

大規模災害時の膨大な被災報告を人工知能で瞬時に整理・要約

～災害状況要約システム「D-SUMM」を試験公開～

【ポイント】

- Twitter上の災害関連の投稿をリアルタイムに自動分析し、要約できるシステムを公開
- 「救助」と「火災」など異なる種類の情報を地図上で同時に確認可能
- D-SUMM 公開アドレス <http://disaana.jp/d-summ/>

国立研究開発法人情報通信研究機構(NICT、理事長: 坂内 正夫)は、耐災害 ICT 研究センター及びユニバーサルコミュニケーション研究所において開発している災害状況要約システム「D-SUMM(ディーサム)」(Disaster-information SUMMMarizer)を10月18日(火)からWeb上に試験公開しました。

D-SUMMは、人工知能を用いて、Twitter*¹に投稿された災害関連情報をリアルタイムに分析し、都道府県単位又は市区町村単位でエリアを指定すると、指定エリア内の被災報告を瞬時に要約し、そのエリアの被災状況の概要が一目でわかるように、コンパクトかつ、わかりやすく提示し、各種救援、避難等を支援します。D-SUMMは、パソコンのWebブラウザ又はタブレット端末から<http://disaana.jp/d-summ/>にてご利用いただけます。

なお、本研究の一部は、内閣府総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム(SIP)「レジリエントな防災・減災機能の強化」(管理法人・JST)の支援を受けて実施したものです。

【背景】

NICTは、既に、対災害SNS情報分析システムDISAANA(ディサーナ)*²のリアルタイム版を2015年4月8日から試験公開し、2016年の熊本地震をはじめとする災害において、Twitter上の災害関連情報へ素早くアクセスできる手段として、ご活用いただいております。しかしながら、DISAANAでは、大規模災害の発生時に膨大な被災報告が出力され、被災状況の概要を一目で把握することは困難でした。そこで、この課題を解決すべく、災害状況要約システムの研究開発を進めてきました。

【今回の成果】

災害状況要約システムD-SUMMは、Twitter上の膨大な災害関連報告をわかりやすく整理し、要約するシステムです。NICTが既に公開しているDISAANAでは、「火災が発生している」「火事が起きている」など、意味的に類似する被災報告が別々に出力されていきました。D-SUMMでは、これらの報告を一まとめにすることで、よりコンパクトに被災報告を要約して提示します。また、被災報告をそのタイプ(地震、道路やインフラの被害、物資の不足等)毎に分類して、必要とする情報へのアクセスを容易にします。さらに、指定エリアの下位のエリア(指定エリアが県の場合は、県下の市町村)単位毎に被災報告を整理し、重大な被災報告が多く挙がっているエリアから順に表示することで、どのエリアの被害が大きいかをわかりやすく提示します。

これらの機能を実現するために、DISAANAで被災報告のタイプを分類するために使用していた意味カテゴリー辞書*³(2,800万語)を機械学習、統計処理を用いて細分化し、より細かい意味カテゴリーを設けました。こうした工夫によって、地図上での表示も含めて、より直感的で分かりやすい被災状況の提示が可能になり、効率的な救援、避難の支援が可能になると期待されます。また、大規模災害時におけるD-SUMMを試していただくことができるように、熊本地震の際のツイートを対象とした「熊本地震試用版*⁴」のD-SUMM(<http://disaana.jp/d-summ-km/>)及びDISAANA(<http://disaana.jp/H28KM/>)を公開しました。

【今後の展望】

今後は、DISAANA同様、実際に救援活動等を行う組織と協力して、実証実験等を行い、更なる機能追加や使い勝手の向上を図ってまいります。



D-SUMMの地図表示動作例

<用語解説>

*1 Twitter

140文字以内の「ツイート」と称する短文を投稿できる情報サービスで、Twitter社により提供されている。

*2 DISAANA

対災害 SNS 情報分析システム DISAANA(ディサーナ)は、NICTがこれまでに培ってきた情報分析技術を活用し、質問応答モード、エリア検索モードの2つの機能を提供するシステムである。2015年4月8日から、現在投稿されているツイート(日本語ツイートの10%ランダムサンプルを対象とする。これは、D-SUMMも同様である。)をリアルタイムに分析し、回答候補を出力するシステムとして試験公開を開始した。2016年熊本地震においては、内閣官房のツイッター分析班において実際に使用され、熊本県の災害対策本部に分析結果が提供された。

