

関西大学審査学位論文

SNS情報構造に着目した人-計算機協調 による災害情報トリアージに関する研究

森野 穰

令和8年3月

関西大学大学院 総合情報学研究科

論文要旨

1995年の阪神・淡路大震災以降、インターネットは災害時における不可欠なライフラインとなり、特に近年では、代表的なソーシャルメディアであるX（旧 Twitter）が、個人の安否確認や被災状況の共有、救助要請などの手段として広く認知され、活用されている。SNSは「今起きていること」を即座に発信できる特性を持ち、社会のリアルタイムな状況をセンシングするための「ソーシャルセンサ」としての機能が期待されている。しかし、SNSの情報空間には、災害発生直後から情報の爆発的な増加（バースト現象）が生じ、被災地からの情報が、災害とは無関係な日常投稿や、被災者に共感する投稿、さらには偽・誤情報といった膨大なノイズ情報の中に埋没するという課題がある。既存の研究では、自然言語処理や機械学習を用いて有益な情報を「自動抽出」するアプローチが試みられてきたが、学習データの不足や未知の災害パターンへの対応の難しさ、そして災害被害情報の偽陰性を排除できないことから、人命に関わる重要情報の「取りこぼし」を完全に防ぐことは困難である。また、実際の災害対応現場では人的・時間的リソースが著しく不足しており、膨大な情報を人手ですべて検証することも現実的ではない。そこで本研究は、「情報トリアージ」の概念を適用し、ノイズ情報を段階的に除外（Subtraction）することで、人が確認すべき情報の効率的に確保する「人と計算機の協調による災害情報トリアージ」の方法論を構築することを目的とする。

本研究ではまず、信憑性の高いSNS災害情報を効率的に収集を行うための指針として、SNS情報の流通構造をモデル化した。災害時のSNS情報は、現場からの「一次情報（第1層）」、報道や転載による「二次情報（第2層）」、そして日常投稿や反応などの「ノイズ情報（第3層）」が層状に重なり、さらに「現場／現場外」および「悪用者」という軸が交差するマトリクス構造となっている。災害対応において真に必要とされるのは、被災者や現地目撃者による被害情報等を発信している第1層の一次情報であることを示した。情報の信頼性を判断するには、テキスト内容だけでなく、添付メディア（画像・動画）、発信者の属性、拡散の経緯（再掲痕など）といった複数の手がかりを総合的に参照する必要がある。しかし、緊急時にこれらを全件確認することは現実的ではないため、計算機が大量データの構造化と可視化を担い、人間が最終的な判断と除外を行うという役割分担が必要となる。本研究では、重要情報の「抽出」よりも「除外」を先行させるアプローチを採用する。具体的には、(1) 投稿の収集、(2) システムによる分析（クラスタリング）、(3) 情報整理（ノイズの除外）、(4) 情報精査（重要情報の確保）、(5) 共有、という段階的峻別プロセスを提案した。このプロセスにおいて、計算機は特徴ごとに投稿を「群（クラスタ）」としてまとめ、人間はその群単位で災害とは無関係な情報を除外していくことで、情報空間を縮小し、必要な情報の可視性を高めることを目指す。

提案モデルの妥当性を検証するため、「令和2年7月豪雨」発生時のX（旧 Twitter）データを対象に、流通した情報の調査を行った。「救助」および「避難」をキーワードに収集された約47万件の投稿に対し、添付された画像を人手により分類し、それぞれのテキスト特徴（TF-IDF）を分析した。調査の結果、収集された投稿のうち、実際に被災者が撮影した被害状況を含む投稿は全体のわずか0.564%に過ぎないことが明らかとなった。また、テキスト分析の結果、被災者による「被害」投稿には特出した単語特徴が見られず、テキスト情報のみに頼った自動抽出は困難であることが示された。一方で、「ゲーム」や「アニメ」、「日常風景」といったノイズ情報には、特定の語彙が高頻度で出現するという特徴が確認された。この結果は、特徴が希薄な

重要情報を抽出しようとする従来のアプローチよりも、特徴が明確なノイズ情報を特定して除外するアプローチの方が、災害時の情報収集において合理的かつ確実性が高いことを定量的に裏付けるものである。

次に、画像とテキストを統合的に扱うマルチモーダルモデルである CLIP (Contrastive Language-Image Pre-training) を用いて、ノイズ情報の機械的な分離可能性を検証した。令和 2 年 7 月豪雨のデータセットに対し、CLIP を用いて画像を言語的特徴量へ変換し、PCA (主成分分析) による次元削減を経て、k-means 法によるクラスタリングを行った。分析の結果、「ゲーム」のようなエンターテインメント関連の画像や、「植物」「動物」といった災害とは無関係なノイズ画像は、特定のクラスタに強く偏在する傾向が示され、機械的に分離可能であることが実証された。一方で、「被害」画像に関しては、「風景」画像との類似性が高く、画像特徴のみでは完全な分離が困難であり、一部のクラスタにおいて混在が見られた。しかし、ゲームや日常投稿などの明確なノイズ群をクラスタとして可視化・分離できたことは極めて重要である。これにより、ノイズ群を人が確認し、群としてまとめて除外することで、残された情報群における「被害」情報の密度と可視性を相対的に高めることが可能となる。

得られた知見を統合し、「災害情報トリアージインタフェース」を設計・実装した。本システムは、CLIP とクラスタリングによって整理された投稿群を 3 次元空間 (PCA 3D マップ) に可視化し、ユーザがクラスタ単位でノイズを能動的に除外できる機能を備えている。本システムと、従来の SNS のようにリスト形式で情報を閲覧・検索するベースラインシステムとの比較実験を行った。被験者に対し、実際の災害シナリオに基づき「被害が確認できる投稿」を収集するタスクを課した。タスク開始直後 (0~10 分) においては、提案システムの習熟に時間を要するため、ベースラインシステムと比較して収集効率が劣る傾向が見られた。しかし、タスク後半 (20~25 分, 25~30 分) においては、提案システムを利用した群が、有意に高い効率で被害情報を収集できることが確認された (20~25 分: $U = 108.500$, $p = 0.0342$, 25~30 分: $U = 108.000$, $p = 0.0369$)。これは、ユーザがシステムを利用する過程で、明確なノイズクラスタを特定し、それらを除外機能によって排除することで情報空間が縮小され、結果として残存する重要情報の可視性が高まったことが考えられる。可視化された情報の空間的配置 (類似性) を手がかりに、ノイズ情報や被害情報を連鎖的に発見できるという、提案システムの有効性が示された。

本研究では、人的・時間的リソースが制限されている大規模災害時において、SNS 上の膨大かつ玉石混交な情報から重要情報を確保するための枠組みとして、「人と計算機の協調による情報トリアージ」を提案し、その有効性を実証した。本研究の貢献は、第一に、災害時 SNS 情報の多層的マトリクス構造をモデル化し、信頼性判断における複合的な手がかりの必要性を示した点。第二に、実データに基づく定量調査により、重要情報の稀少性とノイズ情報の構造的特徴を明らかにした点。第三に、マルチモーダル手法 (CLIP) を用いたクラスタリングにより、ノイズ情報の機械的分離が可能であることを示した点。第四に、情報の「除外 (Subtraction)」を先行させるインタフェースを実装し、人間がノイズを効率的に捨てる支援を行うことで、結果的に重要情報の収集効率が向上することを実験的に証明した点である。本研究が提案したモデルは、計算機と人間を適切に役割分担・協調させるものである。このアプローチは、将来的に SNS のプラットフォームや情報の形態が変化した場合においても、ノイズ情報が混在する情報空間から価値ある情報を確保するための方法論として機能し、迅速な災害対応に貢献することが期待できる。

目次

1	序論	1
1.1	本研究の背景	1
1.2	災害時における偽・誤情報の蔓延	3
1.3	生成 AI によるフェイク情報の混在と対策	4
1.4	災害時 SNS におけるノイズ構造と収集戦略	4
1.5	本論文の構成	5
2	関連研究	7
2.1	SNS を災害時のソーシャルセンサとして活用した研究	7
2.2	災害時における情報膨大化の検出と対策に関する研究	8
2.3	災害時の情報自動抽出・検出に関する研究	9
2.4	マルチモーダル手法による災害情報抽出に関する研究	10
2.5	災害時の偽・誤情報に関する研究	12
2.6	情報トリアージに関する研究	13
2.7	本研究の位置づけ	14
3	SNS を用いた信憑性の高い災害情報収集のための指針	16
3.1	災害時の SNS の情報流通構造について	16
3.2	信頼性判断の手がかり	20
3.3	情報の峻別プロセス	24
3.4	本章のまとめ	26
4	災害時における SNS の情報構造	27
4.1	調査方針と対象災害	27
4.2	対象災害と X データ	28
4.3	ノイズとなる投稿に関する調査	29
4.4	被害投稿の特徴に関する調査	36
4.5	考察	39
4.6	本章のまとめ	39
5	計算機によるノイズ分離の可能性	41
5.1	Contrastive Language-Image Pre-training (CLIP)	41
5.2	対象の災害とデータ	42
5.3	CLIP による言語化と分類	43
5.4	考察	45
5.5	本章のまとめ	49

6	段階的除外型インタフェースの設計と実装	51
6.1	災害データセットとクラスタリング	51
6.2	災害情報トリアージインタフェースのデザイン指針	55
6.3	評価実験	62
6.4	本章のまとめ	65
7	議論	67
7.1	情報抽出アプローチの限界とノイズ構造に関する議論	67
7.2	段階的除外による被害情報の顕在化に関する議論	67
7.3	提案システムの役割に関する考察	68
8	結論	72
A	付録	75
A.1	タスクについて	75
A.2	実験ストーリー	75
A.3	事前アンケート回答結果	76
A.4	事後アンケート回答結果	80

目次

1.1 SNSにおける災害情報のノイズ構造	6
2.1 Notional model of sensemaking loop for intelligence analysis derived from Cognitive Task Analysis (CTA) ([44]より引用).	14
3.1 SNS災害情報マトリクス構造	21
3.2 峻別プロセス	25
4.1 逆L時が確認できる画像例 ³	31
4.2 テレビの報道をカメラで撮影した画像例 ⁴	31
4.3 ニュースに関する画像例 ⁵	32
5.1 CLIPにおける予測モデル ([45]より引用)	42
5.2 CLIP結果を用いたクラスタリング散布図	46
6.1 インタフェース用のデータベース構築に使用した画像のクラスタリング散布図	53
6.2 提案システムのインタフェース(全体像):(A)システムコンソールエリア,(B)投稿可視化エリア,(C)PCA 3D マッピング可視化エリア,(D)Pickした投稿の可視化エリア	56
6.3 データベース読み込み前のインタフェース	57
6.4 システムコンソールエリアと投稿の可視化エリア	58
6.5 PCA 3D マッピングエリア	58
6.6 PCA 3D マッピングエリア:ズーム後と投稿詳細の表示	59
6.7 PCA 3D マッピングエリア:特定クラスタを非表示にした状態	60
6.8 Pickした投稿の可視化エリア	60
6.9 実験で用いたベースラインシステム	63
6.10 提案システムユーザの投稿の累積ピックアップ数	66
6.11 ベースラインシステムユーザの投稿の累積ピックアップ数	66
A.1 事前アンケート	76
A.2 事後アンケート	80

表 目 次

3.1	渋谷らによる偽・誤情報の分類で用いる類型 [75]	20
4.1	人手による画像分類クラス抜粋	30
4.2	TF-IDF 上位 5 単語抜粋	34
4.3	コサイン類似度が高い傾向のあるクラス 抜粋	35
4.4	細分化した“被害”画像付き投稿における TF-IDF 上位 5 単語	37
4.5	“被害”投稿数とその割合（一部抜粋）	38
5.1	人手による画像分類クラス抜粋	43
5.2	分類クラス別 CLIP 結果 出現数上位 5 単語抜粋	44
5.3	CLIP 結果を用いたクラスタリング結果	47
5.4	クラスタ 2・8 における CLIP 算出の頻出単語	49
6.1	インタフェース用データベース構築に使用した画像・動画件数一覧	52
6.2	インタフェース用データベース構築に利用した画像のクラスタリング結果	54
6.3	ユーザが収集した災害被害画像・動画数差の検定結果	64
6.4	ユーザが収集した災害被害画像・動画の正解率	65
A.1	事前アンケート選択設問	77
A.2	SNS における災害情報の閲覧経験と関連回答（設問 3）	78
A.3	SNS における災害情報の閲覧経験と関連回答（設問 3）	79
A.4	事後アンケート選択設問（提案システム群）	81
A.5	事後アンケート選択設問（ベースラインシステム群）	82
A.6	ストレス要因（設問 1）	83
A.7	投稿の確認順序（設問 5）	85
A.8	役立った情報（設問 6）	86
A.9	重要情報（設問 8）	87
A.10	ツールの改善点や感想（設問 9）	88
A.11	期待する機能（設問 11）	89

1 序論

本研究は、人的・時間的制約の大きい災害時において、ソーシャルメディアから得られる災害被害に関連する情報を確保する際に、人と計算機による協調によって災害情報のトリアージを行う方法論の構築を目的とする。本章では、本研究の背景を踏まえ、解くべき課題を整理し、本研究が目指す情報峻別のための枠組みについて説明した後、本研究で取り組む課題について論じる。

1.1 本研究の背景

1.1.1 災害大国

日本は災害が重なって発生する可能性の高い典型的な災害多発国である。複数のプレートがぶつかる境界域に位置し、多数の活断層を抱える。世界のマグニチュード6以上の地震のおよそ2割が日本周辺で発生する地震多発国である [69]。さらに111峰の活火山を擁し [76]、地震や火山噴火に起因する津波等、マルチハザードの常在リスクを抱える。気象面でも、北西太平洋で毎年多くの台風が発生し、平年で年間約12個が日本に接近、約3個が上陸している。さらに、近年は短時間強雨の極端化が統計的に確認されており、1時間あたりの降水量が50mm以上の大雨発生回数は観測初期の10年間に比べて約1.5倍に増加している [70]。地形・土地利用の側面でも、可住地は約30%に限られ国土の約75%が山地で河川は短く急峻であり、洪水・土砂災害が発生しやすい。実際、土砂災害警戒区域は全国で約70万区域に達している [68]。国のデジタル・防災技術ワーキンググループ未来構想チームは、「大規模災害においては人命が刻一刻と失われるため、個人に対するリアルタイム性の高い現状把握と対応が必要である」とし、災害対応のためにデジタル技術の活用可能性について提言している [90]。

1.1.2 災害とインターネット

災害とインターネットの関わりは、インターネットが一般の人に利用され始めた1990年代に端を発する。1995年1月17日に発生した阪神淡路大震災の際、インターネットという情報空間が大きな役割を担った。NTT東日本は、電話通信回線は約28万5千回線が不通になり、復旧には2週間を要したと報告している [59][92]。一方で、インターネット回線は翌日には復旧したが、当時のISDNインターネット回線の普及率は全国で約53万回線しかなく [82]、個人によるインターネットの利用率はわずか1.59%だった¹。そんな中、奈良先端科学技術大学院大学は地震状況を発信する最初のWWWサーバを立ち上げ [63]、神戸外国語大学は被災地の状況をインターネットで公開した [74]。インターネットサービスを提供するニフティ株式会社²が運営していたパソコン通信サービスのNIFTY-Serveでは無料の地震情報掲示板が設置されたり [64] と、被害状況の把握や安否確認、犠牲者リストの公開、ボランティアの募集や志願など、インターネットは幅広く活用された。

その後、通信インフラのブロードバンド化や携帯電話の普及に伴い、インターネットは災害時における不可欠なライフラインとなった。2011年の東日本大震災では、電話回線が

¹World Bank Group: Open Data <https://data.worldbank.org/> (2025/11/29 確認)。

²ニフティ株式会社 <https://www.nifty.co.jp/> (2025/11/29 確認)。

輻輳した中、Twitter や Facebook, Mixi などのソーシャルメディア、Skype といった IP 電話が安否確認や情報共有に大きく貢献した。また、Google は発災からわずか 2 時間以内に「パーソンファインダー」を提供を始めた。パーソンファインダーは被災者本人がアップロードした安否情報、避難所名簿の写真を元にしてボランティア達がテキスト化した情報、マスメディアや関係機関との提携によって登録された情報といったように、各所に散財していた安否確認情報が集約され、行方不明者を検索可能になったことで混乱する被災地における情報収集のハブとなった [91]。この時、大規模匿名掲示板の 2ちゃんねる (2017 年より、5ちゃんねるに改名) の「2ちゃんねる臨時地震掲示板」がハブとなり、派生系として「【私にも】三陸沖地震災害の情報支援【できる】」掲示板が立ち上がり情報入力を行うボランティアの募集が行われたり、東北大震災 まとめ wiki³、東日本大震災 (東北地方太平洋沖地震)@ウィキ⁴ といった情報共有・集約ウェブサイトが立ち上げられたりしたことも報告されている [56]。

1.1.3 ソーシャルメディアの発展と災害利用

ソーシャルメディアは過去十数年の間に大きく普及し、様々なサービスが日常的に利用されるようになってきている [22]。中でも X (2023 年 7 月より、Twitter から改名) は (1) 投稿できるメッセージ (ポスト: 2023 年 7 月より、X への改名に伴ってツイートから改名。) が 140 文字に制限されているため長文を書かなくてもよい、(2) 不特定多数を対象とした即時的な情報発信が可能である、(3) 画像や動画、音声ファイルを添付することも容易にできる、といった特徴から手軽な情報発信手段として認知されている。総務省によると、2012 年に 15.7% だった X (旧 Twitter) の利用率は、2024 年には 45.3% にまで上昇しており [84][83]、代表的なソーシャルメディアサービスと言える。X は、発信の容易さや情報流通の広範さ、即時的な情報流通が可能といった特徴から、単なるコミュニケーション手段としてだけではなく、救助の呼びかけや物資支援依頼など災害時の情報共有・流通手段としても認知され、積極的に活用されるようになってきている。個人が短文と画像・動画等を広域に発信することが容易になったことで、有事の際の状況を伝えるための社会基盤の一つとなり、社会のリアルタイムな状況をセンシングするためのソーシャルセンサとして機能することが期待されている [71]。2011 年の東日本大震災では、発生直後から 1 時間の間に東京都のユーザから毎分 1,200 件以上の投稿があったとされており [97]、2016 年に発生した熊本地震の際には、東日本大震災に比べて 23 倍の地震関連ツイートが発信されネット上で流通していたことも報告されている [96]。

災害時に X の投稿が機能した特に有名な事例として、2011 年の東日本大震災の際にあるユーザによって発信された、「障害児童施設の園長である私の母が、その子供たち 10 数人と一緒に、避難先の宮城県気仙沼市中央公民館の 3 階にまだ取り残されています。下階や外は津波で浸水し、地上からは近寄れない模様。もし空からの救助が可能であれば、子供達だけでも助けてあげられませんか。」という投稿が当時の東京都庁副知事の目に入ったことから、実際に救助された事例 [88] が挙げられる。

SNS が災害に活用された事例は地震にとどまらない。2019 年の長野県の台風 19 号に起因する豪雨災害の際には、長野県が救助要請ツイートを独自に収集して消防や自衛隊につ

³東北大震災 まとめ wiki <https://w.atwiki.jp/acuser001/> (2025/11/29 確認)。

⁴<https://w.atwiki.jp/earthquakematome/> (2025/11/29 確認)。

なげることで約 50 件が実際に救助されたことをジチタイワークス Web が報じている [77]. ただし、膨大な投稿から救助要請ツイートを収集するにあたり、その信憑性を担保するためには、専属職員 6 人による検証を必要としていた。また、その際に収集した対象ハッシュタグは Twitter Japan が指定していた #救助ではなく、「#台風 19 号長野県被害」と独自に設定されたものであった。この事例が示すように、信憑性のある災害情報の収集には人手による検証が不可欠であり、そのための体制をあらかじめ整えておく必要があることも指摘されている [101]. X は災害時に利活用されている一方で、そのメディアの性質上、重要情報以外の災害被害とは無関係な情報（ノイズ情報）も多く投稿されていることから、情報収集をより困難にしている。

このような X を始めとしたソーシャルメディアの隆盛にともない、自治体もソーシャルメディアの防災活用を進めている。内閣官房情報通信技術総合戦略室によると、2014 年には 672 の自治体がソーシャルメディアを発災時の情報発信手段として活用していたのに対し、2018 年には 1090 の自治体が活用するまでになっている [79]. こうした情報発信への活用が進む一方で、情報収集手段としてのソーシャルメディアの活用は期待されていたほど進んでいない [79][66]. X はその特性ゆえに災害時の「現在」の情報を収集するのに適している。特に、被害情報に関するツイートや被災状況の画像・動画は災害状況の確認や支援策の策定に資する有用な情報として期待される。それにもかかわらず、情報収集手段としての利用が進んでいない背景には、(1) 何らかの事件・イベントが起きた際には情報バースト現象が起きることで情報が肥大化すること [36], (2) 多種多様で膨大な情報が重複して流通しているため自治体が必要とする情報のみを整理・選別して取り出すことが難しいこと [38], (3) 投稿されている情報の中には十分に検証されておらず信憑性の乏しい情報が多く含まれていること [58][79], といった SNS の情報構造の複雑さが挙げられる。

1.2 災害時における偽・誤情報の蔓延

インターネットが社会に浸透し、誰もが発信者になり得る状況は、同時に「情報の爆発的な増加」と「信頼性が担保されない情報の拡散」をもたらしたとも言える。特にソーシャルメディアには、たびたび偽・誤情報が投稿される。Muhammed らの定義 [41] によると、誤情報 (Misinformation) とは投稿時点では不確実で未検証の情報であり、情報源の曖昧さや確証バイアス、社会的繋がりなどが発生要因となる。同様に、偽情報 (Disinformation) とは誤情報と同じく虚偽または誤解を招く情報であるが、誤情報と異なり意図的に作成・拡散されるという特徴を持つ。災害時における偽・誤情報の対策に関しては 3 章で詳しく述べるが、災害時のソーシャルメディアで流通した誤情報の例として、東日本大震災の直後には、コスモ石油の千葉製油所が爆発した事に関連して「コスモ石油の爆発により有害物質を含んだ雨が降る」といった根拠のない誤情報がソーシャルメディアやチェーンメールを通じて一瞬にして拡散された事例が挙げられる [62]. また、2016 年熊本地震の際に「動物園のライオンが逃げた」という偽情報が、実際にライオンが道路上に佇んでいる画像付きで X 上で拡散された。これらの事例は、人々の善意に基づく情報の拡散行為が、インターネットを介することで、不確実な情報の流通を加速させ、社会的な混乱を招く結果となった。こういった偽情報は機械的に除去できるものではなく、発信元の把握や現場への問い

合わせ、自治体や有識者による正確な情報の発信など、さまざまな情報を確認してその真偽を確かめる必要がある。つまり、現状では人手による検証が不可欠である。2024年に起きた能登半島沖地震では、救助要請ツイートのコピーや偽情報が大量に発生したことで、オリジナルの投稿を見極めることが一層難しくなった。偽の救助要請に基づいて消防が出動したケースが少なくとも2件あり、被災者の対応に支障が生じた恐れがある[72]。他にも、能登半島沖地震は人工的に起こされたといった人工地震説や、外国人の窃盗団が集まっている等の偽・誤情報が広まる事態となったことが報道されている[57]。

1.3 生成AIによるフェイク情報の混在と対策

近年では、これまで見られた偽・誤情報に加え、比較的誰でも簡単に利用できる大規模言語モデル (Large Language Model; LLM) や Latent Diffusion Model[47] といった画像生成AIモデルの登場により、偽情報や合成画像、偽画像なども確認され始めている[2]。生成AIによるフェイクに関しては、人は最先端の生成文書・合成・生成画像を実写とほぼ区別できないことが複数研究で示されていることから[2][31][37][43][13]、「一見、見分けがつかない」ケースを前提に対策するのが現実的である。対策にあたり、人の目で違和感を拾うための着眼点を述べる。第一に、文字・数字・ロゴなど「規則性の高い人工物」が破綻しやすい[28] (看板の文字がにじむ/一部だけ別書体、車両番号や標識が不自然に歪む)。第二に、反復パターンの均質さ[11] (同じ破片や草むらが微妙に繰り返される、瓦礫が“整いすぎる”等)。第三に、光と影の破綻[49] (夜間でも影が強すぎる/曇天なのに彩度が高すぎる、水面やガラスの反射が被写体と一致しない)。第四に、微細構造の乱れ[49][25][16] (電線・欄干・フェンスなど細い直線が波打つ、人体の指・耳など細部の形が崩れる)。第五に、過剰に劇的な構図や過度に清潔な質感[25] (泥や水滴の汚れ方が均一、衣服の濡れ方が実感に乏しい)。ただし、これらは現時点での「兆候」であって決め手ではない。生成手法や技術の進歩により、多くの痕跡は短期間で目立たなくなる可能性も高い。小さな違和感でも、投稿本文の具体性や他の報告との噛み合わせが悪い、あるいは画像と本文の結びつきが緩い (誰がいつどこで撮ったのかが曖昧) ときは、判断を急がず「保留」扱いとするのが安全である。逆に、違和感があっても他の強い手がかり (物の一致、別角度の写真、現地の時間・天候との整合) が複数確認できるなら、慎重に見直す余地はある。

1.4 災害時 SNS におけるノイズ構造と収集戦略

X 上では被害情報や救助情報といった被災者支援につながる貴重な情報が発信される一方で、災害とは関係のない一般的な投稿や、被災者に共感するといった投稿も多いと指摘されている[14][55]。このような問題を解消するために、現在、ハッシュタグを補完して情報収集しやすくしたり、自然言語処理技術を活用して検出・自動収集・分類したりすることが試みられている (e.g., [6][50])。しかし、玉石混交の投稿から確度の高い災害関連情報を取りこぼすことなく効率的に収集できるまでには至っていない。加えて、システムによる確度の高い情報の自動抽出には限界がある。例えば、TVで災害ニュースを放映しているテレビ画面を撮影した画像や、ウェブメディアによる災害ニュースの記事などが投稿され、そういった情報が混在することで、災害被害情報の効率的な自動収集および分析の妨げと

なっている [39]. SNS には, こういったメディアによって既に発表された情報が, 本来の発信者とは別のユーザが繰り返し再投稿され続ける特性があることも, 情報抽出の困難さに繋がっている.

災害時における情報収集というタスク自体が, 人命救助の可否に直結する可能性があり「情報の取りこぼし」が許されない. そのため, ソーシャルメディアでの情報収集においても, 重要な情報を「取りこぼさない」ことが大前提となる. その検証に相当な人的コストが求められることになる. そうした状況下において重要となるのが「情報トリアージ」の概念である [34][94][99]. 主に医療現場では, 限られた医療資源を傷病者の緊急度や重症度に応じて有限リソースを分配するプロセスである「トリアージ [86]」を行うが, 情報トリアージはその概念を情報技術に適用するアプローチである. 災害時の情報空間において, 膨大な情報の中から緊急性が高く, 対応が必要な情報を迅速に選別することで優先順位をつける処理が不可欠となる. インターネットの普及により, 誰もが情報を発信できる現代社会は, 情報を選び取る技術が問われる社会へと変容したと言える.

上述してきたように, SNS における災害情報空間はその大多数がノイズ情報で構成されている. 一方で, 玉石混交の情報空間においても得られる災害情報があり, 同時に今助けを求めている人の情報を切り捨てることは人道的な観点からもあってはならない. しかし, 重要情報を確保するためには, 信頼性を担保するためにも人目による確認が必要条件となる. こうした SNS における災害情報の構造を図 1.1 に示す. このような状況を鑑みると, 関連情報を絞り込み計算機の情報抽出精度の向上を追求する従来の情報抽出手法では, 抽出する情報の質を上げようとすると漏れが増加する. 特に災害時において過度な精度の追求は, 真の被害情報に対し FalseNegative 判定してしまう第二種の過誤による情報の見落としにつながりかねず, 実際の運用には大きな不安が残る. 一方で, 漏れを少なくすると, 抽出される情報の質が低下しノイズが多く含まれることで, その後の人手検証の負担が増加する. つまり, 一定の信憑性が担保される「実際に被災者が救助を求めている可能性が高い投稿」を漏れなく効率的に収集するために, 従来手法では, 災害時に重要な救助要請の抽出に漏れが発生することから, 人と計算機の各々の長所を活かした協調が欠かせない. システムが確度の高い情報を抽出するとともに無関係な情報のみを除外し, その上で人がグレーゾーンの情報や人手で判断すべき情報を処理するというシステム一人の連携が, 人的負担を軽減させつつ効率を上げる選択肢となるだろう. 本稿では, 計算機が情報の構造化とノイズ整理を担い, 人が情報の精査を行い最終的な信頼性判断に注力する「人と計算機の協調による情報トリアージ」を実現するという方針で, その解決に臨む.

1.5 本論文の構成

本論文の構成は以下の通りである. 1 章では, この研究の背景と課題について述べ, 研究の方針を明確にする. 災害大国日本における, ソーシャルメディアに投稿されているリアルタイム情報から, 被害状況や救助要請を把握するための課題について整理する. 2 章では, これまでの関連研究を概観し, 本研究の位置付けを明確化する. 3 章では, SNS を用いた信憑性の高い災害情報収集のための指針を示す. 災害時の SNS の情報流通構造について整理し, SNS の情報空間の特異性と偽・誤情報の対策について検討し, 災害時に必要な情報

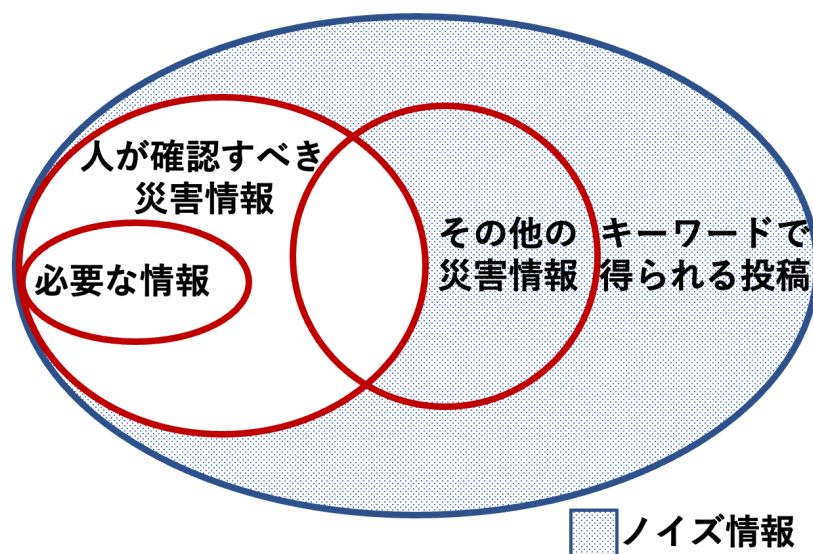


図 1.1: SNS における災害情報のノイズ構造

群の所在を明らかにすることで、災害時の重要情報の確保のために必要となる課題解決モデルを提案する。4章では、災害時のデータ収集と整備、ノイズ情報の定量化を行うことで問題提起を行う。5章では、3章で得られた知見から、膨大な SNS の投稿からノイズ情報群と被害情報群との分離可能性について検討する。6章では、4章および5章で得られた知見から、課題解決のためのシステムを開発し、その有効性について検討する。7章では、ソーシャルメディアを利用した災害時における重要情報の確保のための課題を踏まえ、本研究で提案する、人と計算機の協調による情報トリージのためのノイズ情報の抽出とその峻別システムの貢献について議論する。

2 関連研究

本研究は、災害時という時間的・人的リソースの欠如した状況下において、ソーシャルメディアを利用して迅速かつ確実に災害情報を確保するための枠組みを構築することを目的としている。本研究の枠組みでは、限られたリソースを有効活用するため、膨大でノイズ情報の多いソーシャルメディアの投稿に対して、情報の重要度や緊急度に応じた優先順位づけ（トリアージ）を行う。計算機による支援をもとに優先度の低い情報（ノイズ情報）を効率的に除外するアプローチを採用し、最終的に人手による確認が必要となる情報の純度を高める事を目指す。

本章では、この枠組みの構築において、基盤となる既存の知見について概観する。まず、ソーシャルメディアを災害時のセンサとして捉える初期の研究とその有効性について述べ、次に、災害時に特有の現象である情報の爆発的増加（バースト）に関する研究について整理する。続いて、膨大な情報の中から有益な情報を抽出・分類するための自然言語処理を中心としたテキスト解析技術、および近年精度向上が著しい画像とテキストを組み合わせたマルチモーダル解析手法について解説する。最後に、ソーシャルメディア活用の最大の課題の一つである偽・誤情報の拡散とその対策に関する研究を取り上げ、これらを踏まえた本研究の位置付けと解決すべき課題について述べる。

2.1 SNS を災害時のソーシャルセンサとして活用した研究

大規模災害時において、被災地の状況をリアルタイムに把握することは極めて重要である。X（旧 Twitter）に代表される SNS は、ユーザが「今起きていること」を即座に発信できる特性を持つことから、これを一種の社会観測センサ（ソーシャルセンサ）と見立てて災害検知や状況把握に活用する研究が活発に行われてきた。Sakaki らは X（旧 Twitter）をソーシャルセンサとみなすことで、地震や台風といった実世界の情報をリアルタイムで検知するシステムを構築した [48]。まず、投稿を災害に関連する投稿か、そうでないかを判別するために、SVM を用いて分類器の構築を行った。システムは、Earthquake と Shaking をクエリとして 3 秒ごとに収集し、イベント検知のための時系列解析や、位置情報、ユーザのプロフィールの位置情報を用いた空間解析を行う。具体的には、SVVM による分類と確率モデルによるイベント検出を行った後、カルマンフィルタリング [5] や粒子フィルタリング [8] といった位置推定手法を用いて地震発生地点を推定した。その結果、2009 年 8 月から 9 月に発生した震度 3 以上の地震の 96% を検知し、最短 20 秒でのメール通知を可能にした。これにより、ソーシャルメディアをリアルタイムイベントを観測するためのソーシャルセンサとして活用することの可能性を示した。國友らは X（旧 Twitter）の投稿から、土砂災害の前兆・発生状況把握を試みている [67]。広島県のように人口規模がある程度大きい地域に関して土砂災害の前兆現象の報告が見られる一方で、人口規模の小さい地域からの投稿は少なく、情報過疎化の問題を指摘している。谷口は、2011 年 9 月に近畿地方を通過した台風 12 号に起因する豪雨災害の際に、X（旧 Twitter）がどのように利用されたか整理している [87]。奈良県十津川村ではライフラインの断絶などにより孤立し、マスメディア等の報道も限られていたことから、被災地域に関する情報の流通が限られていた状況下で、十津川村の災害情報収集アカウントが作成され、住民やその関係者からの情報が集約された。

公式発表やマスメディアの報道が限られている災害初期段階において SNS が重要な役割を担ったとしているが、マスメディアがその社名や記者名を記述することで信頼性を担保するのに対し、ソーシャルメディアはその信頼性の担保が困難であるとも指摘している。

これらの研究により、SNS は即時性の高い情報源として災害検知や初期情報の収集に有効であることが示された。しかし、情報の密度が人口規模に依存する地域差の問題や、情報の信頼性担保という課題も浮き彫りとなっている。本研究では、こうしたソーシャルセンサとしての特性を活かしつつも、災害時の情報の信頼性を補完するための仕組みについて検討する。

