロトタイプシステムの構成
Fig. 4 Prototype system architecture.

3.1 言語処理モジュール

言語処理モジュールは、入力されたツイートに対して、以下に説明する各解析器を順次適用する。解析結果は、回答データベースに登録され、質問応答モジュールは、このデータベースを使って質問に対する回答候補を検索する。
文分割 すべての文字を全角に変換し、句点などでツイートを文単位に分割する。

形態素解析 すべての文について、形態素^{*5}に分割し、品詞情報を付与する。解析器には Juman 品詞体系 [12] の辞書を用いた MeCab [13] を用いる。

係り受け解析 すべての文について、形態素列を文節にまとめあげ、その係り受け関係を解析する。解析器には J. DepP [14] を用いる。

地名処理 「どこで火災が起きていますか」といった場所を聞く質問に回答するために、文中に含まれる地名やランドマークについて、それらの完全な住所を辞書から推定し、抽出する。地名やランドマークについてその住所を扱えるようになるため、場所に関する階層性を正しく処理できるようになる。たとえば、「宮城県はどこで～」といった質問に対して、「仙台市で～」 「南三陸町で～」と記述されたツイートも対象として回答候補を抽出できるようになる。また、辞書のエントリーには、それぞれの緯度経度情報も付与されており、地図上への場所の表示も可能になっている。

パターン抽出 パターン抽出は、以上の解析結果から、計算機が扱いやすい形として、述語を含む文節と、それと係り受けの関係にある2つの名詞から構成されるパターン (バイナリ) および、述語を含む文節と、それと係り受けの関係にある1つの名詞から構成されるパターン (ユナリ) を抽出するモジュールである。バイ

*5 言語の意味や文法機能を担う最小の単位と定義され、おおむね「単語」に相当する。

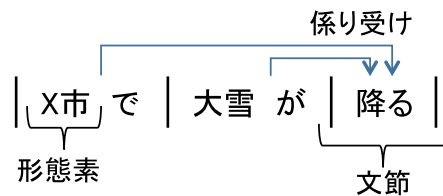


図 5 「X市で大雪が降る」の係り受け解析結果

Fig. 5 Dependency structure of "it is snowing heavily in X city".

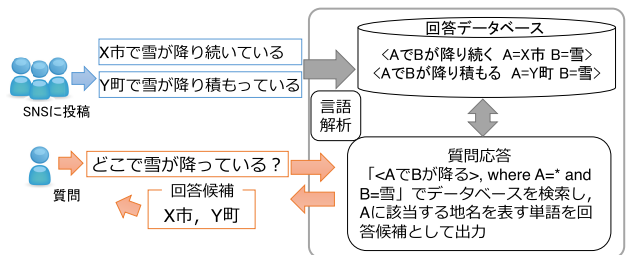


図 6 質問「どこで雪が降っている？」に対する質問応答モジュールの動作例

Fig. 6 Example of question answering process for question "where is it snowing?".

ナリは、たとえば「X市で大雪が降る」という文から、係り受け解析結果 (図 5) に基づいて抽出される (AでBが降る) (ただし、Aは「X市」、Bは「大雪」というパターンである。なお、述語の助動詞などは、助動詞などの違いによりバリエーションが増加することを防ぐために、定められた例外を除きパターンから削除する。同じ文から、ユナリとして (Aで降る) (ただし、Aは「X市」と、(Bが降る) (ただし、Bは「大雪」) の2つのパターンが抽出される^{*6}。回答データベースには、抽出されたパターンと、「こと」や「もの」などのストップワードを除く文中の名詞 (後述する周辺キーワード) が登録される。

3.2 質問応答モジュール

質問応答モジュールは、言語処理モジュールで構築された回答データベースから、質問の回答候補を抽出する。具体的には、入力された質問を簡単なルールにより平叙文に変形し、疑問代名詞に入りうる単語を回答データベースから検索して、回答候補として出力する。たとえば、「どこで雪が降っている？」という質問に対する質問応答モジュールの動作例を図 6 に示す。質問文は、(AでBが降る) というパターンに変形され、疑問代名詞「どこ」がAに入るため、このパターンを、Bが「雪」とあるという制約とともに、回答データベースを検索する。最終的に、Aに入る地名を回答候補として取得する。回答データベースを検索するとき、パターンと言い換えの関係にあるパターンにつ

*6 システム上では、AとBはいずれもAで記述されるが、ここでは分かりやすくするために、AとBで表して区別する。

いても検索する [15]. この例では, (A で B が降り続く) や (A で B が降り積もる) といったパターンでも検索する.

ユナリは, (A が起きている) のように名詞が1つだけのパターンだが, 質問が「何が発生しているか」や「何が止まっているか」などのように, 疑問代名詞と1つの述語からなる場合は, 前述と同様の方法で回答候補を抽出することができる. たとえば, 質問「何が発生しているか」からは, (A が発生している) に変形され, 「地震が起きている」から抽出される (A が起きる) (ただし, A は「地震」) や, 「台風が発生した」から抽出される (A が発生する) (ただし, A は「台風」) などから, 「地震」「台風」を回答候補として抽出する. このとき, (A が起きる) と (A が発生する) は同じ意味を表すと認識されるが, これは「が起きる」と「が発生する」の極性がいずれも「活性」[16]であることから, 言い換えとして認識される.

また, ユナリを抽出した文と同一文中にある他の名詞を回答として抽出する場合がある. たとえば, 場所を聞く質問「どこで大雪が降っている」に対して, ツイート「東京に着きました. こちらは大雪が降っています.」のように, 「東京」と「降る」の間に係り受け関係がない場合においても, ユナリ (A が降る) (ただし, A は「大雪」) と同一ツイート中にある地名「東京」を回答として抽出する. このような名詞を周辺キーワードと呼ぶ. 特に地名については, 文頭のみ書かれ, パターンと係り受けの関係にない場合も少なくないことから, 本戦略が有効である.

3.3 プロトタイプシステムの問題

プロトタイプシステムを用いて, 5章で述べる評価実験と同様の評価を行った結果, 大きく3つの問題があった. また, 平常時のツイートをプロトタイプシステムに入力した場合に生じる問題もある. 本節では, それらについてまとめる.

1つめの問題は, 肯定文以外から頻繁に誤回答が抽出されることである. 後藤らは, 評価実験のエラー分析によって, 誤回答の大半がこの問題が原因であることを報告している. この問題は, たとえば, 質問「どこで水が不足していますか?」に対して, 疑問文「〇〇で水はあるのかな」や否定文「××で水が不足しているというのは誤報です」といった肯定文以外から「〇〇」や「××」を回答候補として抽出してしまうという問題である.