*3 意味カテゴリー辞書

単語に対して意味的カテゴリーを付与した辞書で、例えば、「地震」という単語に対して【災害】というカテゴリーが付与してある。DISAANAでは、このような辞書を2,800万単語について整備しており、回答候補を提示する際に、意味的カテゴリー毎に整理して出力するために使用している。D-SUMMでは、この意味的カテゴリーを2階層にし、「地震」に対しては【災害-地震】のような意味的カテゴリーを持つように拡張した。これにより、「地震が発生した」「余震が起きている」などの表現については、「地震」、「余震」の単語の意味的カテゴリーが【災害-地震】であることから、これらをまとめて、「地震が起きる」という表現へ要約する。

*4 熊本地震試用版 D-SUMM、DISAANA

熊本地震時(2016年4月14日から4月21日までの8日間)のツイート(日本語ツイートの10%ランダムサンプルを対象とする)を用いたシステム。熊本地震時の投稿に基づいて、D-SUMMやDISAANAがどのように動作するかを試すことができる。ただし、元の投稿が削除されている場合には、その投稿から抽出された回答の候補は出力されるが、元の投稿を確認することができなくなる。

<D-SUMM が展示されるイベント>

- NICT オープンハウス 2016 (本部)
 - 日時: 2016年10月27日(木)・28日(金) 9:30-17:00 (28日は16:30まで)
 - 場所: 国立研究開発法人情報通信研究機構(本部) 東京都小金井市貫井北町 4-2-1
 - URL: <http://nict.go.jp/open-house2016/index.html>
- けいはんな情報通信フェア 2016 (NICT オープンハウス 2016 in けいはんな)
 - 日時: 2016年11月10日(木)(13:00-17:00)、11日(金)(10:00-17:00)、12日(土)(10:00-16:30)
 - 場所: けいはんなプラザ 京都府相楽郡精華町光台 1-7
 - URL: <http://khn-fair.nict.go.jp/>

<補足資料について>

D-SUMMの各種機能等を説明した詳細な補足資料は、以下のURLにて公開しております。どうぞ、こちらをご覧ください。

<http://www.nict.go.jp/press/2016/10/18-1.html>

* Twitter は、Twitter,Inc.の商標又は登録商標です。

* その他、各会社名、各製品名及びサービス名などは、各社の商標又は登録商標です。

< 本件に関する問い合わせ先 >

耐災害 ICT 研究センター
応用領域研究室
大竹 清敬
Tel: 0774-98-6329
E-mail: d-summ@khn.nict.go.jp

< 広報 >

広報部 報道室
廣田 幸子
Tel: 042-327-6923
Fax: 042-327-7587
E-mail: publicity@nict.go.jp

【目的】

今回試験公開を開始した D-SUMM の目的は、大規模災害が発生した際に、Twitter に発信される膨大な被災報告を瞬時に、そのタイプ毎、地域毎に整理、要約してわかりやすく提示することです。これによって、被災状況の全体像の把握を短時間でいき、適切な避難や効果的な救援活動を行うためにご活用いただきたいと考えております。

D-SUMM の出力の具体例として、2016 年 4 月の熊本地震の際の D-SUMM(熊本地震試用版)の出力例を図 1、2 に示します。図 1 では、被災報告が熊本県内の自治体毎に提示されており、また、「建物」「ライフライン」「救助」といった被災報告のタイプ毎に整理して提示されています。図 2 では、そうした情報を地図上に整理して表示しています。こうした要約によって、被害が深刻なエリアの特定や、大きな問題の特定を含めて、被災状況の全ぼうの把握が容易になります。これらの情報は、熊本地震本震が発生してから 12 時間の間に、被災者を中心として Twitter 上に発信された数千件に上る被災報告を瞬時に要約、整理して得られたものです。

