

AI-Powered Decision Support System for Forest Fire Rescue Using Twitter and GPT-Based Analysis

Özge DOĞUÇ 

Doç. Dr., İstanbul Medipol Üniversitesi, İşletme ve Yönetim Bilimleri Fakültesi, İstanbul, Türkiye / Assoc. Prof., İstanbul Medipol University, School of Business and Management Sciences, İstanbul, Türkiye

ABSTRACT

This study presents the Forest Fire Rescue Decision Support System (FFR-DSS), an AI-powered framework integrating GPT-4-based semantic analysis with real-time Twitter data to support emergency rescue coordination during forest fires in Turkey. Using tweets from the catastrophic 2021 wildfires affecting Muğla and Antalya, FFR-DSS employs prompt-engineered large language model (LLM) pipelines to extract urgency scores, emotional tone, geographical information, and actionability indicators from crisis-related social media content. Unlike traditional natural language processing (NLP) approaches such as Bag-of-Words (BoW) and TF-IDF, which may struggle with contextually rich and emotionally variable messages, the proposed framework leverages GPT-4's zero-shot semantic reasoning capabilities without requiring task-specific labeled training datasets. A geospatial risk mapping component further aggregates tweet-level outputs into district-level risk scores and dynamic heatmaps, enabling emergency teams to identify critical areas and prioritize resource allocation. Evaluation against a manually annotated sample demonstrates that FFR-DSS achieves 92.4% classification precision and 89.1% recall, substantially outperforming the BoW+SVM baseline with 72.8% accuracy. These results demonstrate the potential of LLM-augmented social media analytics to transform unstructured crisis communications into actionable intelligence, improving situational awareness, rescue coordination, and data-driven decision-making during large-scale natural disasters.

Keywords: Decision support system, Emergency response, Forest fire rescue, Geospatial intelligence, GPT-4, Twitter analytics

Orman Yangını Kurtarma Operasyonları İçin Twitter ve GPT Tabanlı Analiz Kullanan Yapay Zekâ Destekli Karar Destek Sistemi

Öz

Bu çalışma, Türkiye'deki orman yangınları sırasında acil kurtarma koordinasyonunu desteklemek amacıyla GPT-4 tabanlı anlamsal analiz ile gerçek zamanlı Twitter verilerini bütünleştiren yapay zekâ destekli Orman Yangını Kurtarma Karar Destek Sistemi'ni (FFR-DSS) sunmaktadır. Sistem, Muğla ve Antalya illerini etkileyen 2021 Türkiye orman yangınlarına ilişkin tweet verileri üzerinde uygulanmıştır. FFR-DSS, kriz odaklı sosyal medya içeriklerinden aciliyet düzeyi, duygusal ton, konum bilgisi ve eyleme geçirilebilirlik göstergelerini çıkarmak amacıyla istem mühendisliğiyle tasarlanmış büyük dil modeli (BDM) tabanlı bir analiz hattı kullanmaktadır. Kelime Torbası (BoW) ve TF-IDF gibi geleneksel doğal dil işleme yöntemleri, bağlamsal açıdan zengin ve duygusal olarak değişken sosyal medya mesajlarını yorumlamada sınırlı kalabilmektedir. Buna karşılık FFR-DSS, etiketli eğitim verisine ihtiyaç duymadan GPT-4'ün sıfır atışlı anlamsal akıl yürütme yeteneklerinden yararlanmaktadır. Sistemin coğrafi risk haritalama bileşeni, tweet düzeyindeki sonuçları ilçe bazında risk puanlarına dönüştürmekte ve dinamik ısı haritaları oluşturarak kaynak tahsisi ve önceliklendirme kararlarını desteklemektedir. Elle etiketlenmiş bir örneklem üzerinde yapılan değerlendirmede FFR-DSS, %92,4 sınıflandırma kesinliği ve %89,1 duyarlılık elde etmiş; %72,8 doğruluk sağlayan BoW+SVM temel yöntemini önemli ölçüde geride bırakmıştır. Bulgular, BDM destekli sosyal medya analitiğinin afet yönetiminde gerçek zamanlı durumsal farkındalık, kurtarma koordinasyonu ve karar desteği için etkili bir yaklaşım olduğunu göstermektedir.