2つめの問題は, パターン抽出のカバレッジについてである. プロトタイプシステムでは係り受け解析に基づいてパターンを抽出しているが, 述語が省略された場合に有効なパターンが抽出できない. 「仙台で地震!」という文からは, (A で地震) (ただし, A は「仙台」) というパターンが抽出されるが, これは, 質問「どこで地震が起きている」から抽出されるパターン (A で B が起きる) (ただし, B は「地震」で, A に入る単語を回答データベースから取

り出す) とは, 言い換えを含めてマッチしない.

3つめの問題は, 動作速度である. プロトタイプシステムは, 並列処理の導入やデータベースへのアクセス速度が不十分であるため, 東日本大震災時に発生した1秒あたり5,000件のツイートが投稿されるような状況に対して, リアルタイムに解析を行うことは不可能である. また, 災害時には各種検索機能の利用回数も増加することが予想される.

4つめの問題は, 平常時に頻繁に投稿される冗談や過去の災害への言及などから, 回答候補を抽出してしまうことである. プロトタイプシステムは, 想定する入力を, 東日本大震災に関連したツイートのみ限定していた. したがって, 平常時に投稿されるツイートに含まれる冗談 (たとえば「東京で地震が起きる夢を見た」) や, 過去の災害を振り返るようなツイート (たとえば「関東大震災では甚大な被害が出た」) から, 誤って災害情報を抽出してしまい, あたかも現時点で起きている災害であるかのように出力してしまうという問題がある.

4. DISAANA での改善

DISAANA では, プロトタイプシステムの問題を解消するために, 言語処理モジュールに新たなサブモジュールを導入した. その構成を図7に示す. 問題解消の概要は以下のとおりである.

- (1) 否定文や疑問文から回答候補を抽出しないようにモダリティ解析を導入した. これについては4.1節で詳述する.
- (2) 一般名詞からも必要に応じてパターンを抽出する拡張パターン抽出を導入した. これについては4.2節で詳述する.
- (3) 回答データベースにオンメモリの高速な実装を利用し

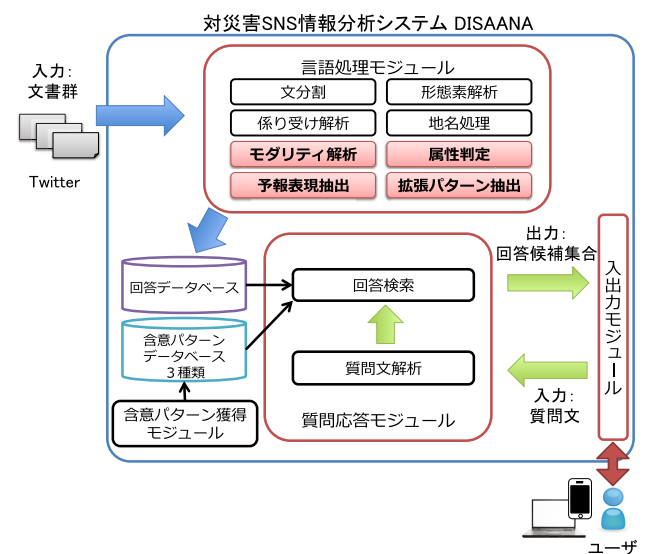


図7 DISAANA の構成

Fig. 7 DISAANA architecture.

たほか、一連の言語処理と質問応答処理を複数の計算機を用いて数百並列で実行し、入力されるツイートや質問を分散させて処理をする。これらの並列実行は、ミドルウェア RaSC [17] を用いて実現した。

(4) 予報情報であるかを認識するために予報表現抽出を、ツイート中の情報が冗談や過去の災害であるかを認識するために属性判定を導入した。これらについては、それぞれ 4.3 節, 4.4 節で詳述する。

また、これらの言語処理モジュールの変更に対応するために施した質問応答モジュールの変更点は 4.5 節で述べる。

4.1 モダリティ解析

モダリティ解析は、一般には事象に関する非常に多くの意味的側面を解析するタスクであるが、本稿では、文中の事象の成否を分類するタスクとする。事象は一般に 1 つ以上の名詞 (何が) と述部 (どうした) によって表現されると考え、本稿ではその核となる述語を含む文節を (<>) で示す。DISAANA で特に重要なのは、否定されている事象 (たとえば、「大雪は (<<降っていない>>)」) と、疑問・要求の事象 (たとえば、「X 学校で携帯は (<<充電できますか? >>)」) である。これらの文節に対して「事実」、「推量」、「否定」、「仮定」、「疑問・要求」の 5 種類のモダリティを表すラベルを人手で付与した学習データを用意し、機械学習によってこの問題を解く。上記の例示はすべて述部の 1 文節内の情報でそのモダリティを判定できたが、「降っていない」といった表現だけでなく、「大雪が (<<降る>>) という予報は当たらなかった」や「X 市が (<<浸水している>>) というチェーンメールが来た」といったより広い範囲の情報を用いて複雑な解析が必要な事例は少なくない。機械学習に基づく日本語を対象としたモダリティ解析の先行研究 [7] には、こういった広い範囲にわたる事例を分類するための素性が含まれていない。また、広い範囲の情報を直接機械学習モデルに入れようとするとデータスパースネスの問題から膨大な量の学習データが必要となる。

そこで、比較的容易に入手可能な膨大な文書に対し、単語クラスタリング、すなわち単語を意味的な類似度に基づいてまとめあげてクラスタを構築することを実施し、そのクラスタの ID を用いることで単語を抽象化し、データスパースネスを回避する。本研究では、単語を形態素 N グラムで表す。形態素 N グラムは、形態素に分割された文から N 個の連続する形態素を取り出したものである。 N は 1~4 と定めた。形態素 N グラムをクラスタリングするために、まず word2vec [18] を用いて、形態素 N グラムのベクトル表現を獲得する。このベクトルは、文書集合における形態素 N グラムの周辺に現れる単語の頻度を表しており、ベクトルの近さ (たとえばコサイン距離) は、形態素 N グラムの意味的な類似度を表す。次に、得られたベクトル表現を k -means クラスタリングによってクラスタリ

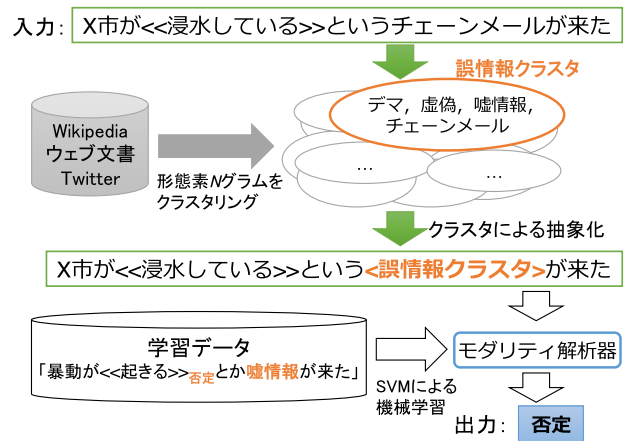


図 8 モダリティ解析例
Fig. 8 An example of modality analysis.