2.2 災害時における情報膨大化の検出と対策に関する研究

災害時において SNS を有効活用するためには、平常時とは異なる情報の発生パターンを理解し、どのように扱うべきか検討する必要がある。特に、災害のような社会的に影響の大きい出来事の際には、情報量が爆発的に増加する「バースト」現象が発生し、重要な情報がノイズに埋没するリスクが高まる。本節では、こうした情報空間の膨大化に関する知見を整理する。

Kleinberg は、電子メールやニュース記事といった時間経過とともに届くような文書には、何らかの話題が立ち上がった時には急激に活動頻度が上昇するバースト現象が発生するとし、それらを形式的にモデル化する方法を提案した [26]。水沼らは、2012 年 3 月から 7 月に収集した約 13.8 億件の日本語 Tweet を用いて、投稿数が平常時の平均より標準偏差の三倍の閾値を超えた時間帯を「バースト」と定義し、その生起要因と類型化を行った。バースト時と非バースト時を比較すると、バースト時には投稿文が短くなり、相手の投稿への返信比率が低下し、リツイート（旧リポスト）と呼ばれる拡散行為の比率が上昇することから、ユーザが即時性を優先し、特定相手との対話よりも不特定多数への情報拡散に行動様式を切り替えることが示された。さらに、日本気象協会の震度データと結びつけることで、震度が大きいことに加えて「東京都心からの距離の近さ」がバースト発生確率を有意に高めることを示し、地震発生位置と SNS 上の情報バーストとの関係を定量的に明らかにしている。

鳥海は、2020 年 5 月に発生した検察庁法改正案に伴い発生した、ツイッターデモと呼ばれる社会現象に着目し、投稿のバースト現象について詳細な調査を行っている [89]。一部のユーザが何度も投稿・拡散することで、情報の水増しが確認できたものの、普段から政治に関心のあるユーザコミュニティのみならず、映画・芸能・スポーツなど、多様なコミュニティの人間が投稿していたことも明らかとなっている。

こういった情報のバースト現象は、災害時の SNS においても顕著に見られる。篠田らは、東日本大震災前後 3 週間の約 4 億件の日本語の投稿を対象に、利用状況とコミュニケーション構造を地域別に分析している [73]。震災直後には Tweet 数・ユーザ数がスパイク状に増加し、その後の推移をみると、被害の大きい岩手・宮城では一時的に投稿が減少した。その一方で、東京・大阪では新規ユーザが急増するなど、地域ごとに異なるバーストパターンが現れることを示した。また、NHK の公式アカウントや災害情報を発信している特定のアカウントが情報集約のハブになっていることや、時間経過に応じて対話的な発信から情

報伝播的なコミュニケーションに変化していく過程を明らかにしている。

これらの研究から、災害時の情報空間ではユーザの行動変容により情報の拡散速度と量が劇的に変化することが明らかである。これは、人的リソースが限られた災害対応の現場において、人手による全ての情報確認が非現実的であることを意味する。したがって、本研究が目指す枠組みにおいては、バーストした情報の中から、人目で処理すべき情報のみをトリアージする枠組みが不可欠となる。

2.3 災害時の情報自動抽出・検出に関する研究

情報のバースト下において、有益な情報のみを選別するためには、自然言語処理 (NLP) や機械学習を用いた自動抽出技術が鍵となる。ここでは、テキスト情報に着目し、災害情報の分類や抽出、さらには学習データの構築に関する主要な研究について概観する。

総務省所管の国立研究開発法人情報通信研究機構 (NICT) は、X (旧 Twitter) 上の膨大な投稿をリアルタイムで分析し、災害時の状況把握や意思決定を支援するシステム DISAANA を開発した [35][61]。「熊本で不足しているものは？」といった 5W1H 形式の問いかけ文に対し、X (旧 Twitter) 上の情報を分析して回答し、同時に矛盾する情報を並列して表示することで、デマ情報の可能性をユーザーに気づかせる機能も備えている。2016 年熊本地震時には、発災直後から活用され、被災地での水や食料の不足状況について可視化された。

Ashktorab らは、災害時のソーシャルメディアにおいて、膨大なノイズ情報の中から、救助活動を行う人が活用可能な情報を抽出するためのマイニングツール Tweedr の開発を行った [4]。Tweedr は分類、クラスタリング、抽出の三段階のフェーズに分かれている。分類にはロジスティック回帰、Support Vector Machine (SVM) を用いてそれぞれ分類器を作成し、投稿が被害や死傷者の報告かどうかの判別を行った。その結果、ロジスティック回帰による分類が優位な結果を示した。クラスタリングでは類似したツイートを統合して重複情報を削除し、抽出では条件付き確率場 (conditional random field; CRF) を用いて、被害を受けたインフラの特定や状況、死傷者数などを示すトークンの特定を行っている。実際の災害時のデータを用いた評価実験では、比較的小規模なラベル付きデータから有益な情報の抽出可能性を示したが、未知データに対する汎用性には課題が残るとした。

Stowe らは、2012 年のハリケーンサンディを対象に、災害対応に有用な情報を抽出・分類するシステムを構築した [52]。ソーシャルメディアのノイズの多さからフィルタリングが不可欠であるとし、Relevance な投稿 (災害に関連している投稿) を、Sentiment (感情)、Action (行動)、Preparation (防災準備)、Reporting (現地報告)、Information (他から得た情報の共有)、Movement (避難)、Other (その他) の 7 分類でラベリングを行った。データセットの作成は専門家 7 名により行われ、分類モデルの作成は SVM を用いた。実験の結果、Relevance な投稿 (災害に関連している投稿) に関しては高い精度が得られたが、詳細な 7 分類に関しては十分な精度が得られなかったとしている。一方で Reporting (現地報告) に関してはパターンが明確な投稿が多く、良好な結果が得られている。

Nguyen らは、X (旧 Twitter) 上の投稿にはプラットフォーム特有の文法や不規則な語彙が存在するとし、それらに対応するための大規模事前学習済み言語モデル BERTweet の提案を行った [42]。従来の BERT[7] の学習データには wikipedia や書籍などの一般的なテキ

スト情報を用いていることから、X（旧 Twitter）のようなプラットフォーム特有の特徴的な文章への適用が難しい。BERTweet は、大規模な英語投稿を用いて追加学習することで、その課題解決を試みている。学習には、より頑健な学習手順を持つ A Robustly Optimized BERT Pretraining Approach (RoBERTa)[29] の手法を用い、事前学習データには 2012 年から 2019 年までの投稿と、Covid-19 のような喫緊のトピックまで含まれたものを用いている。また、投稿内で特定のユーザへの返信を示す @username や URL は特別なトークンを用意することで正規化している。BERTweet は、品詞タグの付与、固有表現の抽出、テキスト分類において、RoBERTa や XLM-R を上回る結果となった。特に、皮肉の検出や感情分析などの分類タスクにおいて有効性が確認され、災害情報を含む X（旧 Twitter）において、強力な基盤モデルであることを示している。

Tekumalla らは、災害時のソーシャルメディアの投稿抽出のために、機械学習用のアノテーション付きデータセットを作成することは、人的・時間的コストがかかることから、災害規模やリアルタイム性を考慮した場合には非現実的であると指摘している [53]。したがって、小規模データセットへの手動アノテーション依存を回避し、大規模データセットへの自動アノテーションへ移行する必要がある。そこで、不正確でノイズの多い SNS データのラベル付けに対して、弱教師あり学習のアプローチとしてヒューリスティックを採用することで、大規模な Silver Standard データ (TweetDIS) の作成を行った。ラベルが自動的に付与されたデータが大量である場合、ラベルの不正確さを補うことができるという仮説を立て、Silver Standard データで訓練・微調整を行ったモデルを用いて、既存の Gold standard データに対して適用することで評価を行った。その結果、Gold Standard データセットに対して、90%以上の精度を達成したことにより、弱教師あり学習で作成された大規模データセットが、災害関連ツイートの識別に一定程度有効であることを示した。

これらの研究は、テキスト解析による情報抽出の精度向上に大きく貢献している。特に、Tekumalla らの研究は、アノテーションコストという「リソースの制約」に対する重要な示唆を与えている。しかし、テキスト情報だけでは被害の規模感や具体的な状況の把握に限界がある場合も多い。本研究では、テキスト解析の技術的基盤を踏まえつつ、次節では画像情報と統合するマルチモーダル手法についてさらに追求する。

2.4 マルチモーダル手法による災害情報抽出に関する研究

これまで述べてきたように、災害時の SNS では、ツイート数やハッシュタグ出現頻度の急増といった情報のバースト現象が顕著に現れ、その時系列パターンやテキスト内容进行分析することで、発災の検知や被災地の把握、人々の関心の集中過程をある程度捉えることができる。しかし実際の危機状況では、被害の具体的な様相や緊急度は、短いテキストだけでなく、同時に投稿される写真・動画といった視覚情報の中にも多く含まれている。テキストの曖昧さやノイズ、逆に画像だけでは文脈情報が抜け落ちてしまうという問題を踏まえると、いずれか一方のモダリティに依存した分析には限界が存在すると考えられる。そこで近年、災害時 SNS 投稿に含まれるテキストと画像を統合的に扱うマルチモーダルな分析手法が注目されている。これらの手法は、テキストと画像の双方から特徴量を抽出し、それらを相互参照させることで、重要な投稿の自動抽出や、被害種別・深刻度の推定などを

高精度に行うことを目指している。以下では、こうしたマルチモーダル系の代表的な研究を取り上げ、モデル構成や学習方法、評価結果を概観することで、災害時 SNS 解析における新たな可能性と課題を整理する。

Alamらは、こうしたマルチモーダルな分析を行うための SNS データセット CrisisMMD を公開している [3]。CrisisMMD にはアメリカやメキシコで発生したハリケーンイルマ、ハリケーンハービー、ハリケーンマリア、カリフォルニア山火事、メキシコ地震、イラン・イラク地震、スリランカ洪水の7つの災害の X (旧 Twitter) の投稿テキスト・画像データの両方が含まれる。これにより、災害状況を把握するための研究の推進が期待されている。Abavisaniらは、災害情報の分類においてテキストと画像の双方を活用し、相互のモダリティのベクトルを参照して特徴選択を行うクロスアテンションモジュールを組み込むことで、マルチモーダルの枠組みを提案している。片方のモダリティに無関係な情報や誤解を招く情報を含んでいる場合があり、それらを効果的に除去しつつ情報を統合することを試みている [1]。さらに、Multimodal Stochastic Shared Embeddings (SSE) を用いて不完全なラベルを持つデータを正則化することで、限られた訓練データでの学習を可能にしている。CrisisMMD を用いた実験の結果、既存のベースラインを上回り、画像とテキストのラベルが一致しないデータや未知の災害データに用いることが期待される。Bijalwanらは、RoBERZa を用いたテキスト特徴量と Vision Transformer (ViT)[9] を用いた画像特徴量を統合し、Middle Fusion モデルを提案している。CrisisMMD を用いた評価実験では、従来研究と比較して約 2% から 5% の精度向上が確認された。これにより、一方のモダリティが不完全または曖昧な場合でも、他方のモダリティが情報を補完することで文脈を考慮した分類の可能性を示唆している。Jainらは、複数の災害に関するツイートデータセット CrisisMMD を利用し、OpenAI embedding モデルでそれらを分類することで従来の手法と比較して精度が向上することを示している [21]。Rezkらも CrisisMMD データセットを利用し、画像とテキストの両方に重要度を付与する Multimodal Channel Attention model を構築して、複数のモダリティを統合した分析手法を提案し、単一モダリティの分類と比較して精度が向上することを示している [46]。

Mandalらは、災害関連ツイートのマルチモーダル分類において、大規模な画像-テキストペアで事前学習された Contrastive Learning モデルの有効性を検証した。これまでの手法では、画像とテキストそれぞれに独立したエンコーダを用い、その後特徴を結合するアプローチが主流であった [32]。しかし、ツイート内のテキストと画像の関係は、一般的なキャプション付き画像よりも複雑で、必ずしもテキストが画像の内容を説明しているとは限らない。Mandalらは、CLIP[45] や ALIGN[23] といった Contrastive Learning 学習 (対照学習) モデルが、テキストと画像の関係が薄くノイズの多い災害データに対しても、頑健な学習ができるかどうか着目している。CrisisMMD および Mouzannarらが公開している Damage Identification Multimodal (DMD)[40] データセットを用いて検証を行い、既存のベースラインモデルを上回り、かつデータが極端に少ない Few-shot 設定においても有効性を示している。これにより、災害発生直後の初期段階において、ラベル付けデータが少ない状況においても実用的であることを示している。

Liらは、これまでの上述した研究と同様の課題に対し、CLIP を拡張する CLIP-BCA-Gated フレームワークを提案している [27]。CLIP-BCA-Gated フレームワークは事前学習済みの

CLIPモデルをベースとし、Bidirectional Cross-Attention (BCA)[30]と入力データに不鮮明な画像やスペルミスを含んでいるような場合に各モダリティの重みを動的に調整する Adaptive Gating を統合したものである。CrisisMMD を用いた評価実験の結果、従来の CLIP モデルよりも 1.55%高い精度を示した。このモデルの特徴として、推論速度が高速であるとし、リアルタイムの災害対応の有効性を示唆している。

以上の研究は、マルチモーダル手法が災害情報の抽出において極めて有効であることを示している。特に、MandalらやLiらの研究は、ノイズの多いデータや少量の学習データ (Few-shot) に対しても頑健なモデルを提案しており、これは「リソースの欠如した状況」を前提とする本研究の目的に合致する。本研究では、これらの先端的なマルチモーダル手法を取り入れつつ、実際の災害対応現場での運用を想定した情報確保の枠組み構築を目指す。

2.5 災害時の偽・誤情報に関する研究

SNSは迅速な情報共有を可能にする反面、情報の信頼性が担保されていないため、偽情報や誤情報が拡散しやすいという側面を持つ。特に災害時の混乱下においては、誤情報が救助活動の妨げや被災者の生命危機に直結する可能性があるため、その対策は不可欠である。本節では、災害時における誤情報の拡散実態とその対策に関する研究について述べる。

Huntらは、災害時のX(旧Twitter)において、誤情報の拡散と、その誤情報の訂正・検証について調査している[20]。2017年に発生したハリケーンハービー、ハリケーンイルマの際に流通した「避難所で移民のステータスチェックが行われる」といった誤情報について、誰がどのように誤情報の収束に貢献したかを分析した。その結果、認証済みアカウントによる投稿は約13から15%だったが、「拡散やいいね」といったエンゲージメントの大部分はこれらの投稿が獲得していた。認証済みアカウントに関しては現在の審査基準と異なり、当時の審査基準は、「著名で信頼に値するアクティブなアカウント」であった¹。特に、DEMAや自治体などの政府関連の認証済みアカウントの拡散力は、ニュースメディアの拡散力を上回り、災害時の誤情報対策において公的機関による迅速な情報発信が、誤情報の訂正に最も重要であることを示唆している。また、HuntらはX(旧Twitter)上の誤情報を手動で監視・訂正することの非効率性を指摘し、投稿の真偽(Veracity)をTrue, False, Neutralの3クラスに自動分類するための機械学習モデルの提案を行った[19]。実験は、ボストンマラソン爆弾テロ、マンチェスターアリーナ爆破事件、ハリケーンハービー、ハリケーンイルマ、ハワイ弾道ミサイル誤警報のX(旧Twitter)投稿データセットを用いてテキスト特徴量とユーザ情報を組み合わせて複数のモデルで検証を行った。その結果、少量のトレーニングデータでも有効的に学習が行えるSupport Vector Machine (SVM)が最も高い精度を発揮した。未知のデータに対して誤情報事例の検出は困難であったが、少量の新しい事例データを追加学習することで精度が向上したことも同時に報告している。これにより、過去の事例モデルと少量のリアルタイムデータを組み合わせることで、誤情報の拡散を検出できる可能性を示した。

川村らは、2018年に発生した北海道胆振東部地震を対象に、SNS上の誤情報の拡散と自治体による対応の課題を調査している[65]。「断水」のような直接的なキーワードではなく、

¹X ヘルプセンター：Xで青いチェックマークを獲得する方法 <https://help.x.com/ja/managing-your-account/about-x-verified-accounts> (2025/11/17 確認)

「らしい」や「聞いた」とった伝聞を表す単語を含む投稿を収集し、それらと「デマ」の共起関係を分析することで、実際に拡散していた誤情報の内容を特定し、その自動抽出の可能性を示した。また、自治体へのヒアリングから、防災部局の人手不足や時間不足による SNS 空間での出来事の把握、情報収集・発信の困難さも課題として挙げられた。実際、誤情報の検知は、地域住民による電話での問い合わせによって認知されていた。福長らは、大阪北部地震の事例等から、マスメディアによる偽・誤情報の打ち消し報道が持つ重要な役割について論じている [93]。藤代は、偽・誤情報とフェイクニュースへの対策には、ネット空間に流通する情報をニュース・コンテンツ・広告のように適切に切り分けることが重要になると述べており、それらへの対策方針を示すことで、伝統メディアの役割の大きさを指摘している [95]。

これらの研究により、誤情報の自動検出の可能性や、公的機関による訂正発信の重要性が示されている。しかし、川村らが指摘するように、実際の自治体現場では人手不足により SNS 監視そのものが困難であるという現実がある。また、Hunt らの研究が示すように、未知の災害における誤情報の自動検出には依然として課題が残る。本研究では、リソースが不足している状況下においても、人的負担を軽減しつつも情報の信頼性を効率的に評価・選別できる支援の枠組みを提案することで、この課題解決に貢献する。

2.6 情報トリアージに関する研究

1 章でも触れたように、主に医療現場で用いられる「トリアージ [86]」と呼ばれる概念がある。日本救急医学会の定義によると「災害時発生現場等において多数の傷病者が同時に発生した場合、傷病者の緊急度や重症度に応じて適切な処置や搬送をおこなうために傷病者の治療優先順位を決定することをいう」と示されている²。この、時間的・人的・資源的リソースが限られている状況下において優先順位を付与するという考え方を情報処理に適応した概念は「情報トリアージ [34][94][78]」と呼ばれる。Marshall らはこの情報トリアージの定義について「Information triage is the process of sorting through (the possibly numerous) relevant materials, and organizing them to meet the needs of task at hand.」と示している。この情報のトリアージを行うことは、単純に機械的にフィルタリングを行うプロセスではなく、人による認知プロセスの一部とも位置付けられている。Pirolli らは、熟練の分析官が行う認知タスク分析 (Cognitive Task Analysis; CTA) を通じて、情報分析のプロセスについて構造化を行っている [44]。Pirolli らの提示した CTA の Foraging (情報の収集) loop プロセスと Sensemaking (意味付け) loop プロセスが互いにループしあう構造を、図 2.1 に示す。この構造によれば、CTA では「Shoobox (一時保管)」に関連データを集め、「Evidencefile (証拠)」へ根拠を抽出し、そこから「Schema (計画)」を形成することで、各タスクにおける仮説検証を行う工程が繰り返し行われる。Pirolli らは、膨大な情報空間から関連情報を効率的に探索するために、Foraging loop を回し、情報を排除 (Filter) し、読み込み (Read)・抽出 (Extract) を行っていく工程の中で重要となるのは、熟練の CTA はある程度ノイズを許容する (Filter を微弱にする) ことで、より微弱な兆候も見逃さないことであると述べている。これは、熟練の分析官は事前知識や暗黙知を持ち、意味を持たないノイズ情報を迅

²医学用語解説集 トリアージ <https://www.jaam.jp/dictionary/dictionary/word/1022.html> (2025/11/29 確認)。

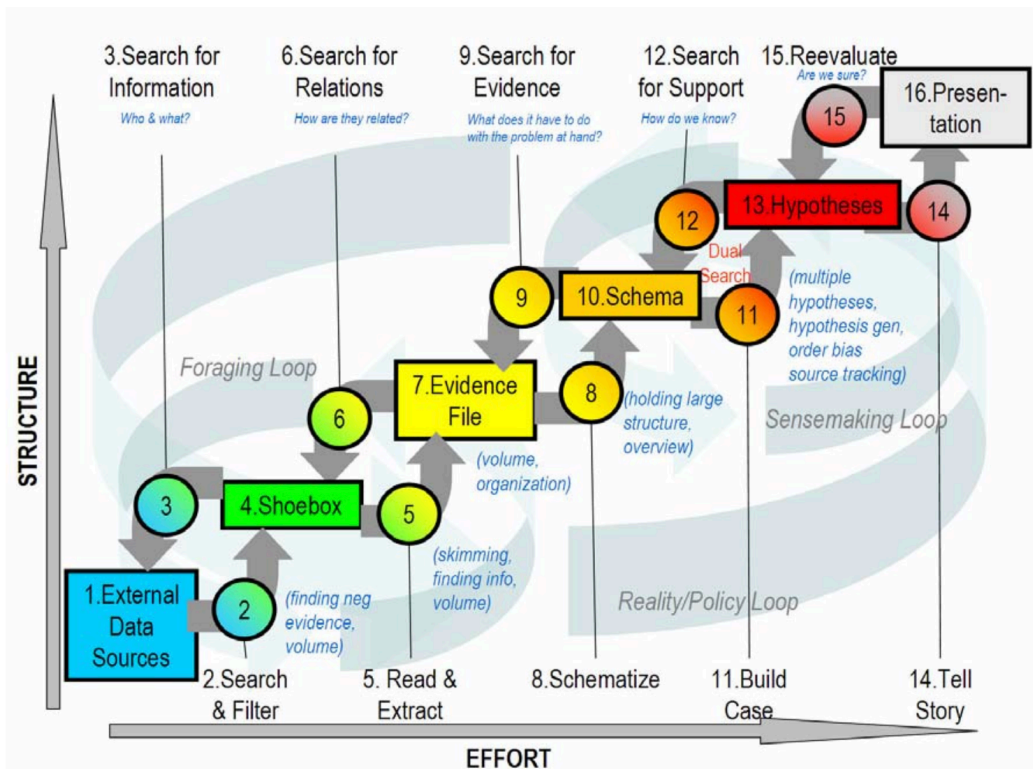


図 2.1: Notional model of sensemaking loop for intelligence analysis derived from Cognitive Task Analysis (CTA) ([44] より引用).

速に排除できるためである。Hercegらは、このPirolliらの提案する情報分析の初期段階に特に有効的な情報トリアージ戦略について提案している[15]。この戦略では、複数人が同期的・探索的に情報検索を行い、その結果からニーズに応じた重要情報の集約を行っていくことで、大規模な情報群の中から時間的制約の中で効率的に情報選別・処理を可能にしている。

2.7 本研究の位置づけ

本章では、ソーシャルメディアが災害時の被害状況等の把握のための情報源として有用である一方で、情報の爆発的増加（バースト現象）や、偽・誤情報の混入といった課題を抱えていることを確認した。また、自然言語処理（NLP）やマルチモーダル手法を用いた自動抽出技術、および誤情報検知に関するこれまでの研究についても概観した。これらの既存研究は、主に「膨大な情報の中から、有益な情報をピンポイントで発見する」というアプローチをとってきている。しかし、前述したTekumallaらやHuntらの研究が示唆するように、学習データの不足や未知の災害パターンに対して、機械学習モデルが完全な精度を保証することは困難である。災害対応の文脈において、救助要請などの重要情報が機械判定の閾値として切り捨てられてしまうと、重要情報の「取りこぼし」につながり、人命に関わる致命的なリスクとなり得る。また、川村らの調査で明らかになったように、実際の災害対応現場では人的リソースが著しく不足していることもあり、完全に人手のみによる

ノイズの多い膨大な情報の真偽を一つ一つ検証することも現実的ではない。すなわち、既存の自動抽出技術を用いたアプローチだけでは、取りこぼし防止と信頼性の担保を、現場の限られたリソース内で両立させることは困難である。

そこで本研究は、主に医療現場で用いられる「トリアージ」の概念を情報処理に適用する「情報トリアージ」のアプローチをもとに、情報の緊急度や重要度に応じて優先順位を決定するための支援を行う。膨大でノイズの多いSNSの情報全てに対して、人間が詳細な優先順位をつけるのは現実的ではない。したがって、機械学習を「正解を見つける」ためではなく、人間が効率的にノイズを捨てるための支援を軸に用いることで、人と機械のそれぞれの得意領域で協調することを目指す。次章からは、災害時の情報構造とその整理、そして最終的な信頼性判断に至るまでの具体的な枠組みと、人と計算機の役割分担モデルについて論じる。

3 SNSを用いた信憑性の高い災害情報収集のための指針

本章では、1章で提示した課題意識と2章で整理した関連研究を踏まえ、災害時 SNS 情報の収集・峻別における人と計算機の協調的枠組みについて検討する。本研究が目指すのは、膨大なノイズを含む情報空間から重要な被害・救助要請情報を「取りこぼしなく、かつ効率的に」確保するための作業構造と人と計算機の役割分担の設計である。既存研究では自動抽出や検出 (extractive) 型のアプローチが主流であったが、現在の情報技術で精度 100% での情報抽出は困難である。特に災害対応における命題として、情報の取りこぼしを防ぐ必要があることから、情報除外 (subtractive) 型の情報処理が現実的である。さらに、情報の信頼性の担保が不可欠であり、人による信頼性判断を最終的に行うことが必要不可欠になる。そこで本章では、(1) 情報の流通構造、(2) 信頼性判断、(3) 峻別プロセス、以上について段階的に整理し、除外型情報処理の有効性について検討する。

3.1 災害時の SNS の情報流通構造について

災害発生直後から、SNS 上には現場からの投稿、ニュース映像の転載、共感的投稿、娯楽的引用に加え、常時と変わらぬ日常的な投稿など、多様な情報が流通する。さらには、偽・誤情報の発信やそれに伴った不必要な不安の煽りが社会を混乱させ、これらの善意による拡散も見られる。本節では、一般にどのような情報が災害時に流通するか構造化を行い、後章の「段階的峻別の必要性」や「協調的検証・共有」について検討する。

SNS 上の投稿は、一見すると単線的な時系列で流れているように見える。しかし実際には、他者の投稿をそのまま保ちつつも自身と繋がっているユーザに流す機能 (例: リポスト (旧リツイート), シェア), 元の投稿を自身の画面内に表示したまま自分のコメントを添える機能 (例: 引用ポスト (旧引用リツイート)), ニュースサイトやブログ記事といった外部のウェブサイトに SNS の投稿を貼り付けて表示する機能 (例: 埋め込み) などが同時に働き、同一内容が複数の経路で繰り返し現れる。これらの機能が、連鎖的に他のユーザに機能することで、急激に情報の可視性が高まる可能性が高い。さらに、多くの SNS では表示アルゴリズムが、ユーザの反応や閲覧履歴などをもとに「誰に・どの投稿を先に見せるか」を自動で選出する。X では、有料プランの特典¹として自分の投稿が他者の画面に優先的に表示されやすくなる仕組みも提供されていることから、ユーザの目に入りやすい情報が必ずしも重要であるとは限らず、可視性と重要性の両者には乖離がある。

このような情報発信・拡散の形態から、災害時の SNS における情報の流通構造は、異なる情報層が重なっていると考えられる。代表的には次のとおりである。

第 1 層： 一次情報 (現場発信)

被災者・現地目撃者・自治体関係者による、被害状況・救助要請・ライフライン断絶等、事実確認の必要性がある投稿。

第 2 層： 二次情報 (報道・転載・啓発)

¹About X Premium <https://help.x.com/en/using-x/x-premium> (2025/11/29 参照)。

マスメディア記事やTV画面の写真，ウェブ記事の要約やURL共有，過去事例の再掲，災害時に役に立つ行動tipsや安全確保行動を促す投稿．投稿された時点では正しい災害情報でも，状況の変化により正当性を失ってしまった投稿．

第3層： 反応・感情・日常

被災者や被災地に対する共感・応援・個人的な感想，被災地とは無関係な日常投稿．

災害時において，この第1層にあたる質の高い一次情報の収集が被害状況の把握や対応において重要となるが，投稿自体が稀有であることも指摘されており，第2層～第3層の情報が大多数を占めることが，これまでの研究からも明らかとなっている [38]．

また，情報の性質は，誰が，何のために投稿したかに強く依存する．災害時における主要なSNSの利用者としてHoustonらは

(i) individuals（個人）：

他の主体（政府や組織など）の代理として利用していない一般市民．

(ii) communities（コミュニティ）：

近隣地域や町といった地理的領域で結ばれた人々の集団および「専門知識，価値観，規範，関心，経験を共有する」人々．

(iii) organisations（組織）：

災害に対応している/影響を受けている，あるいは災害の外部に存在する人々の構造化された集団．

(iv) governments（政府）：

連邦，州，地方レベルの政府および政府機関．

(v) news media（ニュースメディア）：

規模の大小や性質（伝統的／新興）を問わず，ニュースを発信・報道する人々や組織．の5つに分類している [17]．本研究では，この5区分を日本の文脈に当てはめ具体化するとともに，各区分の利用者の主目的，および想定される投稿例を以下に示す．

● 個人（Individuals）：

主目的＝救助要請，安否・所在，被害・危険箇所の一次報告，感情，支援呼びかけ等．
投稿例＝位置・時刻・被害の具体報，避難可否，家族・友人の安否連絡等．

● コミュニティ（Communities）：

主目的＝情報の取りまとめ・要支援情報の集約，地図化，復旧情報の共有，結束等．
投稿例＝配布場所・時間，ニーズ，ボランティアの受け入れ条件，ローカルFAQ等．

● 組織（Organisations）：

主目的＝応急対応・復旧の告知，受援窓口の提示，ソーシャル・リスニング．
投稿例＝停電・断水・通信障害の状況と復旧見通し，医療・受診可否，物資・人員の募集等．

- 政府・公的機関 (Governments) :
主目的＝警報・避難情報，公的決定事項の発表，流言訂正，復旧・支援制度の案内等。
投稿例＝避難指示・高齢者等避難，安全確保行動，通行規制・給水等。
- 報道 (News media) :
主目的＝被害の可視化・検証，公式情報への導線と増幅，長期的な続報等。信頼性が高い。ただし，再掲が伴うことで「二次情報化」する。
投稿例＝被害情報更新，検証記事，解説，公式発表の通達等。

この5区分に横断的に存在する2軸として (A) 現場／現場外，(B) 近年増加する SNS 情報の悪用者の観点を加える。

横断軸 A：現場／現場外 投稿が発災地域内（現場）からか外側（現場外）からか。現場からの情報には被害状況等の一次情報が期待できる。ただし，現場外からの投稿には共感・拡散・解説・支援呼びかけに加え日常的な投稿も多く，再掲（リポスト・引用ポスト・シェア・原文コピー等）を伴うことで「二次情報・三次情報化」し，ノイズ増幅の原因の一つとなる。

横断軸 B：悪用者 悪用者は，投稿を悪用する何らかの意図・動機を持ち，偽情報，なりすまし，過去画像流用，詐欺募金，アフィリエイト誘導，過度の不安煽り等の発信を行う。ノイズ増幅の原因の一つ。人手で最終判断する必要がある。政府や公的機関を装うケースも存在する。悪用者は独立のユーザ区分ではなく，5区分すべてに潜在する。

近年では，マスメディアが第1層に当たる SNS の被害投稿を紹介する形で発信する場合も見られ，それを個人ユーザが再発信などすることでノイズ情報として増幅されてしまう。特に，この悪用者が発信する情報は，第1層から第3層・SNS利用者の5区分を横断し，幅広く混在する。これまでに，第1層に該当する現地被害の様子だと装って虚偽の災害被害を示す画像が投稿された事例が多数確認されている。例えば，2016年に発生した熊本地震の際には，「おいふざけんな，地震のせいでうちの近くの動物園からライオン放たれたんだが熊本」という文言とともに，ライオンが街中を徘徊する画像が投稿された²³。この投稿の発信者は逮捕されたものの，続く2018年の大阪府北部地震の際には，「大阪府北部で震度6弱でシマウマ脱走って #地震速報」という文言とともに山道を徘徊するシマウマの画像が投稿された⁴⁵。また，2020年の台風15号の際，「ドローンで撮影された静岡県の水害。マジで悲惨すぎる...」という文言とともに，市街地が大規模に水没している画像が投稿された。この画像に対して，建物の歪みや配置の歪さなどが指摘され，投稿者も生成 AI を

²ハフポスト：「ライオン逃げた」熊本地震のデマ情報を拡散した疑い 20 歳男を逮捕 https://www.huffingtonpost.jp/2016/07/20/lion-escape_n_11081056.html (2025/11/11 確認)。

³朝日新聞：「ライオン放たれた」ほかネット偽情報 135 件 熊本地震で警察記録 <https://www.asahi.com/articles/AST1B3CY8T1BTIPE019M.html> (2025/11/29 確認)。

⁴ハフポスト：「大阪地震でシマウマ脱走」のデマが拡散。警察が注意呼びかけ https://www.huffingtonpost.jp/entry/zebra-osaka_jp_5c5d72b2e4b0974f75b2abfa (2025/11/29 確認)。

⁵NHK 放送文化研究所メディア研究部部長秀彦：#156 大阪府北部の地震で飛び交った流言・虚偽情報 <https://www.nhk.or.jp/bunken-blog/200/309748.html> (2025/11/29 確認)。

用いて作成したと認めている⁶ ⁷。国外でも同様の投稿が散見される。2022年のトルコ・シリア地震の際には、消防士が赤子を抱く画像が midjourney モデルを用いて作成され、支援金詐欺の拡散の一端を担った⁸。2024年の米国で発生したハリケーンヘレンの際には、テネシー州の街が壊滅的被害を受けたと示す画像が投稿されたが、これも AI による生成画像であることがフランスの大手通信社 Agence France Presse (AFP) が行ったファクトチェックにより明らかになっている⁹。2025年10月にジャマイカ海岸沖に甚大な被害をもたらしたハリケーンメリッサの最中、Facebook に投稿された被害画像が拡散されたが、このケースでも、AFP 社によるファクトチェックで AI による生成画像であることが明らかになっている¹⁰。他にも、公的機関が発信していない情報をあたかも真の情報として発信した事例も存在する。例えば、2012年に発生したハリケーンサンディの際には、電力会社が計画停電を行う、ニューヨーク証券取引所ビルが浸水した、橋が封鎖されている、等の偽情報を流したアカウントも確認されている。2014年のアメリカで発生した大雪の際、バージニア州にある学校の公式アカウントの偽アカウントが大量発生し、偽の休講情報を流したり、2016年に米南部ルイジアナ州で発生した洪水の際には、避難所方針や食料配布に関する Facebook 上の複数の噂や誤情報が拡散され、アメリカ赤十字社はその対応に追われた [51]。

こういった偽・誤情報の投稿や拡散は、2023年に X の利用方針が大きく変更され、投稿が見られた量により報酬が払われる仕組みが導入されたことに伴い、増加する恐れがある。実際、令和6年元旦に発生した能登半島沖地震では、偽・誤情報が相当量流通したと考えられている。これは主に国外ユーザから特に閲覧数稼ぎを企図したインプレッション稼ぎと呼ばれる行為であり、現地情報を引用せずに自分ごととして発信された [94]。