以下では、この D-SUMM の使い方や、要約の見方について解説します。

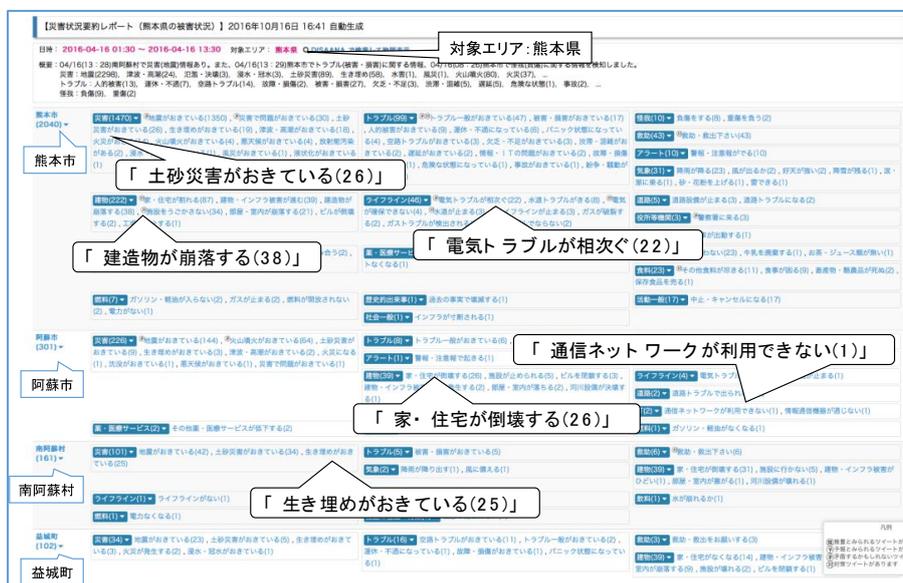


図 1 D-SUMM による熊本地震の被災報告の要約

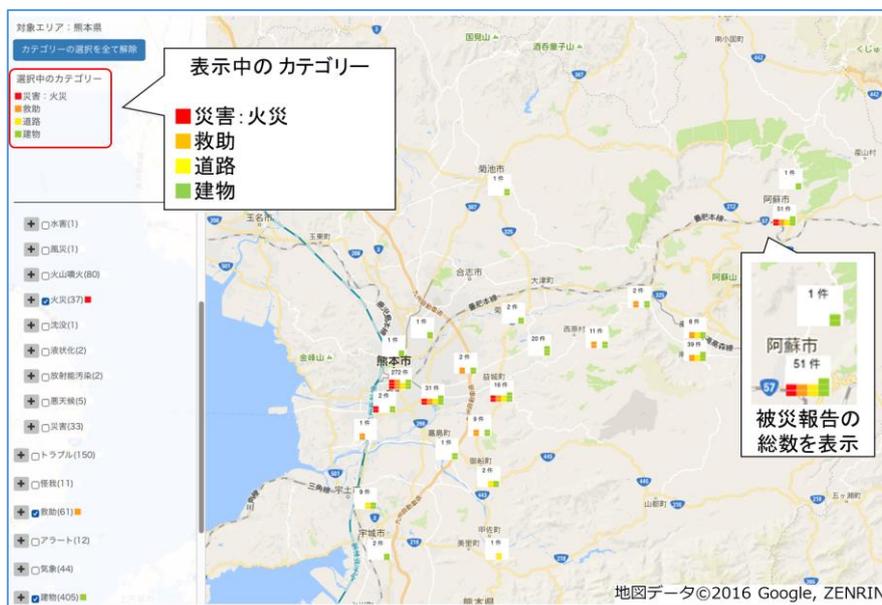


図 2 D-SUMM による熊本地震の被災報告の要約の地図表示

【基本的な使い方】

まず、災害状況要約システム D-SUMM の基本的な使い方を説明します。具体的には、以下の図 3 に示すように、4 ステップの手順になります。

- ① パソコン又はタブレットのブラウザに、以下のいずれかの URL を入力し、D-SUMM のサイトを開く。
 http://disaana.jp/d-summ/（今現在のツイートを要約するリアルタイム版。要約の対象となるツイートは利用する日の 3 日前から利用日までの最長 4 日間から選択可能）
 http://disaana.jp/d-summ-km/（熊本地震の際のツイートを要約する熊本地震試用版。要約の対象となるツイートは 2016 年 4 月 14 日から 4 月 21 日までの 8 日間から選択可能）
- ② Twitter 上の被災情報について要約の対象としたい自治体（都道府県又は市区町村）をメニューから選択する。
- ③ 要約の対象としたいツイートの発信時間帯を指定する。
- ④ 「エリア毎に要約」「カテゴリー毎に要約」「地図表示」のいずれかのボタンをクリックして、被災状況の要約を表示する。