ングする。最後に、モダリティラベルを付与する機械学習には、サポートベクトルマシン (SVM) [19] の実装の 1 つである LIBSVM [20] を用いた。

単語クラスタリングのための文書集合として、3 種類の文書集合 (詳細は 5 章で述べる) を用いた。したがって、3 種類のクラスタリング結果それぞれにおいて、形態素 N グラムをクラスタ ID に抽象化し、その ID を SVM の素性として利用する。したがって、1 つの形態素 N グラムからは、最大 3 つの素性が抽出される。素性の抽出対象となるのは、解析対象の文節に含まれる述語に後続する最大 7 形態素である。なお、述語そのものは抽出対象としない。

図 8 にモダリティ解析の解析例を示す。入力文中の「浸水している」に後続する「チェーンメール」が、「虚偽、デマ」といった誤情報を示唆する表現のクラスタに属していれば、「チェーンメールが来た」は「誤情報クラスタ」が来た」に抽象化される。モダリティ解析の学習データ中に、同様に抽象化される事例が含まれていれば、正しく解析することができる。

word2vec と k -means クラスタリングによって得られる単語クラスタリング結果の ID 以外に、Kazama ら [21] によって得られる単語クラスタリング結果のクラスタ ID も利用する。モダリティ解析のための素性抽出は、同様に述部に後続する最大 7 形態素を対象とする。このクラスタリングは、名詞のみを 2,000 個のクラスタに分類したものであるが、word2vec と k -means クラスタリングで対象とした文書よりも非常に大きな文書から作成されているため、前述のクラスタよりも、カバーされる名詞の種類が多くなることが期待できる。

モダリティ解析の機械学習に用いるその他の素性は、以下のとおりである。

基本素性 解析対象の文節と、その係り先の文節に含まれる形態素の、表層、原形、品詞の 1~3 グラムを素性とする。これらはベクトル表現ではなく、形態素解析結

果をそのまま用いる。解析対象の文節には、その中心となる述語が含まれているが、それ自身は含めず、後続の表現のみを用いる。これは、否定されやすい事象そのものを学習することを防ぐためである。東日本大震災時には、広く拡散された誤情報と、それが誤りであると指摘する情報の両方が存在し、多くの人がそれぞれの情報をツイッターに投稿した。たとえば、「被曝対策にイソジンを《飲む》》というのは間違い。」は、多くの人が投稿したため、モダリティ解析の学習データにも同様の内容のツイートが多く含まれている。このとき、解析対象の文節中の述語「飲む」を素性に加え、学習データ中に多く含まれる同様の事例を学習してしまうと、「給水所で水を《飲む》。》。」といった誤情報ではない事象についても、「飲む」という述語から「否定」と学習してしまう恐れがある。述語そのものを素性に含めないのは、こういった学習を防ぐためである。

後続形態素 述語に後続する最大7形態素の表層をそのまま素性とする。基本素性と同様に、述語そのものは含めない。

否定表現 否定を示唆する表現が解析対象文節の周辺に存在する場合、否定されている可能性が高いと考えた。そこで、人手で33個の誤情報を示唆する形態素（以下、誤情報形態素と呼ぶ）を整備した。述語の前後いずれかにおける誤情報形態素の出現の有無を表す素性と、誤情報形態素との距離を表す素性を用いた。後者の素性は、形態素をその距離として、1から10までの数値、あるいは11以上という11種類の値のいずれかをとる。

誤情報形態素 デマ、でま、ガセ、ガセネタ、がせ、ネタ、風説、流言、流言飛語、流言蜚語、誤報、誤情報、誤解、嘘、うそ、ウソ、偽る、偽り、捏造、ねつ造、虚偽、間違う、間違い、出任せ、でまかせ、誤る、誤り、虚構、違う、違い、チェーンメール、チェンメ、ちえんめ

評価極性 「噴火が《起きていると》》無知な人が言っている」のように、「無知」というネガティブな単語によって否定される事象を認識するために、単語の評価極性を素性にする。基本素性を抽出した文節について、評価極性がポジティブまたはネガティブな単語が含まれる場合、それらを素性とする。評価極性は、人手で整備した辞書 [22] に基づいて判定する。

4.2 拡張パターン抽出

パターン抽出のカバレッジを広げるために、サ変名詞以外の名詞についても災害を示唆する名詞であり、かつ名詞に続く表現が一定の条件を満たす場合は、述語「が発生する」を補ったパターンを抽出する。たとえば、「X市で地震」といった文に対して、述語「が発生する」を補って、〈AでBが発生する〉（ただし、Aは「X市」、Bは「地震」）というパターンを抽出する。

表 1 パターン抽出に用いる災害・被害を示唆する名詞
Table 1 List of nouns indicating disasters and damage for pattern extraction.

ツイート中で現れる名詞	より一般的な名詞
集中豪雨	大雨
土砂降り	大雨
地滑り	土砂災害
強奪	盗難
凍傷	負傷者
靱帯損傷	負傷者
ダニ媒介性脳炎	病人

震」といった文に対して、述語「が発生する」を補って、〈AでBが発生する〉（ただし、Aは「X市」、Bは「地震」）というパターンを抽出する。

具体的な手順を述べる。災害を示唆する名詞は、負担・トラブル表現リスト*7をもとにして、災害や被害を示唆する約1,000種類の名詞を人手で整備した。それらの名詞がツイート中に現れ、かつその名詞で文が終わるあるいは区切れる、または「だ」「である」のような断定表現が続くなどの条件を満たすとき、「が発生する」という述語を補ってパターンを抽出する。これらの名詞の中には「雷火」や「ウイルス性食中毒」といった、あまり一般的ではない名詞も含まれる。「どこで火災が起きていますか」や「どこに病人がいますか」といった質問に対して、「X市で雷火」や「避難所でウイルス性食中毒」といったツイートも検索し、「X市」や「避難所」を回答候補として抽出できるように、「雷火」をより一般的な表現である「火災」に、「ウイルス性食中毒」をより一般的な表現である「病人*8」に置き換えたパターンも抽出する。整備した名詞の一部を表1に示す。

しかし、「北海道で地震はない」といったツイートから、〈AでBが発生する〉（ただし、Aは「北海道」、Bは「地震」）を抽出してはいけぬ。そこで、活性・不活性辞書 [16] を利用する。名詞に続く述語がある場合は、それが活性の述語（たとえば、「が起きる」や「が降る」）である場合は抽出し、不活性の述語（たとえば、「はない」や「は治まる」）である場合は抽出しない。述語が続かず、災害を示唆する名詞で文が区切られている「X市で地震。」や「Y県で地震、…」といったツイートからは、無条件でパターンを抽出する。

4.3 予報表現抽出

「X市は今晚雪になりそう」や「予報では、Y県は明日には台風が上陸するらしい」といった災害の予測や、いわゆる天気予報に基づいた情報発信は、現時点では起きていない事象であるため、他の回答候補とは区別して表示する。