山口らは、偽・誤情報対策には、デマ検証用のプラットフォームの創設が有効であると指摘する [100]。例えば、オンライン上の偽・誤情報対策のために組織された非営利団体 First Draft News には Google や Facebook, Twitter 等のソーシャルメディア企業に加え、欧州の報道機関が参加している。First Draft News は偽・誤情報はどのような意図や動機で作成されたか、が重要であると論じ、それらを7区分にした構造を示した [12]。渋谷らは、能登半島沖地震に流通した偽・誤情報を調査し分類することで、これらの情報がどのような意図や動機を持って投稿されたか精査している [75]。渋谷らがまとめた偽・誤情報の分類で用いられる類型について表 3.1 に示す。偽・誤情報には虚偽や捏造、誤解を生む情報の接続、詐称、陰謀論、うわさ、擬似科学に分類することができ、それらの発信には金銭的、イデオロギー的、心理的、善意といった意図や動機があったと推測している。このことから、第2層、第3層に当たる偽誤情報の発信や拡散行為は、これらインプレッション稼ぎや金銭的な詐欺を目的としている可能性が高い。

⁶朝日新聞：AI を見破る技術もオープンに「ディープフェイクの民主化」危惧 <https://www.asahi.com/articles/ASR6852W0R5SUPQJ006.html> (2025/11/29 確認)。

⁷くろん (27)@kuron_nano: ドローンで撮影された静岡県の水害。マジで悲惨すぎる...https://x.com/kuron_nano/status/1574121450860007424 (2025/11/29 確認)。

⁸e-trend micro 社:Turkey Earthquake Charity Scams Alerthttps://news.trendmicro.com/2023/02/20/turkey-earthquake-charity-scams/?_ga=2.81259863.1502376867.1712787874-1635657832.1699306152 (2025/11/29 確認)。

⁹Agence France-Presse(AFP):Fabricated image of Hurricane Helene flooding spreads online <https://factcheck.afp.com/doc.afp.com.36J668A> (2025/11/29 確認)。

¹⁰AFP: AI-generated image misrepresented as hurricane damage on coastal town in Jamaica <https://factcheck.afp.com/doc.afp.com.82UA49E> (2025/11/29 確認)。

表 3.1: 渋谷らによる偽・誤情報の分類で用いる類型 [75]

偽誤情報類型	虚偽・捏造	手間、虚偽、捏造、真実を装うために使用され、一般の人々や聴衆を欺くために事実として提示される、虚偽または不正確な意図に捏造されたもの。反事実（ハーフ・トゥルース）や虚構（ファクトイド・ストーリー）とも呼ばれる。
	誤解を生む情報の接続	異なる文脈の情報をつなげているもの、誤解を招くような情報の使い方。見出し、ビジュアル、キャプションが内容を裏付けていない場合やソース情報の一部は事実かもしれないが、間違った関連（文脈/内容）を使って提示されている場合など。
	詐称	他人や機関になりすましたもの（ジャーナリストの名前/ロゴの使用/模倣 URL など）。
	陰謀論	ある出来事や強力な陰謀家によって生み出された秘密の計画であるという信念。通常、重要な出来事を政府や権力者による秘密の陰謀として説明する。陰謀論は定義上、真偽の検証が困難であり、通常、それを真実だと信じる人々によって生み出される。陰謀を否定する証拠は、陰謀のさらなる証拠とみなされる。
	うわさ	真偽が曖昧であったり、確認されることのない話（噂話、風説、未確認の主張）。
	疑似科学	実際の科学的研究を、疑わしい、あるいは誤った主張で偽っている情報。専門家と矛盾することが多い。
意図・動機	金銭的（閲覧数）	X の仕様変更により追加。X での投稿が閲覧された回数に応じて広告収益が還元される。
	金銭的（振込/送金依頼・不明・その他）	個人送金への誘導。義援金や寄付金を名目に偽の振込先へ送金させようとするものなど。
	イデオロギー的	
	心理的	例: 愉快犯
	善意	災害時の支援を目的としたものなど
	不明	
トピック	被害、救助、寄附・義援金、犯罪・治安、地震メカニズム（人工地震等）、原発、羽田空港衝突事故、その他	

これらの状況を総括すると、前述した SNS の情報層は、大きく現場ユーザ/現場関係者ユーザと現場外ユーザのそれぞれが発信するマトリクス構造になっていることが考えられる。この構造を図 3.1 に示す。また、悪意がなくとも個人ユーザがこれらの偽・誤情報を善意から投稿したり拡散したりする行動も頻発している。実際、能登半島沖地震の際には、これらの善意と見られる拡散が多く見られた [98]。

3.2 信頼性判断の手がかり

災害時のソーシャルメディアは、現場からの第一報を可視化する一方で、誤情報・再掲・誇張・取り違えが混在する情報空間となる。最低限の「人手による見きわめの基準」を共有しておくことで、判断のばらつきを抑え、誤った拡散や見落としを減らすことができる。Twitter Japan（現 X）は救助要請の際に写真や住所、位置情報を付与する事を推奨しており¹¹、過去の実際に救助に繋がった投稿にはこれらの特徴が存在したが [80]、実際に位置情報が付与されている投稿は少なく [24][18]、投稿の実態にそぐわない場合も多い。本節では、投稿を目で見て読んで信頼性を判断する際の着眼点を、First Draft News がまとめた検証のための 5 つの柱 (1) Provenance（原典）、(2) Source（発信者）、(3) Date（日時）、(4) Location（場所）、(5) Motivation（動機） [54] をもとに、SNS の投稿から得られる情報のうち、内容（テキスト）・メディア（画像/動画）・投稿者情報・拡散の様子に加え、信頼性レベルの 5 つの観点から検討し、現場向けの情報精査の要点として述べる。

¹¹TwitterJapan: 災害時における Twitter の 6 つの活用法 https://blog.x.com/ja_jp/topics/events/2018/howtoutilize_twitter_on_disasterpreventionday (2025/11/29 確認)。

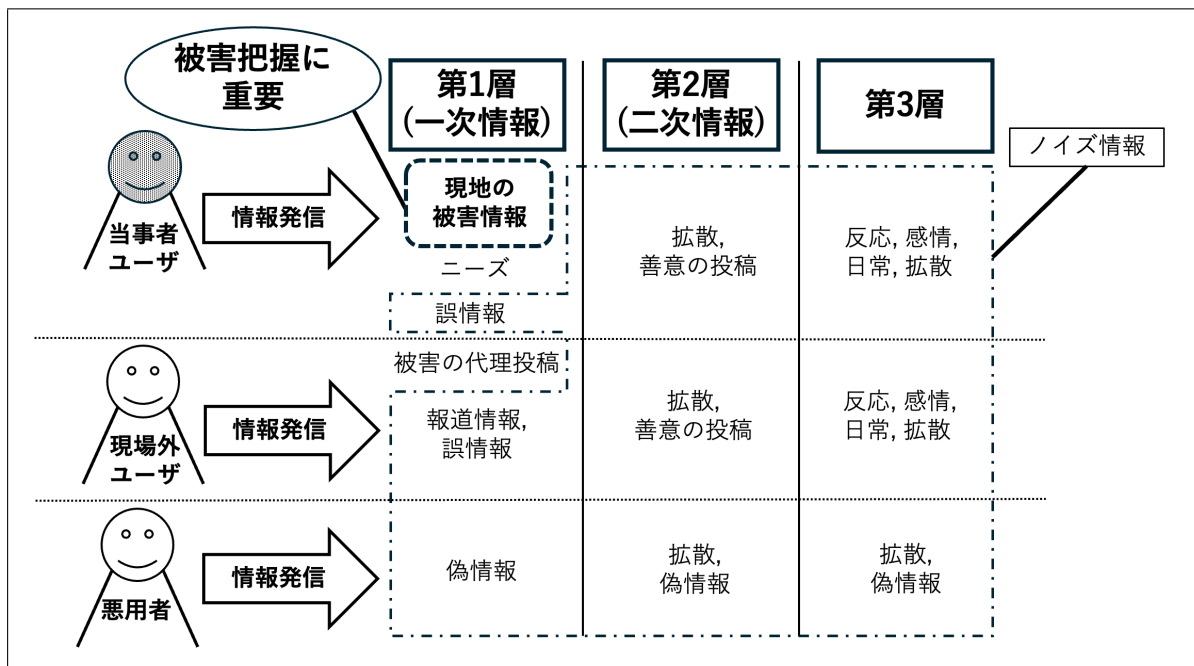


図 3.1: SNS 災害情報マトリクス構造

3.2.1 Provenance (原典)

First Draft News が最も重要であると指摘するのが、その投稿が最初に公開された元のコンテンツかどうかである。投稿者と一次発信者（撮影者）は一致しないことがある。強固な検証のためには、可能な限り一次発信者を特定することを推奨している。

3.2.2 Source (発信者)

投稿者の文脈も重要な指標となり得る。過去の投稿に特定地域の日常が継続的に写り込んでいるか（最寄り駅・地元の店舗等）、普段からその地域の話題に触れているか、当該災害と関係し得る立場にあるか、プロフィールと本文の場所・言語・方言が自然に整合しているかを確認する。新規アカウント、無投稿期間が長いアカウントや外国人アカウントには注意対象だが、長期間運用していない場合や災害を機に作られたアカウントの可能性もあり、停電・避難で更新が途切れたり正当な事情もある。フォロワー数は、アカウント運用の影響を強く受ける指標であり、過信すべきではない。可能であれば直接投稿者とやり取りを行うなどすることが重要になる。

拡散の様子から得られる手がかりも大きい。短時間に同文・同画像が林立している、テンプレート化した文言（「拡散希望」「RT（現在のリポスト）お願いします」）のみが付く、感情語だけが付加され続ける、初期の共有が互いに接点の薄い多くのアカウントによって一斉に行われる、といった場合は特に注意する必要がある。一方で、地理的に近接するユーザからの報告（別角度の写真、小さな違いを含むが実情の記述）が少しずつ積み上がる場合は、実体性の裏づけとして考えられる。ただし、リポストやシェア機能を使っていないとも、再掲の証拠となるような情報が確認できる場合は、一次情報性が損なわれる。例えば、テレビのチャンネルロゴ、ニュース番組名のテロップ、著名サイトの見出し、ロゴの

透かし、別サービスの再生ボタンやコメント欄の断片といった視覚的痕跡があれば、出所を明記した共有として扱い、一次報告として取り違えない。再掲自体は価値を持ち得るが、文脈（いつ・どこで・誰が撮ったか等）の欠落が起きやすいことに留意する必要がある。

3.2.3 Date（日時）

First Draft NewsはSNSのタイムスタンプはアップロード時刻であり、撮影時刻とは限らないと指摘している。ただし、SNSではデジタルカメラやスマートフォンで撮影された画像に記録されているGPS情報やカメラの設定情報が記録されるフォーマット（Exchangeable image file format; Exif）の情報は自動で削除されることが多いこともあり、原本ファイルの入手可否を確認しつつ、太陽高度・過去天気等の外部指標で補う必要性もあるとしている。

3.2.4 Location（場所）

地名は市区町村名だけでなく、交差点・地点名・避難所や店舗名などの固有名詞があるほど検証可能性が高まる[80]。時刻が投稿内に記載してあるのが理想で、投稿時刻とのズレが常識的範囲か確認が必要になる。固有名詞の表記揺れや同名地名の混同している可能性も考慮する必要がある。ただし、先の能登半島沖地震では「石川県 川永市 宮の区 1-5-24 たすけてください #能登地震 #SOS #助けて」のように複数の投稿がみられたが、石川県に川永市は存在せず、架空の地名とともに投稿された[75]。そのため、実在する地なのか検証は必要不可欠である。

3.2.5 Motivation（動機）

動機に関しては、なぜそのコンテンツが作られ、なぜ投稿されたのかは検証しにくいとしている一方で、推測することは十分に有益であるとしている。渋谷らは、偽・誤情報の発信や拡散には、何らかの意図や動機があることを指摘している[75]。したがって、その投稿者は偶発的な目撃者か、特定の主張や目的をもつ活動家か、政府・企業・研究機関の立場か、特定コミュニティに属しているか、金銭的目的があるか等を検証することが必要である。可能なら一次発信者に直接問うことが最も有効だとされている。

3.2.6 添付メディア

画像や動画に対しても、原典、発信者、日時、場所、動機に即して検討すべきだとしている。例えば、「その時間・その場所の現実と矛盾していないか」という状況整合性を確認する。天候（雨滴の付着、濡れた路面、雲量）、時間帯、季節指標（街路樹の葉、季節装飾）、物（道路標識、店舗看板、公共物の銘板、電柱プレート）といった要素が、本文や既知の地理・時間の文脈と自然に噛み合っているかを検討する。Google Street Viewなどを利用して検証することも有効としている。添付メディア内に、テレビ局ロゴやニュースのテロップ、ウェブ記事の見出し断片、他プラットフォームのインタフェースやロゴらしき要素が画面内にある場合は「再掲」や「二次流通」の可能性が高く、一次報告であると誤解しない配慮が要る。写真の縁に不自然な黒帯、極端な縦長・横長、重要部分だけを切り出した痕跡は、何らかの意図と動機を持って編集がされている場合の手がかりになる。また、InVidやgoogle画像検索などの外部ツールを利用することで、過去に別のプラットフォーム等で発信されていたものかどうかを検証することも推奨されている。

3.2.7 信頼性ラベル

近年、プラットフォーム側が信頼性を判断するための補助的な指標や仕組みを導入する動きも見られる。これは、ユーザが情報の信頼性や真偽、文脈を判断する上で重要な手がかりとなる。例えば、日本最大のニュースポータルサイトである Yahoo!ニュース¹²では、「Yahoo!ニュース エキスパート¹³」と呼ばれる認定制度を導入している。これは、各分野の専門家や有識者が実名でニュース記事に対する解説・見解等を「エキスパートコメント」として投稿する仕組みである。エキスパートコメントは、他のコメントより優先的に上位に表示される。これにより、偽・誤情報の混入を極力防ぎ、専門知識に基づいた信頼性の高い情報の可視化を図っている [85]。こういった専門家によるコメントは、ニュース閲覧者に対して、一定の信頼性を担保する指標となり得る [37]。Xプラットフォームには、2023年に導入された新しい機能であるコミュニティノート¹⁴と呼ばれる仕組みが導入されている。これは誤解を招きうる投稿に対し、プラットフォーム側が選定した“協力者”ユーザが注釈や根拠を付与するものであり、文脈補足・誤解訂正・出典提示に有効である [10]。ただし、米国非営利通信社の AP 通信社は付与までの時間差と網羅性の限界があることも指摘している¹⁵。信頼性ラベルが付与されていないからといって真偽不明が直ちに否定されるわけではなく、付いたからといって個々の派生投稿すべてに適用できるとも限らない。あくまで複数ある手がかりの一つとして扱う必要がある。

3.2.8 信頼性評価のまとめ

3.2節で述べてきたように、単一の決め手で即断できる事例は少なく、内容・メディア・投稿者・拡散様式・再掲痕・1章で例示した生成 AI の兆候といった弱い手がかりを突き合わせて矛盾や整合を確かめる逐次的な検証が不可欠である。信頼性において重要なのは、複数のモダリティの相補性である [94]。テキストは検索性と凝縮性に優れるが、災害時は短文・比喩・省略が増えることも考えられるため、単独での真偽判断は難しい。一方、画像や動画は状況を直観的に示すが、撮影の意図や前後関係が見えにくく、過去映像の再掲・テレビ画面の再撮影・スクリーンショットの混在などにより一次性が不明瞭になることも考えられる。そのため、本文の具体性と添付メディアの示す状況が互いに補い合っているかを、人の目で重ね合わせて確認する必要がある。この作業は、時間・注意・専門知の高コストを要し、特に災害下ではそれらの時間的・人的資源が最も逼迫している。そのため、最終的な信頼性判断を広い母集団に対して高頻度で行っていくことは実務上困難であり、見落とし回避と判断の質を両立させるには、その前段で人の目に触れる情報量そのものを戦略的に減らす必要がある。つまり、災害時における現場/現場関係者が発信する第1層の重要情報へ到達するには、第2層～第3層に属するノイズ情報の確実な排除を行うことが現実的である。

¹²Yahoo!ニュース <https://news.yahoo.co.jp/> (2025/11/29 確認)。

¹³Yahoo!ニュース エキスパートについて <https://support.yahoo-net.jp/ScsNews/s/article/H000011258> (2025/11/29 確認)。

¹⁴X: About Community Notes on X <https://help.x.com/en/using-x/community-notes> (2025/11/29 確認)。

¹⁵The Associated Press: Report says crowd-sourced fact checks on X fail to address flood of US election misinformation <https://apnews.com/article/x-musk-twitter-misinformation-cddh-0fa4fec0f703369b93be248461e8005d> (2025/11/29 確認)。

そこで本研究では、重要情報の検出・抽出よりも除外を先に据える立場をとり、明確なノイズを“まとまり（群）”単位で段階的に捨て、曖昧なものは保留として後段の人手精査に回す。除外は常に可逆とし、曖昧性を含む保留域からの繰り返し確認していくことで誤除外を防ぐ。さらに、必要な情報の抽出結果（Pick）と過程（操作ログ等）の共有により、重要情報の共有後に実施される最終的な信頼性判断のための材料と説明可能性を確保する。つまり、人が確認すべき候補集合を小さく、かつ整える前工程こそが、信頼性判断の負担を現実的な水準に保つために重要な役割を果たす。この観点から、次節では本研究が採る段階的峻別プロセスと協調作業の枠組みを具体化する。

3.3 情報の峻別プロセス

前節までに、災害時の SNS 情報の流通構造について、その信頼性を担保するためのプロセスについて論じた。本節では、3.1 の多層構造を前提に、雑多な投稿群から、人が確認すべき情報群の候補を漸次的に確保していくための情報峻別プロセスを定義する（図 3.2 参照）。本研究が想定する人と計算機間の協調とは、計算機による広域の収集・構造化・可視化・可逆な操作基盤・履歴管理を担い、人による収集情報の整理・決定・結果と過程の共有・信頼性判断を担うといった適切に役割分担がされた枠組みである。以下では、収集、分析、整理（除外）、精査（Review / Pick）、共有の 5 段について順に述べる。また、前述の最終的に行うべきである情報の信頼性判断は、本研究で行う一連のプロセスの後に行う工程であり、本節のプロセスはその前段として被害情報の候補集合を整えるためのものである。

まず SNS 情報の収集プロセスでは、災害時に重要情報が含まれる可能性のあるキーワード（e.g.、“救助”や“避難”）を起点に取得を行う。この段階では自動除外は行わず、最低限の網羅性の毀損を避けることを原則とする。

次に、分析プロセスでは、個々の投稿の良否を判定するのではなく、作業単位として扱える程度に整形し、傾向の似た投稿を群として整理する。その際、テキストとメディアの手掛かりを併用してクラスタリングを行い、テレビ画面や記事スクリーンショット、地図キャプチャ、エンタテインメント関連画像や、日常写真などの典型的パターンを視覚的に把握できる状態にする必要がある。ただし、これらの可視化は直接的な情報の信頼性の判断手段ではなく、あくまでも次のプロセスの人による情報整理を行うための支援的情報である。

情報の整理プロセスでは、明確に不要と判断できる情報群から優先的にまとめて捨て、探索空間を縮小することを目指す。情報の段階的除外は、人が「災害とは関連性の低い」と「被害情報の一次情報性が極めて低い」といった条件を満たすかどうかで行う必要がある。例えば、災害とは無関係なエンターテインメント関連コンテンツや、日常的なコンテンツ、もしくは災害に関連し非日常的ではあるものの、テレビ局ロゴやニュースのテロップ、記事タイトル断片、他サービスの UI 断片といった再掲痕が明瞭な場合を典型とする。誤判断による除外を避けるためには、除外操作は常に可逆性であることを要件とする。曖昧な情報群は「保留」として残し、保留と除外済みの双方から情報の再提示することで点検するといったプロセスを行うことが必要になると考えられる。その際、ユーザが情報群の優先順位を付け、判断を「保留」した情報群があることを示すことで、曖昧な情報群の

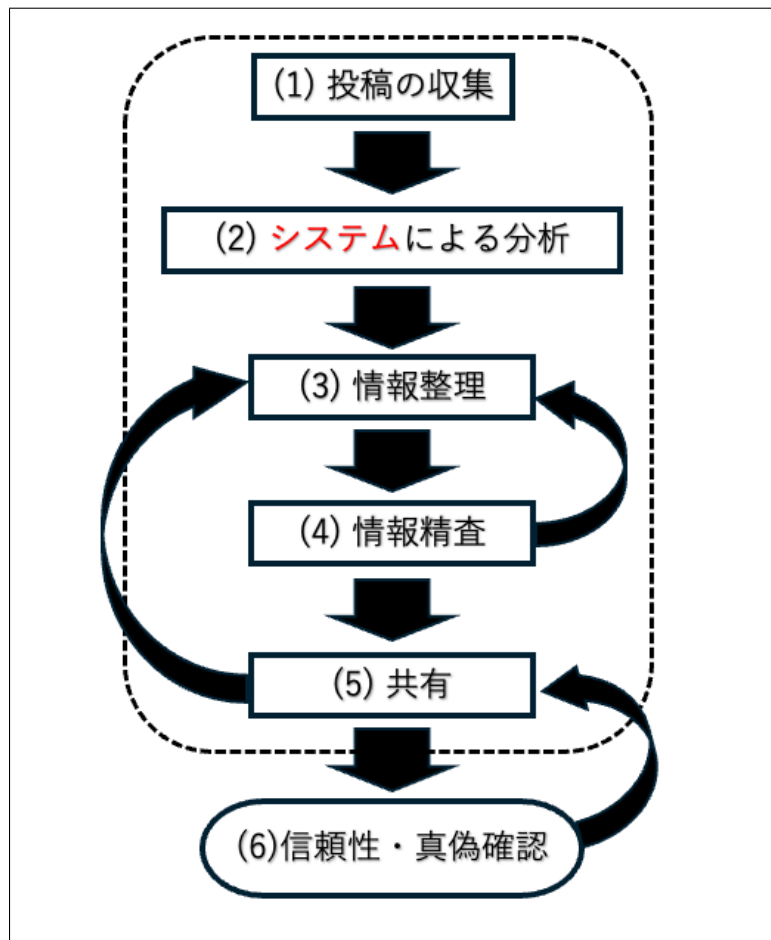


図 3.2: 峻別プロセス

再精査の必要性を忘却することを防ぐ必要がある。また、各除外・保留の決定時には時刻と操作対象とともに操作ログへ自動記録する。

精査プロセスでは、前の整理プロセスで縮小された SNS 情報空間から、人が確認すべき候補を確保する工程である。ここでの判断は暫定であり、候補は「要対応」「要確認」であることを前提に確保する。この確保行為は、テキストには何が書かれているか、メディアに何が映っているか、再掲痕の有無、といった事実ベースで判断することが必要になる。このプロセス中には、誰がいつ何を Pick したかを記録する。この精査プロセスは、情報整理プロセス途中段階においても精査可能であることから、情報の整理プロセスと審査プロセスはループ構造となっている。

最後に、共有プロセスでは情報の整理プロセスおよび精査プロセスの結果と過程の両方が共有可能であることとする。Pick や保留、除外の履歴を保存することで、第三者が容易に追跡・確認できる状態をつくる。また、この後に行う投稿の信頼性判断では、順次共有されたものから判断されていくことが想定でき、両者はループ構造となっている。

これらの一連のプロセスが完了した後、共有した情報の信頼性判断フェーズに入ることが可能になる。本研究で行うこれら一連の情報の峻別プロセスは、最終的な真偽判断に資するための準備工程であり、本モデルの将来的な運用は、役割の異なる複数人による互助

を想定する.

3.4 本章のまとめ

この章では、災害時という人的・時間的リソースが限られた状況下で、SNS からユーザ投稿情報から災害被害に関連する情報を確保するために、計算機によって投稿を特徴ごとに群としてまとめ、人手によって災害とは無関係な情報を除外していくことで必要な情報の可視性を高め、情報確保を容易にするための人と計算機による協調モデルを提案した.

本章で示した枠組みとして、Houston らがまとめた災害時の SNS 利用者 5 区分に加え、現場/現場外・悪用者という横断軸を加えることで、災害時の SNS 情報の層構造と発信主体のマトリクスにより情報空間を把握するためのモデルを検討した. 次に、情報の信頼性評価のためには、内容・メディア・投稿者・拡散・再掲痕・コミュニティノートの観点に加え生成 AI の兆候を加えた手がかりを総合的に参照する必要があることを示し、人的・時間的に制約の大きい災害時には、いかに人が確認すべき情報を減らすことが必要かを検討した. また、災害時の重要情報を確保するためのモデルとして、収集、分析（クラスタリング等）、整理（除外／保留）、精査（確保）、共有という段階的峻別プロセスを提案した. これにより、重要情報の計算機処理による取りこぼしを最小限にすることを目的とした情報除外手法の有効性を概念・手順・役割の観点から検討した. 本章のモデルは、後続章で順次検証される. まず 4 章では、令和 2 年 7 月豪雨データにおいて災害と関係のないノイズとなる情報について検討した上で、「ユーザが撮影した被害」投稿は全体の 0.564% に過ぎないことを示し、抽出より除外を先行させる必然性を定量的に裏づける. 5 章では、マルチモーダル特徴に基づく可視化・クラスタリングにより、ノイズ群がクラスタとして偏在し情報除外が機能し得る構造を示す. 6 章では、クラスタの提示・3D 可視化・クラスタ単位の段階的除外・検索・Pick/操作ログ等を備えたインタフェースを実装し、時間制約下で被害が確認可能な投稿の収集効率を上げることを目指す.

4 災害時における SNS の情報構造

これまでに、災害時のソーシャルメディア活用における課題として、情報の信頼性担保と、時間的・人的リソースの限界を指摘し、重要情報を直接抽出するのではなく、ノイズ情報を段階的に除外することで、人が確認すべき情報を確保する枠組みについて提案した。この枠組みの有効性を検証するためには、実際の災害時において、ソーシャルメディア上にどのような情報が流通し、人が確認すべき情報と、除外すべきノイズ情報がどのように存在しているかを定量的に把握する必要がある。もし、重要で有益な情報が十分に多く、特徴が明確であれば、既存の検索技術や抽出手法で対応可能だと考えられる。しかし、有益な情報が極めて稀少であり、かつノイズ情報構造が複雑であれば、単純なキーワード検索や NLP 技術のみを用いての抽出は機能不全に陥る可能性が高く、これまでに提案した除外型アプローチの必要性が裏付けられる。

そこで本章では、「令和 2 年 7 月豪雨」発生時に投稿された Twitter (X) データを対象に、画像分類、およびテキスト分析を用いてどのような情報が実際に流通していたか悉皆的に調査を行う。収集された膨大な投稿に含まれる画像を人手によって分類し、それぞれのカテゴリにおけるテキスト特徴および類似度を算出することで、災害時の情報空間におけるノイズ情報について明らかにする。また、被災者が発信する「被害」を示す情報の稀少性を定量的に示し、本研究が目指すノイズ除外による支援の必要性について論じる。

4.1 調査方針と対象災害

4.1.1 本章の位置づけ

ノイズを取り除くアプローチを採る場合ノイズ情報の特定が必要になるが、災害時の投稿にどのようなものがあるのかは十分に明らかになっていない。「救助」や「避難」といったキーワードやそれを意味するハッシュタグで検索しても、災害や被害には無関係な情報が多く含まれる。本研究ではこうした無関係な情報を判断する手がかりとして投稿に添付される画像に着目した。投稿の収集は災害に関するキーワードマッチング方式で行い、それらの投稿に添付されている画像の種類を分類した上で、それに伴う投稿テキストの傾向分析することで、どのような情報が災害時に流通しているか調査を行う。

本研究では、収集した投稿にどのような画像が添付されているかの調査を行い、その特徴を明らかにする。その上で、それらの画像がどのような文章と同時に投稿されているかを調査し、残すべき情報・取り除くべき情報の各投稿の文章特徴を明らかにすることで、災害や被害とは無関係な情報を機械的に抽出して除外する方法を検討する。X には画像や映像、音声、テキストなど複数のモダリティの情報が含まれている。常時であれば、こういった複数のモダリティが存在することで、ユーザ間のコミュニケーションが促され SNS の発展につながるが、災害時においてはこれらの膨大な情報がノイズとして認識される場合がある。これらの投稿データを災害情報の収集源として扱うためには、文章トピックやユーザ属性、画像の分類や位置情報などを多面的に考慮していく必要がある。本研究は、複数のモダリティを含む情報を対象に、モダリティごとの特性を相補的に利用することで、効率的に救助情報や被害情報を抽出することを志向している。そのためには、多面的な情報を入力とし、組み合わせで処理するマルチモーダルな技術が必要になってくる。マルチモー

ダルな技術を扱うためには、それぞれのモダリティを含む情報を各々定量的に特徴化する必要がある。Xに投稿された画像の特徴とテキストの特徴を定量的に特徴量化するためには、各モダリティの情報が分類されている必要がある。前提として、Xには災害や被害とは無関係の画像が多く投稿されていることが想定される。そのため、収集した投稿にどのような画像が添付されているかの調査を行い分類することで、災害と関係するの否か、災害とは関係するが不必要な情報なのか、二次的・三次的に発信されたものなのか、一見災害に関連する情報であるが実際に関連するものか、といった、判断材料として活用できるか検証する。具体的には、投稿中で供出する画像と文章を調査対象とし、各投稿の内容特徴を明らかにすることで、どのような情報として扱うべきか考察する。

4.1.2 調査方針

これまで上述してきたように、人手による迅速な災害情報トリアージを実現するためには、適切なノイズ情報の排除を経る必要がある。これらの無関係な投稿を排除するためには各段階でノイズとなる情報の特徴を明らかにする必要がある。まず、災害時のXに投稿された画像および動画の調査を行う。全ての画像・動画を人手で参照し、どのようなものが写っている画像か調査し分類する。次に、先で行った画像の分類をもとに、画像が添付された投稿の文章特徴の調査を行う。画像分類調査では、抽出すべき災害被害に関する情報がどの程度存在するか、不必要な情報がどの程度存在するかを明らかにする。文章特徴調査では、分類した画像が添付された投稿の排除可能性を明らかにする。調査は以下の手順で行う。まず、人手によって画像の分類を行う。その後、テキスト特徴を算出し抽出することで、被害に関連する投稿や、ノイズとなる投稿の抽出が可能になるか調査を行う（ステップ1）。次に、ステップ1では分類しきれない画像に対し人手で投稿本文と画像を確認することで、さらに細分化された分類が可能になるか調査を行う（ステップ2）。ステップ1については4.3節で、ステップ2については4.4節で各々述べる。

4.2 対象災害とXデータ

本章では、令和2年7月3日から31日にかけて発生した「令和2年7月豪雨」を分析対象とする。気象庁の発表によると、この豪雨により九州地方や岐阜県周辺で記録的な大雨となり、球磨川や筑後川、最上川といった大河川での氾濫が相次ぎ、土砂災害によって甚大な被害をもたらされた¹。

「令和2年7月豪雨」発生の際のX（旧Twitter）データはTwitter API²を用いてキーワード「救助」および「避難」をクエリとした収集を行った。収集キーワードの設定は、これまでの先行研究を参考に設定した[80][81]。データ収集対象期間は、全国を通して最も降水量の多い期間であった令和2年7月4日～7日を含むように、7日から前後1週間を目処に令和2年7月1日～15日とした。これらのキーワードを用いて得られた「令和2年7月豪雨」の投稿総数は476,827件であった。そのうち、キーワード“救助”で得られた投稿は110,261件、“避難”で得られた投稿は370,531件となった。この投稿の中から何らかの画像・動画

¹国土交通省気象庁：令和2年7月豪雨 令和2年(2020年)7月3日～7月31日（速報）<https://www.data.jma.go.jp/stats/data/bosai/report/2020/20200811/20200811.html>（2025/11/29 確認）。

²Twitter API<https://developer.x.com/ja/docs/x-api>（2025/11/29 確認）。

が含まれている投稿は、キーワード“救助”で得られた投稿には18,197件、キーワード“避難”で得られた投稿には34,777件、“救助”と“避難”両方のキーワードが含まれ重複した投稿は5,540件となり、合計で47,434件の投稿となった。これらの投稿の中から画像および動画を抽出した結果、合計94,111枚の画像・動画となった。なお、画像・動画は投稿のIDと紐付いている。

4.3 ノイズとなる投稿に関する調査

まず、Xで発信される情報の特徴を明らかにするため、投稿に含まれる画像の分析を行う。

4.3.1 投稿画像の分類

Xでは一投稿あたり4枚まで画像を添付することが可能であり、投稿によってはそれぞれの画像が異なる内容を示すケースがある。その際は、それぞれ異なる画像クラスとして個別に分類を行った。分類にあたっての観点は以下の通りである。

- 災害に関連するか
- メインに写っているものが何か
- ニュースに関連するか

本研究では、災害時にソーシャルメディアから収集すべき情報は、実際に被災地にいる人が災害状況や被害に関して発信した投稿であり、新聞やTVなどの報道機関の発信や被災地以外の人々が被災地を心配したり共感したりする発信ではない。そのため、本研究で一次災害情報として収集したい画像は、投稿者自身が被災地で自ら撮影した写真である。その写真が撮れることは、「実際に被災地にいる」、「被害を実際に確認した」、ことを含意しており、被災地において実際に問題に直面している証拠として確度が高いと考えられる。以下に分類の手順を述べる。分類は情報系学部に通う大学生4名に依頼した（以下、分類担当者）。はじめに、分類の基準を統制するために、分類担当者に対してニュース報道に関連する画像例、および事前のサンプリングにより多数観測できたゲームに関連する画像例として、以下の4つのクラスを例示した。各クラスに分類すべき画像の例を図4.1から図4.3に示す。これらの例に挙げた画像は計算機による画像認識を行った場合、被害画像として認識される可能性があるため、ノイズとして扱う必要がある。特に逆L字は、災害発生中のテレビ番組やニュース放送の上から重ねられたL字型帯の枠内に災害速報等を表示するため、他クラスと混同しやすいことが考えられる。

- 逆L字と呼ばれる災害被害発生中のニュース速報などが放送される際に付けられる帯のようなものが映り込んでいる画像（図4.1）³
- テレビのニュースをカメラで撮影した画像（図4.2）⁴
- メディア媒体を問わず、ニュースに関する画像（図4.3）⁵

³<https://x.com/tasklong/status/1278562348823797760>（2025/11/29 確認）。

⁴<https://x.com/TAKAYA05378399/status/1279181014711930880>（2025/11/29 確認）。

⁵<https://x.com/amasehimika147/status/1150073316863959041>（2025/11/29 確認）。

表 4.1: 人手による画像分類クラス抜粋

クラス	基となった投稿数
Any ニュース	1546
Twitter	1538
ゲーム	8230
テレビのニュース	1471
二次元画像	2115
人	2389
地図	1205
救助	691
被害	4293
避難	3826
風景	3168
飲食物	1445
天気	842
防災グッズ	618
自衛隊	439
逆L字	324

- ゲーム画面やアニメ・漫画などの画像

ここで、ある画像が複数の分類クラスに該当すると判断した際は、その画像を該当するすべての分類クラスに割り当て、その際は画像を複製することとした。その結果、分類担当者が実際に分類した画像・動画の複製を含めた総数は 292,466 となった。その後、クラスの名称に表記揺れが起きたクラスや同一概念のクラスが複数作成される可能性を考慮し、分類担当者に聴取を行った上で、同一の分類クラスであると判断できるものを著者の判断により統合した。その分類結果の一部を表 4.1 に記す。4.2 節の手順で収集した際、リポスト（旧リツイート）を排除していなかったため、同一の投稿 ID を持つ画像が多く確認できた（重複投稿）。また、同一の画像でも異なる投稿 ID を持つ画像も確認できた（他者による再投稿）。これらの画像を人手で分類した結果、分類担当者によっては同一の画像を異なるクラスに分類するケースが多く存在した。