①まず、パソコンまたはタブレットのブラウザで
 http://disaana.jp/d-summ (今現在のツイートを要約するリアルタイム版)
 もしくは
 http://disaana.jp/d-summ-km (熊本地震の際のツイートを要約する熊本地震試用版) を開きます。(まずお試しは熊本地震試用版をお勧めします。)

D-SUMMのサイトを開いた直後の初期画面 (図は熊本地震試用版)

②自治体を選択します

ツイートにある被災情報を要約したい地域（県、その他の自治体）をメニューから選択。都道府県のみ選択でも要約できます

③要約の時間帯を設定します

要約の対象となるツイートが発信された時間帯をメニューで設定します。（時間帯の開始時間と終了時間を選択します。今現在のツイートを要約するリアルタイム版では、現在時刻の直近数時間を選択しての設定もできます）

④要約を表示します

「エリア毎に要約」「カテゴリー毎に要約」「地図表示」のボタンのいずれかをクリックすることで要約が表示されます

要約の見方、操作法の詳細は後ほど説明します

地図データ©2016 Google, ZENRIN

図 3 D-SUMM 使用例

【要約の見方】

次に、先の手順を踏んで表示された要約の見方について説明します。

先の手順では、「エリア毎に要約」「カテゴリー毎に要約」「地図表示」のボタンのいずれかを押すことで要約が表示されますが、ボタン毎に要約の表示が異なりますので、各ボタンの解説を以下にします。

「エリア毎に要約」のボタンを押した場合

図 4 は、エリア毎に要約した場合の表示例です。この表示は、熊本地震試用版のサイトを開いて、「熊本県」をまず選択し、熊本地震の本震から 12 時間後までに(4 月 16 日 1:30 から 13:30 まで)発信されたツイートを対象として要約した場合です。D-SUMM のサイトを開いて、最初に指定した自治体が都道府県の場合は、その都道府県下の市区町村毎に被災状況を要約して表示します。図 2 の場合は、最初に「熊本県」を指定していますので、熊本県下の自治体、つまり、熊本市、阿蘇市、南阿蘇村等の市町村毎に被災状況を要約しています。

また、最初に指定した自治体が市区町村の場合は、その中にある地域毎(字、集落毎)に被災状況を要約します。

自治体、字、集落等は、重要な被災報告が多い方から順に並べて表示されます。図 4 の場合は、熊本県全体を最初に指定して要約しているため、重要な被災報告の多い、熊本市、阿蘇市、南阿蘇村、益城町が順に並びます。また、各エリアの要約としては、被災報告のカテゴリー毎(災害、ライフライン、救助、建物等)に被災報告が整理されて表示されます。こうした表示法により、どのエリアの被害が相対的に深刻であるか、また、どのエリアで、どのような被災報告が出ているかを短時間で把握することができます。