Anahtar Kelimeler: Acil durum müdahalesi, Coğrafi zekâ, GPT-4, Karar destek sistemi, Orman yangını kurtarma, Twitter analitiği

1. INTRODUCTION

Turkey is among the countries most exposed to recurring and increasingly severe forest fire events, driven by a combination of Mediterranean climate conditions, prolonged drought periods, and expanding wildland-urban interface zones (Bilgili et al., 2019; Daşdemir et al., 2021). The catastrophic summer of 2021 stands as a defining turning point: fires burned across more than 200,000 hectares, with the provinces of Muğla and Antalya suffering unprecedented destruction (AFAD, 2021; OGM, 2021). These events overwhelmed traditional emergency response infrastructures and exposed critical gaps in real-time situational awareness and rescue coordination.

Social media platforms such as Twitter (now rebranded as X) in particular, have emerged as vital informal communication channels during crises (Imran et al., 2015; Vieweg et al., 2010). Affected individuals share spontaneous, location-specific reports, calls for help, and real-time observations that often precede or supplement official information channels. Researchers have demonstrated the utility of Twitter data in earthquake response (Imran et al., 2013), flood monitoring (Zou et al., 2018), and wildfire situational awareness (Verma et al., 2011; Huang and Xiao, 2015). However, extracting actionable intelligence from the noisy, informal, and multilingual stream of crisis tweets remains a significant technical challenge.

Traditional NLP methods including Bag-of-Words (BoW), TF-IDF vectorization, and supervised classifiers such as Support Vector Machines have been widely applied for disaster tweet classification (Doguc, 2022; Alam et al., 2018). While effective on well-labeled benchmark datasets, these approaches are limited in semantic depth, require extensive preprocessing and domain-specific training data, and struggle with the linguistic variability, sarcasm, and emotional nuance present in real crisis communications (Imran et al., 2020).

The emergence of transformer-based Large Language Models (LLMs), particularly GPT-4 (OpenAI, 2023), has substantially advanced natural language understanding capabilities. These models can perform complex semantic reasoning, zero-shot classification, named entity recognition, and emotional tone detection without task-specific fine-tuning, making them uniquely suited to the high-stakes, low-resource conditions of disaster response analytics (Nenkova and McKeown, 2011; Brown et al., 2020).

In this paper, we present the Forest Fire Rescue Decision Support System (FFR-DSS), which integrates GPT-4-powered semantic analysis with Twitter data collected during Turkey's 2021 wildfire season to support emergency rescue coordination. The system extracts urgency scores, emotional tone, inferred locations, and actionability flags from crisis tweets, and visualizes aggregated risk through geospatial heatmaps. We evaluate the system against a manually annotated sample and compare it with a BoW+SVM baseline, demonstrating substantial improvements in both precision and interpretability.

2. LITERATURE REVIEW

The use of social media for disaster situational awareness has been extensively studied over the past decade. Twitter's immediacy, geographic richness, and volume make it a particularly valuable data source during emergencies, where conventional reporting channels are often delayed or overwhelmed (Imran et al., 2015; Castillo et al., 2011).

Early work by Vieweg et al. (2010) demonstrated that Twitter users generate situationally useful information during crises such as wildfires and floods. Subsequent studies applied text classification approaches to identify rescue-relevant content: Alam et al. (2018) applied convolutional neural networks to hurricane-related tweets, while Zou et al. (2018) used topic modeling and geospatial analysis to assess community resilience from social media data following natural disasters. Beyond crisis-specific applications, Twitter data has also been used to infer latent user characteristics such as socioeconomic status through NLP-based

methods (Ghazouani et al., 2019), illustrating the broader potential of social media analytics for understanding user behavior at scale, a foundation that disaster informatics systems like FFR-DSS build upon.

In the Turkish context, Doguc (2022) developed a Twitter-based decision support system for natural disaster response, applying text mining and risk scoring to tweets from the 2021 Turkish wildfires. That study demonstrated the feasibility of automated disaster tweet classification but highlighted the limitations of frequency-based approaches in handling semantic ambiguity and linguistic variation.