最も多くの画像が分類されたクラスは“ゲーム”であった。“ゲーム”に分類される画像群には「Identity V 第五人格」と呼ばれるゲームに関する画像や動画が多く確認できた。この原因として、ゲーム内で一般的に利用される「救助」という語が、検索結果として収集されたことが考えられる。次点で多かったのは“被害”のクラスに分類された画像であった。“被害”に分類された画像には、個人が撮影したと思われる被災画像の他に、メディアの報道で流れた被災画像、ウェブニュースのサムネイルに設定された被災画像、日本各地に設置されたライブカメラからの被災画像などが確認できた。分類を行った結果、被災したことを示唆する画像を添付した投稿は 4135 件となり、全体の 0.867% となった。



図 4.1: 逆L時が確認できる画像例³

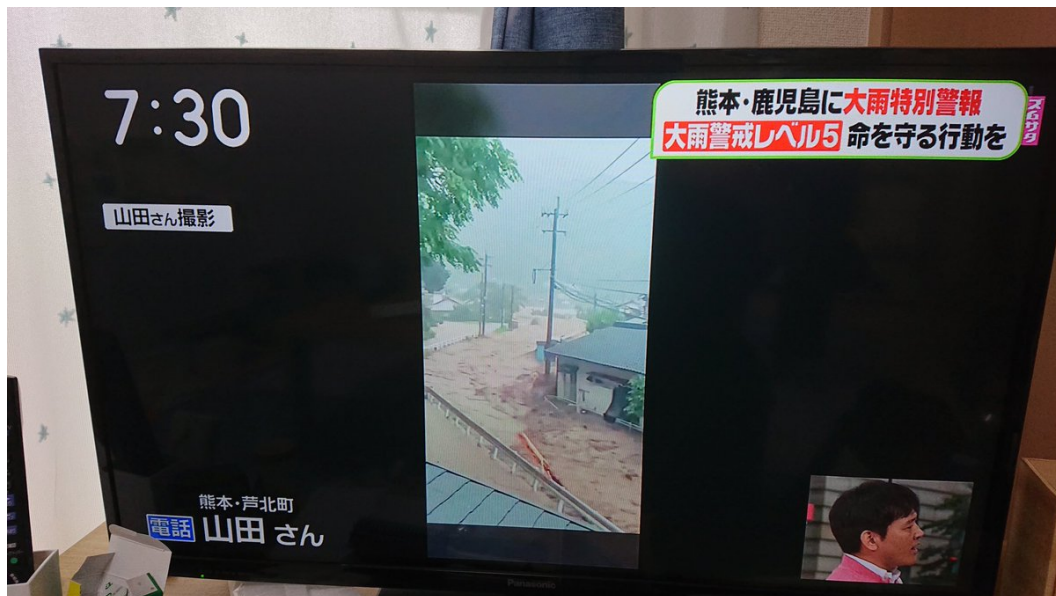


図 4.2: テレビの報道をカメラで撮影した画像例⁴

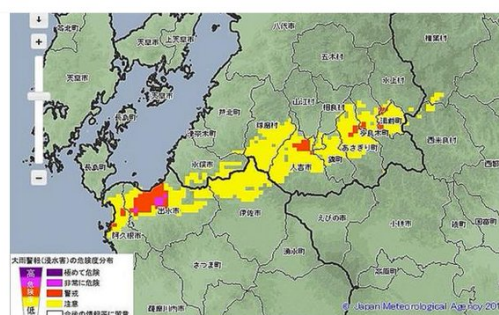
九州 非常に危険 線状降水帯発生中

2019年7月13日 21時43分 tenki.jp



九州、危険な雨の降り方。発達した雨雲が同じような場所にかかり続ける、線状降水帯が発生中。熊本、鹿児島に警戒レベル4相当の土砂災害警戒情報発表。

同じ場所に滴のような雨



13日(土)20時30分現在、熊本県から鹿児島県にかけて、線状に活発な雨雲が連なる線状降水帯が発生しています。熊本県から鹿児島県の県境付近では、夕方から同じような場所で1時間に50ミリ以上の滴のような雨が降り続けています。1時間雨量は、熊本県球磨郡山江村(山江)で73.0ミリ、鹿児島県出水市(出水)では、3時間で125.5ミリの大雨となっています。(1時間雨量、3時間雨量ともに13日20時30分まで)

土砂災害警戒情報発表中

熊本県や鹿児島県には、警戒レベル4にあたる土砂災害警戒情報が発表されています(13日20時30分現在)。同じような地域で、非常に激しい雨が長く続くため、土砂災害の危険が高まっています。安全な場所へ避難、夜間のためかえって避難が危ない場合は、自宅が2階建ての場合は2階へ避難。崖に近い場所は崖から離れた部屋へ移動するなど早め早めの行動をお願いします。

図 4.3: ニュースに関する画像例⁵

4.3.2 投稿画像に付随する投稿テキストの処理

前節で挙げた災害情報の収集に当たっては、精度の点から画像自体を画像処理で解析してその内容を理解するのではなく、その画像に付随するテキストに着目して収集することが現時点で適切だと考えている。そこで本節では、収集すべき災害画像に付随する投稿文章の特徴について調査する。この調査にあたり、本稿では文章の特徴把握の手段として Term Frequency-Inverse Document Frequency (TF-IDF) 値を用いる。TF-IDF 値を算出することで、分類別の画像群を代表する特徴的な単語の抽出が可能になる。投稿テキストの傾向分析を行うため、4.3.1 節で得られた各画像クラスを一群として TF-IDF 値を算出する。

はじめに、4.2 節で取得した投稿データから、他者の投稿を再発信するリポスト（旧リツイート）を排除した。また、投稿本文に含まれる URL を排除した。その後、全角英数字を全て半角英数字に変換した。次に、形態素解析器 MeCab (ver.0.966.2)⁶を用いて形態素解析を行い、単語ごとに分割した。最後に、それらの単語から名詞のみを抽出した。その際、SlothLib プロジェクト [60] でストップワードに設定された単語と、投稿取得の際に設定したキーワードである救助・避難をストップワードとして設定した。分析は TF-IDF を用いる。Term Frequency (TF) とは、ひとつの文章中におけるある単語の出現頻度を指す。その単語の出現回数が多ければ多いほどこの数値は大きくなる。算出方法は、以下の通りである。

$$TF_{i,j} = \frac{N_{i,j}}{X} \quad (4.1)$$

X はある文章 j に出現する単語の総出現回数であり、 $N_{i,j}$ はある文章 j における単語 i の出現回数を示している。Inverse Document Frequency (IDF) とは、とある単語が、いくつの文章で出現したかを示す指標であり、特定の文章でしか出現しない場合に数値は大きくなる。

⁶<https://taku910.github.io/mecab/> (2025/11/29 確認)。

算出方法は以下の通りである。

$$\text{IDF}_i = \log_e \frac{N}{df_i} + 1 \quad (4.2)$$

N は総文章数を示す。 df_i は単語 i を含む文章数である。以上の TF 値、IDF 値をもとに TF-IDF 値を算出する。算出方法は以下の通りである。

$$\text{TF-IDF}_{i,j} = \text{TF}_{i,j} \cdot \text{IDF}_i \quad (4.3)$$

本稿では、表 4.1 のクラス 23 種類をもとに、以上の式から算出した。算出の際、各画像分類クラスの投稿テキストに出現する単語を Term、各画像分類クラスの投稿テキスト群を Document として扱う。算出の際、全ての文章のうち 8 割を超えるクラスで出現した単語はクラスの特徴として不適であるため、それらの語を除外した。算出した TF-IDF 値の結果の一部を表 4.2 に示す。この分析によって、各画像分類クラスの投稿テキストの傾向から、画像を特徴づける単語の傾向を見つけることが可能になる。TF-IDF 値算出の際、全てのクラスのうち 8 割を超えるクラスで出現した単語を除外し、各々のクラスで固有の単語が上位を算出した。“ゲーム”で TF-IDF 値が高く上位に算出された単語は、前節で挙げたゲーム固有の表現であった。“TV ニュース”で上位に算出された単語に関しては、“【2018 年 7 月】避難勧告対象 101 万 4930 世帯・232 万 1947 人。大雨特別警報が 11 府県にわたって発令される前代未聞の状況のなか宴会を楽しむ安倍さん／東洋経済オンライン”という投稿の単語が大部分を占め、繰り返し投稿されていたことが確認できた。この投稿は Bot に依るもので、付随した画像は、株式会社 TBS テレビ⁷の報道番組である「news23⁸」で報道されたニュースの切り抜き画像であった。この原因として、本来、同一クラス内であれば投稿 ID の重複により、同内容の投稿は排除されるが、繰り返し投稿されることで異なる投稿 ID が付与されていたため重複したと考えられる。“被害”で上位に算出された単語である“坂本”および“八”は熊本県にある八代市坂本町の一部であった。“逆 L 字”では 7 月 8 日に岐阜県で大雨特別警報が発表された際の投稿が収集されており、早めの避難を呼びかける投稿が確認できた。

4.3.3 コサイン類似度の算出

同一 ID が TF-IDF 算出の際に影響を与えている可能性を考慮し、テキスト群間のコサイン類似度の算出を行った。コサイン類似度は、ベクトル同士の角度の近さを表現することができる。これを文書特徴の一つである TF-IDF を用いて算出することで、文書同士の類似度を表現することができる。類似度が高く算出される傾向の高かったクラスを表 4.3 に示す。類似度が高かった語群として、“救助”と“自衛隊”、“地図”と“天気図”と“天気”などがあった。また、“テレビニュース”は、“被害”や“避難”と類似度が高く、一般的な報道と被害情報や避難情報が混交する可能性が考えられる。“被害”では複数のクラスと類似度が高く算出され、災害情報全体を象徴する語である可能性が示された。また、当初の懸念通り、同一 ID の投稿が他クラス間で存在している場合、クラス間のコサイン類似度が高く算出される傾向にあった。分類担当者への聴取から、画像を分類するクラスの判断が困難だった場合の画像分類先は、本来分類すべきクラスではなく、結果としてコサイン類似度

⁷TBS テレビ <https://www.tbs.co.jp/> (2025/11/29 確認)。

⁸news23 <https://www.tbs.co.jp/news23/> (2025/11/29 確認)。

表 4.2: TF-IDF 上位 5 単語抜粋

ゲーム	TF-IDF	テレビニュース	TF-IDF	救助	TF-IDF
チェイス	0.3242	前代未聞	0.1966	師団	0.4122
傭兵	0.3007	安倍	0.1903	派遣	0.3462
解読	0.2681	府県	0.1640	陸上	0.2264
祭司	0.2384	東洋	0.1640	神瀬	0.2065
即死	0.1950	八	0.1575	実施	0.1883

被害	TF-IDF	逆L字	TF-IDF	防災グッズ	TF-IDF
坂本	0.2052	早め	0.5147	セット	0.6452
八	0.1825	岐阜	0.3666	グッズ	0.3340
筑後川	0.1754	放流	0.1693	楽天	0.3184
ダム	0.1611	ダム	0.1585	袋	0.2343
増水	0.1414	日田	0.1580	用品	0.2254

が高く算出される傾向のあったクラスへ割り振るケースを確認した。そのため、同じ画像でも分類担当者によっては、違うクラスに割り振られた場合も存在する。この結果から、類似度が高く算出される傾向のあるクラス同士の画像は、人間の判断による画像分類においても分類ミスを起こす可能性が高いことが示唆された。

表 4.3: コサイン類似度が高い傾向のあるクラス抜粋

テレビニュース	Cos 類似度	人	Cos 類似度	地図	Cos 類似度	天気	Cos 類似度
被害	0.4728	避難	0.4206	被害	0.5291	天気図	0.9031
避難	0.4519	風景	0.3697	天気	0.5234	地図	0.5234
地図	0.4254	テレビニュース	0.3173	天気図	0.5183	避難	0.3455
人	0.3173	被害	0.2904	避難	0.4818	風景	0.3294

救助	Cos 類似度	被害	Cos 類似度	避難	Cos 類似度	風景	Cos 類似度
自衛隊	0.7939	風景	0.5341	地図	0.4818	被害	0.5341
被害	0.3499	地図	0.5291	テレビニュース	0.4519	避難	0.3759
テレビニュース	0.3141	テレビニュース	0.4728	被害	0.4499	地図	0.3728
SNS	0.2139	避難	0.4499	人	0.4206	人	0.3697

4.4 被害投稿の特徴に関する調査

4.1.2 節での懸念通り、調査1の結果から“被害”では複数のクラスと類似度が高く算出される傾向があった。そこで“被害”に分類された画像について、筆者によるランダムサンプリングを行った結果、“被害”画像には他の分類クラスの画像が混交していることが確認された。X Corp.Japan 株式会社⁹（旧 Twitter Japan 株式会社）は、救助を求める場合、自身が置かれている状況を説明するために、その状況を示す写真を投稿することを推奨している¹⁰。したがって、“被害”には実際に被災者が投稿した画像が含まれている可能性が高い。そのため、“被害”画像には、他の分類クラスの投稿が比較的多く混交していることを考慮し、再度人手による画像分類を行うことで、細分化の可能性に関する調査を行う。この混交を、画像を参照することのみで分類することの限界点とみなし、テキスト情報加えた分類を行う。

4.3 節で行った画像特徴に着目した分類と、TF-IDF の特徴では分類できなかった画像群を、画像とテキストを人目で確認することで調査する。まず、“被害”に分類された画像を再度確認する。何を意図した画像か判断が難しいため、画像の元となった投稿を参照することで、その画像の出典や投稿の意図まで確認する。それにより、画像を見ただけでは判断が難しい分類を行うことが可能になる。その後、分類担当者は各自の判断で被害画像のみを細分化して分類を行う。最後に、細分化した“被害”に分類された投稿文同士を4.3.2 節と同様の手順で TF-IDF 値の算出を行う。被害に関する画像を細分化し、各クラスの持つ文章的な特徴を明らかにする。

4.4.1 被害画像の分類と細分化

“被害”に分類された投稿には、明らかな分類ミスを起こしていた画像を除くと“テレビのニュースで報道された被害の画像”（66 投稿），“ネットニュースで報道された被害の画像”（204 投稿），“ユーザが撮影した被害の画像”（2689 投稿），“ライブカメラで撮影された被害の画像”（209 投稿），“自治体が発信した被害の画像”（95 投稿），“分類不可能だった被害の画像”（153 投稿），“元投稿が消されていた被害の画像”（101 投稿）に分類された。“分類不可能だった被害の画像”は、中文や英文で投稿されていた場合、日本国以外で発生した災害について言及されている場合（e.g., 2019 年 9 月 28 日韓国南東部での石油タンカーの爆発、2020 年 7 月 6 日に始まった中国江西省での洪水災害）、元投稿を辿っても画像の撮影状況や画像の出典が不明な場合に分類されている。“自治体が発信した被害の画像”は地方自治体によって運営されているアカウントや政治家のアカウントから投稿された被害の画像が分類されている。再分類の結果と再分類された投稿のみで算出した TF-IDF 値の上位 5 単語を表 4.4 に示す。

⁹X Corp.Japan<https://about.x.com/ja>

¹⁰Twitter Japan: 災害に備えるための Twitter 活用法 https://blog.x.com/ja_jp/topics/company/2019/how-to-utilize-twitter-during-naturaldisaster

表 4.4: 細分化した“被害”画像付き投稿における TF-IDF 上位 5 単語

テレビニュース	TF-IDF	ネットニュース	TF-IDF	ユーザが撮影	TF-IDF	ライブカメラ撮影	TF-IDF
熊本	0.3606	熊本	0.3952	川	0.2912	氾濫	0.3950
当時	0.3114	豪雨	0.2803	雨	0.2708	水位	0.3402
園	0.2854	大雨	0.2021	氾濫	0.2482	カメラ	0.3112
寿	0.1986	指示	0.1823	熊本	0.2354	ライブ	0.2847
千寿	0.1631	災害	0.1792	大雨	0.2232	川	0.2581

自治体発信	TF-IDF	分類不可能	TF-IDF	元投稿消去	TF-IDF
災害	0.4000	峡	0.2617	熊本	0.2230
駐屯	0.2539	ダム	0.2580	お願い	0.2088
豪雨	0.2461	洪水	0.2051	地区	0.2036
派遣	0.2377	大雨	0.1864	坂本	0.1998
熊本	0.2256	中国	0.1764	連絡	0.1978

表 4.5: “被害” 投稿数とその割合（一部抜粋）

	投稿数	被害画像投稿 (%)	画像投稿 (%)	投稿総数 (%)
被害-自治体発信	95	2.297%	0.200%	0.020%
被害-分類不可能	153	3.700%	0.323%	0.032%
被害-元投稿消去	101	2.443%	0.213%	0.021%
被害-テレビニュース	66	1.596%	0.139%	0.014%
被害-ネットニュース	204	4.933%	0.430%	0.043%
被害-ユーザ撮影	2689	65.030%	5.669%	0.564%
被害-ライブカメラ撮影	209	5.054%	0.441%	0.044%
“被害” 画像投稿総数	4135	100.000%	8.717%	0.867%
画像投稿総数	47434		100.000%	9.948%
投稿総数	476827			100.000%

4.4.2 細分化された分類のテキスト特徴

“テレビのニュースで報道された被害の画像” で上位に算出された単語である“園”，“寿”，“千寿”はいずれも熊本県球磨郡の特別養護老人ホーム“千寿園”に関するニュースであった。また，TF-IDF 値の第6位は“取材” (0.1565)であったが，その他の分類において“取材”のTF-IDFは最高で0.0233となった。“ライブカメラで撮影された被害の画像”で上位に算出された単語はいずれも河川の氾濫に関連する単語であった。これはライブカメラの設置場所が主に河川の観測所にあるためだと考えられる。“自治体が発信した被害の画像”の最上位に算出された“災害”は，“ネットニュースで報道された被害の画像”では第5位に確認できたものの，その他の分類におけるTF-IDF値は比較的低い傾向が確認できた。自治体や政治家によって投稿される投稿には“豪雨”や“雨”などによる被害であることに加えて，“災害”による被害であることをより強調している。また，“災害”に伴う自衛隊の“派遣”状況を伝える投稿が多く確認された。“分類不可能だった被害の画像”で上位に算出された“峡”，および“ダム”は中国の長江中流にある“三峡ダム”を指していた。分類不可能とする判断基準の一つが中国語による投稿であった。“元投稿が消されていた被害の画像”で上位に算出された“お願い”は，“「救助を」お願いします”などといった救助を要請する場合に多く使われていた。救助要請投稿は具体的な地名を入れるという特徴があり [81]，“地区”や“坂本”といった単語はその救助要請投稿の特徴を反映したものだと考えられる。一方で，被災者の救助要請や被害の状況報告が最も反映されていると想定される“ユーザが撮影した被害の画像”のTF-IDFにおいて，特出した値や単語は認められなかった。

4.4.3 細分化により得られた知見

4.4.2節で得られた結果から，キーワードによる投稿抽出には，共通の文字を扱う中国語の投稿も抽出されるため，こうした無関係な中文による投稿を除去する必要がある。“元投稿が消されていた被害の画像”には，具体的な地名が書かれており，X Corp.Japan社が想定している救助要請の形式を保った投稿や，実際に救助されたケースなどが確認できる可能性が高い。この101件の投稿は今後定性的に調査する必要がある。“被害”に分類された

投稿の細分化が可能であった一方で、被災者の救助要請や被害の状況報告が最も反映されていると想定される“ユーザが撮影した被害の画像”のTF-IDFにおいて、特出した値や単語は認められなかった。つまり、“被害”に関する投稿には特徴となる単語は見受けられず、投稿された文章のみを用いた被害投稿の抽出は困難である。ユーザが撮影した被害画像の投稿を抽出するためには、テキスト特徴に加えて画像特徴といった異なるモダリティの特徴を考慮したモデルを構築することで、抽出精度向上に繋がるか、今後確認する。

これまでの結果から、収集した投稿からノイズとなる投稿を排除した場合の結果として、関連する各投稿数を表4.5に示す。“救助”および“避難”で投稿検索を行った場合、その投稿数は476,827件となる。それらの投稿から、災害時の実被害に該当する、ユーザ自身によって撮影されたと想定される“ユーザ撮影”投稿に絞った場合、2,689件となり、全体の0.564%となった。

4.5 考察

本章では、令和2年7月豪雨の期間に投稿された投稿を収集し、投稿された画像を人手で分類した分類クラスをもとに投稿テキストの特徴を調査した。4.3章で行った画像分類の結果、投稿される画像の傾向や、誤分類が発生しやすい分類クラスが明らかになった。その画像分類クラスを一つの文書とみなして投稿テキスト群を作成し、分類クラスの文書間のTF-IDFを算出した結果、各分類クラスで特徴的な単語が確認できた。つまり、画像特徴のみに着目した分類クラス別のTF-IDFという単純なテキスト特徴抽出によって“被害”に関する画像以外ではある一定の特徴抽出が可能であった。しかし、ある一定の投稿がノイズとして混合し、それらが大きく数値に影響したケースも確認できた。算出したTF-IDFを基に文書間のコサイン類似度を算出した結果、誤分類を起こしやすいクラス間が明らかになり、同一の投稿を複数クラスへ分類する可能性が示唆された。4.4章で行った、災害時に必要になるであろう“被害”に分類された画像に対して元投稿のテキストを参照することで、画像と投稿の詳細な状況などを確認するといった質的調査を行った結果、“被害”画像をさらに細分化して分類できることが確認できた。一方で、“被害”を細分化したクラスの一つである“ユーザが被害を報告する投稿”には特徴的な単語は見られなかった。被災者自身が投稿した画像をピンポイントで抽出することは困難である。そのため、被災者自身の被害投稿をピンポイントで抽出するのではなく、無関係な投稿や情報をノイズとして扱い除去することで、災害時における迅速な情報収集を目指す必要がある。

4.6 本章のまとめ

本章では、実際の災害データ（令和2年7月豪雨）を用いた定性的・定量的な調査により、SNS上の災害情報におけるノイズ構造と、そこから重要情報を抽出する際の困難性について議論した。4章で行った調査の結果、災害に関連するキーワード（「救助」「避難」）で収集された投稿のうち、実際に被災者が撮影した被害状況を含む投稿は全体のわずか0.564%に過ぎないことが明らかとなった。また、テキスト特徴（TF-IDF）を用いた分析からは、ゲームや特定のノイズ情報には固有の単語特徴が見られた。一方で、肝心の「ユーザが撮影した被害報告」には際立った単語特徴が見られなかった。このことから、投稿テキスト

情報に頼った自動抽出や、人手による情報の全確認といったアプローチは、災害時の膨大な情報量において非効率的であることを示唆している。そのため、特徴が顕著に現れやすいゲーム画像や転載画像といったノイズ情報を機械的に峻別し、それらを群として除外する事で、SNSの情報空間を縮めるアプローチが有効であると考えられる。次章では、本章で得られた知見に基づき、画像とテキストのマルチモーダル特徴を用いた機械学習手法を適用し、これらのノイズ情報が機械的に分離可能であるかを検証する。

5 計算機によるノイズ分離の可能性

前章の調査により、災害時の SNS 上には被災者が発信する有用な情報（被害報告や救助要請）は極めて稀少であり、その大部分が災害とは無関係なノイズ情報で占められていることが定量的に示された。また、これらの投稿はテキスト情報に顕著な特徴が見られないため、キーワード検索のみによる抽出には限界があることも確認された。SNS 投稿における本文テキストは必ずしも添付画像の内容を直接的に説明するものではないことも指摘されている [1]。このことから、投稿本文の言語表現に依存せず、現地の被害状況を直感的に把握可能であり、情報の信頼性評価において極めて重要な役割を果たす画像情報そのものに着目し、画像から得られる情報のみを用いたクラスタリングがどの程度機能するか検証する必要がある。そこで本章では、3 章で提案した「ノイズ情報の段階的除外」を実現するための技術的基盤として、画像と自然言語を統合的に扱うマルチモーダルモデルである CLIP (Contrastive Language-Image Pre-training) を用いた情報分離の可能性について検証する。SNS に投稿された画像を CLIP により image to text 処理を行い、それらを特徴量とした教師なしクラスタリングを行うことで、ノイズ情報群（ゲーム、アニメ、日常投稿等）が分離可能であるかを明らかにする。もし、これらのノイズ情報が特定のクラスに偏在することが確認できれば、それらを群として除外していくことで、人が確認すべき情報空間を効率的に縮めることが可能となる。本章では、その機械的な分離可能性と、現時点での限界について議論する。

5.1 Contrastive Language-Image Pre-training (CLIP)

本研究では、SNS 上の画像に含まれる多様なノイズ情報を機械的に抽出・分離するために、OpenAI によって開発されたマルチモーダルモデルである Contrastive Language-Image Pre-training (CLIP) を採用する。本節では、CLIP の学習構造、特徴、および本研究における採用の妥当性について論じる。

CLIP の予測プロセスは、画像エンコーダ (Image Encoder) とテキストエンコーダ (Text Encoder) の二つのモジュールから構成され、両者の正しい対応関係を予測する。この構成を図 5.1 に示す。CLIP は、画像とテキストのペアを用いて両者の対応関係を学習したモデルである。インターネット上に流通している 4 億の画像とテキストのペアデータセット (WebImageText; WIT) を用いて学習されている。学習手法は Contrastive Learning (対照学習) が用いられ、 N 個の画像と、ペアになっている N 個のテキストのコサイン類似度を最大化し、ペアではない $(N \times N - N)$ 個のテキストとの類似度を最小化するように学習している。その後、これらの類似度に対して、画像からテキスト、テキストから画像の交差エントロピー誤差を用いて最適化している。各エンコーダから得られた表現から、Linear Projection (線形射影) で埋め込み空間へと射影している。これにより、意味的に近い画像とテキストのベクトル同士は距離が近く、意味的に異なるものは距離が遠くなるように配置される。

本研究の目的である災害情報トリアージにおいて、CLIP を採用した理由は主に以下の 3 点である。第一に、CLIP 最大の特徴である、zero-shot transfer 能力が高い点が挙げられる。これにより、特定のタスクに応じたファインチューニングを行わなくても高い分類性能が

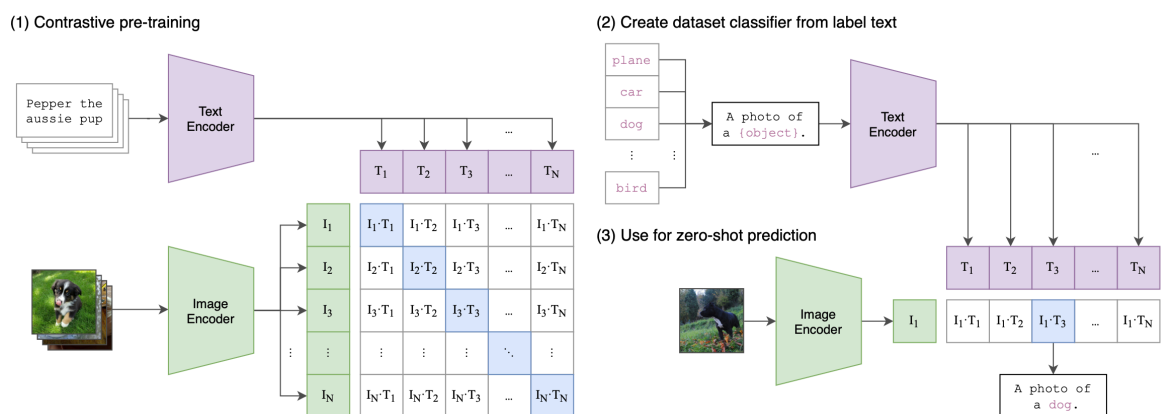


図 5.1: CLIP における予測モデル ([45] より引用)

得られることが期待できる。この特徴は、時間的制約の大きい災害時において、新たに学習データを用意する必要は必ずしもない。本研究で扱う画像データには分析・検証のために人手で分類した正解データが付与されているが、CLIP を用いる場合、別の災害時のデータを扱う際に新たに正解データを付与するプロセスを省くことができるという点からも有効だと考えられる。

第二に、多様なノイズ情報の分類可能性が高い点が挙げられる。災害時の SNS には、ゲームや日常の風景やペットなど、極めて多種多様な画像が投稿されることが、4 章の調査でも明らかとなっている。これらを ImageNet 等で学習した分類モデルを利用した場合、学習データに含まれないエンターテインメント的な画像や、Web 特有の画像に対して適切な特徴抽出ができない可能性がある。CLIP の学習データは多種多様なインターネット上のデータを利用しており、これらの分類可能性が高い。また、2 章で述べた Mandal らの研究 [32] の結果からも、災害時の SNS 情報に対して頑健性が高いとされている。

第三に、意味的特徴にもとづくクラスタリングへの親和性がある点が挙げられる。CLIP で用いられた画像とテキストの埋め込み空間は、画像の内容をテキストの意味的解釈をして構成されている。そのため、CLIP を用いることで、視覚的には異なる画像内の物体も、意味的に共通していた場合に、同一のクラスタとして集中させることが期待できる。

以上から本章では、これらの CLIP が持つ特性を活かし、SNS 上に投稿された画像を言語化することで特徴量化し、得られた特徴量からクラスタリングのアプローチを通じて、ノイズ情報の機械的な分離可能性を検証する。

5.2 対象の災害とデータ

本章では、令和 2 年 7 月 3 日から 31 日にかけて発生した「令和 2 年 7 月豪雨」を分析対象とする。第 4.2 節で準備した同様のデータを扱うものとする。ただし、前回からの変更点として、人手作業の誤分類を排除するため、4 章での分類担当者とは別の 12 名の大学生に対して、分類クラス内の画像が適切なクラスに割り振られているかの確認を依頼した。これらの分類結果の一部を表 5.1 に記す。本研究では表 5.1 で示した合計 32,571 枚の画像を

表 5.1: 人手による画像分類クラス抜粋

分類クラス	画像数
SNS	244
Twitter	654
グッズ	222
ゲーム	8539
二次元画像	2515
デバイス緊急通知	1932
救助	628
テレビ	333
TV ニュース	504
逆L字	525
Web ニュース	653
新聞	491
地図	584
天気図	1104
防災グッズ	847
動物	3674
植物	1809
本	304
飲食物	1774
風景	2868
ユーザ撮影被害	2367

対象に分析を行う。

5.3 CLIP による言語化と分類

テキスト情報と画像情報の複数モダリティを扱うことができ、自然言語特徴を用いた画像分類が可能になる CLIP[45](ver.1.0.)を用いて、分類した画像の image to text 処理を行う。ある画像を入力とし、その画像に写っている可能性の高いものを単語で出力する。その際に出力される単語は New General Service List¹内にある単語とし、上位 10 件を抽出した。その CLIP 結果と、その単語の出現数を表 5.2 に示す。その後、CLIP で算出した確率値を全て 1 に変換し正規化することでバイナリベクトルとして扱うこととした。これは、CLIP で算出した確率値の分散を排除したうえで分析を行うためである。

¹New General Service List Project<https://www.newgeneralservicelist.com/> (2025/11/29 確認)。

表 5.2: 分類クラス別 CLIP 結果 出現数上位 5 単語抜粋

SNS	出現数	%	Twitter	出現数	%	グッズ	出現数	%	ゲーム	出現数	%	緊急通知	出現数	%	テレビ	出現数	%
translate	219	4.632	translate	592	6.254	toy	108	2.078	tournament	5799	2.527	translate	1843	3.395	episode	141	1.814
remind	88	1.808	account	404	3.306	collection	35	5.157	menu	4889	5.809	text	1111	1.215	broadcast	120	4.014
text	79	1.265	mention	355	1.580	shelf	27	2.156	rank	4463	4.473	application	1016	1.200	video	114	1.616
invite	67	2.134	report	257	1.316	bag	26	6.564	carry	3772	2.057	announcement	974	1.537	movie	98	3.261
conversation	63	5.773	reply	206	0.848	display	24	2.448	game	3640	1.205	report	922	1.098	translate	75	1.703
TV ニュース	出現数	%	WEB ニュース	出現数	%	飲食物	出現数	%	逆L字	出現数	%	植物	出現数	%	新聞	出現数	%
broadcast	361	3.470	translate	404	2.330	food	1172	3.411	broadcast	448	3.943	plant	1313	4.005	newspaper	481	10.429
situation	229	2.101	report	324	1.576	lunch	1031	5.014	situation	318	2.378	flower	933	6.897	article	479	13.066
news	208	2.462	article	306	4.316	meal	1017	8.238	news	296	2.475	spring	813	2.226	weekly	447	1.903
incident	199	2.351	news	182	1.996	dinner	889	4.242	flood	212	14.580	grow	764	5.029	publication	354	1.560
reporter	159	3.774	situation	169	1.558	eat	776	1.662	incident	209	2.486	bunch	724	1.709	edition	346	1.339
地図	出現数	%	天気図	出現数	%	動物	出現数	%	二次元画像	出現数	%	被害	出現数	%	風景	出現数	%
map	525	4.467	forecast	1064	18.671	cat	1531	2.496	translate	849	1.815	flood	2066	24.531	weather	529	4.585
nearby	401	4.656	weather	1051	25.621	lucky	1466	1.803	episode	472	2.079	situation	1254	1.099	sky	422	4.901
area	371	1.924	rain	952	7.168	spot	1222	3.223	story	380	1.049	severe	1153	1.173	rain	421	3.909
route	361	6.903	situation	942	1.162	sit	1098	1.599	trigger	338	1.149	stream	1005	4.318	outside	321	1.097
location	343	4.087	map	790	1.731	dog	855	2.790	index	321	1.695	river	977	3.349	wet	287	2.233
防災グッズ	出現数	%	本	出現数	%	救助	出現数	%									
supply	503	3.059	translate	180	2.310	operation	441	5.869									
commercial	319	3.374	publication	121	2.175	rescue	422	15.023									
product	316	1.824	cover	111	4.051	incident	344	2.370									
package	312	2.931	novel	107	1.587	army	295	3.099									
consumer	285	2.496	book	104	3.986	military	287	2.807									

次に、これらの結果からk-平均法を用いてクラスタリングを行う。その際、元々の特徴量からより低次元の特徴で表現するPCA分析を用いて、累積寄与率が0.9を超えるように次元削除を行い、エルボー法とシルエットスコアを用いて最適なクラスタ数を算出した。累積寄与率はその値が大きいほど、より情報の損失を防いだ分類が期待できる。PCA分析の結果、次元数は751となり、最適クラスタ数は10となった。各クラスタに割り振られた画像数を表5.3に示す。また、PCAにより特徴抽出されたクラスタリング結果を視覚的に理解するための散布図を図5.2に示す。各点は1画像データを表し、クラスタごとに色分けされている。クラスタ3や9は第1主成分（横軸）に広がっており、PCA Component1が分散の主要な要因であると考えられる。クラスタ5や2は第2主成分（縦軸）に広がっており、PCA Component2がクラスタ内のばらつきを捉えていることが確認できる。