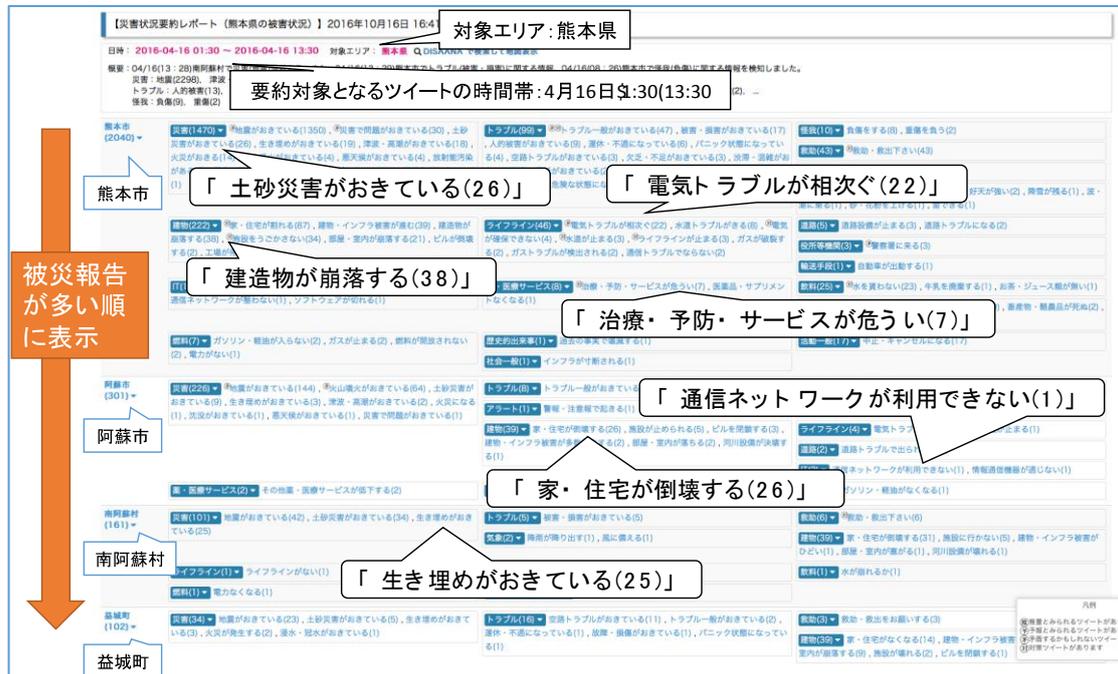


図 4 熊本地震試用版での「エリア毎に要約」ボタンを押した場合の要約の表示

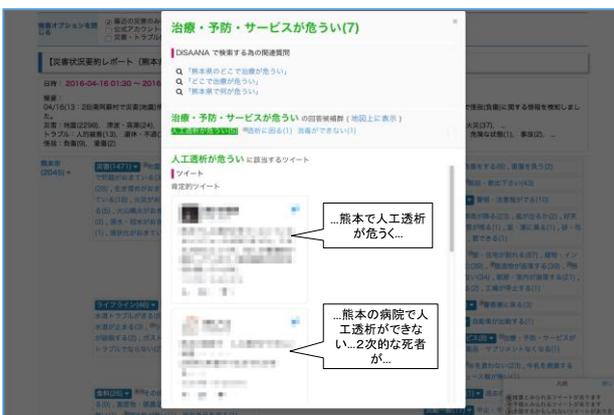


図 5 オリジナルのツイートの確認



図 6 「対」マークが付いた被災報告

要約に表示された「建造物が崩落する(38)」のような各項目は、「熊本城の石垣が崩れた」といったツイートからその意味が似通ったものをまとめることで表示されます。括弧の中の数字は、その項目の情報源となったツイートの数

を表しています。更に例を挙げると、「治療・予防・サービスが危うい」といった項目は、「人工透析ができない」「消毒ができない」といったツイートが基になって表示されます。また、各項目をクリックすると、図 5 に示すように、その情報の基となったツイートを表示したり、それに関する質問を、NICT が公開している災害に関する質問応答システムである DISAANA に問い合わせ、更に情報を取得することができます。DISAANA は、D-SUMM よりも確度の低い情報も発見するように設計されており、D-SUMM が発見していない情報を DISAANA では発見できる可能性があります。

また、D-SUMM で要約された項目には、図 6 の左側の「水を貰わない」の場合のように、「対策」の「対」の字をマルで囲ったマークが付いていることがあります。これは、報告された被災情報「水を貰わない」に対して、事態が沈静化した、もしくはその事態に対して対策が取られたことを報告するツイートが存在する可能性を示します。例えば、図 6 の右側は、左側の「水を貰わない」をクリックして表示したツイートですが、上半分が「熊本の友達が水が買えないと言っている」と被災状況の報告であるのに対して、下半分があるスーパーで水が無料配布されていること、つまり、水がない、という事態に対する対策が講じられていることを報告しています。こうした対策に関する報告をチェックすることで、救援活動のフォーカスを決定したり、あるいは被災者側からすれば、避難生活で有用な情報を取得することが可能になります。また、同様に、「矛盾」の「矛」をマルで囲ったマークが付いた被災報告もありますが、これは、被災報告と矛盾している報告がツイートされている可能性を示しており、この報告をクリックすることで、矛盾している可能性のあるオリジナルのツイートを見ることができます。仮に、矛盾した報告がツイートされていれば、元々の情報がデマである可能性がわかりますので、情報の信憑性を判断する際にご利用ください。