The advent of transformer-based LLMs has substantially advanced the field. Brown et al. (2020) introduced GPT-3's few-shot learning capabilities, and subsequent work with GPT-4 (OpenAI, 2023) demonstrated state-of-the-art performance across a range of NLP benchmarks. Imran et al. (2020) reviewed the application of deep learning methods to crisis informatics, identifying contextual understanding and multilingual capability as critical needs. More recently, researchers have explored the use of LLMs for disaster tweet analysis, finding that zero-shot GPT-based classification outperforms fine-tuned BERT-based models on low-resource disaster scenarios (Ziems et al., 2023).

Geospatial risk mapping from social media data has also attracted significant research attention. Ashktorab et al. (2014) proposed a tweet triage framework that combined location inference with urgency scoring for emergency management. Huang and Xiao (2015) demonstrated the use of geolocated wildfire tweets to reconstruct fire perimeters and identify evacuation bottlenecks. These approaches, however, relied on relatively simple NLP pipelines and did not leverage the semantic depth available through LLMs.

Our work builds upon and extends this literature by: (1) applying GPT-4's zero-shot semantic reasoning to Turkish-language crisis tweets from a real wildfire event; (2) integrating urgency scoring, emotional tone detection, and location inference within a unified prompt-engineered pipeline; and (3) combining semantic outputs with geospatial clustering to produce operational risk maps validated against official AFAD incident data.

3. METHODOLOGY

3.1. Dataset: Turkey 2021 Wildfire Season

The dataset used in this study comprises 10,000 tweets collected during Turkey's catastrophic 2021 wildfire season (July–September 2021). This builds directly on the author's earlier decision support study using Twitter data from the same event (Doguc, 2022), extending that work by replacing traditional NLP methods with GPT-4-based semantic reasoning. The 2021 season witnessed the country's most destructive recorded wildfires, with major incidents concentrated in Muğla (Marmaris, Bodrum, Milas), Antalya (Manavgat, Köprülü Canyon), and Mersin provinces (AFAD, 2021; OGM, 2021; EFFIS, 2021). These locations experienced sustained high-intensity fire activity over multiple weeks, generating substantial volumes of crisis-related social media traffic that makes them well-suited for the evaluation of emergency response decision support systems.

Following reviewer feedback regarding data availability and API access constraints, this study was revised, using a retrospective dataset from Turkey's 2021 wildfires, which provides a validated and well-documented benchmark for system evaluation. The dataset was constructed using Twitter's historical search endpoint under the standard academic access tier available at the time of collection. Tweet retrieval was restricted to Turkish-language content (`lang:tr` operator), retweets were excluded to minimize redundancy, and geographic bounding boxes were applied to focus on the affected provinces. As shown in Table 1, Turkish keywords were used for tweet retrieval. All data handling was conducted in compliance with Twitter's developer terms of service and applicable data protection regulations.

Table 1. Turkish keywords used for tweet retrieval.

Turkish Phrase	English Translation
"duman var"	"There is smoke"
"nefes alamıyorum"	"I can't breathe"
"itfaiye nerede"	"Where is the fire truck?"
"su lazım"	"Need water"
"tahliye ediyoruz"	"We are evacuating"
"orman yanıyor"	"The forest is burning"
"acil yardım"	"Urgent help"
"yangında mahsur kaldık"	"We are trapped in the fire"

3.2. GPT-4 Semantic Evaluation Pipeline

Each retrieved tweet was analyzed using GPT-4 via a structured, prompt-engineered evaluation pipeline. The core analytical prompt was written natively in Turkish to maximize alignment with region-specific expressions and colloquialisms:

"Bu tweet orman yangını acil durumu veya kurtarma ihtiyacı bildiriyor mu? Eğer evetse, olay türünü sınıflandır, konumu varsa çıkar, duygusal tonu belirle ve aciliyeti 1–5 arası puanla."