5.4 考察

5.4.1 人手による画像分類について

本研究では、大学生4名による人手分類を実施した後、分類クラスの統合・整理を行うことで分類を行った。これは、SNSにはデマや不確定な情報が出回る点・人命救助の観点から機械処理による取りこぼしを防ぐ必要がある点を鑑み、最終的な情報の確認には人手が必須条件であるためである。その結果、直接災害には関係のないエンターテインメント関連の情報も多く混在することも確認した。これは、ゲームのコンテンツとして“救助”を行ったり、ゲーム内の敵から“避難”するといった行為が見られるためだと考えられる。また、一部の画像では明確に区分することが困難であることも明らかになった。特に、「被害」クラスに関しては、テレビニュース、ネットニュース、ユーザ撮影、ライブカメラといった別クラスの画像が多く混在した。つまり、被災者が自ら撮影した画像とマスメディアが撮影した画像、さらにテレビニュースやネットニュースのスクリーンショット等が被害を示す画像だとして判断される場合があることを示している。どれも被害を示している情報ではあるものの、実際にユーザが置かれている状況やマスメディア等で報道されていない被害の把握のためには、ノイズ情報となる可能性が高い。特に、被災者自らが撮影したと思われる被害状況等を優先的に確認したい場合には、メディアによる二次的な報道画像や過去の報道画像をどのように排除するかが課題となる。このような混在は、災害時の情報の信頼性を評価する上で重要な要素となるため、例えば出典を識別するためのメタデータの作成や、より詳細な分類基準の作成、ノイズ情報の明確化および排除が必要になる。別メディアからの情報を適切に切り分けることは、情報の信頼性の低下を防ぐだけでなく、ソーシャルメディアとマスメディアの情報を相補的に利用者に届けることを可能にする。

本研究では、人手分類の誤りを排除するために別途12名の大学生による検証を実施したが、個々の分類担当者の主観的判断による影響は完全には排除できていない。今後、特定のクラス（例：被害・救助など）のみを精度高く選別する際には、複数名のアノテーション結果を統合するといったアンサンブル手法を活用することで、分類の精度向上が期待できる。また、被害の深刻度や救助の緊急性といった定量化が必要な局面では、人手による判断基準の明確化や客観的な評価指標の策定が不可欠である。

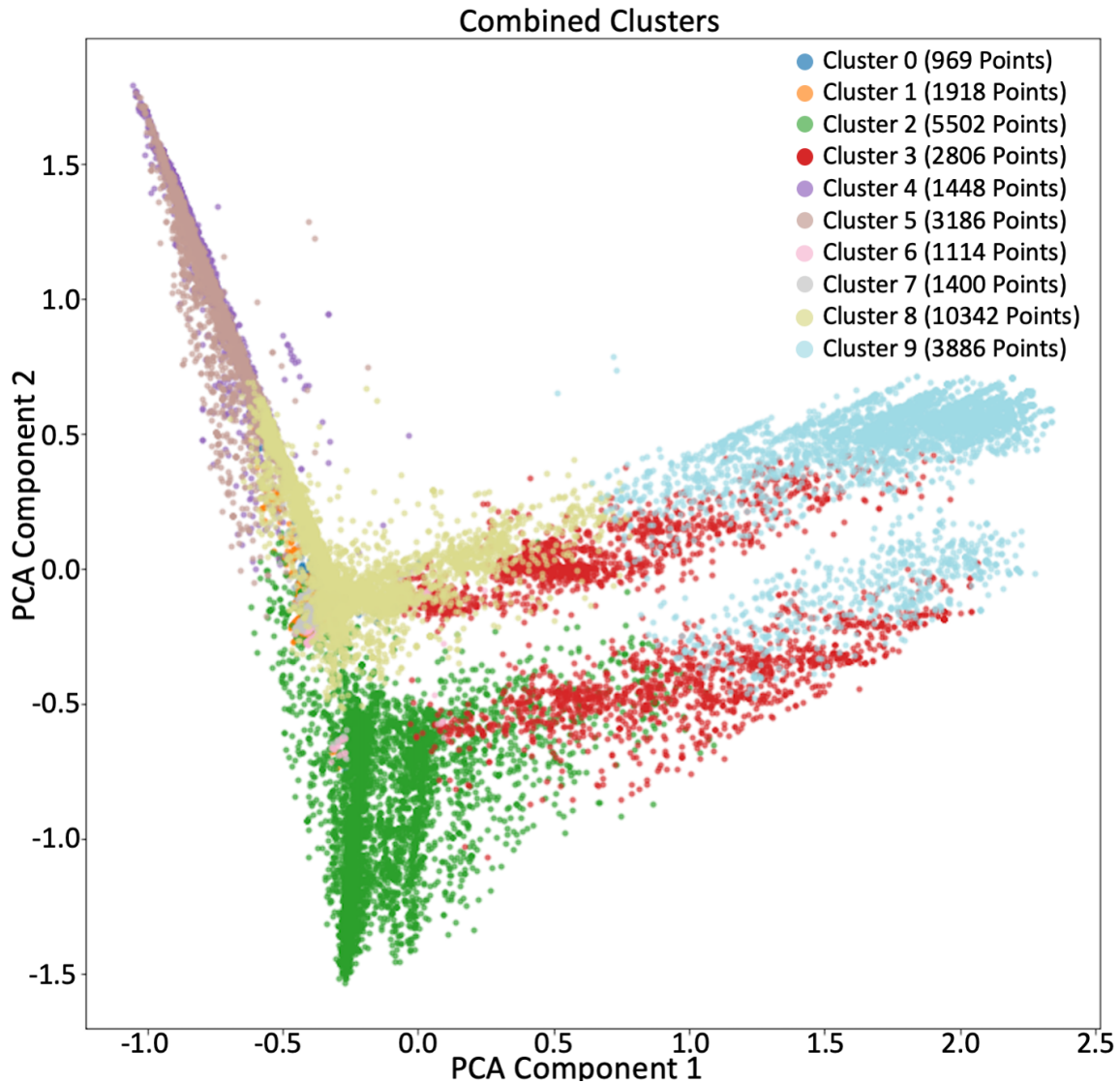


図 5.2: CLIP 結果を用いたクラスタリング散布図

5.4.2 CLIP を用いた画像の text 化について

本研究では、CLIP を用いて画像に写っている可能性の高い単語を抽出し、上位 10 件の単語を用いた。結果を概観すると、一般的な単語（例：「translate」、「report」、「text」）が頻出しており、画像の内容を具体的に反映していない可能性がある。このことは、画像内に文字情報が入り込むことが多いこと、CLIP のモデルが災害時の画像に特化した学習を行っていないこと等により、そういった汎用的な単語を出力しやすいと考えられる。

“SNS” や “Twitter” といった、ソーシャルメディアのスクリーンショットが多く含まれている分類クラスでは、「translate」、「text」といったテキスト情報を示す単語が多く含まれ、「account」等のソーシャルメディアに起因する語が確認できた。ほぼテキスト情報のみで構成されている“緊急通知”では「translate」、「text」、「announcement」、「report」から確認できるように、テキスト情報の画像であることを示している。

表 5.3: CLIP 結果を用いたクラスタリング結果

	Cl:0	Cl:1	Cl:2	Cl:3	Cl:4	Cl:5	Cl:6	Cl:7	Cl:8	Cl:9	合計
グッズ	6	1	7	3	0	0	0	1	204	0	222
地図	0	0	25	0	14	15	0	0	530	0	584
SNS	0	0	205	1	2	21	1	0	14	0	244
飲食物	0	30	63	0	0	0	1066	0	615	0	1774
新聞	0	0	417	0	0	12	0	0	62	0	491
防犯グッズ	1	0	129	0	0	1	1	0	715	0	847
本	0	0	185	1	2	0	0	1	115	0	304
救助	0	0	41	0	0	299	0	0	288	0	628
TV	0	0	71	6	3	7	1	0	244	1	333
天気図	0	0	1	0	1064	1	0	0	38	0	1104
Twitter	0	0	570	9	12	10	0	0	53	0	654
被害	0	7	5	0	23	2042	0	0	290	0	2367
植物	3	1525	8	0	1	15	6	1	250	0	1809
TV ニュース	0	0	115	4	74	146	0	0	163	2	504
動物	929	30	33	0	3	26	33	1373	1247	0	3674
逆L字	1	0	78	3	75	211	0	1	153	3	525
Web ニュース	0	0	366	1	50	93	1	0	139	3	653
風景	5	315	59	3	100	274	3	4	2105	0	2868
二次元画像	18	8	832	18	22	8	2	17	1588	2	2515
緊急通知	0	0	1858	6	2	2	0	0	64	0	1932
ゲーム	6	2	434	2751	3	1	0	2	1465	3875	8539
合計	969	1918	5502	2806	1450	3184	1114	1400	10342	3886	32571

“ゲーム”, “グッズ”, “二次元画像” といった, エンターテインメントに関する分類クラスでは, コンテンツ特有の単語が上位に出現している. 最も画像数の多かった “ゲーム” 分類クラスの上位単語には「tournament」, 「menu」, 「rank」, 「game」といった, ビデオゲームに関連する画像であることが特定可能な単語が多くを占めている. 特に「tournament」は全 “ゲーム” 画像数 8,539 枚に対して 5,799 枚 (67.9%), 「menu」は 4,889 枚 (57.2%), 「game」は 3,640 枚 (42.4%) で出現していることから, 比較的特定が容易であると考えられる. ただし, これらの単語が特定のゲームでのみ出現する可能性もあるため, あらゆるジャンルのゲーム画像に対して比較していく必要がある. “グッズ” についても「toy」や「collection」, 「display」といった, グッズを飾っていることを示している単語が確認できた. 漫画やアニメキャラクターのイラストや, アニメ映像のスクリーンショット・漫画の切り抜き画像が多く含まれている “二次元画像” では, 「translate」が最上位なものの, 「episode」, 「story」といった, エンターテインメントや映像作品を示す単語が確認できた.

“テレビ”, “TV ニュース”, “逆L字” といった, テレビ番組やニュースをカメラで撮影した画像が多く含まれている分類クラスでは, 「broadcast」や「episode」, 「news」等の, テレビ映像やニュースに関係する語が確認できた. ただし, 災害被害発生中に放送映像の外側

に別途付けられるL字型の帯が映り込んでいる“逆L字”分類クラスに関しては、逆L字自体を示す語は確認できなかった。また、“新聞”、“Webニュース”、“TVニュース”といった、ニュースに関する分類クラスでは、「newspaper」、「article」、「news」、「report」といったマスメディアの発信であることを示す単語が多く確認できた。“地図”、“天気図”といったGoogle Map²やApple社の天気アプリ³、Yahoo!天気アプリ⁴などのスクリーンショット画像が多く含まれている分類クラスでは、「map」が共通して多く出現しているものの、“天気図”では「forecast」、「weather」が最上位と差別化されており、適切に分離できていることが確認できた。

特に、“被害”分類クラスの画像では、「flood」、「situation」といった単語が最上位に出力されており、具体的な被害の様態を示す単語が抽出されていた。「stream」や「river」といった河川を示す単語も上位に確認できることから、豪雨災害時の増水状況や洪水状況を示している可能性が高い。しかし、“被害”分類クラスと混交しやすいと考えられる“風景”分類クラスにおいて、「weather」や「sky」、「rain」といった単語が上位ではあったものの、第7位に「flood」、第9位に「river」、第16位に「stream」が出現したように、出現した単語自体は“被害”と類似していることが確認できた。つまり、画像のみを用いた分類では両者は混交しやすく、“被害”と“風景”画像の分離は困難であった可能性がある。

CLIPでの単語抽出は、抽出する単語が少なければ、より特徴的な単語を抽出しやすく、画像をより簡潔に表現することが可能になるが、クラスごとの差異が小さくなりクラス判別がかえって困難になることも考えられる。一方で、抽出する単語数が多ければ、より細かな分類が可能になるが、低確率な単語が含まれることでノイズ情報が混入する確率が高くなるという特徴がある。そのため、クラスタリングを行うにあたり十分な特徴抽出が可能かつクラスタ特徴が明確になるといった、適切な単語数を決定することが、精度向上に繋がると考えられる。今後、実際の単語数の決定は、異なる単語数を抽出した上でクラスタリング精度の比較を行うことが有効と考えられる。

5.4.3 クラスタの特徴

CLIPを用いて出力した出現単語に対してバイナリベクトル化を行い、それを用いたクラスタリングの結果、クラスタ2（図5.2内の緑色点）およびクラスタ8（図5.2内の黄色点）には全ての分類画像種が含まれることが確認された。これらのクラスタには画像の多様性を原因とした、ノイズ情報が多くなっていることが考えられる。そのため、どのような単語がクラスタリングに寄与しているか確認するため、クラスタ2と8の頻出単語を抽出した。その結果を表5.4に示す。

クラスタ2の頻出単語上位には「translate」、「report」、「text」などを確認した。特に「translate」は、なんらかの日本語文字情報が含まれる場合に抽出されていた。つまり、テキスト情報を含む画像が多く分類されている可能性が高く、テキストのみの画像が多かった“デバイスの緊急通知”がクラスタ2に多く分類されていることもそれを裏付けている。一方、クラスタ8に関しては、第一位に出現した「location」は648回であり、第十位の「route」も456回と、分散が小さい傾向が見られ、特徴的な単語は確認できなかった。「location」、

²<https://www.google.co.jp/maps> (2025/11/29 確認)。

³<https://apps.apple.com/jp/app/weather/id1069513131> (2025/11/29 確認)。

⁴<https://weather.yahoo.co.jp/weather/promo/app/> (2025/11/29 確認)。

表 5.4: クラスタ 2・8 における CLIP 算出の頻出単語

クラス 2	出現数	クラス 8	出現数
translate	5323	location	648
report	1899	episode	640
text	1634	supply	608
announcement	1565	map	584
application	1110	nearby	584
invite	1075	operation	531
story	974	commercial	499
article	970	video	484
weekly	875	broadcast	471
paragraph	867	route	456

「map」, 「route」といった地理的情報に関連する単語は頻出しているが, 「location」は 52.9% が, 「map」は 89.9% が, 「route」は 79.2% が “地図” 分類クラスによるものだった. つまり, このクラスタ 8 は, “地図” に関連する単語が基準となりつつも, 他分類クラスにおいて分類基準とは外れた画像や, 地理的関連単語が出現した場合に割り振られることで, ノイズ情報が多く混交した可能性がある.

クラスタ 2 及び 8 に多種の画像が分類されたのは, クラスタリングを行う際に CLIP 出現単語をバイナリベクトルとして扱ったことに起因するとも考えられる. これは, CLIP 技術の結果に依存しすぎないための処置であったものの, 出現単語の組み合わせ次第では適切にクラスタリングが困難だったことが考えられる. これは, 5.4.2 節で上述したように, 各画像を決定するにあたり十分な出力単語数の検討が必要不可欠であると考えられる. ただし, より詳細にクラスタの解釈性を高めるためには, 各クラスタの代表的な画像を可視化した上で実際にどのような単語が出力されているか検証し, どのような単語が分類に与しているのか分析することが求められる. また, クラスタ間の類似度の計算や, 近いクラスタの統合可能性について検討することで, より直感的な解釈が可能になると考えられる.

また, “被害” 分類画像が最も多く割り振られたクラスタ 5 (図 5.2 内の灰桜色) には, 災害とは無関係な分類クラスの画像がクラスタリングされることは少なかったことが示された. しかし, “被害” の他に “救助”, “TVnews”, “逆 L 字”, “風景” といった画像が多く含まれる結果となった. これは, マスメディア等で既に報道された情報と, 一見 “被害” と見間違ふような “風景” の画像が誤ってクラスタリングされている可能性があることから, クラスタ 5 に分類された画像は人目での確認が必要不可欠な情報だと考えられる.

5.5 本章のまとめ

本章では, 災害時における SNS 情報から, ノイズ情報の計算機による分離可能性を検証するため, CLIP を用いた画像の言語化と, その特徴量を用いたクラスタリングを行った. CLIP を用いた分析の結果, “ゲーム” のようなエンターテインメント関連の画像や, “植物”,

“動物”といった災害とは無関係なノイズ画像に対しては、そのカテゴリを象徴する特定の単語である「tournament」や「character」等が高頻度で抽出されることが確認された。また、それらを用いたクラスタリングの結果、これらのノイズ情報は特定のクラスタに強く偏在する傾向が示され、機械的な分離が可能であることが実証された。一方で、“被害”画像に関しては“風景”画像との類似性が高く、画像特徴のみでは完全な分離が困難であり、混在するクラスタが存在するという課題も明らかとなった。しかし、本研究の目指す膨大なノイズ情報の除去という観点において、ゲームや日常投稿などの明確なノイズ群がクラスタとして可視化・分離できたことは極めて重要な成果である。ノイズ群を人が確認し除外するための支援情報として十分に機能すると考えられる。本章で得られた知見から、機械学習によるクラスタリングは、人間が情報を峻別する前に情報を整理する上で有効であると結論づけた。次章では、本章で確認された「ノイズのクラスタ化」を利用し、ユーザが能動的にノイズクラスタを除外しながら重要情報を探索できる、除外型トリアージインタフェースの設計と実装を行い、その有効性について検証する。

6 段階的除外型インタフェースの設計と実装

4章までの調査の結果、災害時のSNS上には被害者自身が投稿する被害情報といった重要な情報が極めて稀少であることが明らかとなっている。5章では4章の結果を踏まえて、ノイズ情報を特定のまとまり（クラス）として機械的に分類が可能であることから、ノイズ情報の分離可能性が示された。これらの知見は、膨大なSNSの情報空間から、重要な情報を直接抽出するような従来の検索的アプローチではなく、機械学習によって整理されたノイズの塊を、ユーザが能動的かつ段階的に除外していくアプローチの妥当性を示唆している。機械学習によるクラスタリングはあくまで情報整理の一環であり、最終的な情報の信頼性判断や重要情報の確保は、災害対応について適切に理解している人間が行う必要がある。そのため、本研究が目指す「人と計算機の協調による情報トリアージ」を実現するためには、計算機によるクラスタリングを人間が直感的に理解し、効率的に操作できるインタフェースの存在が不可欠となる。そこで本章では、5章で検証したCLIPとクラスタリングを用いた分析結果を基盤とし、ユーザが視覚的の手がかりをもとにノイズ情報を段階的に除外しながら、膨大な情報空間に埋もれた有用情報の探索が可能な「災害情報トリアージインタフェース」の設計と実装について述べる。また、実装したシステムを用いた被験者実験を行い、従来のSNSのように情報をリストで見えていくことを軸としたベースラインシステムとの比較を通じて、段階的除外アプローチが時間制約下における情報収集効率にどのような影響を与えるか評価する。

6.1 災害データセットとクラスタリング

本章では、対象災害のデータとそのクラスタリングについて説明する。本研究では、令和2年7月3日から31日にかけて発生した「令和2年7月豪雨」を分析対象とする。自身の被災状況を伝える際には画像・動画を含めた投稿をすることが推奨されている点から、分析対象をツイートに添付された画像・動画とした。まず、収集したツイート画像は人手で分類を行う。次に、自然言語と画像のマルチモーダルな処理が可能なCLIPを用いて画像のテキスト化を行い、k-means法を用いてクラスタリングを行う。これらのクラスタリング手法について5章と同様の方法で行う。

6.1.1 対象災害とXデータ

本章では、令和2年7月3日から31日にかけて発生した「令和2年7月豪雨」を分析対象とする。5.2節で準備した同様のデータを扱うものとする。ただし、前回からの変更点として、人手作業の誤分類を排除するため、4章での分類担当者とは別の12名の大学生に対して分類クラス内の画像・動画が適切なクラスに割り振られているかの確認した際に、適切に分類された画像に加え動画を含める。また、全ての分類クラスを含めることで、不用意な情報の埋没を避けるため、表6.1に提示したクラスと画像・動画の合計33,651枚に対してクラスタリングを行いインタフェースの構築を行う。

前章までの知見を踏まえ、これらのデータからk-平均法を用いてクラスタリングを行う。その際、元々の特徴量からより低次元の特徴で表現するPCA分析を用いて、累積寄与率が0.9を超えるように次元削除を行い、シルエットスコアを用いて最適なクラスタ数を算出し

表 6.1: インタフェース用データベース構築に使用した画像・動画件数一覧

分類クラス	画像・動画数
ゲーム	10743
動物	4084
被害	2836
風景	2817
植物	2060
緊急通知	1936
飲食物	1758
天気図	1139
防災グッズ	894
Web ニュース	803
救助	693
Twitter	675
地図	586
逆L字	543
TV ニュース	533
新聞	491
テレビ番組	459
本	333
SNS	279
合計	33651

た. PCA 分析の結果, 次元数は 659 となり, 最もシルエットスコアが高く算出された最適クラスタ数は 17 となった. 各クラスタに割り振られた画像・動画の合計数を表 6.2 に示す. また, PCA により特徴抽出されたクラスタリング結果を視覚的に理解するための散布図を図 6.1 に示す. 各点は 1 画像データを表し, クラスタごとに色分けされている. この散布図からは, 特定のクラスタの形が細長い方向が, そのクラスタ内のデータを区別する主要な要素であることが読み取れる. 例えば, クラスタ 1 や 8 は第 1 主成分 (横軸) に広がっている. これは PCA Component1 の値の違いによって生じており, PCA Component1 がデータの分散の主要な要因であることが確認できる. また, クラスタ 5 や 13 は第 2 主成分 (縦軸) に広がっている. これは, PCA1 の値の違いよりも, PCA Component2 の値の違いのほうが大きく, PCA Component2 がデータの特徴づけに利用されていることが確認できる.

6.1.2 CLIP による言語化と分類

テキスト情報と画像情報の複数モダリティを扱うことができ, 自然言語特徴を用いた画像分類が可能になる CLIP[45](ver.1.0.) を用いて, 分類した画像の image to text 処理を行う. その際, 分類した動画に関しては, moviepy(ver1.0.3) を用いて動画開始後 1 秒時点での映像をサムネイル画像として保存することで, image to text 処理を行う. ある画像を入

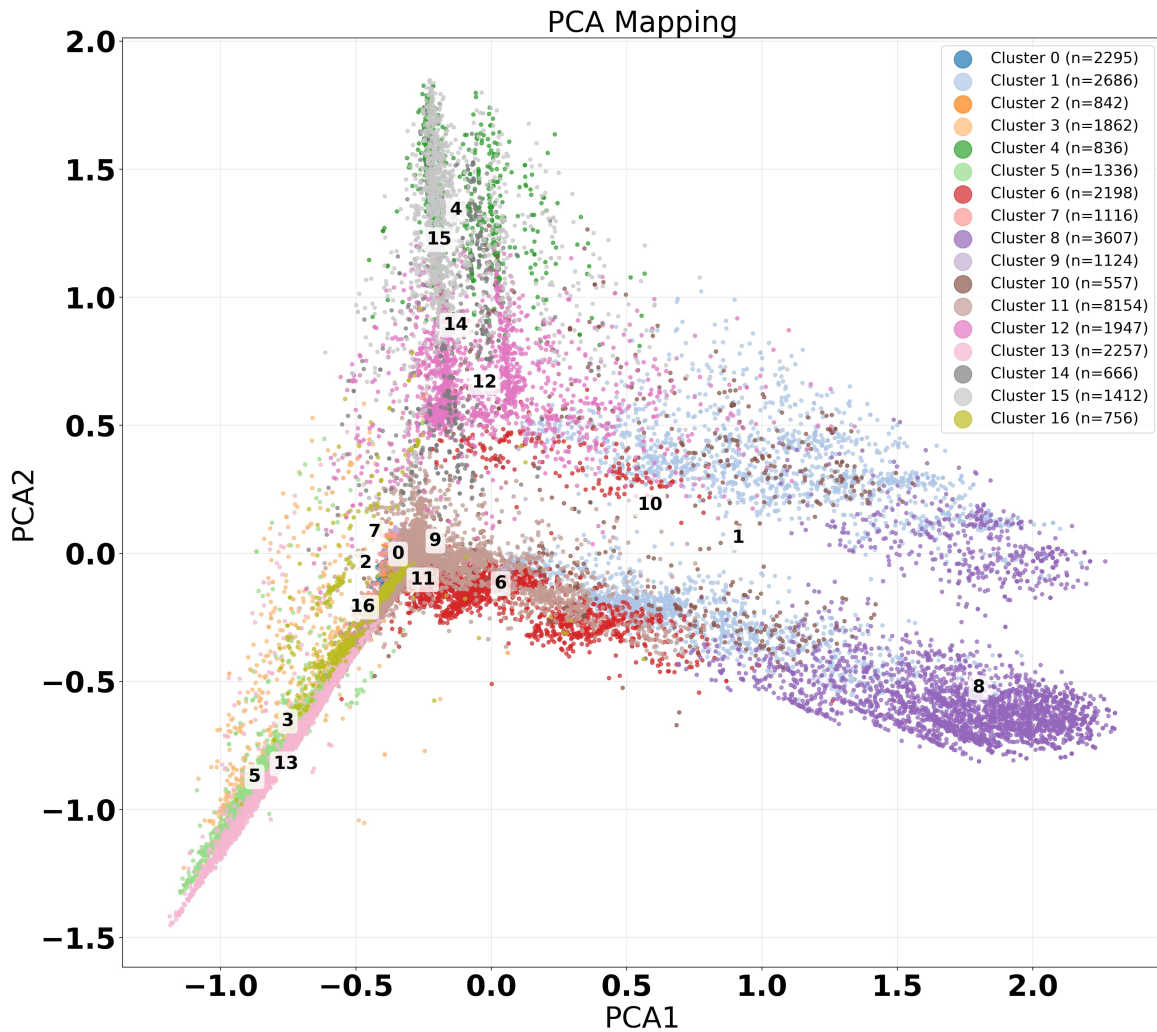


図 6.1: インタフェース用のデータベース構築に使用した画像のクラスタリング散布図

力とし、その画像に写っている可能性の高いものを単語で出力する。その際に出力される単語は New General Service List¹内にある単語とし、上位 10 件を抽出した。その後、CLIP で算出した確率値を全て 1 に変換し正規化することでバイナリベクトルとして扱うこととした。これは、CLIP で算出した確率値の分散を排除したうえで分析を行うためである。

¹<https://www.newgeneralservicelist.com/>(2024/1/22 確認)

表 6.2: インタフェース用データベース構築に利用した画像のクラスタリング結果

	Cl:0	Cl:1	Cl:2	Cl:3	Cl:4	Cl:5	Cl:6	Cl:7	Cl:8	Cl:9	Cl:10	Cl:11	Cl:12	Cl:13	Cl:14	Cl:15	Cl:16	合計
ゲーム	4	2659	0	3	14	2	2174	0	3606	0	412	1390	418	0	0	41	9	10732
動物	2279	0	35	24	1	3	4	3	0	44	0	1617	44	26	2	2	0	4084
被害	0	0	4	747	1	20	4	0	0	0	0	310	5	1740	1	0	4	2836
風景	5	3	59	42	4	89	4	106	0	4	1	2103	61	318	3	0	15	2817
植物	2	0	716	0	0	1	0	1007	0	7	0	293	7	25	1	1	0	2060
緊急通知	0	4	0	2	569	2	0	0	0	0	34	32	91	0	9	1188	5	1936
飲食物	0	0	28	0	0	0	0	0	0	1064	0	602	64	0	0	0	0	1758
天気図	0	0	0	1	0	1064	0	0	0	0	1	2	0	0	0	0	71	1139
防災グッズ	0	0	0	2	10	0	0	0	0	1	17	718	126	0	10	10	0	894
Webニュース	2	0	0	135	42	41	3	0	0	1	20	146	158	40	110	71	34	803
救助	0	0	0	483	0	0	2	0	0	0	0	179	19	7	0	1	2	693
Twitter	0	8	0	6	74	7	0	0	0	0	15	52	455	8	18	28	4	675
地図	0	0	0	2	0	5	1	0	0	0	1	40	12	1	0	0	524	586
逆L字	0	2	0	217	0	44	1	0	1	0	17	104	64	38	8	5	42	543
TVニュース	0	4	0	166	2	53	1	0	0	0	11	120	89	34	3	13	37	533
新聞	0	0	0	12	6	0	0	0	0	0	4	11	3	0	454	0	1	491
テレビ番組	0	5	0	7	1	3	4	0	0	1	10	327	89	5	4	1	2	459
本	1	0	0	2	43	0	0	0	0	0	6	88	145	0	43	3	2	333
SNS	2	1	0	11	69	2	0	0	0	2	8	20	97	15	0	48	4	279
合計	2295	2686	842	1862	836	1336	2198	1116	3607	1124	557	8154	1947	2257	666	1412	756	33651

6.2 災害情報トリアージインタフェースのデザイン指針

本インタフェース実装の目的は、上述のクラスタリング結果を提示した上で、災害に直接関連する情報を抽出することである。本研究では、災害時という人的・時間的制約の大きい状況下において、SNSを用いて災害被害情報を集約し関係部署への迅速な情報共有・要否判定を行うための自治体職員をユースケースとして設定した。3章で論じたように、投稿の必要性・真偽判断等は人による判断を行うことを前提に、救助や被害把握に直結する投稿の可視性を高めつつ、それらが確認できる候補となる投稿を取りこぼさないためには、直接投稿を抽出するのではなく不必要情報の除外過程を経る必要がある。本節では、システムが満たすべき要件および実装機能について説明する。

6.2.1 システムが満たすべき要件と実装機能

本節では、災害時にSNS上の災害被害情報の収集を支援するインタフェースの要求事項として以下の7つと、それらを満たす機能として以下の10項目を定めた。

- (1) 出力されるデータに再現性があること (図 6.4 中①)

機能 (a): リレーショナルデータベースに格納された投稿・分析結果の読み込み機能

- (2) 必要情報・不要情報の判断の支援がされること (図 6.5 中②)

機能 (b): PCA 分析結果の3次元マッピング機能

- (3) 情報要求に応じた投稿の提示が可能なこと (図 6.4 中③)

機能 (c): 投稿の検索機能

- (4) 災害被害が確認可能な情報を確保し、後から検討することが可能なこと (図 6.4 中④)

機能 (d): 投稿の表示機能

機能 (e): 投稿のピックアップ機能

機能 (f): ピックアップした投稿の保存・出力機能

機能 (g): 操作ログの出力機能

- (5) 段階的な情報の除外が可能なこと (図 6.4 中⑤)

機能 (h): 選択したクラスタに属する投稿の除外機能

- (6) ユーザが優先順位を記憶できること (図 6.4 中⑥)

機能 (i): 優先すべきクラスタのメモ機能

- (7) 表示される情報が多すぎないこと。 (図 6.4 中⑦)

災害情報トリアージ

最優先: [e.g. 3.5] 要検討: [13] 不要: [e.g. 4.6]

DBファイル: [ファイルを選択] tweets_images.db [Show SQL] Cluster: [12] [選択クラスターのみサブプリン] [除外追加] [除外解除] [全除外クリア] 除外中: 11,0,6,14,4,12 [実行]

[Picked CSV出力] [操作ログ出力]

全 29343 件 / 対象クラスター 全 / 表示 500 件 / 残 8280 件 / 対象後 15579 件

Show [10] entries Search: []

tweet_id	text	media	cluster
1278709309329690631	7人がバツ! 縄救助出来たし縄で板まで行けたよ! https://t.co/yNa1jXVtyR		10
1280780602363404288	「もう第五人格向いてないのかな...やめどきかな...」って思ってたハンター練習しに行ったらめっちゃ救助狩り決まる接待第五人格やめて。 https://t.co/5S0lc0UC7W		1
1279454298347204608	5分チェイス久しぶりにできた…。肘当て残ってなかったから脱出できなかったよ、風船救助してくれたじゅじゅありがとう…。ゴメンな…。 https://t.co/PSZVld1czG		8
1279188226914148352	#熊本 #球磨川 市房ダム この後8:30ごろ 緊急放流!! さらに氾濫地域が増える可能性大 ただちに高い所に避難して下さい!! https://t.co/kG6ZxR5rmR		3
1280609911357845504	暖かな避難所より 今朝がたは雷様のお出ましで3時から起きて居ます... 小さな子供さんには 堪りませんよね~ 3人くらい、あやして 居ます... 今は小降りですが雷神様も 移動されました 今日もうまく 命を守る行動を #愛で繋がる #雨の降り方に注意 #早めの避難 https://t.co/gjTfqGF7gr		2

Showing 1 to 10 of 500 entries Previous 1 2 3 4 5 ... 50 Next

PCA 3D マッピング可視化エリア

Cluster 1 (75)
Cluster 2 (31)
Cluster 3 (60)
Cluster 5 (34)
Cluster 7 (26)
Cluster 8 (99)
Cluster 9 (30)
Cluster 10 (11)
Cluster 13 (65)
Cluster 15 (41)
Cluster 16 (28)

Picked Tweets

Show [] entries Search: []

tweet_id	text	media	cluster
1279188226914148352	#熊本 #球磨川 市房ダム この後8:30ごろ 緊急放流!! さらに氾濫地域が増える可能性大 ただちに高い所に避難して下さい!! https://t.co/kG6ZxR5rmR		3
1280273009010302977	2年前の六時半頃の状況 一緒に避難していた仕事の関係者を31号線まで送り届けたあと、サンスターホールまで戻る所です。サンスターホールに居る人たちに今の現状を見てもらうために録った動画です。 https://t.co/UJ2pgpamFP		3

Showing 1 to 2 of 2 entries Previous 1 Next

Clear Operating Log

図 6.2: 提案システムのインターフェース (全体像): (A) システムコンソールエリア, (B) 投稿可視化エリア, (C) PCA 3D マッピング可視化エリア, (D) Pick した投稿の可視化エリア

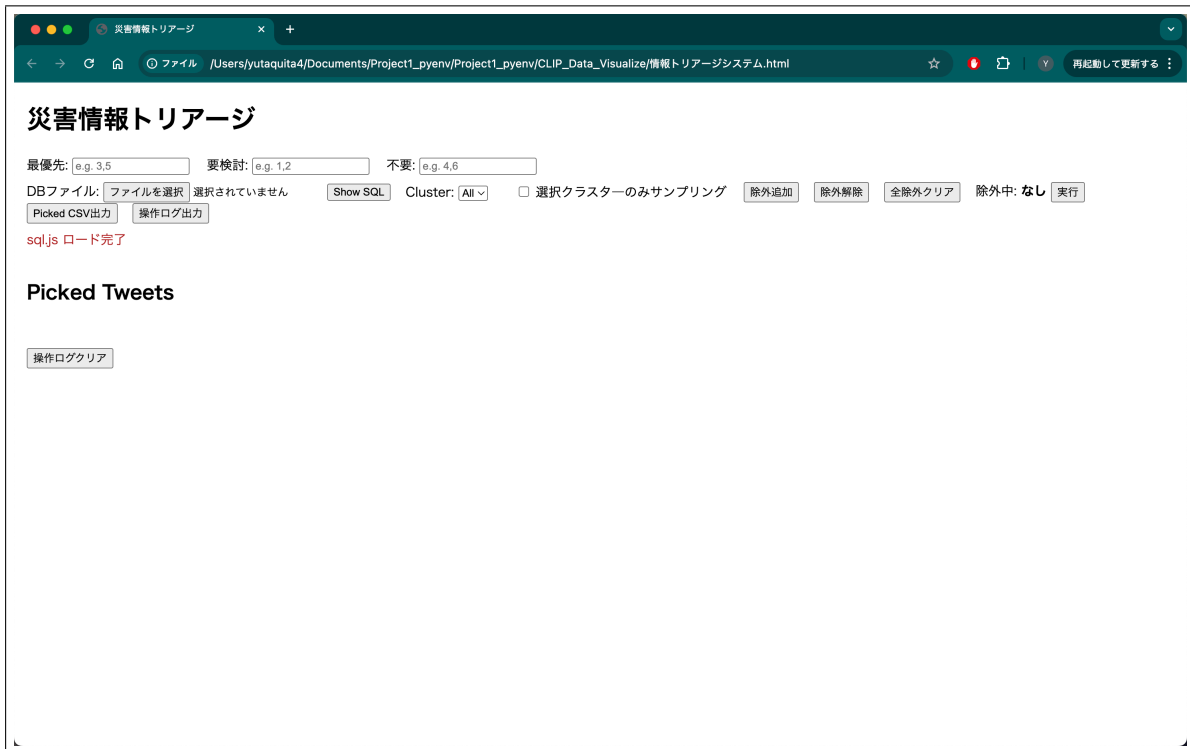


図 6.3: データベース読み込み前のインターフェース

機能 (j): 提示する投稿の表示数切り替え（最大 500 件）

機能 (k): 提示する投稿の最大サンプリング数の設置（最大 500 件）

以下ではこの項目に従い、インターフェースのデザインと情報探索手順・操作方法について説明する。インターフェースは、HTML5.0 を用いて実装した。まず、インターフェースの全体像を図 6.2 に示す。提案のシステムは、図 6.2 中の①システムコンソールエリア（画上段）、②投稿の可視化エリア（画面中央上段）、③PCA 3D マッピング可視化エリア（画面中央下段）、④Pick した投稿の可視化エリア（画面下段）の 4 つのエリアからなる。システムコンソールエリアには、機能 (a)、機能 (c)、機能 (f)、機能 (g)、機能 (h)、機能 (i)、機能 (j) の操作ボタンが配置されている。投稿の表示エリアでは、機能 (c) および機能 (d) が配置されている。PCA 3D マッピング表示エリアでは、機能 (b) の操作が可能になっている。Pick した投稿の表示エリアには、機能 (d)、機能 (f) がそれぞれ配置されている。