なお、D-SUMM で用いている各種の日本語解析技術は必ずしも完全ではありません。実際に避難や救援活動で利用される場合は、元のツイートまで表示して、D-SUMM の解析が正しいかどうかをご判断の上、ご利用ください。

「カテゴリ毎に要約」のボタンを押した場合

The screenshot shows the D-SUMM interface with the following elements:

- Header:** 南阿蘇村(6) and 熊本県阿蘇郡南阿蘇村の災害状況を要約
- Search/Filter:** Q DISAANAで検索, Q 場所毎に要約, Q カテゴリ毎に要約
- Left Panel (Categories):**
 - 災害(2696)
 - 地震がおきている(2330)
 - 水害がおきている(2)
 - 火山噴火がおきている(86)
 - 放射能汚染がある(2)
 - トラブル(169)
 - 人的被害がおきている(15)
 - 故障・損傷がおきている(2)
 - 欠乏・不足がおきている(4)
 - 危険な状態になっている(3)
 - カテゴリー: 救助
 - 救助(76)
 - アラート(33)
- Main Content:**
 - 救助・救出 > 南阿蘇村の回答候補群 (地図上に表示)
 - 助けて下さいに該当するツイート
 - ツイート
 - 肯定的ツイート
 - 熊本市(49) 南阿蘇村(9) 阿蘇市(7) 益城町(6) 西原村(3) 熊本県(2)
- Callout Box:** わずか数クリックで重要なツイートも容易に見発見

図 7 熊本地震試用版で「カテゴリ毎に要約」ボタンを押した場合の要約の表示

図 7 は、最初に熊本県を指定して、熊本地震本震発生直後 12 時間を時間帯として設定し、「カテゴリ毎に要約」した場合の要約の表示です。「エリア毎に要約」ボタンを押した場合には、画面左側には自治体の名前が並びましたが、「カテゴリ毎に要約」した場合には、被災報告のカテゴリー(「災害」「トラブル」「怪異」「救助」等)が並びます。例えば、「救助」の項を見ると、「救助・救出ください」という項目が表示されており、その横にそのタイプの被災報告が発信されているエリアが「熊本市(49)」「南阿蘇村(9)」「阿蘇市(7)」「益城町(6)」のように、該当する被災報告の数が多い順にリストアップされます。

「エリア毎に要約」の場合と同様に、要約の中の各項目をクリックすると、元のツイートや DISAANA への質問等が表示されます。図 7 では、実際に南阿蘇村で緊急の救助を求めているツイートを表示している例を示していますが、このような重要な情報を D-SUMM では数クリックで表示することができます。通常の検索エンジンで、「救助」等のキーワードを入れても、大量の有用でないツイートが表示されます。その中から、こうした有用なツイートを発見するには、大量のツイートを読む必要があり、短時間での作業は非常に難しくなりますが、D-SUMM では、必要な情報に短時間でアクセスすることが可能になります。