*7 <https://alaginrc.nict.go.jp/resources/nict-resource/li-info/li-outline.html#A-3>

*8 「ウイルス性食中毒」の一般的な表現は「病気」だが、述語「発生する」を補うことも考慮して、「病人」とする。

それらを判定するために人手でキーワードリストを整備した。キーワードリストには、「来年、明日、将来」といった未来を示唆する単語だけでなく、「見込み、かも、恐れ」といった予測・予報を示唆する単語が含まれる。また、ツイート中に含まれる「7月8日」や「7/5」といった日付情報について、年月日の3値で正規化し、ツイートの投稿日より未来であるかを判定する。日付の表現には「8日」のように年月の情報を省略したものもあるが、その場合は、投稿日と同じ年月を補完する。「来月の8日」といった表記の場合は、投稿日の月を1つ進めた月の「8日」として判定する。未来を示唆するキーワードや日付情報が含まれる場合、そのツイートに含まれる情報は未来を示唆する情報として取り扱われ、ユーザには予報情報であることを明記して提示する。

4.4 属性判定

ツイッターには、「京都で大雪が降る設定のドラマ」といったフィクションを題材にしたツイートや、「東京では昔大きな地震があった」といった過去の事実を題材にしたツイートが投稿されることがある。こういったツイートを情報源として、「京都」や「東京」を災害の起きた場所として回答候補とするのを避けるため、ツイート単位で「冗談」、「過去」、「広告」といった属性を付与するのが属性判定である。いずれも人手で整備したキーワードリストによって判定する。キーワードリストは、単語リストと正規表現のリストから構成される。現時点で46,996個の単語および正規表現を整備しており、今後も拡張していく。

「冗談」属性と判定するためのキーワードには、「～という夢を見たんだ」といったいわゆる冗談だけでなく、「物語」や「舞台」といったフィクションを示唆するキーワードや、絵文字や顔文字も含まれる。フィクションを示唆する表現は、小説やアニメ番組のタイトルや登場人物なども含まれるが、これらはWikipediaなどを参照して人手で収集した。

「過去」属性は、過去の災害や過去の情報であることを示唆する表現によって判定する。過去の災害やそれに関連する情報は、上位下位関係抽出ツール^{*9}によりWikipediaから抽出した上位下位概念辞書から、「地震」などの災害を表す単語の下位概念に相当する単語を列挙して人手で精査した。過去の情報であることを示唆する表現には、「昔」「～年前」といったものもあるが、これらは人手で整備した。キーワードリストだけでなく、ツイート中に現れる時間表現と、ツイートの投稿時間とを比較した結果も用いる。ツイートの投稿時間よりも3日以上前の日付がツイート本文に含まれる場合は、その内容は過去の情報である、すなわち「過去」と判定する。ツイート中に現れる時間表現は、

4.3章の予報表現抽出によって認識する。たとえば、7月8日に投稿されたツイートについて、その本文中に「7月1日にX市に台風が上陸した」といった記述が含まれる場合、2つの時間表現を比較することで7日前の情報であることから、「過去」と判定される。

「広告」属性の判定には、「～した人RT」^{*10}や「【無料】」といった集客を目的とした表現や、会社名などが手がかりとなる。会社名は、それが災害時のツイートに現れた場合は広告の一部ではなく被害の発生場所を意味する可能性があるため、会社名单体で「広告」と判定するのではなく、URLまたは電話番号が同一ツイート内に含まれる場合のみ、「広告」として判定する。

DISAANAのユーザには、ツイートが「冗談」や「過去」と判定され、かつ災害関連の用語もツイート本文中に含まれる場合に限って、それらを表示するかを選択できる。DISAANAは、回答候補は「災害」や「トラブル」「気象」などのカテゴリに分類して表示する。後者の条件は、「災害」や「トラブル」といった災害関連情報のカテゴリの検索結果に対して冗談や過去の情報が混在することは大きな問題だが、その他のカテゴリについては緊急性が低いため、冗談や過去の情報であっても表示されてかまわないという判断によるものである。

4.5 質問応答の改善

質問応答モジュールは、プロトタイプシステムと比べて、基本的な戦略は変更しないが、モダリティ解析、属性判定、予報表現抽出の結果を考慮できるように変更した。また、パターンの言い換えを認識するための知識については、文献[15]に加えて、文献[23],[24],[25]により獲得された言い換えパターンを追加した。

モダリティ解析について、解析結果に基づいて回答候補を抽出するかを選択する。たとえば、「X市では雪は降っていない」というツイートがあるとき、「降る」は否定事象と判定されるので、「どこで雪が降っているか」という質問に対して「X市」は回答候補として抽出されなくなる。また、「疑問・要求」と判定された事象が含まれるツイートからも回答候補を抽出しない。たとえば、「Y市でも雪が降るかな?」というツイートがあるとき、同様に「Y市」を回答候補として抽出しない。属性判定について、判定された属性は、モダリティ解析と同様に取り扱い、判定された属性が「冗談」や「過去」となったとき、回答候補を抽出しない。予報表現抽出について、質問応答では抽出結果を利用せず、DISAANAのユーザインタフェースにおいて、「予報」と判定された回答候補には、予報情報であることを明記してユーザに提示する。

^{*9} <https://alaginrc.nict.go.jp/hyponymy/>

^{*10} 投稿内容に同意する人に対してリツイートすることを求めるツイート。

表 2 モダリティ解析の 5 分割交差検定による評価結果

Table 2 Results of modality analysis by 5-fold cross validation.

モダリティ	適合率		再現率		F 値
事実	0.735	(39053/53117)	0.874	(39053/44672)	0.799
推量	0.690	(4137/5999)	0.484	(4137/8546)	0.569
否定	0.833	(9848/11827)	0.731	(9848/13480)	0.778
仮定	0.611	(10785/17651)	0.561	(10785/19239)	0.585
疑問・要求	0.7908	(9710/12278)	0.650	(9710/14935)	0.714
マクロ平均	0.732		0.659		0.694

5. 東日本大震災データにおける性能評価

モダリティ解析と、質問応答モードの性能を、東日本大震災時のツイートを用いて評価する。利用したツイートは、2011年3月9日から2011年4月4日までのツイート((株)ホットリンク提供)である。リツイートによって重複するツイートをおおよそ削除するため、ツイートの文字列比較によって同一と判断されたものは、時刻が最も古いもののみを残すという処理を施し、その結果残った約8,500万件のツイートを実験に用いた。