6.2.2 分析結果のデータベース化

6.2.1 節で提示したシステムへの要求事項 (1)、(2)、(3) に対応するため、上述の分析結果を格納するためのデータベースの作成を行う。PCA・クラスタリングの結果を保持するために、Python 標準ライブラリの sqlite3 (ver.3.43.2) を用いて二つの主テーブル (“tweets” と “images”) を構成した。“tweets” はツイート本文の原情報を格納するテーブルであり、主キー tweet_id (TEXT) と text (TEXT) からなる。データの前処理として、“tweets” テーブルの tweet_id (TEXT) キーに格納されている情報がないケース (投稿文が取得できていなかった投稿)4,308 件を除外とし、合計 29,343 の画像・動画を対象にインターフェースでの

災害情報トリアージ

最優先: [e.g. 3.5] 要検討: [13] 不要: [e.g. 4.6]

DBファイル: [ファイルを選択] tweets_images.db [Show SQL] Cluster: [12] [選択クラスターのみサンプリング] [除外追加] [除外解除] [全除外クリア] 除外中: 11,0,6,14,4,12 [実行]

Picked CSV出力 [操作ログ出力] fg [4] f g [1]

全 29343 件 / 対象クラスター 全 / 表示 500 件 / 残 8280 件 / 対象後 15579 件

Show [10] entries [7] [4] d [3] Search: [3]

tweet_id	text	media	cluster
1278709309329690631	アマガハバカ! 縄救助出来たし縄で板まで行けたよ! https://t.co/yNa1jXVtyR	[Image]	10 [3]
1280780602363404288	「もう第五人格向いてないのかな…やめどきかな…」って思ってハンター練習しに行ったらめっちゃめっちゃ救助狩り決まる接待第五人格やめて。 https://t.co/5S0lc0UC7W	[Image]	1
1279454298347204608	5分チェイス久しぶりにできた…。 肘当て残ってなかったから脱出できなかったよ、風船救助してくれたじゅじゅありがと…。 ゴメン…。 https://t.co/PSZVld1czG	[Image]	8
1279188226914148352	#熊本 #球磨川 市房ダム この後8:30ごろ 緊急放流!! さらに氾濫地域が増える可能性大 ただちに高い所に避難して下さい!! https://t.co/kG6ZxR5rmR	[Image]	3
1280609911357845504	暖かな避難所より 今朝がたは雷様のお出ましで3時から起きて居ます… 小さな子供さんには堪りませんよね〜3人くらい、あやして居ます… 今は小降りですが雷神様も移動されました! 今日もうまくも宜しくです! #命を守る行動を #要で繋がる #雨の降り方に注意 #早めの避難 https://t.co/gjTfqGF7gr	[Image]	2

Showing 1 to 10 of 500 entries Previous [1] 2 3 4 5 … 50 Next

図 6.4: システムコンソールエリアと投稿の可視化エリア

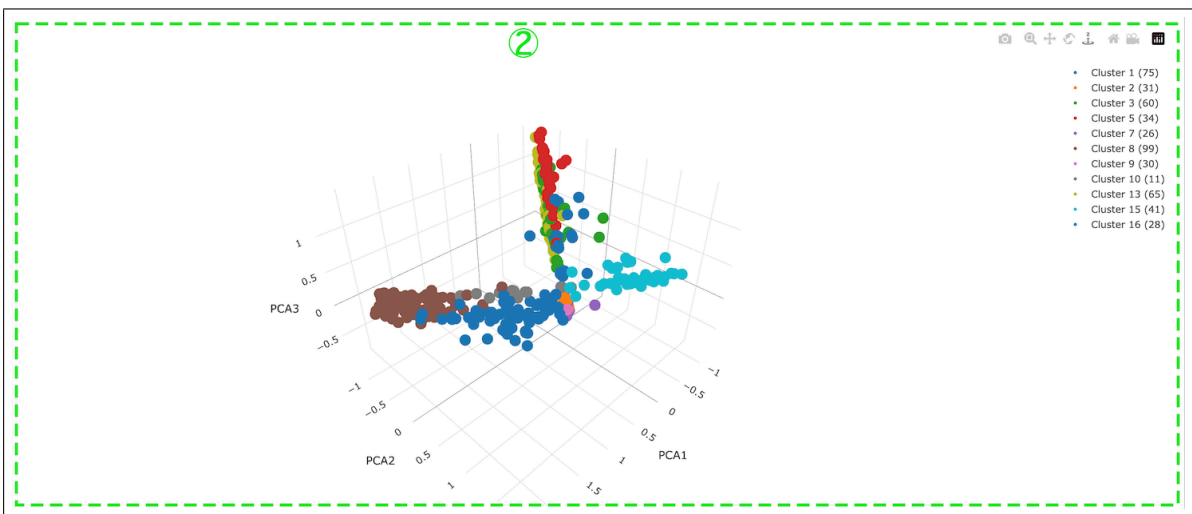


図 6.5: PCA 3D マッピングエリア

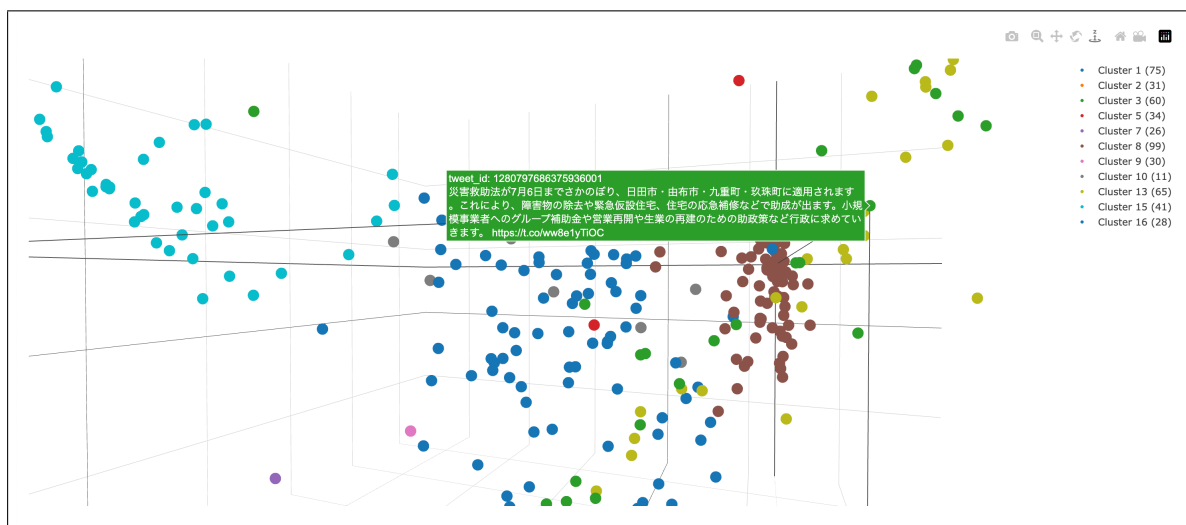


図 6.6: PCA 3D マッピングエリア: ズーム後と投稿詳細の表示

提示を行う。“images”は各メディア単位（画像または動画から抽出した代表フレーム画像）に対する分析結果を保持するテーブルで、主キーid（INTEGER, AUTOINCREMENT）のほか、tweet_id（TEXT）を外部キーとして“tweets”テーブルのtweet_idと連結し、ファイル名 image_file（TEXT）、PCA後の座標 pca1・pca2・pca3（REAL）、K-meansによるクラスタラベル cluster（INTEGER）、および手動分類カテゴリ manual_class（TEXT）を格納する。manual_classは手動で分類した結果に基づくラベルである。なお、このmanual_classは分析・検証に用いるためのもので、インタフェース上での可視化は行わない。なお、PCA分析の結果は、3次元空間に可視化するために利用する、分類に最も寄与率の高いpca1・pca2・pca3のみを格納した。この構造により、要求事項（1）“images”と“tweets”によりメディアと原文脈の対応付けが可能になり、出力されるデータに整合性が認められる、要求事項（2）必要に応じてtweet_id・text・clusterに索引を付与することで検索・集計性能を高められること、要求事項（3）pca値を用いた散布図描画等の可視化が直接行えること、の三点が実現される。

6.2.3 インタフェースデザインと探索手順

6.2.1節で提示したシステムへの要求事項に対応した機能と、実インタフェース上での提示について説明する。

まず、上述のリレーショナルデータベースの作成により、要求事項（1）（2）（3）に対応した。要求事項（1）出力されるデータに整合性があること、に対して機能（a）（データベースの読み込み機能）を実装した（図6.4中①）。データベースが読み込まれる前の画面を図6.3に示す。画面上左部の“ファイルを選択”ボタンから作成したデータベースファイルを読み込み、画面上右部の“実行”ボタンを押すことでデータが読み込まれる。データベースを読み込んだ後の画面の全体図は図6.2の通りである。一度実行ボタンを押すと、機能（k）と連動し、データベース内の全29,343件から500件の投稿がランダムにサンプリングされる。もう一度実行ボタンを押すことで、一度目の500件とは別の500件の投稿がランダム

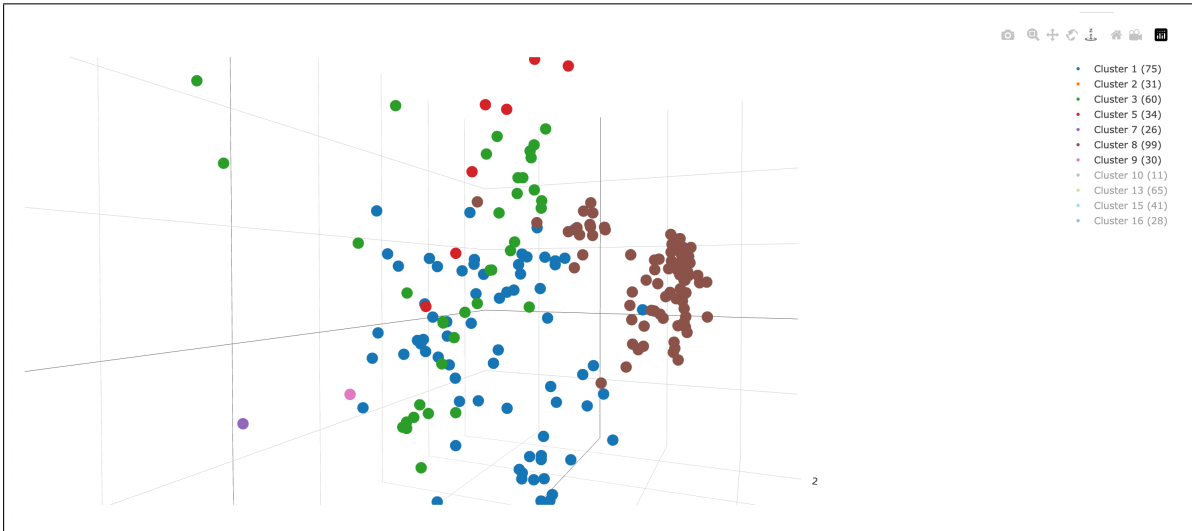


図 6.7: PCA 3D マッピングエリア: 特定クラスタを非表示にした状態

Picked Tweets 4f

Show entries Search:

tweet_id	text	media	cluster
1279188226914148352	#熊本 #球磨川 市房ダム この後8:30ごろ 緊急放流!! さらに氾濫地域が増える可能性大 ただちに高い所に避難して下さい!! https://t.co/kG6ZxR5rmR		3
1280273009010302977	2年前の六時半頃の状況一緒に避難していた仕事の関係者を31号線まで送り届けたあと、サンスターホールまで戻る所です。サンスターホールに居る人たちに今の現状を見てもらうために録った動画です。 https://t.co/UJ2pgpamFP		3

Showing 1 to 2 of 2 entries Previous Next

図 6.8: Pick した投稿の可視化エリア

サンプリングされる。なお、一度読み込んだ投稿は、再度実行ボタンを押しても表示されることはないが、機能 (h) を用いて情報を除外することで情報提示条件が変更された場合には、再表示される。

要求事項 (2) 必要情報・不要情報の判断の支援がされること、に対して機能 (b) (PCA 分析結果の 3D マッピング機能) を実装した (図 6.5 中②)。この 3D マッピング機能の実装は、Plotly(ver.2.24.1)²を用いて表示している。クラスタリングの結果を色分けして表示し、各クラスタの表示/非表示が切り替えられる。3D マッピング上でスクロールを行うことで、マッピング結果のズーム/ズームアウト表示され、3D マッピングにプロットされているドット上にマウスオーバーすることで投稿の詳細が確認できる。その変化を図 6.6 に示す。この投稿の詳細が表示されている状態で、マウスクリックを行うと、tweet_id がクリップボード上にコピーされる。その投稿の詳細や投稿されている画像を確認したい場合は、その tweet_id を後述の機能 (c) で確認可能である。画面右側では、機能 (a) で現在読み込まれているクラスタおよびその投稿数が確認できる。例えば、Cluster1(75) という表記は、Cluster

²Plotly <https://plotly.com/python/> (2025/11/25 確認)。

番号1から75件の投稿が読み込まれていることを示している。また、この Cluster1(75)の表記をクリックすることで、非表示化することも可能である。いくつかのクラスターを非表示にした状態を図6.7に示す。この機能(b)の実装により、ユーザは3D空間に配置された投稿・クラスターの距離などを参考にして類似度の高い情報の探索が可能となっている。例えば、ノイズ情報が多く含まれているクラスターを一つ発見した場合、そのクラスターと空間内の距離が近いクラスターや投稿はノイズ情報である可能性が高い。同様に、被害情報が多く含まれているクラスターを一つ発見した場合、そのクラスターと空間内の距離が近いクラスターの投稿には、被害情報が含まれている可能性が高い。これにより、連鎖的なノイズ情報や被害情報の発見が可能となっている。

要求事項(3) 投稿の検索機能、に対して機能(c) (投稿の検索機能)を実装した(図6.4中③)。画面中央部 tweet_id,text,cluster の下部にあるテキストボックスへの入力での投稿の検索を可能にした。各々のテキストボックスは、それぞれ対応しているキー内での検索となる。例えば、tweet_id の直下に設置してあるテキストボックスからは、tweet_id の検索のみ可能である。各投稿の tweet_id は、機能(b)を用いることで取得可能である。また、画面上右部の Search: 横のテキストボックスからの検索も可能となっている。このテキストボックスからの検索は、全てのキーから横断的に検索が行われる。この機能(c)は、ユーザが探索的にノイズ情報や被害情報の獲得を可能にしている。

要求事項(4) 災害被害が確認可能な情報を確保し、後から検討することが可能なこと、に対して、機能(d) (投稿の表示機能)・機能(e) (投稿のピックアップ機能)・機能(f) (ピックアップした投稿の保存・出力機能)・機能(g) (操作ログの出力機能)を実装した(図6.4中④, および図6.8中④f)。機能(d)では、データベースに格納されている投稿を表示した。投稿の表示は、1件の画像または動画につき1フィールドでの提示とした。Xの投稿は最大4枚までの画像・動画を添付することが可能なため、データベースが格納している tweet_id は最大4件までの重複が存在する。一度のサンプリング(500件)中に同一の tweet_id が確認できた場合は、連続して提示することとした。機能(e)では、tweet_id の上をクリックすることで投稿のピックアップを可能にした。ピックアップ状態にある投稿の tweet_id をもう一度クリックすることでピックアップ状態を解除することも可能にした。同一の tweet_id が確認された場合は、自動的にピックアップ状態とした。図6.4中④eで提示したように、ピックアップ状態となった投稿は tweet_id の箇所が黄色くハイライトされることで、ピックアップされたかどうかの視覚的な確認が可能となっている。機能(f)では、ユーザがピックアップした投稿を、画面最下部にリストで提示している。また、画面上左部の“Picked CSV出力”ボタンから、ピックアップした投稿のリストをCSV形式で保存が可能になっている。機能(g)では、画面上左部の“操作ログ出力”ボタンを押すことで、ユーザの操作ログを json 形式で出力することが可能になっている。操作ログは、何時何分に、機能(h)をどのクラスターに対して実行したか、機能(e)をどの投稿に対して実行したか、機能(i)のテキストボックス内の情報が更新されたか記録される。これにより、ユーザがどのクラスターを除外し、どの投稿をピックアップしたのかを時系列で把握し、後に検証することを可能にする。機能(f)および機能(g)は複数人での情報峻別を想定し、他ユーザとの情報共有を容易にするための機能の一つとして実装した。

要求事項(5) 段階的な情報の除外が可能なこと、に対して機能(h) (選択したクラスター

に属する投稿の除外機能)を実装した(図6.4中⑤)。まず、画面上中央部の“Cluster:”タブからクラスタ番号を選択する。次に、“除外追加”ボタンを押すことで、“除外中:”のエリアにクラスタ番号が追加される。最後に、“実行”ボタンを押すことで、除外選択したクラスタに属する投稿が除外されて、投稿が表示される。除外追加は一度に複数のクラスタを選択することも可能である。また、“Cluster:”タブからクラスタを選択している状態で、“除外解除”を押すことで、除外の解除が可能であり、“全除外クリア”ボタンを押すことで、除外選択しているすべてのクラスタを解除することも可能となっている。この機能(h)により、ユーザが判断した不必要情報が多く含まれているクラスタの投稿の段階的排除が可能になる。なお、ここで除外したクラスタは、機能(b)にも適用される。

要求事項(6)ユーザが優先順位を記憶できること、に対して機能(i)(優先すべきクラスタのメモ機能)を実装した(図6.4中⑥)。画面上左部に、最優先・要検討・不要のテキストボックスを表示し、ユーザは自由に書き込むことが可能である。このテキストボックス内のメモは常時表示される。このメモ機能により、ユーザは最優先に確認すべきクラスタ、その次に確認すべきクラスタというように優先順位の付与が可能となる。

要求事項(7)表示される情報が多すぎないこと、に対して機能(j)(提示する投稿の表示数切り替え(最大500件))、機能(k)(提示する投稿の最大サンプリング数の設置(最大500件))を実装した(図6.4中⑦)。機能(j)は、一度に画面上に表示される最大投稿数の切り替え機能であり、10件、50件、500件(全件)に対応している。機能(k)は、データベースから一度にサンプリングされる件数を最大500件に設定することで、膨大な量の投稿の情報を確認する必要がなくなる。なお、このサンプリングの最大数の設置は、その他の機能(b)や(c)、(d)にも適用される。これらの機能によりユーザの情報確認コストを下げること、ユーザの負担低減を図る。

6.3 評価実験

6.3.1 実験の目的

本実験では、提案のシステム(全ての機能を保持)と、機能を制限したベースラインシステム(図6.9参照)を用いて、実際に被災したユーザが投稿したと想定される「災害被害が確認できる画像・動画」をどれだけ収集できたかを評価する。ベースラインシステムで採用した機能は機能(a)、機能(c)、機能(d)、機能(e)、機能(f)、機能(g)、機能(j)、機能(k)であり、提案システムの持つ機能(b)、機能(h)、機能(i)を除いた。これは、一般にソーシャルメディア内の投稿を検索する際と同等の機能を再現するためである。

6.3.2 実験概要

実験参加者は、情報系の大学の学生および卒業生とし、提案システム13名、ベースラインシステム13名の合計26名とした。ただし、そのうち提案システムを利用した2名は、実験中に全く別のタスクを始める等の行動が見られ、課したタスクの理解が十分と言えなかったため、分析からは排除した。実験参加者には、実際の災害時に利用を想定した実験ストーリーを提示し、タスクとして災害被害が確認できるSNSの投稿の収集を課した。実験ストーリーは、1)「今現在、大規模な自然災害(豪雨災害)が発生している」、2)「ユーザは自治体の災害対策担当職員であり、「実際に被害が発生している」「住民が救援を求めている」等の投

SQLite + Plotly 3D (操作ログ) x +

新しい Chrome をご利用いただけます

ツイート

DBファイル: tweets_images.db

全 29343 件 / 表示 500 件 / 残 29343 件 /

Show entries Search:

tweet_id	text	media
1276274791473807360	. Good morning 🌞。雨の朝です 🌧️! お昼から晴れになってるけど ほんとなな ~~~。きょうこそ 避難訓練だよ。皆さま体調気をつけて いらっしやい♡ https://t.co/Y1LEUnk4X	
1276353531272818688	地震 震災 防災セット 一人用 【8月下旬~9月中旬発送】 防災グッズ 1人用 防災リュック 防災 グッズ 非常用持ち出し袋 1人 一人暮らし 防災ラジオ 家族 防災リュックセット 避難セット 災害グッズ 災害セット 懐中電灯 非常食 防災用品 おすすめ 送料無料 [楽天] https://t.co/BxJsYehKGC #rakuaf1 https://t.co/ajbqy8DQld	
1276373502086414336	元々多かったウーバーの外人配達員さんたち。母国に避難してたのか知らんけど、コロナで急減してたのが戻り始めました。でも今はウーバーの需要が減ってるので稼げないと嘆いてる。先月、先々月、日本にいれば沢山稼げたのに残念(>_<) https://t.co/9Bw87aOBxu	
中略		
1276424423071838208	今日から始まった #横浜元町 #チャージングセール 横浜元町の洋菓子店 #霧笛楼 のお楽しみ袋と、いつも行く #ウチキパン 買って来た 🍞 もう少しゆっくり見たかったけど、暑過ぎて、みなとみらいにそそくさ避難 🏠 知ってて行ってみたものの、やっぱり伊東屋は何もセールは無かった 😞 https://t.co/UVxOcjF6HH	

Showing 31 to 40 of 500 entries Previous 1 2 3 **4** 5 ... 50 Next

Picked Tweets

Show entries Search:

tweet_id	text	media
1277839673889419265	ヤバい〜 嵐だぁ〜笑 避難場所確認しとこっ!!!!!! https://t.co/WyviK4q5e5	

Showing 1 to 1 of 1 entries Previous **1** Next

図 6.9: 実験で用いたベースラインシステム

表 6.3: ユーザが収集した災害被害画像・動画数差の検定結果

Period	提案_平均	提案_標準偏差	ベースライン_平均	ベースライン_標準偏差	u_stat	u_p
0-5	3.6364	4.3422	8.0769	5.2195	37.5000	0.0508
5-10	16.7273	17.7994	22.0000	12.9808	46.5000	0.1553
10-15	37.9091	26.2277	35.5385	17.8868	73.5000	0.9307
15-20	60.7273	35.4432	47.0000	25.1992	88.5000	0.3388
20-25	97.4545	43.5692	60.3846	28.2682	108.5000	0.0342
25-30	128.1818	63.8229	74.4615	36.7641	108.0000	0.0369

稿を、迅速かつ漏れなく発見する」、3)「判断に困った投稿は、とりあえずピックアップする（他ユーザと相談して判断することを想定しているため）」の3つの観点から作成した。

6.3.3 実験手順

実験は、(1) 実験概要説明、(2) 実験同意書への署名、(3) 事前アンケート (4) タスクについての説明、(5) 実験ストーリーの説明、(6) システムの動作説明、(7) システムの使用練習、(8) 本実験、(9) 事後アンケートおよびインタビュー、の順に行った。(3)の事前アンケートでは、日常的に使う SNS、災害時の SNS の利用経験、コンピュータ操作の習熟度について質問した。(4) および (5) では、システムを用いて SNS から実際に被災者が投稿している“被害状況”や“救助要請”といった災害被害が確認できる投稿をできるだけ多く見つけて収集することを指示した。その際、ニュースやウェブ、テレビで発信されたであろう情報の再発信だと分かる投稿は含まないように指示した。(6) および (7) では、実際にシステム画面を提示し、ユーザが触れることができる状態で行った。提案システムを利用する実験参加者群には、クラスタの説明および、PCA 分析結果の 3D マッピング図の見方や操作方法、得られる知見についての詳細な説明も行った。また、システム使用練習は 2 分間に設定した。(8) の本実験は 30 分間とし、進行度に関わらずその時点で終了とした。時間の計測は、データベースを読み込み、投稿内容が表示された瞬間から開始とした。(9) の事後アンケートでは、システムの改善点や、クラスタ表示による円滑さ、投稿を確認する際にどのような情報を優先的に確認したか、どのような情報があれば災害被害と判断できたかを質問した。

6.3.4 実験結果

実験参加者がシステムの操作を始めてから 30 分で収集した「災害被害が確認できる画像・動画」を集計し、提案システムを利用したユーザの結果（以後、提案システム群）と、ベースラインシステムを利用したユーザの結果（以後、システム群）を比較する。提案システムの複雑さを考慮し、収集した画像・動画の集計は、開始後 5 分まで、10 分まで、15 分まで、20 分まで、25 分まで、最後まで、の六段階で行う。その後、u-検定 [33] を行い、収集した数の差について検証した。その結果を表 6.3 に示す。提案システムを用いたユーザと、ベースラインシステムを用いたユーザの、時間ごとのピックアップ数の推移とその平均を図 6.10 および図 6.11 に示す。また、ユーザが収集した投稿の正解率を表 6.4。正解率は、?? 章で人手で付与した“ユーザが収集した被害”を正解ラベルとし、ユーザが収集した投稿とその一致率を算出したものである。

表 6.4: ユーザが収集した災害被害画像・動画の正解率

Period	提案システム_平均	ベースラインシステム_平均
0-5	0.4089	0.5088
5-10	0.6017	0.6733
10-15	0.6822	0.6905
15-20	0.7983	0.6860
20-25	0.8375	0.7016
25-30	0.8434	0.7096

6.4 本章のまとめ

本章では、災害時における情報の「段階的除外」を支援するインタフェースを設計・実装し、その有効性を検証するための評価実験を行った。実装したシステムには、CLIP による画像の言語化とその特徴量を用いたクラスタリングによって整理された投稿群を3次元空間に可視化し、ユーザがクラスタ単位でノイズを除外できる機能を実装した。比較実験の結果、提案システムを利用したユーザ群は、単純なキーワード検索のみのベースラインシステムを利用した群と比較して、タスク開始直後こそ操作習熟に時間を要したものの、後半の20分から30分ピリオドにおいては有意に高い効率で被害情報を収集できることが確認された。これは、ユーザがシステムを利用する過程で「ゲーム」や「動物」といった明確なノイズクラスタを特定・除外することで情報空間が縮小され、結果として残存する重要情報の可視性が相対的に高まったことに起因すると考えられる。この結果は、本研究が提案する、情報の直接的な抽出を行うよりも、「情報の除外を先行させる」というトリアージモデルが実際のインタフェースを通して人間の認知コストを低減することを実証するものである。一方で、一部のクラスタにおいては、依然としてノイズ情報の混在が見られ、機械学習モデルやバイナリベクトル化による情報の単純化等に起因する課題も明らかとなった。次章では、これまでの研究の全体を通して得られた知見を統合し、実際の災害対応現場での運用や技術的な課題について総合的に議論を行う。

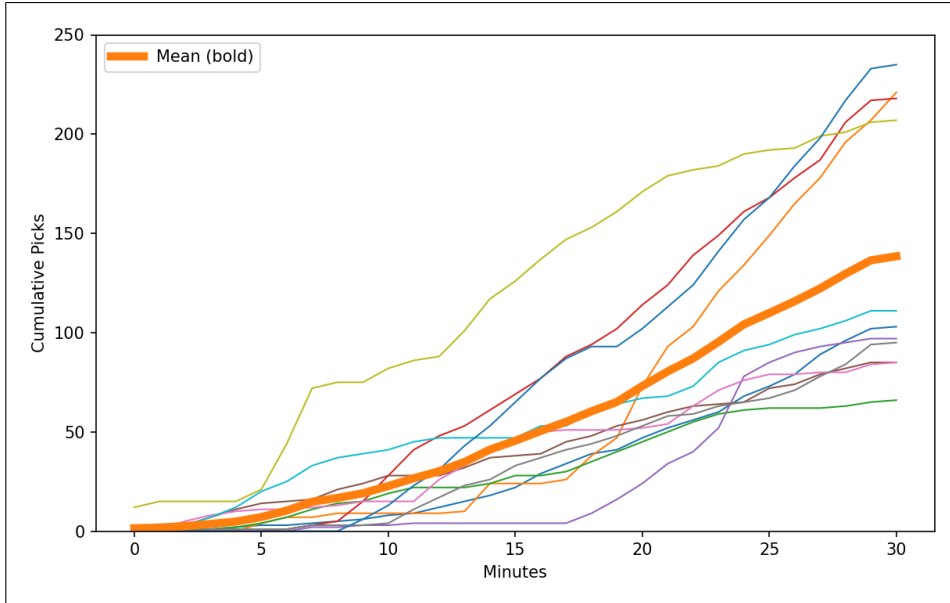


図 6.10: 提案システムユーザの投稿の累積ピックアップ数

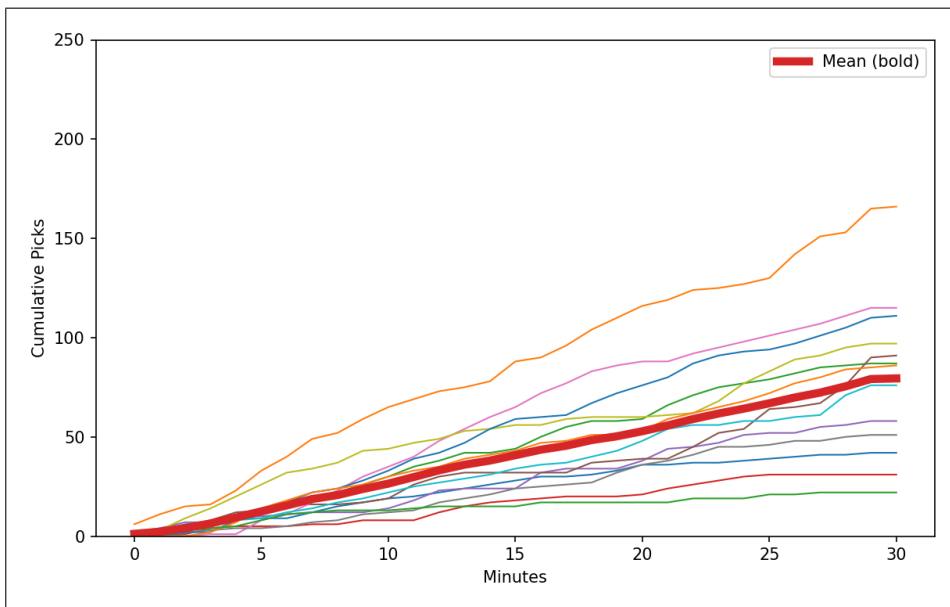


図 6.11: ベースラインシステムユーザの投稿の累積ピックアップ数

7 議論

本章では、本研究が提案する「人と計算機の協調による情報トリアージのためのノイズ情報抽出と峻別システム」について、4章から6章までの調査・実験結果を統合し、その有効性と到達点について議論する。本研究の最大の貢献は、従来の災害時 SNS 利活用研究で主流であった「重要情報の直接抽出」というアプローチに対して、「無関係情報の確実な排除を先行させる」というある種、逆向きの設計原理の有効性を、時間制約付きの実験タスクにおいて実証した点にある。以下では、4章での災害時の SNS 情報におけるノイズ構造の理解、5章でのマルチモーダル手法を用いたノイズ情報の機械的分離の可能性提示、6章での画像情報を中心としたクラスタリングと段階的な情報除外を基盤とした情報峻別インタフェースによる情報収集効率の変化について総括し、本システムが時間的・人的リソースの限られた災害時の情報収集について、どのような貢献を果たすか議論する。

7.1 情報抽出アプローチの限界とノイズ構造に関する議論

4章の調査において、キーワード検索（「救助」「避難」）によって収集された投稿のうち、実際に被災者が撮影した被害画像を含む投稿は、全体のわずか 0.564% に過ぎないことが明らかとなった。テキスト情報の TF-IDF 分析の結果、有用な被害報告には特出すべき単語特徴が見られない一方で、ゲームや日常投稿などのノイズ情報には語彙的特徴が存在することも確認された。これは、これまで多くの研究で行われてきた重要語句によるフィルタリングや、重要情報を直接抽出するアプローチが災害時の SNS においては困難であることを定量的に示すものである。重要情報の特徴が希薄であるため、適合率 (Precision) を上げようとすると、再現率 (Recall) が低下してしまうことは、災害時における情報の取りこぼしを防ぐことを命題とした場合に致命的な欠点となり得る。

特に災害時には、信頼できる情報の可視化は、自治体など支援や助けを求めるユーザとの信頼関係の基礎となる。ノイズが低減された情報には、ニュースや SNS 情報の可視性と信頼性を高めることが期待される。これは、迅速な判断が求められる災害下において、タイムリーでより正確な情報が効果的な意思決定支援のために重要であると考えられる。また、本研究で提案する人—計算機協調による情報トリアージのアプローチは、膨大な情報の中から真に重要な情報が希少であるといった、同様の問題構造を持つ社会問題において特に有効であると考えられる。

7.2 段階的除外による被害情報の顕在化に関する議論

5章では、CLIP を用いたマルチモーダル手法により、テキスト特徴だけでは判別困難だったノイズ情報が、画像特徴をテキスト特徴化しクラスタリングすることで、特定のクラスターに偏在し、機械的に分離可能であることを示した。表 5.3 のクラスター別の画像分布で示したように、人手による分類でノイズ情報だと判断されたゲームや動物、テレビ番組、天気図、飲食物、といった画像群は特定のクラスターに集中している。これにより、災害の被害とは無関係なノイズ投稿の多くが、ノイズ情報群の除外を行うことで非可視化され、結果として情報の残集合から“被害”画像が偏在するクラスター（例：Cl:3/Cl:13 など）が相対的