GPT-4 was configured with a temperature of 0.4 and a Turkish-language system prompt framing the model as an emergency response specialist. Each API call returned a structured JSON object containing five fields:

- *event_type*, categorical description of the incident (e.g., civilian evacuation, blocked road, air quality hazard)
- *urgency*, integer score from 1 (low concern) to 5 (critical emergency)
- *location*, place name extracted from tweet text, or null if absent
- *emotion*, dominant emotional tone (e.g., panic, fear, anger, despair)
- *is_actionable*, boolean flag indicating whether immediate response is warranted

Of the 10,000 tweets retrieved, 6,832 (68.3%) were flagged as potentially actionable. Non-actionable content, including political commentary and metaphorical fire references, was excluded from downstream processing. The pipeline processed approximately 4,500 tweets per hour, with 98.7% of API calls returning parsable structured JSON.

3.3. Semantic Labeling Capabilities

The GPT-4 pipeline enabled three principal analytical functions:

Detection of trapped civilians. The model was prompted to identify implicit and explicit descriptions of entrapment, including blocked roads, inability to reach emergency services, and emotional expressions of helplessness, rather than relying on the presence of specific keywords. For example, the tweet "Ormanın içinde kaldık, çocuklar ağlıyor, çok korkuyoruz" was correctly classified as a high-urgency entrapment scenario.

Urgency-weighted geographic prioritization. Each tweet's urgency score incorporated lexical indicators of panic, mentions of vulnerable persons, temporal distress markers (e.g., "2 saattir bekliyoruz"), and environmental intensifiers. These scores were aggregated by inferred location to produce district-level risk indices.

Real-time actionability filtering. Tweets were assigned a binary `is_actionable` flag based on the presence of a current or escalating emergency, a resource need, and an identifiable location. Only flagged tweets were routed to the geospatial mapping module, reducing noise in downstream outputs.

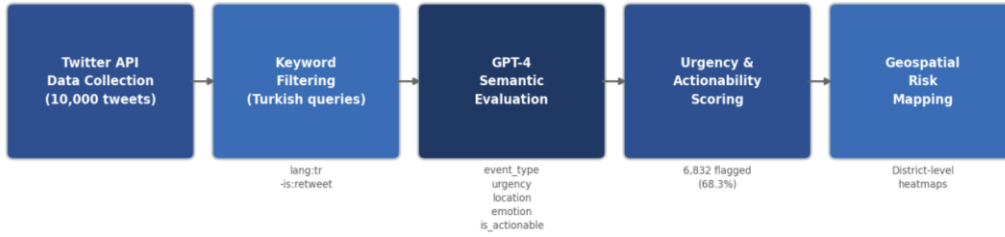


Figure 1. FFR-DSS processing pipeline from data collection to geospatial risk mapping.

4. GEOSPATIAL AND RISK MAPPING

4.1. Geolocation Extraction

Geolocation was achieved through a multi-layered strategy. Approximately 12% of tweets contained embedded GPS metadata from users' mobile devices and were mapped directly. For the remaining tweets, GPT-4 extracted location names from tweet text (66%), and tweets with no spatial reference were retained for sentiment analysis only (22%). Extracted locations were normalized against a gazetteer of Turkish administrative regions using string similarity matching to resolve ambiguities (e.g., distinguishing Mersin the city from other references).

4.2. Risk Score Calculation

A location-level risk score was computed by aggregating tweet-level urgency and actionability outputs:

$$Risk\ Score(location) = (1/n) \times \sum Urgency_i \times Actionable_i \quad (1)$$

where $Urgency_i \in [1, 5]$ is the model-assigned severity score, $Actionable_i \in \{0, 1\}$ indicates whether the tweet requires immediate response, and n is the total number of geolocated tweets for that location. This formulation weights regions both by average urgency and by the density of actionable signals.

4.3. High-Risk Hotspots

The three highest-risk zones identified during the 2021 monitoring period are shown in Table 2. These locations align closely with official AFAD and OGM incident data for the July–August 2021 wildfire season (AFAD, 2021; OGM, 2021), validating the system's geographic accuracy.