5.1 モダリティ解析の性能評価

まず、単語 N グラムクラスタの構築に用いた文書集合と、パラメータ選択について述べる。単語 N グラムクラスタは、日本語の Wikipedia の全記事 (2015/1/18 版)、2007 年頃にクロールした 6 億のウェブ文書のうち 0.125%、2015/2/14 から 2015/2/28 までに投稿されたツイートの 3 種類からそれぞれ構築した。これらのテキストはそれぞれ、4.2 GB, 4.5 GB, 4.3 GB 程度の大きさであり、言語処理モジュールと同じ形態素解析器で解析した。次に、word2vec のツールに含まれる、出現頻度に基づいて隣接する単語を結合する word2phrase を 2 回適用し、1~4 グラムにまとめあげる。その結果を word2vec に適用し、最後に k -means クラスタリングによってクラスタリングする。パラメータとして、word2vec のベクトルの次元数と、クラスタ数を選択する必要がある。ベクトルの次元数 8 種類 (50, 100, 150, 200, 250, 300, 350, 500) と、クラスタ数 6 種類 (100, 500, 1000, 2000, 5000, 10000) の組合せを、学習データの 5 分割交差検定によって選択した。結果として、Wikipedia に対しては 250 次元の 10000 クラスタが、ウェブ文書に対しては 350 次元の 10000 クラスタ、ツイートに対しては 300 次元の 10000 クラスタがパラメータとして選択された。

モダリティ解析の学習に用いるデータは、ツイートからランダムサンプルした 96,824 事例、および人手で作文した否定表現をとまなう文の 4,048 事例である。各事例は、述語を含む文節に対してそのモダリティを表現するラベルとして「事実」、「推量」、「否定」、「仮定」、「疑問・要求」の 5 種類のいずれかが付与されたものである。

学習データの 5 分割交差検定の結果を表 2 に示す。いず

れも再現率がやや低いものの、「否定」、「疑問・要求」は、実用に十分な適合率が得られた。

5.2 質問応答モードの性能評価

DISAANA の質問応答モードの性能を、約 8,500 万件のツイートを用いて評価する。プロトタイプシステムの性能評価には、後藤らと同様に 345 個の災害関連単語を含むものに限定した結果残る約 5,400 万件のツイートを利用した。後述のように、質問に対する正答は約 5,400 万件のツイートから抽出されたものであるが、差分となる約 3,100 万件のツイートから抽出された回答候補が正解となる場合があるため、検索対象を増やしたことは再現率の向上に寄与することが期待できる。また、プロトタイプシステムの性能評価では、質問入力から回答候補までの処理時間に制限を設けなかったが、DISAANA は現実的な状況を想定して、10 秒の上限時間を設けた^{*11}。

評価対象として、後藤らの評価実験と同様に、川田ら [26] が構築した 300 個の質問のうち、5,400 万件のツイートから正答を見つけることができた 192 個の質問と、それらに対する 17,524 個の正答を用いる。192 個の質問それぞれに対して DISAANA で回答候補の集合を得る。再現率は、出力結果が正答に含まれるかどうかによって計算する。このとき、正答に含まれるかの判定は、システムが出力した回答候補が、正答の文字列に含まれるかによって判断する。正答は、各質問に対して最大 1,000 件の関連ツイートを人手で見ても、抽出することで構築しているため、システムの出力のうち、正答には含まれていないが人手で確認をすると正解であると判定できる場合がある。後藤らと同様に、これらは正解として取り扱わない。適合率は、出力結果である質問・回答候補ペアからランダムに 250 ペアを選択して、人手で正解判定を行う。

評価結果を表 3 に示す。まず、再現率が大きく向上 (有意水準 1% で有意差あり) している。これは、パターン抽出の改善によって回答候補の抽出数を増やすことができたことが有効に働いたためであると考えられる。また、システムの高速化によって検索対象となるツイートを増やすこ

*11 公開している DISAANA の時間制限は、より短い 1 秒としているが、災害時などで再現率が優先される場合は、制限時間を緩めることを考えている。

表 3 東日本大震災データにおける質問応答モードの評価結果
 Table 3 Results of question answering using tweets posted during Great East Japan Earthquake.

	適合率		再現率		F 値
プロトタイプシステム	0.608	(152/250)	0.519	(9,099/17,524)	0.560
DISAANA	0.568	(142/250)	0.707	(12,382/17,524)	0.630

とができたことも、再現率の向上に寄与している。次に適合率について、モダリティ解析は、否定、疑問、要求を示唆する文からの回答候補抽出を抑制できるため、適合率の向上に寄与する。その結果、回答候補の抽出数が大幅に増えたにもかかわらず適合率が大きく下がることはなかった。F 値（調和平均）では、後藤らに比べて7ポイントの改善が見られた。適合率が下がった原因の1つは、ランダムサンプルした250の質問・回答候補ペアに含まれる質問の種類である。後藤らの方には、比較的正しく回答しやすい「何が発生していますか」という質問と回答候補のペアが82件含まれていたのに対して、本稿では2件であった。これは、本稿ではパターン抽出などの改善により、質問に対する回答候補の数が増えていることと、前述の質問応答の動作上限時間の違いが原因である。

適合率の評価に用いた250質問・回答候補ペアのうち正解であった112事例のうち、パターンによって回答候補が抽出されたのは48事例、周辺キーワードによって回答候補が抽出されたのは64事例であった。一方で、138の誤り事例については、それぞれ36事例、102事例であった。パターンによる回答候補抽出の誤りについて誤り分析を行った。

「どの X が〜」質問への誤り 12 事例 (33.3%) 「どの病院が開いていますか」「どんな危険がありますか」といった、疑問詞と名詞の組合せからなる質問文を解析する際、「どの病院」「どんな危険」を抽象化して、それぞれ(Aが開いている)、(Aがあります)のパターンにマッチするAが回答候補として抽出される。本来であれば、それらの候補に対して「病院」や「危険」の下位概念であるかを判定する必要があるが、「病院」は表層情報からある程度判定できることが期待できる一方で、「危険」は「放射線量〇〇シーベルト以上」のように数値表現を含めて判断する必要がある場合もあり、その下位概念を表層情報から判定することは難しい。特に災害に特化して上位下位概念を整備する必要がある。

パターン言い換えの誤り 9 事例 (25.0%) 言い換えたパターンが誤っていた事例で、(Aを提供)と(Aを提案)は言い換え可能となっているため、質問「何を提供していますか?」に対して、「Xを提案している」からXを回答候補として抽出しているが、誤りである。言い換えパターンは自動獲得されているため、言い換

え不可能なパターンを人手で除外していくことが必要である。

質問の解析誤り 6 件 (16.7%) 「流行っている病気は何ですか」や「ボランティアの作業は何になりますか」といった、疑問代名詞が文の後方に現れる質問は、(Aが流行っている)のように解析され、「病気」以外に流行っているものも回答候補として抽出される。誤り分類1と同様に、抽出された回答候補が「病気」や「作業」の下位概念であるかを精査する必要がある。

モダリティ解析の誤り 3 事例 (8.3%) 「どこで安否確認ができますか」といった質問に対して、安否確認ができないとされる場所を回答候補としてしまった、モダリティ解析の誤り。誤った3事例を示す。それぞれ、下線部が抽出された回答候補である。