「地図表示」のボタンを押した場合

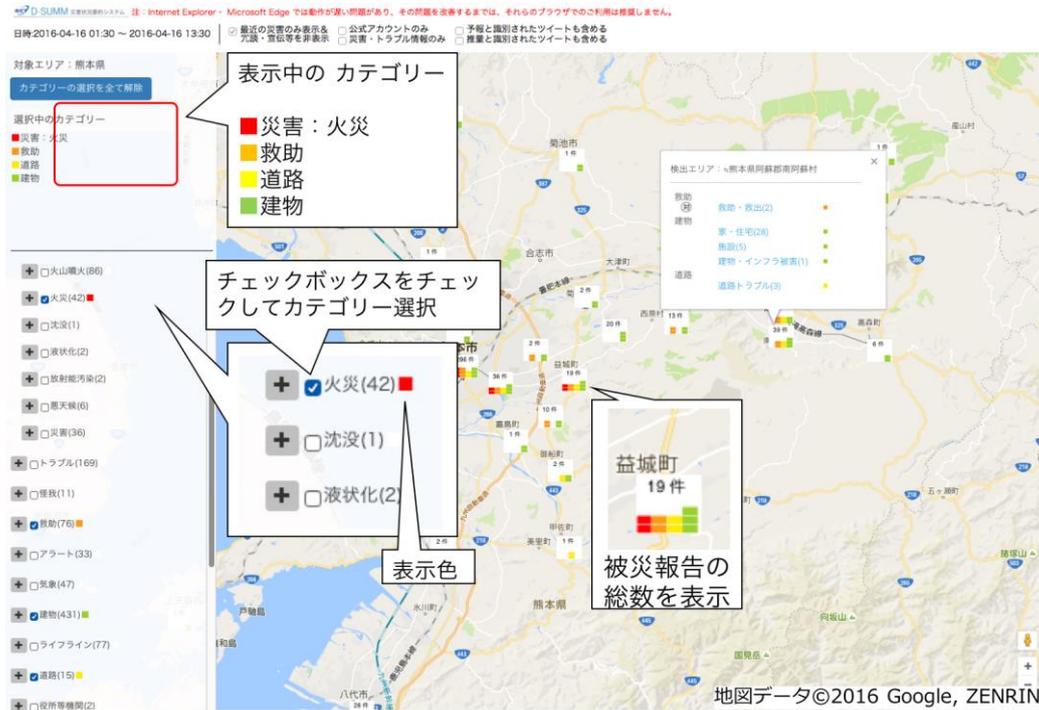


図 8 熊本地震試用版で「地図表示」ボタンを押した場合の要約の表示

図 8 は、「地図表示」ボタンを押した際の要約の例を示しています。地図表示ボタンを押すと、これまで操作していた画面では、カテゴリー毎の要約が表示され、さらに、新たにタブが開かれて、そこに地図を表示する画面が表示されます。なお、新規タブはポップアップとして扱われることが多く、また、ポップアップが制限されていることも多いので、地図表示を実行するためにポップアップを許可してください。この例の対象ツイートは、これまでと同様、熊本地震の本震から 12 時間分、つまり、4 月 16 日 1:30 から 13:30 までに発信されたものです。画面左側には、被災報告のカテゴリー（「災害」、「トラブル」、「怪我」、「救助」等）がリストアップされています。各々のカテゴリーには、チェックボックス（白い四角）が左横に付いていますが、それをクリックすると、該当するツイートに関係しているエリア毎に、地図上にツイート件数の大小を表すグラフと、ツイートの件数そのものが表示されます。チェックボックスをもう一度クリックすると、地図上の表示は消えます。グラフは、カテゴリー毎に色分けされて表示されます。どのカテゴリーにどの色が割り振られるかはチェックボックスをクリックした順序で決まります。例えば、図 8 では、画面左上に表示されているように、「災害：火災」に赤色、「救助」にオレンジ、「道路（の被害）」に黄色、「建物（の被害）」に緑色が割り当てられており、例えば、南阿蘇村周辺では救助要請（オレンジ）が出ているものの、道路の被害（黄色）や建物、インフラの被害（緑色）も出ているため、救援活動を行うには通行可能な道路を確認するなど、入念な準備が必要になります。また、広い範囲で建物の被害が報告されていますが、火災の報告は一部地域に集中していることも分かります。また、図 8 は、地図中のグラフの一つをクリックした直後の画像であり、吹き出しで、グラフでカウントされているカテゴリーが表示されています。それらのカテゴリーをクリックすると、被災報告の抽出元のツイートが表示されます。

なお、地図表示ボタンを押した直後は、「災害」というカテゴリーのチェックボックスがクリックされた状態で地図が表示され、地震、水害等の自然災害そのものに関する報告の件数等が赤いグラフで表示されます(図 3 右下参照)。上で説明した例を表示するためには、まず、「災害」というカテゴリーのチェックボックスをクリックして一旦表示を消し、上で述べたカテゴリーのチェックボックスを一つずつクリックする必要があります。