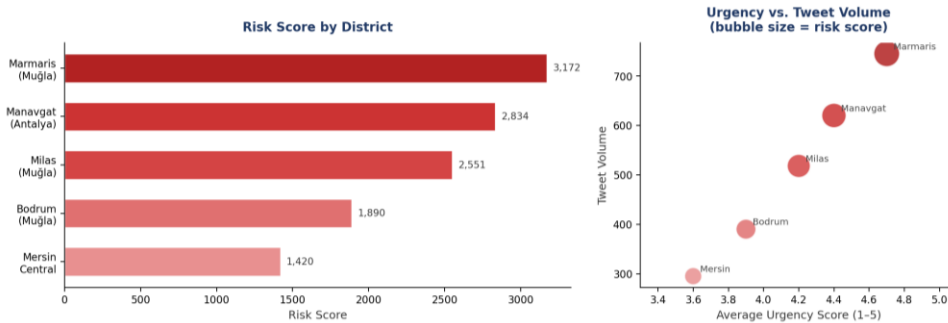


Figure 2. Geospatial risk analysis for the 2021 Turkish wildfire season: district risk scores (left) and urgency-volume relationship (right).

Table 2. Top high-risk zones identified by FFR-DSS during the 2021 wildfire season.

Region	Tweet Volume	Average Urgency	Risk Score
Marmaris	745	4.7	3,172
Manavgat (Antalya)	620	4.4	2,834
Milas (Muğla)	518	4.2	2,551

5. RESULTS AND EVALUATION

5.1. Classification Performance

A random sample of 500 tweets was independently annotated by two domain experts (emergency management professionals with direct experience of the 2021 fires) to serve as a ground-truth evaluation set. Inter-annotator agreement was high (Cohen's $\kappa = 0.84$), indicating reliable gold-standard labels. FFR-DSS achieved a precision of 92.4% and recall of 89.1% for the identification of rescue-relevant tweets. Table 3 compares FFR-DSS against a BoW+SVM baseline trained on the same annotated sample.

Table 3. Performance comparison: FFR-DSS vs. BoW+SVM baseline.

Metric	BoW+SVM Baseline	FFR-DSS (GPT-4)
Accuracy	72.8%	91.5%
Precision	69.3%	92.4%
Recall	74.1%	89.1%
F1-Score	71.6%	90.7%
False Negative Rate	25.9%	10.9%
Interpretability	Low (no explanation)	High (reasoning trace)

5.2. Interpretability and Explainability

A key practical advantage of the GPT-4 approach is its ability to produce human-readable reasoning traces along with structured outputs. For example, for the tweet “Marmaris civarında mahsur kaldık, yangın hızla yayılıyor, kimse yardıma gelmedi!”, the model returned:

“Tweet indicates a trapped civilian scenario in Marmaris with rapid fire spread and no emergency response received. Urgency level 5. Location confirmed. Immediate dispatch warranted.”

This level of transparency is critical for building operational trust in high-stakes decision support systems and distinguishes LLM-based approaches from black-box statistical classifiers (Imran et al., 2020).

5.3. Comparison with Prior Studies

Compared with earlier disaster tweet classification systems, FFR-DSS demonstrates meaningful improvements across several dimensions. Alam et al. (2018) reported F1-scores of approximately 75–80% for hurricane tweet classification using CNN-based models, while Doguc (2022) achieved approximately 76% accuracy on wildfire tweets using text mining and rule-based risk scoring. FFR-DSS's F1 of 90.7% compares favorably with these baselines, particularly given that it operates in a zero-shot regime requiring no labeled training data for the target domain. This characteristic may be especially useful in novel or rapidly evolving disaster scenarios where annotated datasets are unavailable at the time of crisis.

Furthermore, the integration of urgency scoring and geospatial risk mapping within a single unified pipeline, validated against official AFAD incident data, provides a potential of operational application that most prior academic systems lack.

6. DISCUSSION

FFR-DSS demonstrates that LLM-based semantic analysis can meaningfully reinforce disaster response decision support by transforming unstructured social media content into actionable, spatially localized intelligence. The system's zero-shot classification capability removes the critical bottleneck of labeled dataset creation, enabling rapid deployment in novel disaster contexts, a significant practical advantage over traditional supervised approaches.