- 質問「どこで安否確認ができますか」に対する「安否確認のために 高速道路 を利用しようという考えだけは絶対に避けて欲しい。」
- 質問「何が発生していますか」に対する「地震で揺れた。最初 目眩 が起こったかと思った。」
- 質問「どこに遺体はありましたか」に対する「藤沢市が江ノ島に鳥葬の施設を作って遺体を受け入れようとしたが、周辺の鎌倉市と 茅ヶ崎市 の猛反対で中止になったらしい。」

それぞれ、述語「利用する」「起こる」「受け入れる」が後続の表現によって否定されているが、正しく解析することができなかった。このような複雑な表現は、大量のデータからも効率良く集めることができないため、誤った事例を学習データに加えていくことが有効である。

その他 6 事例 (16.7%) その他の誤り事例は、係り受け解析の誤りによるものや、地名を聞く質問に対して、正答ではない地名を出力した誤りなどである。

後藤らは、誤り分析において、単純な否定文や疑問文から回答候補を抽出してしまう誤りが散見されたと報告しているが、モダリティ解析を導入することによって、それらの誤りを大幅に減らすことができた。その結果、質問文の解析や、パターン言い換えに改善の余地が大きいことが明らかとなった。

6. DISAANA のの有効性の検証

我々は、DISAANA を一般の利用者が利用する場合の問

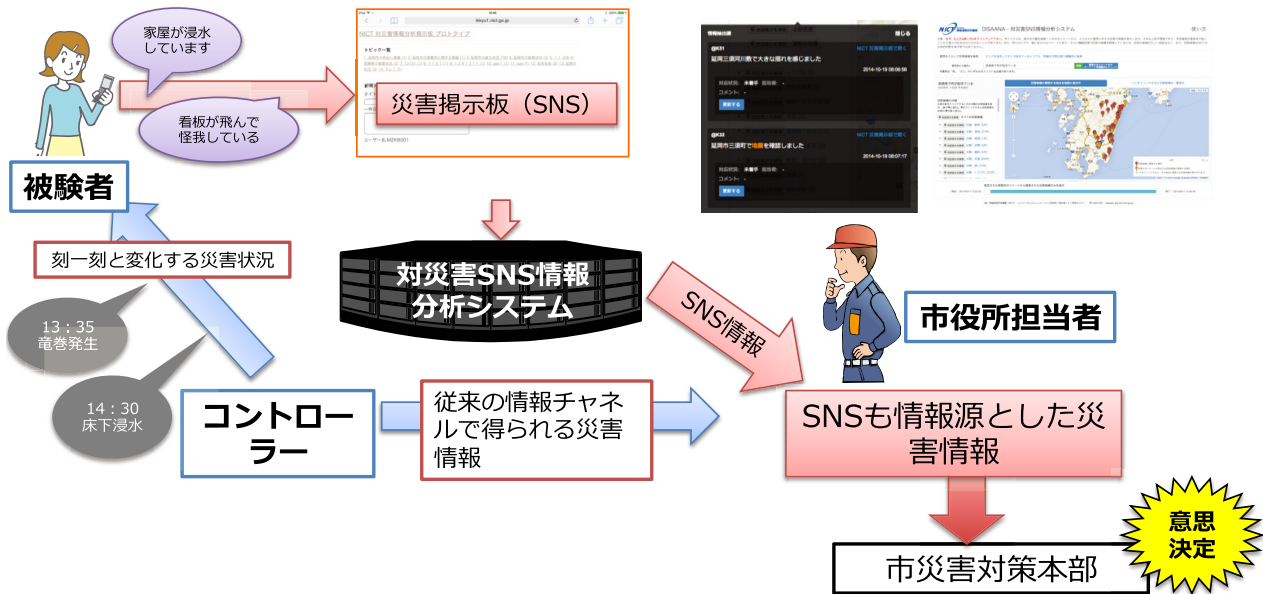


図 9 DISAANA を用いた実験概要

Fig. 9 Outline of DISAANA demonstration.

題点や、自治体などで実際の災害対応を前提として用いる場合の問題点を明らかにするために、宮崎県において DISAANA の有効性を検証した。

宮崎県は、台風銀座と呼ばれるように地理的・自然的条件などから台風の接近にともなう風水害や土砂災害が発生しやすいところである。台風や集中豪雨による大きな土砂災害も近年では 1997 年、2004 年、2005 年と頻繁に発生し、県内各地に大きな爪痕を残している。また、南海トラフによる巨大地震の発生も懸念される。そのため県全体として防災意識が高く、防災士の育成に力を入れている。

これらの防災士や、防災意識の高い一般市民が SNS 上に災害関連情報を提供することを想定する。DISAANA を用いて提供された情報を分析することで、自治体における災害対応の意志決定に有用な情報を提供できるかどうかについて、机上訓練形式の防災訓練を通じた実験により検証した。

実験の概要を図 9 に示す。実験では、コントローラと呼ばれる人員を配置し、訓練開始後の時間経過にあわせて被験者および市役所担当者に被害状況などを与える形式をとった。すなわち、シナリオは存在するが、被験者および市役所担当者にそれは知らされておらず、訓練開始後に逐一被害状況が伝達される形式である。防災士を含む一般市民役の被験者は、設定された状況下において想定される被害状況を SNS を模した掲示板に自由に書き込んだ。本掲示板には Twitter のような文字数制限は設けなかったが、インストラクションで「Twitter のように被害状況を簡潔に書き込む」よう依頼したため、投稿は平均 73.5 文字であった。また、投稿内容に関する制限はなく、本訓練とは無関係の書き込みも散見された。そして、DISAANA がそれらの書き込みを解析する。市役所担当者は、コントロー

ラより与えられる従来の情報チャネルで得られる被害情報に加えて DISAANA を用いて得られる SNS からの情報もあわせて、避難、救援などの意志決定を行った。また、被験者も、DISAANA を利用して自らの書き込みが期待するように処理されているかどうかを確認した。

実験は 2015 年 1 月 18 日に延岡市で、2015 年 2 月 7 日に宮崎市で実施し、延べ 115 名の被験者、15 名の自治体関係者、2 名のコントローラによる 5 時間半の訓練を通して計 4,400 件以上の書き込みと 1,760 回を超える DISAANA での質問応答を得ることができた。被験者は、延岡市、宮崎市在住の防災士資格を持つものを含む 18 歳以上の一般市民である。被験者は、NICT で用意したタブレット端末もしくはノート PC を用いて掲示板への書き込み、DISAANA の利用を行った。災害対策本部の市役所関係者は、すべてノート PC を利用した。被験者全員を会議室に集め、2 時間程度操作方法などの説明を行ってから机上訓練（実験）を実施した。災害対策本部の市役所担当者にも同じく 2 時間程度の説明を行い、被験者とは別の部屋に設置した災害対策本部で実験を行った。

災害対策本部で用いた DISAANA は、被験者が利用した通常のものとは異なり、自治体向けの機能を組み込んである。それは、書き込みに対して、(1) 着手、(2) 未着手、(3) 解決、(4) デマと認定という 4 状態を設定できるようになっており、さらに特定の状態だけに絞り込んで書き込みを表示できる。また、状態の設定に加えてコメントを書き込むことができ、書き込まれたコメントは掲示板に災害対策本部からの書き込みとして反映される。なお、この自治体向けの機能は、開発中であり、現在試験公開している DISAANA には組み込まれていない。

実験後に、災害対策本部で災害対応にあたった市役所担

表 4 「災害時に DISAANA は役立つと感じるか」のアンケート結果
 Table 4 Questionnaire results: Is DISAANA useful for disaster situations??.