に可視化されやすい構造になっていることが考えられる。

6章の提案システムの評価実験では、この構造を利用してユーザが能動的にノイズを除外していくシステムが、投稿をリストで閲覧していくベースラインシステムと比較して、実験後半で優位に高い収集効率を示した。「被害が確認可能な投稿」のピックアップ数は、提案システム群の実験終盤（20分から25分ピリオド）において平均97.4545件、ベースラインシステム群の終盤（20分から25分ピリオド）において平均43.5692件となり、平均差は約53.9件であった。マン・ホイットニーのU検定の結果、 $p=0.0342$ となり、提案システム群が有意に上回った。また、実験最終盤の25分から30分ピリオドにおける提案システム群の平均ピックアップ数は128.1818件、ベースラインシステム群の平均ピックアップ数は74.4615件となり、平均差は53.7件であった。マン・ホイットニーのU検定の結果、 $p=0.0369$ となり、20分から25分ピリオドと同様に提案システム群が有意に上回った。一方で、実験開始直後の0分から10分ピリオド間ではベースラインシステム群に劣後した。10分から15分ピリオドで優位傾向に転じた後、最終盤で差が拡大するという変化が確認できた。これらの結果は、重要情報の直接的な抽出が困難な状況下において、比較的特徴が明確であったノイズ情報の除外を行う本研究のアプローチが有効であったことを示唆するものである。

また、ユーザが収集した投稿の正解率を概観すると、提案システムとベースラインシステムの両方において、ユーザーは妥当な精度を達成した。本研究の主目的は、ユーザーが災害被害と判断し、さらなる信憑性や詳細の確認が必要な投稿に値すると判断する投稿の数を検証することである。したがって、精度を1.0に最大化することが目的ではない。両システム間で精度に有意な差は認められなかったため、ユーザーはいずれのシステムでも関連情報を識別できたと結論付ける。本研究の目的である「人と計算機の協調」において、本システムの実装と評価実験をすることで、計算機と人の得意領域を適切に分担するモデルを提示した。次節では、実験結果の詳細な分析から、提案システムが有効的に機能した理由を考察する。

7.3 提案システムの役割に関する考察

提案システムが有効に機能した理由の一つとして、それぞれの機能が相互に作用したことが考えられる。提案システムでは、まず、CLIPを用いて画像を言語特徴化しPCA分析を経てクラスタリングを行った。次に、PCA分析の結果をPCA 3Dマップ（機能(b)）として実装することで、クラスタの空間的な配置を可視化した。これにより3D空間内で投稿/クラスタ間の距離を手がかりにして類似度の高い情報を視覚的に確認することで、ノイズ情報と他のノイズ情報の発見することが可能であった。同様に、被害画像が含まれているクラスタについても3D空間内の分布を手がかりにすることで発見することが可能となっている。これらの機能と、投稿の検索機能（機能(c)）を併用することで、候補を確保するワークフローを意図的に支える設計である。ノイズとなる情報群（クラスタ）の探索が容易になり、選択したクラスタの情報を除外する機能（機能(h)）で重要な情報を確保することを容易にした。この情報の「可視化」「除外」を反復的に繰り返すことで、関連クラスタが“浮き上がる”状況が作られ、結果的に重要情報が含まれている情報の可視性が上がる。

ユーザ観察の結果、PCA 3Dマップ機能を頻繁に確認した提案システム群の実験参加者

は11名中6名となり、この機能(b)の利用頻度の低い参加者に比べると、比較的ピックアップの開始が早く、ノイズ情報が偏在するクラスタを先に発見・除外することが確認できた。これは、ユーザが直接撮影した「被害」画像が含まれているクラスタの可視性を相対的に高めるといふ、本研究で提案している段階的な情報除外が、可視化機能と除外機能の相互作用によって機能したことを示唆している。

この可視化、除外、再可視化の反復により、関連クラスタが“浮き上がる”状況が作られる。PCA 3D マップでは、近似しているクラスタを視覚的・直感的に把握することができ、一度ノイズとなるクラスタを見つけたら、連鎖的に他のノイズとなりえるクラスタの発見が可能であった。被害画像が含まれているクラスタも同様の方法により発見が容易になる。ユーザ観察では、PCA 3D マップ機能を頻繁に確認したシステム群の実験参加者は11名中6名おり、利用していない群と比較してpickし始めるまでに要した時間が短く、除外したクラスタも多い傾向が見られた。つまり、ノイズが偏在するクラスタを先に除外することが信頼性の高い“被害”画像が含まれているクラスタの可視性を相対的に高めるといふ本研究で提案している段階的情報の除外が、可視化と操作系の相乗効果により機能したことを示唆している。

7.3.1 時間制約化における学習コストに関する課題

6章での実験結果は、提案システムの有効性を示す一方で、その習熟に一定のコストを要することも明らかにした。提案システム群は、被害が確認できる情報を初めてピックアップするまでに最長554.937秒を要した一方で、ベースラインシステム群での最長は224.447秒となり、始めのピックアップまでに200秒以上を要したユーザは、提案システム群で4名、ベースラインシステム群で1名のみとなった。これは、この実験初期段階の0分から10分ピリオドでの遅れは、提案システムの抱える学習コストを要する高度機能（クラスタ除外・PCA3Dマップ・検索併用）が理解・定着した後に効果が逡増する学習曲線が存在したことを示唆している。ベースラインシステムは、日常的なSNS利用やWeb検索行為と検索モデルが共通しており、直感的に利用可能であったが、提案システムで提示しているクラスタリングの結果を理解する行為や不要情報を群除外していく行為はユーザにとって既存の概念ではなく、その効果を実感し、情報を確保するための戦略として確立するまでに時間を要したと考えられる。この、システム群の実験後半20分から30分ピリオド間結果は、災害発生直後の被災地が混乱している状況において、初動数分間の遅れについては考慮する必要がある。一刻を争う災害時において即座に使えるシステム構築の必要性もあるが、平時からシステムに触れるなどのトレーニングを行うことで習熟させたり、これまでの知見を生かした既知のノイズに対しては、ノイズ情報である可能性を明確にすることで判断の一助とすることで、初動の遅れを解消できる可能性は高い。例えば、5章で特定された「ゲーム」といったエンターテインメント関連のコンテンツのような、日常的/恒常的に存在するノイズ情報に対して、除外候補として提示することが考えられる。ユーザの、システム操作の初期段階の負担を軽減することで、本実験で確認された学習曲線をより短縮できると考える。

7.3.2 CLIP を用いた画像解析に関する課題

本研究ではCLIPとクラスタリングの有用性を示したが、同時に技術的および実装上の課題も浮き彫りとなった。本システムでは、再現性を高めるためにCLIPによる出力単語をバイナリベクトル化して確率分散を考慮しない設計を採用した。この単純化は、計算コストの低減や処理の安定に寄与する一方で、情報の単純化につながった可能性が高い。具体的には、6.1.2節のCl:11およびCl:12にはほぼ全ての種類の画像が存在する結果となった。“被害”が含まれてはいるが、ゲームや動物、風景といった災害とは無関係な情報も多く含まれていたことから、これらのクラスを除外した実験参加者も確認している。これは、情報の取りこぼしを防ぐという観点で課題が残る。この問題に対しては、多くの種類が含まれているクラスに絞って再度クラスタリングしたり、recall（再現率）100を目指したクラスタリングを行う等の対策が考えられる。これにより、クラスタ数が増加してしまったとしても、上述のように学習コストの軽減をあらかじめ行ったり、人同士で分担するために複数人で同時に情報処理を行うことが可能な仕組みを導入することで、ある程度のクラスタ数の増加には十分対応可能だと考えられる。ただし、人的・時間的リソースの分配については十分に考慮する必要があるため、クラスタ数の決定を最適化するのではなく、あくまで災害時に最適なクラスタ数の調査が必要になる。

7.3.3 インタフェースの操作性と協調作業に関する課題

事後アンケートやユーザ観察から、インタフェースの操作性に関する具体的な課題も得られた。まず、誤動作の観点から、データベースの読み込みの待ち時間の間に、実行ボタンを何度も連打することで、読み込むデータが変わってしまうという現象や、除外クラスタの全除去、ピックアップ状態の意図しない解除といった操作ミスが見られた。ピックアップした投稿をまとめて表示している図6.8画面で、ピックアップ状態を直接解除してしまうと、再度同じ投稿が画面上に流れてくるまで、図6.4の画面で確認し続ける必要があった。また、投稿の一覧性を高めるために、一度の表示件数を多く設定したユーザが、ページ送りの際に画面スクロール位置が維持されることで、次のページを確認するために再度最上部までスクロールしなければならないというUI上の不備も見られた。これらは、システムの実運用上において、ストレス要因となり得るため改善の必要がある。

また、本システムの実験では、実験参加者一人が30分の間に全ての作業をさせ個人作業を前提としており、他ユーザとの共同作業機能を搭載していない。2章で述べたHercegらが、情報を探索するユーザと、それらの結果から情報を収集するユーザを分担させたように、本研究のプロセスも、クラスタリングの結果を参照しどのような情報を除外し、どのような情報の優先度を高くするか設定するユーザと、実際にクラスタの中身を精査し、情報を収集していくユーザによる分担作業化することも考えられる。実際の災害対応はチームで協調的に行われることを考慮すると、交代制で長時間にわたる情報の監視・収集を前提にし、操作ログの共有だけではなく、リアルタイムでの除外状態の同期機能や優先順位のコセンサスを得る機能、ピックアップ状態の共有やそれらの承認フローの実装など、ステークホルダー達との共有までを見据えたComputer-Supported Cooperative Workシステムの開発が必要となる。

7.3.4 データセットの限界に関する課題

本研究で使用したデータセットは「令和2年7月豪雨」という単一の災害に限って収集されたものである。地震や火災などの他の種類の災害や、日本国外の災害事例に対して、同様のノイズ構造やクラスタリング傾向が見られるかは未検証である。より汎用的なシステムにするためには、異なる災害種別や異なる文化圏のデータを用いた検証は必要不可欠である。また、本研究ではTwitter（現 X）APIを利用して過去の災害データを収集・分析を行ったが、2023年に実施されたAPI仕様の変更および、研究目的利用の有償化に伴い、従来のような大規模データの収集は極めて困難となっている。そのため、Enterprise APIへのアクセス権を持つプラットフォーム事業者や関連企業との産学連携するなど、持続可能なデータ共有の枠組みが必要不可欠となっている。また、Xといった単一のサービスプロバイダに一極集中してしまうリスクを回避するため、Bluesky や Mastodon, Threads などの分散型ソーシャルメディアや、Facebook といった API の公開がされているソーシャルメディアへの分析対象の拡大も視野に入れる必要がある。ただし、ノイズ情報の除外を行なっていく情報トリージングを元にした設計思想と、限られたリソースを有効活用するために無関係情報の迅速な把握・除外する技術の重要性は高い。仮に対象のプラットフォームが変化したとしても、災害時における日常投稿やエンターテインメント関連投稿といった、ノイズが混入する構造自体は変わらず、それらをクラスタとして捉えて除外する本研究のアプローチは、将来的な情報環境においても有効に機能すると考えられる。

8 結論

本研究では、人的・時間的リソースが制限されている大規模災害時という状況下において、ソーシャルメディア上の膨大かつ玉石混交な情報の中から、被害状況の把握や迅速な救助活動に資する重要情報を、効率的かつ確実に確保するための枠組みである、「人と計算機の協調による情報トリージ」について検討し、その有効性を実証した。

1章では、災害大国である日本において、災害時のインターネットとソーシャルメディアの関わりについて概観した。SNSがリアルタイムな情報源（ソーシャルセンサ）として極めて高い期待を寄せられている現状と、SNSの持つ深刻な課題について整理した。具体的には、災害発生時の情報の爆発的な増加（バースト現象）や、善意・悪意を問わず拡散される偽・誤情報等により、現場での迅速な状況把握が困難になっている実態を指摘した。特に、これまでの情報抽出技術の多くが「正解を見つける」ことに注力しており、第二種の過誤による重要情報の取りこぼしが避けられない点、および膨大な情報を人手で全て確認することが現実的ではない点が、災害時における実務上での大きな課題となっていることを示した。

2章では、本研究の学術的・社会的意義を明確化するために、災害時のソーシャルメディア活用に関する研究動向を包括的に概観し、既存アプローチの到達点と未解決の課題を整理した。まず、SNSを社会の状況を検知するソーシャルセンサとして捉える一連の研究においては、災害の即時検知や状況把握における有効性が示されてきた一方で、人口密度の低い地域における情報の空白や、発信される情報の信頼性担保という課題が残されていることを確認した。また、情報の爆発的増加（バースト）に関する研究動向からは、発災直後に情報量が急増し、ユーザー行動が対話から拡散へと変容し、有益で重要な情報がノイズ情報に埋没するSNS構造が明らかにされている。これらに対し、自然言語処理や画像解析技術を用いて重要な情報を「自動抽出・検出」しようとする研究が多く試みられてきた。しかし、これらの「抽出型」アプローチは、学習データに依存するため未知の災害パターンへの適応が難しく、重要情報の取りこぼしを招きかねないという課題がある。さらに、偽・誤情報対策に関する研究においても、機械的な自動検知には限界があり、最終的な真偽判断には依然として高度な判断力を持つ人的リソースが必要であることが示されている。これらの課題にもとづき、情報の「抽出」のみに頼るのではなく、「情報トリージ」の概念を応用することで、計算機を「正解を見つける」ためではなく人間が効率的にノイズを捨てるための支援を行うことの必要性について論じた。

3章では、SNSを用いた信憑性の高い災害情報収集のための指針として、情報の流通構造と峻別プロセスをモデル化した。災害時のSNS情報は、現場からの一次情報、報道やそれらの拡散による二次情報、そして日常投稿や反応などのノイズ情報が層状に混在しているマトリクス構造を持つことを示した。その上で、情報の信頼性判断にはテキスト情報だけでなく、画像、発信者情報、拡散の経緯などの複合的な手がかりが必要であることを論じた。この複雑な判断プロセスを支援するために、計算機が大量データの構造化と可視化を担い、人間が最終的な判断と除外を行うという、役割分担に基づいた協調モデルを提案した。

4章では、実際の災害データ（令和2年7月豪雨）を用いて、実際に流通した情報の調査を

行った調査を行い、SNS上の災害情報構造を定量的に明らかにした。キーワード検索（「救助」「避難」）で収集された約48万件の投稿のうち、実際に被災者が撮影した被害画像を含む投稿は全体のわずか0.564%に過ぎないという事実は、従来の単一のモダリティに着目した情報抽出アプローチの限界を示唆するものであった。また、テキスト特徴量（TF-IDF）の分析から、被害報告には際立った単語特徴が見られない一方で、ゲームや日常投稿などのノイズ情報には明確な語彙的特徴が存在することを確認した。この結果は、重要情報の特徴を用いて抽出するよりも、ノイズ情報の特徴を抽出して除外する方が、災害時の情報収集において合理的かつ確実性が高いことを裏付けるものであり、本研究のアプローチの妥当性を支持する根拠となった。

5章では、ノイズ情報の機械的な分離可能性を検証するため、画像とテキストを統合的に扱うマルチモーダルモデルであるCLIPを用いた分析を行った。画像を言語的特徴量へ変換しクラスタリングを行った結果、ゲームや動物、日常風景といったノイズ情報が特定のクラスターに偏在し、機械的に分離可能であることが実証された。被害画像と風景画像の一部に混在は見られるものの、明確なノイズ群をクラスターとして可視化・分離できたことは、人が確認すべき情報空間を縮小する上で重要な知見である。これにより、計算機による事前処理が、人間の認知負荷の低減に有効的に機能することが示された。

6章では、これまでの知見を統合し、段階的な情報除外を支援する「災害情報トリアージインタフェース」を設計・実装し、その有効性を検証した。実装したシステムは、クラスタリングされた情報群を3次元空間に可視化し、ユーザがクラスター単位でノイズを能動的に除外できる機能を備えている。比較実験の結果、提案システムを利用した群は、従来のリスト形式で閲覧するベースラインシステムと比較して、タスクの後半において有意に高い効率で被害情報を収集できることが確認された。これは、可視化されたノイズクラスターをユーザが自身の判断で除外していくことで、探索空間が縮小され、残存する重要情報の可視性が相対的に高まるという「除外型」のトリアージモデルが、実際の作業において機能し、人間の認知コスト低減に寄与することを実証している。

7章では、本研究の総括として、情報抽出アプローチの限界とノイズ構造、段階的除外による被害情報の可視化、および提案システムの役割について多角的な視点から議論した。本研究が提案したモデルとシステムによって、計算機が得意とする「大量データの整理・可視化」と、人間が得意とする「信頼性判断・最終決定」を適切に分担・協調させることが可能となり、人的・時間的制約の大きい状況下における災害情報の収集効率を飛躍的に向上させることが示された。また、本システムの導入における学習コストや、将来的なプラットフォームの変化への対応可能性についても議論し、実運用に向けた課題と展望を示した。

結論として、本研究は、災害時のSNS活用において、従来主流であった「抽出(Extraction)」手法から「除外(Subtraction)」手法への転換を提案し、その有効性を実証したものである。SNSの情報構造に応じた「人と計算機の協調モデル」と「情報トリアージ」の必要性は、将来的にSNSのプラットフォームや情報の形態が変化した場合においても、ノイズ情報が混在する情報空間から価値ある情報を確保するための普遍的な方法論として機能し、災害対応における意思決定支援に貢献することが期待できる。

謝辞

本研究を纏めるにあたり、長きにわたり多大なるご指導、ご鞭撻を賜りました関西大学大学院総合情報学研究科総合情報学専攻の松下光範教授に、深甚の謝意を表します。研究室に配属されてから9年の間に、研究に関することのみならず、様々なことを教わり成長する機会を沢山いただきました。また、数多くのご助言をいただきました関西大学大学院総合情報学研究科総合情報学専攻の堀雅洋教授、ならびに林貴宏教授に心より御礼申し上げます。研究会や研究室交流を通して、数多くのご助言をいただきました関西大学大学院総合情報学研究科総合情報学専攻の山西良典教授に心より御礼申し上げます。

本研究の遂行にあたり、数多くのご助言、ご協力が欠かせないものでありました。社会的意義についての議論や研究遂行について多数のご助言をいただきました法政大学社会学部メディア社会学科教授の藤代裕之教授に心より感謝いたします。藤代教授には、これまでの研究や「RISTEX ニュース発信者と受信者間における「トラスト」形成プロジェクト」を通して、様々な交流機会をいただきました。合同研究会を通じて多数のコメントやご助言、また研究者としての心構えを教えてくださいました明治大学総合数理学部先端メディアサイエンス学科の中村聡史教授に心より感謝いたします。学部生の頃から論文の試筆や実験設計、研究遂行に関わるあらゆるご指導を賜りました立命館大学立命館グローバル・イノベーション研究機構専門研究員の安尾萌氏に心より感謝いたします。

2018年に研究室に配属されてから同期として切磋琢磨し、時には遊興にふけあった松下研究室10期生の玄道俊氏、福元颯氏、小林光氏、宮本誠人氏、樋口友梨穂氏、ならびに井上須美氏、今江柚実氏、坂下奈々美氏、建田伸氏、三溝俊介氏にも深く感謝いたします。日頃より議論を通じて様々な観点から助言いただき、また多くの時間を共にし、心の拠り所となっただきました樋口亮太氏、高橋知奈氏、田中瑠慧氏、そして松下研究室のメンバーに深く御礼申し上げます。関西大学総合情報学部総情工房 MonoLab の皆様をはじめ、関西大学大学院総合情報学研究科の教職員の皆様には、日頃より研究における多大なご協力をいただきました。深く御礼申し上げます。その他、実験へのご協力を賜った方々、学会発表を通じてご意見をご教示いただきました方々等、各者各様の立場からのご協力によって本研究を完成させることができました。深くお礼申し上げます。

本研究は、家族の支えなくしては遂行しえませんでした。筆者の生活を長きにわたって支援していただいた父、森野高晴と、母、森野光子に心から感謝します。研究者として常に私の目標として存在し続ける、故森野一高、故石本泰雄に心から感謝します。また、筆者の学位取得を長年にわたって応援してくれた姉、森野絢子、森野蘭子に心から感謝します。

A 付録

付録 A では、6 章の実験で使用した指示書や、アンケート結果を記載する。

A.1 タスクについて

これから本研究で作成したシステムで、検索等の機能を駆使し、SNS の投稿から実際に被災者が投稿している“被害状況”や“救助要請”といった被害状況が確認できる投稿をできるだけ多く見つけてピックアップしてもらいます。その際、一次情報を見つけることを前提に、ニュース等で流されている情報は除いてください。また、判断に迷った投稿があった場合は、それも含めてピックアップしてください。本システムでは、全にランダムに 500 件ずつ完投稿が表示され、SNS の投稿は、必ず写真が添付してあることを条件に提示しています。なお、本研究では意図的にユーザー情報や投稿日時を表示していません。制限時間が来ましたら進行度に関わらず、その場で終了となります。作業途中で操作方法等がわからなくなった際は、手を上げて実験実施者をお呼びください。

A.2 実験ストーリー

あなたは、ある自治体（市役所や県庁など）の災害対策担当職員です。現在、大規模な自然災害が発生しており、多くの住民が SNS（X（旧 Twitter）など）を通じて、被害状況や救助を求める情報を投稿しています。本タスクでは、SNS 上に投稿された大量の災害関連情報の中から、「実際に被害が発生している」「住民が救援を求めている」などの重要な投稿を、できるだけ迅速かつ漏れなく発見することが求められます。あなたが迷った投稿（本当に被害・救援要請なのか判断に困るもの）があった場合、最終的には他の職員と相談して判断できます。実際の災害現場では、こうした SNS 分析によって、救助の優先順位決定や現場対応の迅速化を目指しています。本実験においても、「見落としを防ぐ」「素早く見つける」ことが最も大切です。また、判断に迷う投稿については、実際の現場では他の職員と相談することもできます（本実験では迷ったものもピックアップして OK です）。

A.3 事前アンケート回答結果

事前アンケートに使用した回答用紙は以下の通りである。

参加者 ID: _____

事前アンケート

1. あなたの日常的な SNS (LINE や X, Facebook 等) 利用頻度を教えてください。
 - ① 毎日利用している
 - ② 2～3 日に数回利用している
 - ③ 週に数回利用している
 - ④ ほとんど使わない
2. あなたが日常的に使う SNS を教えてください。

3. SNS で災害情報を見た経験はありますか？
 - ① ある
 - ② 全くない
 - 「ある」と答えた方：どのような災害情報を見かけましたか？また、どのような情報発信を期待していますか？ (自由記述)

 - 「全くない」と答えた方：どのような災害関連の情報発信なら見たいと思いますか？ (自由記述)

4. 災害時に SNS 等で災害関連の情報を投稿した経験はありますか？
 - ① ある
 - ② 全くない
 - 「ある」と答えた方：どのような情報を何回程度投稿しましたか？ (自由記述)

 - 「全くない」と答えた方：実際に被災した場合、投稿したいと思いますか？理由もあれば教えてください (自由記述)

5. あなたのコンピュータ操作の習熟度を教えてください。(1=全く慣れていない、5=とても慣れている)
 - ① 1 2 3 4 5

図 A.1: 事前アンケート

A.3.1 設問の集計結果

各設問における回答件数は以下の通りである。

表 A.1: 事前アンケート選択設問

設問番号	選択肢	回答数
問1 (SNS利用頻度)	毎日利用している	24名
	2~3日に数回利用している	0名
	週に数回利用している	0名
	ほとんど使わない	0名
問2 (使用SNS)	LINE	20名
	Instagram	17名
	YouTube	2名
	TikTok	3名
	X	14名
	Discord	2名
	Slack	1名
問3 (災害情報の閲覧経験)	ある	23名
	全くない	1名
問4 (災害情報の投稿経験)	ある	1名
	全くない	23名
問5 (操作習熟度)	1(全く慣れていない)	0名
	2	1名
	3	3名
	4	14名
	5(とても慣れている)	6名

A.3.2 問3：SNSでの災害情報閲覧経験について

表 A.2: SNSにおける災害情報の閲覧経験と関連回答（設問3）

【「ある」と回答した方の詳細】
<ul style="list-style-type: none">・地震直後の電車の運行状況を確認した。公式アカウントによる正確な情報を期待する。・どこで災害があったかの情報、避難地域・どこで何時頃に、震度情報を添えた情報。また何をしたらいいか知りたい・土砂崩れとか見ました。自分の身に危険があるような災害情報を期待します。・地震、豪雨、津波など状況を把握できるような内容・地震の震源地と震度の分布、今後の地震の状況（余震が来るかどうか）・地震、火災、道路の陥没、落雷、津波、熱波・台風等の自然災害による影響、自分の地域は安全かどうかを期待する。・地震、津波情報 安全に対する情報を期待します。・台風に関しての情報をよく見かける。どこの地域に警報が出ているのかを知りたい。・地震の震源地や強さの情報、信頼性のある情報。・大雨、洪水、地震、身内が安全かどうか知れる、フェイクではない情報、自分ごとと思わせてくれるような発信。・津波、台風の被害、公共交通機関の状況。・気象庁など公式の災害情報を見る。・地震、震度や被害状況の発信・安否の確認情報・救助要請や交通情報、物資のある場所の情報共有を見たことがあり、それを期待する・過去の災害の津波や洪水の映像・土砂崩れや豪雨災害、期待する情報発信は、現在危険な状況になっており、近寄ってはいけないところ。・地震の被害状況、津波の状況、交通状況・避難所に関する情報・東日本大震災の映像がよく流れている。フェイクじゃない情報がほしい・地震、震源地と震度・災害の状況を知らせる発信
【「全くない」と回答した方の詳細】
<ul style="list-style-type: none">・一番被害がひどいところ。

A.3.3 問4：SNSでの災害情報発信経験について

表 A.3: SNSにおける災害情報の閲覧経験と関連回答（設問3）

【「ある」と回答した方の詳細】
<ul style="list-style-type: none">・見かけた災害の投稿をリポスト，共有
【「全くない」と回答した方の詳細】
<ul style="list-style-type: none">・ 思いません。自分の情報が正しいかわからないため。・ 思わない，居住地が知られる可能性があるため・ 助けになりそうな情報（ここで水もらえるなど）は投稿したい・ 正確な情報と判断できれば投稿したいと考えます。一人でも助かってくれたら良いと思う。・ 自分が居る場所が危険な場合には，近づかないようにして警告する意味で投稿したい・ 投稿したいと思わない。知り合いに知らせる程度でいいし，緊急を要する場合は然るべき機関に連絡することを優先するため。・ 被災地の現状を発信するために投稿したいと思う・ 情報を早く見つけてもらえるので，必要であれば投稿したい・ 思わない，普段からSNSを投稿しない。情報集めのために利用しているから・ したいとは思わない。発信よりも自分たちの命の方が優先だから。・ 必要があればしたい。・ 思わない。日常から投稿をあまりしないから・ 誰にでもためになるような情報は投稿したい・ 思わないです。・ しないと思います。身内だけの確認でいっぱいいっぱいになると思うからです。・ したくない。余裕がないから，投稿がエンタメとして見られるかもしれないのがイヤだから・ 大きな被害があれば，友人の安否確認のためにも投稿したい・ 投稿したいと思いません，公的機関への通報が適切と考えるからです。・ 投稿したいとは思わない。誤った情報を発信したくないから。・ 思わない，SNSで発信するという選択肢がない・ もともとSNS投稿への関心がないことに加えて，自身が投稿する内容に必要性を感じないため。・ 思わない，個人情報になりそうで怖いから・ 知り合いの安全確認のために投稿したい

A.4 事後アンケート回答結果

事前アンケートに使用した回答用紙は以下の通りである。

参加者 ID: _____

事後アンケート

1. 使いにくいと思った点やストレスの要因はありましたか？（自由記述）

2. ピックアップ操作や画面遷移に迷うことがありましたか？

- 1 (全くなかった)
 2
 3
 4
 5 (頻繁にあった)

3. (A群)clusterの表示で作業は捗りましたか？

- 1 (全く捗らなかった)
 2
 3
 4
 5 (とても捗った)

4. 投稿を確認する際、どの情報から優先的に確認しましたか？（複数選択可）

- 投稿 ID
 テキスト (本文)
 メディア (画像・動画)
 クラスタ情報 (※Aのみ)

5. また、投稿を確認する際、情報をどの順番で確認しましたか？

6. 情報を峻別する際、役に立った情報を教えてください。

裏面にも質問項目があります

図 A.2: 事後アンケート

A.4.1 提案システム群における設問の集計結果

各設問における回答件数は以下の通りである。

表 A.4: 事後アンケート選択設問 (提案システム群)

設問番号	選択肢	回答数
問 2 (画面遷移の迷い)	1 (全くなかった)	1 名
	2	5 名
	3	1 名
	4	4 名
	5 (頻繁にあった)	0 名
問 3 (cluster 表示の有用性) ※ A 群のみ対象 (11 名)	1 (全く捗らなかった)	0 名
	2	0 名
	3	0 名
	4	3 名
	5 (とても捗った)	8 名
問 4 (優先した情報) ※複数選択可	投稿 ID	0 票
	テキスト (本文)	8 票
	メディア (画像・動画)	11 票
	クラスタ情報 (※ A のみ)	4 票
問 7 (重要情報の発見度)	1 (全くできなかった)	0 名
	2	2 名
	3	1 名
	4	6 名
	5 (十分できた)	1 名
問 10 (今後の活用意欲)	はい	9 名
	いいえ	0 名
	わからない	1 名

A.4.2 ベースラインシステム群における設問の集計結果

各設問における回答件数は以下の通りである。

表 A.5: 事後アンケート選択設問（ベースラインシステム群）

設問番号	選択肢	回答数
問 2 (画面遷移の迷い)	1 (全くなかった)	4名
	2	6名
	3	1名
	4	2名
	5 (頻繁にあった)	0名
問 4 (優先した情報) ※複数選択可	投稿 ID	0票
	テキスト (本文) メディア (画像・動画)	5票 12票
問 7 (重要情報の発見度)	1 (全くできなかった)	0名
	2	2名
	3	3名
	4	8名
	5 (十分できた)	0名
問 10 (今後の活用意欲)	はい	9名
	いいえ	1名
	わからない	3名

A.4.3 使いにくい点やストレス要因

事後アンケート設問1における回答文は以下の通りである。

表 A.6: ストレス要因 (設問1)

【提案システム群の回答詳細】
<ul style="list-style-type: none">・ Search ボックスがわかりにくかった・ 画像が別のタブで開くのが少し面倒くさいと思った・ 実行を押してからちょっと時間かかって、押せているか一瞬不安になった。写真押しに行くのにちょっとめんどくさく感じた。特に動画・ クラスタの色がどんどん変更されていく点・ 動画や画像を拡大する際に、同じウィンドウの別タブで開かれるせ知恵そのタブを消す作業が発生したのが使いにくかったかもです。せめて別ウィンドウなら良かったと思います。・ ページを更新したときに、上から見れるのではなく同じ位置で固定され更新されたため、ピックアップしようとしたツイートが既にピックアップされたものだったことが何度もあった。・ 次のページに行ったときに上から表示して欲しかった・ 実行ボタンが改行される。除外とリフレッシュの実行ボタンが同じで少し混乱した・ 除外したクラスタの3D マッピングを確認して、離れた位置関係を見たかったが、除外を先に実行してしまうと3D マッピングの方のクラスターも消えてしまうのが扱いづらかった。・ 出てくる写真やテキストがランダム枚出てくるから既にピックアップしたようなものも出てきたところ。・ クラスタの [A11] 選択がうまくできなかった。
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・ 上部に次のページに行くボタンが欲しかった・ 並び替えとか少しわかりにくかった (正しく機能していないように感じる)・ 画像をもっと大きくした方が見やすい・ ゲームに関する内容が多かったこと・ ページを飛ぶときに更新しても下のページからになるのが少しストレスだった・ 特定のアカウントの投稿をブロック、非表示にする機能が欲しい・ 画像が拾い画かどうか判断に迷った、ピックしたツイートが再度表示されるのが勿体無いと感じた。・ ペットの話とゲームの話がストレスだった (関係ないけど多いため)・ 特になし。簡単に操作できた。・ DBD(ゲームコンテンツ) 多い・ 画像や動画が見づらい点、再生が別画面に飛ぶ点・ ゲームの情報ばかりが出てきたのがみづらかった

【ベースラインシステム群の回答詳細（続き）】

・動画の操作や、表示件数が少ない場合に Picked Tweet が表示されるので誤操作をしてしまうことがあった

A.4.4 投稿情報の確認順序

事後アンケート設問5における回答文は以下の通りである。

表 A.7: 投稿の確認順序（設問5）

【提案システム群の回答詳細】
<ul style="list-style-type: none">・メディア → テキスト → クラスタ情報 → ID・メディア, クラスタ情報, テキスト, 投稿ID・メディアで論外じゃないか確認して, クラスタ情報で見るべきかの目星をつけて, 最終判断をテキストで行なった.・画像や動画で判断し, その後本文の確認, ある程度進んでからはクラスタで判断.・クラスタ → メディア → テキストの順. ただし初めはメディア → テキスト → クラスタの順・メディアをみて, 災害に関連していそうだと判断したらテキストを軽く読んだ.・メディア → テキスト・クラスタ情報 → メディア → テキスト・写真 → テキストの順・クラスタ → テキスト → メディア・メディアのサムネ → 投稿本文 → メディアの全画面表示
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・画像か動画を見て, 文を見て判断した・1. テキスト 2. メディア・画像 → テキストの順番・画像 → テキスト・画像 (メディア), テキストで確認した・メディア → テキスト・画像・動画が災害に関するものか, 本文で被害を報告しているか+救助要請を出しているか.・テキストからメディアの順で確認しました・メディアを先に確認し, テキストを次に確認した.・メディア → テキスト・メディア → テキスト・メディア, テキスト, ID の順・場所 → ハッシュタグ → メディア

A.4.5 情報を峻別する際に役に立った情報

事後アンケート設問6における回答文は以下の通りである。

表 A.8: 役立った情報（設問6）

【提案システム群の回答詳細】
<ul style="list-style-type: none">・テキストを読んで絵文字が多かったり、ふざけた文章・県（どこに住んでいるか）、川の水の量、川が茶色・クラスタ情報で、目星をつけることで論外な情報の除外が捗った。・クラスタ情報は役に立ちました。・クラスタ情報（確実に関係のない投稿を一気に省けたから）、メディアの字幕（一次情報かどうかの判断に役立った）・ピックアップしたツイートが同じクラスタに何回も出てくることが確認できたので、そのクラスタを中心にみるのができたため、クラスタ情報は役に立った。関連しないと判断できたクラスタを除外するのも役立った。・地名（川の名前）や救助を求める文章・ハッシュタグ・クラスタ、画像、textに「救助」と含まれる情報を役立てた。クラスタの検索機能、あからさまに氾濫してる街の画像はピックアップした。・地名や写真・テキスト、その中での「家族」や「友人」といった第三者ワードの有無
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・画像や動画、町などのはっきりした名前・拡散、実際の被害情報、避難所の情報・画像や避難、助けて、状況、近くなどの言葉・住所が書いていたとき、生の写真があるとき、救助要請であることが明記されているとき・画像が役に立った、大体似たような救助画像だったので、画像から探した・メディア（茶色っぽいか、壊れてそうか）、テキスト（#救助、孤立）・#救助要請の文字、OOニュース、NHK等のメディアは排除、画像にテロップがない・風景の画像があるかどうか・メディアの情報が一番役にたった。その反面、メディアだけで判断してしまったものも多くある。・メディア、今回は、画像が茶色・テキストの長さや伝聞などの表現・テロップやアイコンがあればピックアップすべきではない情報であること・ハッシュタグとメディア情報