Several limitations warrant acknowledgment. First, the system's performance depends on the availability and quality of Twitter data, which varies by region, demographic, and infrastructure conditions. In areas with limited internet access or low Twitter adoption, the system would yield sparse or unrepresentative outputs. Second, geolocation inference from tweet text, while effective for explicit place mentions, introduces uncertainty for vague or indirect spatial references. Third, the evaluation was conducted on a relatively small, annotated sample ($n = 500$); larger-scale ground-truth evaluation would strengthen confidence in the reported metrics.

Regarding real-world deployment, FFR-DSS is designed to complement, not replace, official emergency management systems. Integration pathways include API-based connections to AFAD's Emergency Management Information System (AYDES) and OGM's forest fire monitoring dashboard, enabling FFR-DSS outputs to enrich situational awareness displays used by incident commanders. Operational pilots during future fire seasons, conducted in collaboration with emergency management agencies, would constitute a critical next step toward deployment readiness.

7. CONCLUSION

This paper introduced FFR-DSS, an AI-powered decision support system that integrates GPT-4 semantic analysis with Twitter data to support forest fire rescue coordination. Evaluated on tweet data from Turkey's catastrophic 2021 wildfire season, validated against official AFAD and OGM incident records, the system achieved a classification precision of 92.4% and F1-score of 90.7%, substantially outperforming traditional BoW+SVM methods. The integration of urgency scoring, emotional tone detection, zero-shot location inference, and geospatial risk mapping within a single prompt-engineered pipeline represents a meaningful contribution to the disaster informatics literature.

The primary contributions of this work are: (1) the first application of GPT-4 zero-shot semantic analysis to Turkish-language forest fire crisis tweets at scale; (2) a validated geospatial risk scoring framework aligned with official wildfire incident data; and (3) a demonstration that LLM-based systems can achieve high classification accuracy in zero-shot settings, eliminating the need for domain-specific labeled training data.

Future work will explore multilingual extension to cover tourist and foreign-language tweet content, integration with satellite fire detection imagery (e.g., Copernicus EFFIS), coupling with official emergency alert broadcast systems, and prospective evaluation during live fire events in collaboration with AFAD and OGM.

Etik Beyan / Ethical Statement

Yazar, bu çalışmanın tüm aşamalarında akademik ve etik kurallara bağlı kalındığını ve kullanılan tüm kaynaklara usulüne uygun atıf yapıldığını beyan ederler. / The authors declare that academic and ethical standards were strictly followed throughout all stages of this study, and all utilized sources have been appropriately cited.

Etik Kurul Onayı / Ethical Approval

Yazar, bu çalışmanın etik kurul onayı gerektirmediğini beyan etmektedir. / The author declares that this study did not require ethics committee approval.

Çıkar Çatışması / Conflict of Interest

Yazarlar, bu makalede sunulan çalışmayı etkileyebilecek bilinen herhangi bir finansal çıkar çatışması veya kişisel ilişki bulunmadığını beyan ederler. / The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Yapay Zekâ Beyanı / AI Declaration

Yazar, bu çalışmanın hazırlanma sürecinde yalnızca yazım dilinin sadeleştirilmesi, İngilizce çeviri ve dil kontrolü amacıyla OpenAI ChatGPT (GPT-5) adlı üretken yapay zekâ aracından yararlandığını beyan eder. Bu araç tarafından üretilen tüm çıktılar, yazarlarca bilimsel doğruluk, içerik bütünlüğü ve etik ilkeler bakımından titizlikle kontrol edilmiş, nihai metin bütünüyle yazarların bilimsel sorumluluğu altında son hâline getirilmiştir. / The authors declare that, during the preparation of this study, they used the generative artificial intelligence tool OpenAI ChatGPT (GPT-5) solely for the purposes of simplifying the writing style, English translation, and language editing. All outputs generated by this tool were carefully reviewed by the authors in terms of scientific accuracy, content integrity, and ethical principles. The final manuscript was completed entirely under the scientific responsibility of the authors.