また、各カテゴリーには「+」の記号がチェックボックスの横にあります(図 8 左側参照)。これをクリックすると、カテゴリーの更に詳細な分類がリスト中に表示されます。例えば、「災害」の横の記号をクリックすると、「災害」カテゴリー内の更に詳細な分類(例えば、「地震」「津波・高潮」「氾濫・決壊」「土砂災害」「火災」)等がリストアップされます。図 8 では、「災害:火災」を赤色で地図中に表示しましたが、これは、「災害」カテゴリーの表示を一旦消去した後、その詳細な分類の一つである「火災」に関してチェックボックスをクリックして、地図中に表示しております。また、「+」の記号をクリックすると、「-」の記号に変化します。この「-」の記号をクリックすると、表示された詳細な分類が閉じられ、元の表示に戻ります。

【DISAANA と D-SUMM の違い】

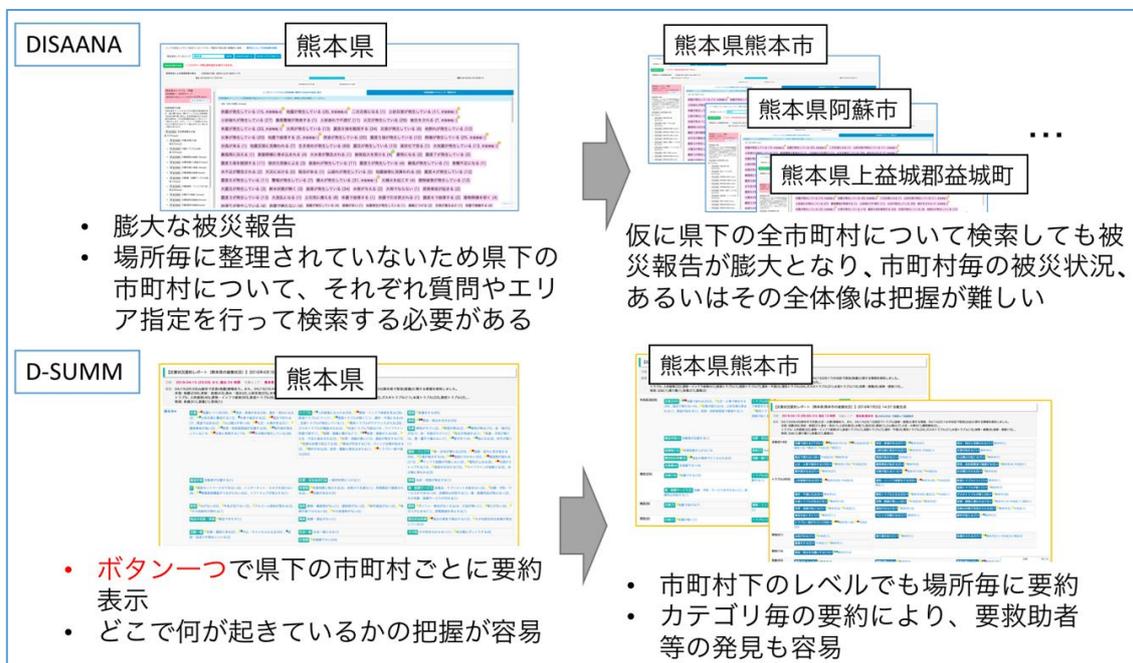


図 9 DISAANA と D-SUMM の比較

図 9 は、NICT が 2015 年に試験公開を開始した対災害 SNS 情報分析システム DISAANA と、今回試験公開する D-SUMM の違いを説明しています。熊本地震のような規模の災害になると、DISAANA のエリア検索モードでは、膨大な被災報告が出力されます。しかも、それらの被災報告は、場所毎に整理されていないため、例えば、各市区町村等の一つずつ指定して調べる必要がありました。D-SUMM では、このような煩わしさを解消し、ボタン一つで、県下の市区町村毎に要約を作成することができます。これにより、どこで何が起きているかを把握することが短時間で可能になります。また、カテゴリー毎に要約することで、要救助者等の発見も非常に容易に行えます。

一方で、DISAANA は、特に、質問応答モードで「熊本県で何が不足していますか?」のように質問を入力した場合、D-SUMM の提示する情報より、確度の低い情報まで発見するように作られており、特に重要な情報は DISAANA でも質問を入力され、検索されることをお勧めします。また、D-SUMM で提示された情報をクリックすると、DISAANA への問い合わせができるリンクが提示されますが、それらも、そうした目的でのご活用を考え、提供しているものです。