	強く感じる	感じる	感じない	まったく感じない	合計
被験者	53	56	6	0	115
市役所関係者	3	12	0	0	15
合計	56	68	6	0	130

当者と一般市民役の被験者にアンケートを実施した。災害時に DISAANA が役立つと感じるかを「強く感じる」「感じる」「感じない」「まったく感じない」の4段階で聞いた結果を表 4 に示す。「強く感じる」「感じる」を合わせると、被験者の 94.8%、市役所関係者は全員が「役立つ」と感じたという結果が得られた。次に、自由記述形式のアンケートをとったところ、被験者側からは、DISAANA に対する期待や、様々な改善すべき点の指摘があった。市役所関係者からも災害対応に役立つという好評を得る一方で、災害対応を実務とする者の視点から非常に貴重な改善点の指摘、コメントがあった。その一部を以下に示す。

被験者からのコメント・要望

- スマホから使う際には GPS と連動して欲しい (20 代女性)。
- 登録されていない地名があった (60 代男性)。
- 入力が災害ごとにカテゴリ分けされ、エリアや人員などの記入で済むと楽 (40 代男性)。
- 災害に限らず、防犯、交通情報、不審者情報にも使えるのでは (60 代男性)。
- 災害時には是非利用したい (10 代女性)。
- 慣れてしまうと意外と簡単だった (40 代女性)。

市役所関係者からのコメント・要望

- 同一の災害事象に関する書き込みに一括して返信できるとよい。
- 結果が負傷者・火災・倒壊などの区分で分けられているとよい。
- 結果に対して単語検索ができるとよい。
- 写真付きの情報は、状況把握がよりしやすくなるのでよい。
- 意志決定上、写真で把握できる場合はすぐに動ける。
- 結果の表示の際に最新の書き込みが上に来るようにして欲しい。

要望のいくつかは、現在試験公開している DISAANA に反映させることができた。また、実験を通して情報の信憑性をいかにとらえるべきかという課題も示された。

7. おわりに

本研究では、災害時に Twitter 上に投稿される情報を効率的に検索できるシステムとして DISAANA を開発し、誰でも利用可能な Web アプリケーションとして試験公開した。本システムは、後藤らが東日本大震災を契機に開発し

たシステムを改善したものであり、モダリティ解析、属性判定、予報表現抽出が新たに導入されている。質問応答モードの性能を、東日本大震災のツイートを用いて、後藤らのプロトタイプシステムと比較したところ、適合率はやや下がるものの、災害時に重要な再現率は大幅に向上し、F 値では 7 ポイントの改善が見られた。後藤らが大きな問題として報告していた、否定事象からの回答候補の誤抽出は、モダリティ解析を新たに導入することでほぼ解消された。誤り分析によって、パターン言い換えの適用や質問文解析に改善の余地があることが明らかとなった。今後は、これらの改善を行っていく。

謝辞 本研究で利用したツイートは、株式会社ホットリンク様よりご提供いただきました。ここに記して感謝いたします。

参考文献

- [1] Varga, I., Sano, M., Torisawa, K., Hashimoto, C., Ohtake, K., Kawai, T., Oh, J.-H. and De Saeger, S.: Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster, *Proc. 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp.1619–1629 (2013).
- [2] 後藤 淳, 大竹清敬, Stijn, D.S., 橋本 力, Julien, K., 川田拓也, 鳥澤健太郎: 質問応答に基づく対災害情報分析システム, *自然言語処理*, Vol.20, No.3, pp.367–404 (2013).
- [3] Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A.A., Lally, A., Murdock, J.W., Nyberg, E., Prager, J., Schlaefel, N. and Welty, C.: Building Watson: An Overview of the DeepQA Project, *AI Magazine*, Vol.31, No.3, pp.59–79 (2010).
- [4] 吉村 健: シャベってコンシェルと言語処理, *IPSJ SIG Technical Report*, Vol.2012-SLP-93, No.4, pp.1–6 (2012).
- [5] 乾健太郎, 松吉 俊: 言語情報編集のための広義モダリティ解析に向けて, *Japio year book*, pp.128–133 (2009).
- [6] 松吉 俊, 江口 萌, 佐尾ちとせ, 村上浩司, 乾健太郎, 松本裕治: テキスト情報分析のための判断情報アノテーション, *電子情報通信学会論文誌. D, 情報・システム*, Vol.93, No.6, pp.705–713 (2010).
- [7] 江口 萌, 松吉 俊, 佐尾ちとせ, 乾健太郎, 松本裕治: モダリティ, 真偽情報, 価値情報を統合した拡張モダリティ, *言語処理学会第 16 回年次大会*, pp.852–855 (2010).
- [8] Sutton, C., McCallum, A. and Rohanimanesh, K.: Dynamic Conditional Random Fields: Factorized Probabilistic Models for Labeling and Segmenting Sequence Data, *The Journal of Machine Learning Research*, Vol.8, pp.693–723 (2007).
- [9] Saurí, R. and Pustejovsky, J.: Determining Modality and Factuality for Textual Entailment, *Proc. 1st IEEE International Conference on Semantic Computing*, pp.509–516 (2007).