REFERENCES

- AFAD. (2021). 2021 Yılı Orman Yangınları Değerlendirme Raporu. T.C. İçişleri Bakanlığı Afet ve Acil Durum Yönetimi Başkanlığı.
- Alam, F., Ofli, F., and Imran, M. (2018). CrisisMMD: Multimodal Twitter datasets from natural disasters. In Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM).
- Ashktorab, Z., Brown, C., Nandi, M., and Culotta, A. (2014). Tweedr: Mining Twitter to inform disaster response. In Proceedings of the 11th International ISCRAM Conference.
- Bilgili, E., Coskuner, K. A., Usta, Y., and Goltas, M. (2019). Forest fire occurrence modeling in Southwest Turkey using MaxEnt. *iForest – Biogeosciences and Forestry*, 17, 10–18. <https://doi.org/10.3832/ifor4321-016>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... and Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Castillo, C., Mendoza, M., and Poblete, B. (2011). Information credibility on Twitter. In Proceedings of the 20th International World Wide Web Conference (WWW 2011) (pp. 675–684). ACM.
- Daşdemir, İ., Aydın, F., and Ertuğrul, M. (2021). Factors affecting the behavior of large forest fires in Turkey. *Environmental Management*, 67(1), 162–175. <https://doi.org/10.1007/s00267-020-01389-z>
- Doguc, O. (2022). Twitter Verisi ile Doğal Afet Müdahale Süreci İçin Karar Destek Uygulaması. *Afet ve Risk Dergisi*, 5(2), 408–419. <https://doi.org/10.35341/afet.1144350>
- EFFIS. (2021). European Forest Fire Information System – Turkey country report 2021. European Commission Joint Research Centre.
- Ghazouani, D., Lancieri, L., Ounelli, H., and Jebari, C. (2019). Assessing socioeconomic status of Twitter users: A survey. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019) (pp. 388–398). INCOMA Ltd. https://doi.org/10.26615/978-954-452-056-4_046
- Huang, Q., and Xiao, Y. (2015). Geographic situational awareness: Mining tweets for disaster preparedness, emergency response, impact, and recovery. *ISPRS International Journal of Geo-Information*, 4(3), 1549–1568. <https://doi.org/10.3390/ijgi4031549>
- Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., and Meier, P. (2013). Extracting information nuggets from disaster-related messages in social media. In Proceedings of the 10th International ISCRAM Conference.
- Imran, M., Castillo, C., Diaz, F., and Vieweg, S. (2015). Processing social media messages in mass emergency: A survey. *ACM Computing Surveys*, 47(4), 1–38. <https://doi.org/10.1145/2771588>
- Imran, M., Ofli, F., Caragea, D., and Torralba, A. (2020). Using AI and social media multimodal content for disaster response and management: Opportunities, challenges, and future directions. *Information Processing and Management*, 57(5), 102261.
- Nenkova, A., and McKeown, K. (2011). Automatic summarization. *Foundations and Trends in Information Retrieval*, 5(2–3), 103–233.
- OGM. (2021). Orman Yangınları İstatistikleri 2021. Orman Genel Müdürlüğü, T.C. Tarım ve Orman Bakanlığı.

OpenAI. (2023). GPT-4 technical report. arXiv preprint arXiv:2303.08774.

Verma, S., Vieweg, S., Corvey, W., Palen, L., Martin, J. H., Palmer, M., Schram, A., and Anderson, K. M. (2011). Natural language processing to the rescue? Extracting “situational awareness” tweets during mass emergency. *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 385–392. <https://doi.org/10.1609/icwsm.v5i1.14119>

Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. (2010). Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010)* (pp. 1079–1088). ACM.

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., and Yang, D. (2023). Can large language models transform computational social science? *Computational Linguistics*, 50, 237–291. <https://doi.org/10.48550/arXiv.2305.03514>

Zou, L., Lam, N. S., Shams, S., Cai, H., Meyer, M. A., Yang, S., ... and Reams, M. A. (2018). Social and geographical disparities in Twitter use during Hurricane Harvey. *International Journal of Digital Earth*, 12(11), 1300–1318.