A.4.6 被害情報以外の重要情報

事後アンケート設問8における回答文は以下の通りである。

表 A.9: 重要情報（設問8）

【提案システム群の回答詳細】
<ul style="list-style-type: none">・投稿された時間をみたい・地名や写真・地元住民による大きな災害が到来する前の予兆のつぶやき（こんなことになっているの初めてみた等）・避難所の情報，どれくらい人がいるかとか・救助隊が派遣されているか否か．致命や位置情報などを用いて都道府県のどこにツイートが集中しているか知りたい．・避難情報，危険区域情報，避難していない人の情報・救助が終了していると思われる情報があると，そのほかを優先的に見やすい．・どの場所の被害かわからない投稿は使いものにならないのではと思った．・ない・なし
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・特になし・インフラに関する情報・危険が迫っているなどの次の被害を防ぐことができる情報・位置情報（タグとか？）・なし・身内が被災した情報・避難所の情報（物資など）・ハザードマップ，現在の被害状況（通れる道，通れない道，空いている避難場所）・避難所がどこにあるか，機能しているかなども重要だと思った・住所や被害に遭っている人数とその人の特徴（高齢など）・救助活動の有無・投稿内容で♡や✖などの絵文字がついているものはイヤに感じた

A.4.7 画面・ツールの改善点や感想

事後アンケート設問9における回答文は以下の通りである。

表 A.10: ツールの改善点や感想（設問9）

【提案システム群の回答詳細】
<ul style="list-style-type: none">・動画の再生をしやすくしてほしい・なし・被災者か、第三者の分配は自動でできないのかなと思った。救助が必要かどうかの判断をこっちがしてほしい。・ページを変更した時などでページの一番下に移るのが少し扱いづらいと感じました。・ピックするのが優先か、絞り込むのが優先か迷ってしまった。今回はピックするのを優先したため、見落としている重要なクラスタがありそうなのが気になった。・ページ更新の際に、ツイートが一番上まで移動してほしい・選択した投稿のクラスタの割合・選択した投稿を絞り込んで表示できる機能があると、とりあえず一気に見てから後から精査しやすいと感じる・「Previous」「next」の一番最後のページに飛ぶこと、三次元マップに写ってしまうこと。・本文がハッシュタグと区別しづらかったので、ハッシュタグは斜体にするといい？
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・検索ワードが定まってきたら素早く救助情報を選択できる場所がいいと感じた・ピックした情報が検索結果のすぐしたにあったため、既にピックした情報を誤ってクリックしてしまったこと・写真や動画が詳細情報まで見られるサイズだと良いと感じた。・ゲームの画像が少なくなったら良いと思った。・動画を再生する際に1クリックで再生できたら良いと感じた。・下まで行って次のページに行くと、下のまま、画像が小さい、文字サイズを変える・マイナス検索の追加、メディア欄のサイズアップ・アカウント単位のマイナス検索、キーワードのマイナス検索・1で回答した部分を改善してほしいです。・次のページに遷移したとき、したから始まるのではなく、上に自動でいってくれるといいなと思いました。・IDは表示する意味ないと思う・操作説明とウィンドウ説明が画面内にほしい・探しているところと、自分が選択してためっているところが見ただけで区別できなかった

A.4.8 期待する機能等や理由

事後アンケート設問 11 における回答文は以下の通りである。

表 A.11: 期待する機能（設問 11）

【提案システム群の回答詳細】
<ul style="list-style-type: none">・ニュース記事や陸上自衛隊など、公式の情報は個人の一次情報ではないのでいらないと思った。・被災レベル（何も食べ物がない、家が流された）を設定して、発信できるようになれば良い。・ハッシュタグ、そのものを選択できると入力時間が省けて良いと感じる・地震発生地域の地名で検索できる機能・キーワードで除外できる機能があれば良いと思った。「訓練」というキーワードを含むツイートが見られたので、そういったワードを含むツイートは消したかった。・3次元にプロットされたクラスタをタッチすることで投稿の絞り込みができる機能、クラスタに自分でタグつけする機能・画像に文字が入っているものを除く機能・動画と画像の要約
【ベースラインシステム群の回答詳細】
<ul style="list-style-type: none">・画像を優先して見るのももう少しサイズを大きく捨て見やすくすることや動画はループ再生していてもいいと思った。・特になし・画像を見てゲーム画面であれば排除する機能・明らかに災害と関係ないものは初めから内容が探しやすいなと思いました。・機能：文字画像やゲーム画像を排除する機能（自動）すぐに判別できる画像を人間が時間をかけて判別するのは勿体無いと思ったから。・拾い画かどうかの判別、地域の絞り込み・特定のワードや画像をフィルタリングする・画面内に“カテゴリー”のような分類があれば良いと感じた。・特になし・不要な言葉を除外する機能・「父からもらった写真」としてアップする人もいるため、被災者本人の投稿かわかるようになる嬉しい。・場所を特定できる機能。指定された場所以外の投稿がみられたため。

参考文献

- [1] Abavisani, M., Wu, L., Hu, S., Tetreault, J. and Jaimes, A.: Multimodal Categorization of Crisis Events in Social Media, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, DOI: 10.1109/CVPR42600.2020.01469 (2020).
- [2] Aïmeur, E., Amri, S. and Brassard, G.: Fake news, disinformation and misinformation in social media: a review, *Social Network Analysis and Mining*, Vol. 13, No. 1, p. 30, DOI: 10.1007/s13278-023-01028-5 (2023).
- [3] Alam, F., Offi, F. and Imran, M.: Crisismmd: Multimodal twitter datasets from natural disasters, *Proceedings of the international AAAI conference on web and social media*, Vol. 12, No. 1, DOI: 10.1609/icwsm.v12i1.14983 (2018).
- [4] Ashktorab, Z., Brown, C., Nandi, M. and Culotta, A.: Tweedr: Mining twitter to inform disaster response, *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management*, pp. 269–272 (2014).
- [5] Bishop, G. and Welch, G.: An introduction to the kalman filter, *SIGGRAPH 2001, Course 8* (2001).
- [6] Chowdhury, J. R., Caragea, C. and Caragea, D.: On identifying hashtags in disaster twitter data, *Proceedings of the AAAI conference on artificial intelligence*, Vol. 34, No. 01, pp. 498–506, DOI: 10.1609/aaai.v34i01.5387 (2020).
- [7] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding, *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, Vol. 1 (long and short papers), pp. 4171–4186, DOI: 10.18653/v1/N19-1423 (2019).
- [8] Djuric, P. M., Kotecha, J. H., Zhang, J., Huang, Y., Ghirmai, T., Bugallo, M. F. and Miguez, J.: Particle filtering, *IEEE signal processing magazine*, Vol. 20, No. 5, pp. 19–38, DOI: 10.1109/MSP.2003.1236770 (2003).
- [9] Dosovitskiy, A.: An image is worth 16x16 words: Transformers for image recognition at scale, *arXiv preprint arXiv:2010.11929*, DOI: 10.48550/arXiv.2010.11929 (2020).
- [10] Drolsbach, C. P., Solovev, K. and Pröllochs, N.: Community notes increase trust in fact-checking on social media, *PNAS Nexus*, Vol. 3, No. 7, p. pgae217, DOI: 10.1093/pnas-nexus/pgae217 (2024).
- [11] Durall, R., Keuper, M. and Keuper, J.: Watch your up-convolution: Cnn based generative deep neural networks are failing to reproduce spectral distributions, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7890–7899, DOI: 10.1109/CVPR42600.2020.00791 (2020).

- [12] FirstDraftNews: Fake news. It's complicated. (2017). <https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79> (2025/11/29 確認) .
- [13] Frank, J., Herbert, F., Ricker, J., Schonherr, L., Eisenhofer, T., Fischer, A., Durmuth, M. and Holz, T.: A Representative Study on Human Detection of Artificially Generated Media Across Countries, *2024 IEEE Symposium on Security and Privacy*, IEEE Computer Society, pp. 55–73, DOI: 10.1109/SP54263.2024.00159 (2024).
- [14] Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A. and Chakraborty, B.: A review on application of data mining techniques to combat natural disasters, *Ain Shams Engineering Journal*, Vol. 9, No. 3, pp. 365–378, DOI: 10.1016/j.asej.2016.01.012 (2018).
- [15] Herceg, P. M., Allison, T. B., Belvin, R. B. and Tzoukermann, E.: Collaborative exploratory search for information filtering and large-scale information triage, *Association for Information Science and Technology*, Vol. 69, No. 3, pp. 395–409, DOI: 10.1002/asi.23961 (2018).
- [16] Hoque, M. A., Ferdous, M. S., Khan, M. and Tarkoma, S.: Real, Forged or Deep Fake? Enabling the Ground Truth on the Internet, *IEEE Access*, Vol. 9, pp. 160471–160484, DOI: 10.1109/ACCESS.2021.3131517 (2021).
- [17] Houston, J. B., Hawthorne, J., Perreault, M. F., Park, E. H., Goldstein Hode, M., Halliwell, M. R., Turner McGowen, S. E., Davis, R., Vaid, S., McElderry, J. A. and Griffith, S. A.: Social media and disasters: a functional framework for social media use in disaster planning, response, and research, *Disasters*, Vol. 39, No. 1, pp. 1–22, DOI: 10.1111/disa.12092 (2015).
- [18] Huang, B. and Carley, K. M.: A large-scale empirical study of geotagging behavior on Twitter, *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 365–373, DOI: 10.1145/3341161.3342870 (2019).
- [19] Hunt, K., Agarwal, P. and Zhuang, J.: Monitoring misinformation on Twitter during crisis events: a machine learning approach, *Risk analysis*, Vol. 42, No. 8, pp. 1728–1748, DOI: 10.1111/risa.13634 (2022).
- [20] Hunt, K., Wang, B. and Zhuang, J.: Misinformation debunking and cross-platform information sharing through Twitter during Hurricanes Harvey and Irma: a case study on shelters and ID checks, *Natural Hazards*, Vol. 103, No. 1, pp. 861–883, DOI: 10.1007/s11069-020-04016-6 (2020).
- [21] Jain, T., Gopalani, D. and Kumar Meena, Y.: Informative task classification with concatenated embeddings using deep learning on crisisMMD, *International Journal of Computers and Applications*, pp. 1–18 (2025).

- [22] Java, A., Song, X., Finin, T. and Tseng, B.: Why we twitter: understanding microblogging usage and communities, *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pp. 56–65, DOI: 10.1145/1348549.1348556 (2007).
- [23] Jia, C., Yang, Y., Xia, Y., Chen, Y.-T., Parekh, Z., Pham, H., Le, Q., Sung, Y.-H., Li, Z. and Duerig, T.: Scaling Up Visual and Vision-Language Representation Learning With Noisy Text Supervision, *Proceedings of the 38th International Conference on Machine Learning* (Meila, M. and Zhang, T.(eds.)), Proceedings of Machine Learning Research, No. 139, pp. 4904–4916 (2021).
- [24] Jiang, J., Thomason, J., Barbieri, F. and Ferrara, E.: Geolocated social media posts are happier: Understanding the characteristics of check-in posts on twitter, *Proceedings of the 15th ACM Web Science Conference 2023*, pp. 136–146, DOI: 10.1145/3578503.3583596 (2023).
- [25] Kamali, N., Nakamura, K., Kumar, A., Chatzimparmpas, A., Hullman, J. and Groh, M.: Characterizing Photorealism and Artifacts in Diffusion Model-Generated Images, *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–26, DOI: 10.1145/3706598.3713962 (2025).
- [26] Kleinberg, J.: Bursty and hierarchical structure in streams, *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 91–101, DOI: 10.1145/775047.775061 (2002).
- [27] Li, S., Liu, Q., Pan, Z. and Wu, X.: CLIP-BCA-Gated: A Dynamic Multimodal Framework for Real-Time Humanitarian Crisis Classification with Bi-Cross-Attention and Adaptive Gating, *Applied Sciences*, Vol. 15, No. 15, p. 8758, DOI: 10.3390/app15158758 (2025).
- [28] Liu, R., Garrette, D., Saharia, C., Chan, W., Roberts, A., Narang, S., Blok, I., Mical, R., Norouzi, M. and Constant, N.: Character-aware models improve visual text rendering, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, Vol. 1 (Long Papers), pp. 16270–16297, DOI: 10.18653/v1/2023.acl-long.900 (2023).
- [29] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V.: Roberta: A robustly optimized bert pretraining approach, *arXiv preprint arXiv:1907.11692*, DOI: 10.48550/arXiv.1907.11692 (2019).
- [30] Liu, Z., Wang, S., Jia, Z., Tang, Y., Zhi, G. and Wang, X.: Bidirectional Cross-Attention Domain Adaption Transformer Network for Aircraft EMA Fault Diagnosis Under Varying Working Conditions, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 61, No. 4, pp. 9584–9600, DOI: 10.1109/TAES.2025.3556808 (2025).
- [31] Lu, Z., Huang, D., Bai, L., Qu, J., Wu, C., Liu, X. and Ouyang, W.: Seeing is not always believing: Benchmarking human and model perception of ai-

- generated images, *Advances in neural information processing systems*, Vol. 36, pp. 25435–25447 (2023). https://proceedings.neurips.cc/paper_files/paper/2023/file/505df5ea30f630661074145149274af0-Paper-Datasets_and_Benchmarks.pdf (2025/11/29 確認).
- [32] Mandal, B., Khanal, S. and Caragea, D.: Contrastive learning for multimodal classification of crisis related tweets, *Proceedings of the ACM Web Conference 2024*, pp. 4555–4564, DOI: 10.1145/3589334.3648143 (2024).
- [33] Mann, H. B. and Whitney, D. R.: On a test of whether one of two random variables is stochastically larger than the other, *The annals of mathematical statistics*, Vol. 18, No. 1, pp. 50–60 (1947).
- [34] Marshall, C. C. and Shipman III, F. M.: Spatial hypertext and the practice of information triage, *Proc. 8th ACM Conference on Hypertext*, pp. 124–133, DOI: 10.1145/267437.267451 (1997).
- [35] Mizuno, J., Tanaka, M., Ohtake, K., Oh, J.-H., Kloetzer, J., Hashimoto, C. and Torisawa, K.: WISDOM X, DISAANA and D-SUMM: Large-scale NLP systems for analyzing textual big data, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pp. 263–267 (2016).
- [36] Mizunuma, Y., Yamamoto, S., Yamaguchi, Y., Ikeuchi, A., Satoh, T. and Shimada, S.: Twitter bursts: Analysis of their occurrences and classifications, *Proc. 8th International Conference on Digital Society*, pp. 182–187 (2014). https://www.thinkmind.org/library/ICDS/ICDS_2014/icds_2014_7_40_10139.html (2025/11/29 確認).
- [37] Mogi, N., Morino, Y. and Matsushita, M.: Analysis of the changes in the attitude of the news comments caused by knowing that the comments were generated by a large language model, *12th International Conference on Informatics, Electronics & Vision*, No. 32, pp. 1–6 (2025).
- [38] Morino, Y., Matsushita, M. and Fujishiro, H.: Vocabulary cross-contamination between entertainment content and disaster-related social media posts, *2024 International Conference on Information and Communication Technologies for Disaster Management*, pp. 1–6, DOI: 10.1109/ICT-DM62768.2024.10798944 (2024).
- [39] Morino, Y., Megumi, Y., Matsushita, M. and Fujishiro, H.: Investigating the influence of Web-Media in disaster situations by analyzing diffusion of tweets, *The 34th Annual Conference of the Japanese Society for Artificial Intelligence*, No. 3G5ES103 (2020).
- [40] Mouzannar, H., Rizk, Y. and Awad, M.: Damage Identification in Social Media Posts using Multimodal Deep Learning, *Proceedings of the 15th International Conference on Information Systems for Crisis Response and Management* (2018). <https://idl.iscram.org/>

files/husseinmouzannar/2018/2129_HusseinMouzannar_etal2018.pdf (2025/11/29 確認) .

- [41] Muhammed T, S. and Mathew, S. K.: The disaster of misinformation: a review of research in social media, *International journal of data science and analytics*, Vol. 13, No. 4, pp. 271–285, DOI: 10.1007/s41060-022-00311-6 (2022).
- [42] Nguyen, D. Q., Vu, T. and Nguyen, A.-T.: BERTweet: A pre-trained language model for English Tweets, *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pp. 9–14, DOI: 10.18653/v1/2020.emnlp-demos.2 (2020).
- [43] Nightingale, S. J. and Farid, H.: AI-synthesized faces are indistinguishable from real faces and more trustworthy, *Proceedings of the National Academy of Sciences*, Vol. 119, No. 8, p. e2120481119, DOI: 10.1073/pnas.2120481119 (2022).
- [44] Pirolli, P. and Card, S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis, *Proceedings of international conference on intelligence analysis* (2005). https://www.researchgate.net/publication/215439203_The_sensemaking_process_and_leverage_points_for_analyst_technology_as_identified_through_cognitive_task_analysis (2025/11/29 確認) .
- [45] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G. and Sutskever, I.: Learning transferable visual models from natural language supervision, *International conference on machine learning*, pp. 8748–8763 (2021). <https://proceedings.mlr.press/v139/radford21a> (2025/11/29 確認).
- [46] Rezk, M., Elmadany, N., Hamad, R. K. and Badran, E. F.: Categorizing crises from social media feeds via multimodal channel attention, *IEEE Access*, Vol. 11, pp. 72037 – 72049, DOI: 10.1109/ACCESS.2023.3294474 (2023).
- [47] Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B.: High-resolution image synthesis with latent diffusion models, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695 (2022).
- [48] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors, *Proceedings of the 19th international conference on World wide web*, pp. 851–860, DOI: 10.1145/1772690.1772777 (2010).
- [49] Sarkar, A., Mai, H., Mahapatra, A., Lazebnik, S., Forsyth, D. A. and Bhattad, A.: Shadows don’t lie and lines can’t bend! generative models don’t know projective geometry... for now, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 28140–28149, DOI: 10.1109/CVPR52733.2024.02658 (2024).

- [50] Shimauchi, T., Taguchi, N., Nambo, H. and Kimura, H.: A study on extracting disaster information from tweets, *Journal of Global Tourism Research*, Vol. 2, No. 2, pp. 93–98, DOI: 10.37020/jgtr.2.2_93 (2017).
- [51] Social Media Working Group for Emergency Services and Disaster Management: Countering False Information on Social Media in Disasters and Emergencies, Technical report, United States Department of Homeland Security (2018). <https://www.dhs.gov/archive/science-and-technology/publication/st-frg-countering-false-information-social-media-disasters-and-emergencies> (2025/11/29 確認) .
- [52] Stowe, K., Paul, M., Palmer, M., Palen, L. and Anderson, K. M.: Identifying and categorizing disaster-related tweets, *Proceedings of The fourth international workshop on natural language processing for social media*, pp. 1–6, DOI: 10.18653/v1/W16-6201 (2016).
- [53] Tekumalla, R. and Banda, J. M.: TweetDIS: A Large Twitter Dataset for Natural Disasters Built using Weak Supervision, *2022 IEEE International Conference on Big Data (Big Data)*, pp. 4816–4823, DOI: 10.1109/BigData55660.2022.10020214 (2022).
- [54] Urbani, S.: Verifying Online Information Essential Guides, <https://firstdraftnews.org/long-form-article/verifying-online-information/> (2020). First Draft. Unless otherwise noted, content licensed under CC BY 4.0.
- [55] Verma, S., Vieweg, S., Corvey, W., Palen, L., Martin, J., Palmer, M., Schram, A. and Anderson, K.: Natural language processing to the rescue? extracting “situational awareness” tweets during mass emergency, *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 5, No. 1, pp. 385–392, DOI: 10.1609/icwsm.v5i1.14119 (2011).
- [56] 相田慎, 新堂安孝, 内山将夫: 「東日本大震災関連の救助要請情報抽出サイト」による救助活動支援, 自然言語処理, Vol. 20, No. 3, pp. 405–422 (2013).
- [57] 朝日新聞: 「人工地震」「窃盗団」繰り返される災害デマ 削除だけでは不十分, <https://www.asahi.com/articles/ASS2Q4R9FS2FUTIL00M.html> (2024). (2025/11/29 確認) .
- [58] 白井翔平, 鳥海不二夫, 平山高嗣, 榎堀優, 間瀬健二: なぜ震災後デマが拡散したのか — ネットワーク構造の影響分析 —, 電気学会論文誌 C (電子・情報・システム部門誌), Vol. 133, No. 9, pp. 1796–1805 (2013).
- [59] NTT 東日本: 通信サービスの使命 第 1 回 大規模災害と通信ネットワーク, <https://www.ntt-east.co.jp/universal/web1-1.html> (2004). (2025/11/29 確認) .
- [60] 大島裕明, 中村聡史, 田中克己: SlothLib: Web 検索研究のためのプログラミングライブラリ, 日本データベース学会, Vol. 6, pp. 113–116 (2007).

- [61] 大竹清敬: 災害時における DISAANA、D-SUMM の活用～DISAANA・D-SUMM と防災チャットボット SOCDA～, https://www.soumu.go.jp/main_content/000672984.pdf (2020).
- [62] 小笠原盛浩, 川島浩誉, 藤代裕之: マスメディア報道は Twitter 上の災害時流言を抑制できたか? : 2011 年東日本大震災におけるコスモ石油流言の定性的分析, 関西大学社会学部紀要, Vol. 49, No. 2, pp. 121–140 (2018).
- [63] 株式会社インプレス R&D: インターネット白書 1996 (1996).
- [64] 川上善郎: 阪神・淡路大震災はどのように語られたのか: パソコン通信における会話, コミュニケーション紀要, Vol. 14, pp. 61–80 (2001).
- [65] 川村壮, 佐々木優二: 大規模災害発生時の SNS における誤情報拡散と自治体による対応の課題, 情報処理学会研究報告, Vol. 2022-IS-162, No. 2, pp. 1–8 (2022).
- [66] 川村壮, 佐々木優二, 戸松誠: 大規模災害発生時の自治体による SNS を利用した情報収集・情報発信に関する研究, 北海道立総合研究機構建築研究本部調査研究報告, Vol. 43 (2023).
- [67] 國友優, 神山嬢子: Twitter 情報を活用した土砂災害の前兆・発生状況把握の可能性, 土木技術資料, Vol. 57, No. 9, pp. 18–21 (2015).
- [68] 国土交通省: 全国の土砂災害警戒区域等の指定状況推移 (R7.3.31 時点), <https://www.mlit.go.jp/mizukokudo/sabo/content/001888481.pdf> (2025). (2025/11/29 確認).
- [69] 国土交通省気象庁: 国土交通白書 2021, 第 1 節 現在直面する危機, 2 災害の激甚化・頻発化, <https://www.mlit.go.jp/hakusyo/mlit/r02/hakusho/r03/html/n1112000.html> (2021). (2025/11/29 確認).
- [70] 国土交通省気象庁: 日本の気候変動 2025 —大気と陸・海洋に関する観測・予測評価報告書— (HTML 版), https://www.data.jma.go.jp/cpdinfo/ccj/2025/html_honpen/cc2025_honpen_5.html (2025). (2025/11/29 確認).
- [71] 榊剛史, 松尾豊: ソーシャルセンサとしての Twitter: ソーシャルセンサは物理センサを凌駕するか? (<特集>Twitter とソーシャルメディア), 人工知能学会誌, Vol. 27, No. 1, pp. 67–74 (2012).
- [72] 信濃通信毎日デジタル: 能登地震直後、偽情報で消防出動 SNS で虚偽の救助要請, <https://www.shinmai.co.jp/news/article/CNTS2024020801000>. (2025/11/29 確認).
- [73] 篠田孝祐, 榊剛史, 鳥海不二夫, 風間一洋, 栗原聡, 野田五十樹, 松尾豊: 東日本大震災時における Twitter の活用状況とコミュニケーション構造の分析, 知能と情報, Vol. 25, No. 1, pp. 598–608 (2013).
- [74] 芝勝徳: 阪神・淡路大震災とインターネット, 情報管理, Vol. 41, No. 8, pp. 593–601 (1998).

- [75] 澁谷遊野, 中里朋楓: 令和6年能登半島地震におけるデジタル空間の偽誤情報流通状況の報告, 総務省: デジタル空間における情報流通の健全性確保の在り方に関する検討会資料, Vol. 10-2 (2024).
- [76] 島田政信: 火山観測における現状や未来について, 日本リモートセンシング学会誌, Vol. 41, No. 2, pp. 217-223, DOI: 10.11440/rssj.41.217 (2021).
- [77] ジチタイワークス Web: 台風時に SNS で住民を励まし迅速かつ無事に約 50 件救助へ, <http://jichitai.works/articles/330> (2020). (2025/11/29 確認) .
- [78] 杉原健一郎, 石野航平, 松下光範: 災害情報を対象とした意思決定支援システムの検討: 複数人による協調的情報トリアージを目指して, Vol. 4, pp. 77-80 (2013).
- [79] 杉山正平: 地方公共団体の災害対応における SNS 活用, 電子情報通信学会 通信ソサイエティマガジン, Vol. 13, No. 4, pp. 289-295 (2020).
- [80] 宋晨潔, 藤代裕之: 災害時における信頼性の高い救助要請の見つけ方～西日本豪雨「救助」ツイートの検証～, 電子情報通信学会技術研究報告, Vol. 118, No. 439, pp. 7-12 (2019).
- [81] 宋晨潔, 藤代裕之: 救助要請ツイートの特徴の検証 —令和2年7月豪雨を対象に—, 電子情報通信学会技術研究報告, Vol. 120, No. 166, pp. 18-23 (2020).
- [82] 総務省: インターネットと IT 革命, 平成27年版 情報通信白書, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc111210.html>. (2025/11/29 確認) .
- [83] 総務省情報通信政策研究所: 令和4年度情報通信メディアの利用時間と情報行動に関する調査報告書<概要>, https://www.soumu.go.jp/main_content/000887659.pdf. (2025/11/29 確認) .
- [84] 総務省消防庁: 大規模災害時におけるソーシャル・ネットワーキング・サービスによる緊急通報の活用可能性に関する検討会報告書(案), https://www.fdma.go.jp/singi_kento/kento/items/sns04/shiryo02.pdf (2013). (2025/11/29 確認) .
- [85] 総務省: プラットフォームサービスに関する研究会 最終報告書, https://www.soumu.go.jp/main_content/000765135.pdf (2020).
- [86] 高橋隆雄: トリアージの倫理, 人間と医療, Vol. 5, pp. 4-11 (2015).
- [87] 谷口慎一郎: 災害時における Twitter の有用性について—2011年9月の台風12号による豪雨災害を例に, 災害情報, Vol. 10, pp. 56-67 (2012).
- [88] 東北放送株式会社: ロンドンからの投稿が、気仙沼で避難する446人の命を救った…あの日の教訓は「5W1Hと信頼関係」【DIG 防災】, <https://newsdig.tbs.co.jp/articles/tbc/1082042?display=1> (2024). (2025/11/29 確認) .
- [89] 鳥海不二夫: バースト現象における拡散の定量分析—ツイッターデモはどう広がったか—, マーケティングジャーナル, Vol. 40, No. 4, pp. 19-32 (2021).

- [90] 内閣府: 防災・減災、国土強靱化新時代の実現のための提言, デジタル・防災技術ワーキンググループ 未来構想チーム提言, https://www.bousai.go.jp/kaigirep/teigen/pdf/teigen_03.pdf (2021). (2025/11/29 確認) .
- [91] 林信行, 山路達也: 東日本大震災と情報, インターネット, Google: パーソンファイnder、東日本大震災での進化(1), https://www.google.org/crisisresponse/kiroku311/chapter_06.html (2012). (2025/11/29 確認) .
- [92] 広谷徹: 阪神大震災と放送・通信, 安全工学, Vol. 35, No. 1, pp. 57–67 (1996).
- [93] 福長秀彦: 流言・デマ・フェイクニュースとマスメディアの打ち消し報道, 放送研究と調査, Vol. 68, No. 11, pp. 84–103, DOI: 10.24634/bunken.68.11.84 (2018).
- [94] 藤代裕之, 松下光範, 小笠原盛浩: 大規模災害時におけるソーシャルメディアの活用—情報トリアージの適用可能性, 社会情報学, Vol. 6, No. 2, pp. 49–63 (2018).
- [95] 藤代裕之: 偽・誤情報とフェイクニュース対策の方向性, 情報処理学会研究報告, Vol. EIP-104, No. 20, pp. 1–5 (2024).
- [96] 毎日新聞: ツイッター投稿、1週間で2610万件, <https://mainichi.jp/articles/20160519/k00/00m/040/059000c> (2016). (2025/11/29 確認).
- [97] 三浦麻子: 東日本大震災とオンラインコミュニケーションの社会心理学: そのときツイッターでは何が起こったか, 電子情報通信学会誌, Vol. 95, No. 3, pp. 219–223 (2012).
- [98] 宮本聖二: 能登半島地震をめぐる偽・誤情報: ファクトチェックとデジタルアーカイブ, デジタルアーカイブ学会誌, Vol. 8, No. 4, pp. 157–162 (2024).
- [99] 安尾萌, 藤代裕之, 松下光範: 協調的情報トリアージにおけるメンバ間のコミュニケーションが情報峻別の結果に与える影響の検討, 電子情報通信学会論文誌 D, Vol. 103, No. 5, pp. 382–392 (2020).
- [100] 山口真一, 彌永浩太郎: 災害時のソーシャルメディア活用と課題—民間サービス・教育・マスメディアの変革による解決を—, *GLOCOM OPINION PAPER*, No. 6 (2016). https://www.glocom.ac.jp/wp-content/uploads/2016/11/OpinionPaper2016_No.6.pdf (2025/11/29 確認).
- [101] 山本佳世子: 災害発生時におけるソーシャルメディアの利活用と課題, 消防防災の科学, No. 152, pp. 38–41 (2023).

業績一覧

学術論文

- [1] 森野 穰, 藤代 裕之, 松下 光範. 信憑性の高い救助要請ツイートを抽出するためのノイズ除去手法の検討, *情報通信学会誌*, Vol.43, No.1, pp.55-62 (2025)
- [2] 森野 穰, 藤代 裕之, 松下 光範. 災害時におけるソーシャルメディアからの情報収集を目的としたノイズ情報分離に関する調査, *ヒューマンインタフェース学会論文誌*, Vol.27, No.3, pp.133-140 (2025)
- [3] Nanase Mogi, Megumi Yasuo, Yutaka Morino, Mitsunori Matsushita. Analysis of the changes in the attitude of the news comments caused by knowing that the comments were generated by a large language model, *International Journal of Activity and Behavior Computing*, Vol.2025, No.3, pp.1-13 (2025)

国際会議論文

- [4] Kanna Miyagawa, Yutaka Morino, Mitsunori Matsushita. Interactive visualization of comic character correlation diagrams for understanding character relationships and personalities, *Proc. the 24th International Conference on Information Integration and Web Intelligence*, pp.561-567 (2022)
- [5] Yutaka Morino, Hiroyuki Fujishiro, Mitsunori Matsushita. Vocabulary Cross-Contamination between Entertainment Content and Disaster-Related Social Media Posts, *2024 International Conference on Information and Communication Technologies for Disaster Management*, pp.1-6 (2024)
- [6] Nanase Mogi, Megumi Yasuo, Yutaka Morino, Mitsunori Matsushita. Human or LLM? Distinguishing Online Comments by Emotion and Tone, *The 30th International Conference on Technologies and Applications of Artificial Intelligence* (2025)

口頭発表等

- [7] 森野 穰, 安尾 萌, 松下 光範, 藤代 裕之. ウェブメディアが SNS に与える影響の調査—2019 年台風 19 号のツイートデータを対象に—, 第 12 回データ工学と情報マネジメントに関するフォーラム, No.C8-4(2020)
- [8] Yutaka Morino, Megumi Yasuo, Mitsunori Matsushita, Hiroyuki Fujishiro. Investigating the influence of Web-Media in disaster situations by analyzing diffusion of tweets, *The 34th Annual Conference of the Japanese Society for Artificial Intelligence*, No.3G5-ES-1-03(2020)

- [9] Yutaka Morino, Megumi Yasuo, Hiroyuki Fujishiro, Mitsunori Matsushita. How Web-Media impacts on twitter in disaster situation —Analysis of Actual Tweets during Typhoon Hagibis Attacking Japan—, *Asian student seminar round table*, No.2E-1, pp.345-351(2020)
- [10] 森野 穰, 安尾 萌, 松下 光範, 藤代 裕之. Twitter に投稿された画像の分類に基づくツイート文の傾向分析, 第 13 回データ工学と情報マネジメントに関するフォーラム, No. I25-1(2021)
- [11] 森野 穰, 松下 光範. 災害情報収集におけるエンタテインメントコンテンツによるコンタミネーションの調査, 情報処理学会研究報告, Vol.2022-EC-65, No.33, pp.1-2(2022)
- [12] 安尾 萌, 森野 穰, 松下 光範. 災害情報収集における SNS のメディア特性に関する一検討, 人工知能学会第 30 回インタラクティブ情報アクセスと可視化マイニング研究会予稿集, pp.47-54 (2023)
- [13] 畑 玲音, 森野 穰, 松下 光範. 皮肉文検出のための皮肉状況の検出, 2023 年度人工知能学会全国大会 (第 37 回) 論文集, No.3M1-GS-10-03 (2023)
- [14] 中谷 智生, 田口 潤智, 松下 光範, 森野 穰. オンライン学会と対面学会参加者の発信内容の特徴の分析 —Twitter データにおける頻出語句の比較—, 第 21 回日本神経理学療法学会学術大会, No.P36-7, pp.379 (2023)
- [15] 森野 穰, 藤代 裕之, 松下 光範. 災害情報トリアージのための SNS 情報峻別システムの構築, 情報処理学会研究報告, Vol.2025-HCI-216 (2025)

学位論文に関連する既発表論文

本論文における各章の主たる内容については、下記に示す論文において公表済みである。

1章

- 森野 穰, 藤代 裕之, 松下 光範. 信憑性の高い救助要請ツイートを抽出するためのノイズ除去手法の検討, 情報通信学会誌, Vol.43, No.1, pp.55-62, 2025.

4章

- 森野 穰, 松下 光範. 災害情報収集におけるエンタテインメントコンテンツによるコンタミネーションの調査, 情報処理学会研究報告, Vol.2022-EC-65, No.33, pp.1-2, 2022.
- 安尾 萌, 森野 穰, 松下 光範. 災害情報収集における SNS のメディア特性に関する一検討, 人工知能学会第 30 回インタラクティブ情報アクセスと可視化マイニング研究会予稿集, pp.47-54, 2023.
- Morino, Yutaka and Matsushita, Mitsunori and Fujishiro, Hiroyuki. “Vocabulary cross-contamination between entertainment content and disaster-related social media posts”, 2024 International Conference on Information and Communication Technologies for Disaster Management, pp.1-6, 2024.
- 森野 穰, 藤代 裕之, 松下 光範. 信憑性の高い救助要請ツイートを抽出するためのノイズ除去手法の検討, 情報通信学会誌, Vol.43, No.1, pp.55-62, 2025.

5章

- 森野 穰, 藤代 裕之, 松下 光範. 災害時におけるソーシャルメディアからの情報収集を目的としたノイズ情報分離に関する調査, ヒューマンインタフェース学会論文誌, Vol.27, No.3, pp.133-140, 2025.

6章

- 森野 穰, 藤代 裕之, 松下 光範. 災害情報トリアージのための SNS 情報峻別システムの構築, 情報処理学会研究報告, Vol.2025-HCI-216, 2026.