- [10] Pustejovsky, J., Verhagen, M., Saurí, R., Littman, J., Gaizauskas, R., Katz, G., Mani, I., Knippen, R. and Setzer, A.: *TimeBank 1.2* (2006).
- [11] Saurí, R. and Pustejovsky, J.: FactBank: a corpus annotated with event factuality, *Language Resources and Evaluation*, Vol.43, No.3, pp.227-268 (2009).
- [12] Kurohashi, S., Nakamura, T., Matsumoto, Y. and Nagao, M.: Improvements of Japanese morphological analyzer JUMAN, *Proc. International Workshop on Sharable Natural Language*, pp.22-28 (1994).
- [13] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis, *Proc. 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)*, pp.230-237 (2004).
- [14] Yoshinaga, N. and Kitsuregawa, M.: Polynomial to Linear: Efficient Classification with Conjunctive Features, *Proc. 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP 2009)*, pp.1542-1551 (2009).
- [15] Kloetzer, J., Saeger, S.D., 鳥澤健太郎, 佐野大樹, 後藤淳, 橋本 力, 呉 鍾勲: Supervised Recognition of Entailment Between Patterns, 言語処理学会第 18 回年次大会, pp.431-434 (2012).
- [16] Hashimoto, C., Torisawa, K., De Saeger, S., Oh, J.-H. and Kazama, J.: Excitatory or Inhibitory: A New Semantic Orientation Extracts Contradiction and Causality from the Web, *Proc. 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, pp.619-630 (2012).
- [17] 田仲正弘, 大竹清敬, 鳥澤健太郎, 田浦健次朗: RaSC: 高速なストリーム通信をサポートする言語処理プログラムの高速化・高並列化ミドルウェア, 言語処理学会第 20 回年次大会, pp.705-708 (2014).
- [18] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J.: Distributed representations of words and phrases and their compositionality, *Advances in neural information processing systems*, pp.3111-3119 (2013).
- [19] Vapnik, V.: *The nature of statistical learning theory*, Springer Science & Business Media (2000).
- [20] Chang, C.-C. and Lin, C.-J.: LIBSVM: A library for support vector machines, *ACM Trans. Intelligent Systems and Technology*, Vol.2, pp.27:1-27:27 (2011).
- [21] Kazama, J. and Torisawa, K.: Inducing Gazetteers for Named Entity Recognition by Large-Scale Clustering of Dependency Relations, *Proc. ACL-08: HLT*, pp.407-415 (2008).
- [22] Oh, J.-H., Torisawa, K., Hashimoto, C., Kawada, T., De Saeger, S., Kazama, J. and Wang, Y.: Why question answering using sentiment analysis and word classes, *Proc. 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, pp.368-378 (2012).
- [23] Kloetzer, J., Torisawa, K., Hashimoto, C. and Oh, J.-H.: Large-Scale Acquisition of Entailment Pattern Pairs by Exploiting Transitivity, *Proc. Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pp.1649-1655 (2015).
- [24] Sano, M., Torisawa, K., Kloetzer, J., Hashimoto, C., Varga, I. and Oh, J.-H.: Million-scale Derivation of Semantic Relations from a Manually Constructed Predicate Taxonomy, *Proc. COLING 2014, the 25th International Conference on Computational Linguistics: Tech-*

nical Papers, pp.1423-1434 (2014).

- [25] 川田拓也, Kloetzer, J., 鳥澤健太郎, 橋本 力: 質問応答システムのための含意パターンペアの生成, 言語処理学会第 21 回年次大会, pp.159-162 (2015).
- [26] 川田拓也, 大竹清敬, 後藤 淳, 鳥澤健太郎: 災害対応質問応答システム構築に向けた質問・回答コーパスの構築, 言語処理学会第 19 回年次大会, pp.480-483 (2013).



水野 淳太 (正会員)

1984年生。2006年豊橋技術科学大学工学部情報工学課程卒業。2009年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。2012年同博士課程修了。博士(工学)。同年より東北大学大学院情報科学研究科研究員, 2013年より独立行政法人情報通信研究機構研究員, 2015年より国立研究開発法人情報通信研究機構研究員, 現在に至る。自然言語処理の研究に従事。言語処理学会, 人工知能学会各会員。



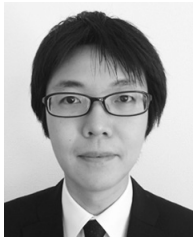
後藤 淳

1993年徳島大学大学院工学研究科博士課程前期修了。同年日本放送協会入局。2006年ニューヨーク大学客員研究員, 2011年より2015年まで情報通信研究機構専門研究員。現在, NHK放送技術研究所上級研究員。博士(情報学)。自然言語処理の研究に従事。映像情報メディア学会, 電子情報通信学会, 日本データベース学会各会員。



大竹 清敬 (正会員)

2001年豊橋技術科学大学大学院博士後期課程修了。博士(工学)。同年より株式会社ATR音声言語コミュニケーション研究所。2006年より独立行政法人情報通信研究機構。2015年より国立研究開発法人情報通信研究機構を経て, 現在, 同機構情報分析研究室主任研究員および情報配信基盤研究室室長を兼務。音声言語処理, 自然言語処理の研究に従事。言語処理学会, 人工知能学会各会員。



川田 拓也

1979年生。2003年国際基督教大学教養学部卒業。2005年京都大学大学院文学研究科博士前期課程修了。2010年同博士後期課程修了。博士(文学)。2011年より独立行政法人情報通信研究機構研究員, 2015年より国立研究開発法人情報通信研究機構研究員, 2016年よりNEC情報・ナレッジ研究所主任, 現在に至る。言語資源の設計と構築に関する研究に従事。日本語用論学会, 言語処理学会各会員。



橋本 力 (正会員)

1999年福島大学教育学部卒業。2001年北陸先端科学技術大学院大学博士前期課程修了。2005年神戸松蔭女子学院大学院博士後期課程修了。同年より京都大学情報学研究科産学官連携研究員, 2007年より山形大学大学院助教, 2009年より独立行政法人情報通信研究機構, 2015年より国立研究開発法人情報通信研究機構を経て, 現在, 同機構情報分析研究室研究マネージャー。博士(言語科学, 情報学)。文部科学大臣表彰若手科学者賞, 情報処理学会論文賞等受賞。



鳥澤 健太郎 (正会員)

1968年生。1992年東京大学理学部卒業。1995年同大学院中退。同大学院助手, JST さきがけ研究 21 研究員兼任, 北陸先端科学技術大学院大学助教授を経て, 現在, 国立研究開発法人情報通信研究機構情報分析研究室室長。博士(理学)。自然言語処理の研究に従事。日本学術振興会賞, ドコモ・モバイル・サイエンス賞等受賞。言語処理学会会員。



奥村 明俊 (正会員)

1986年京都大学大学院工学研究科修士課程修了。同年日本電気(株)入社。自然言語処理, 音声翻訳, メディア処理技術の研究開発に従事。博士(工学)。情報処理学会平成20年度喜安記念業績賞, 第21回独創性を拓く先端技術大賞経済産業大臣賞, 第1回AAMT長尾賞等受賞。現在, (株)NEC情報システムズ執行役員, 国立研究開発法人情報通信研究機構協力研究員。



クロエツェー ジュリアン

2006年パリ第6大学卒業。2010年北陸先端科学技術大学院大学博士課程修了。博士(情報科学)。2011年より独立行政法人情報通信研究機構研究員, 2015年より国立研究開発法人情報通信研究機構研究員, 現在に至る。



田仲 正弘 (正会員)

1981年生。2004年京都大学工学部情報科学科卒業。2005年同大学大学院情報学研究科修士課程修了。2009年同博士課程修了。博士(情報学)。同年より独立行政法人情報通信研究機構研究員, 2015年より国立研究開発法人情報通信研究機構主任研究員, 現在に至る。大規模情報分析のための基盤ソフトウェアの研究に従事。電子情報通信学会会員。