



Adaptive Machine Learning Pipelines for Instantaneous Disaster Triage and Supply Chain Coordination

Vinod Battapothu

Independent Researcher

vinod.battapothu.researcher@gmail.com

Abstract

AI-based big data systems support near real-time disaster response and resource allocation. These systems shorten reaction times by leveraging scarce resources (data, bandwidth, processing, and decision-making) more efficiently in disaster management. A layered architecture provides distributed computing, data fusion, and decision management for near real-time deployment. Future implementations will prototype alarm generation using social-media and sensor data.

Large-scale disasters place severe and often insurmountable demands on emergency agencies. Under such conditions, resource allocation plays a critical role in minimizing disaster impact, but conventional methods based on historical data often cannot meet real-time requirements. The integration of geographical information with data from social networks, sensor networks, and other sources can facilitate the relevant prediction and decision-making processes, but the available data, computing, and human resources are all severely constrained—with the same scarcity of responses and alerts as for other data types. For these reasons, proposed methods shorten response times by applying AI-based big-data systems.

Keywords : Big Data; Disaster; Streaming Data; Real-Time; Predictive Modelling; Resource Allocation; Command-and-Control; Decision Management; Social and Open Data; Sensor Networks; Internet of Things.

1. Introduction

Disaster management relies heavily on data sources and computing infrastructure that enable powerful analytics for pre-disaster forecasting, impact assessment during the event, and post-disaster recovery. The data can be generated by geolocated social media content, open datasets published by various agencies, sensor networks, and mobile phones. Effective integration of these sources on big data platforms with advanced real-time capabilities and support for streaming systems is crucial. AI-based predictive models can estimate the influence of hazard events in near real time, while allocation models can optimize the resources dedicated to the response across multiple agencies with high speed and full pre-event reliability through a map-based decision-response framework.

Command-and-control interfaces visualize the situation with geolocated information and dedicated algorithms for field operations, while data-generated cues can also support workflow analytics. The overall infrastructure includes a dedicated cloud service, capable of managing the flexible combination of these specific systems and enabling a coherent orchestration of system decisions and responses that can also take advantage of external AI models, where appropriate.

1.1. Background and Significance

Societal dependence on information and communication technologies (ICT) continues to grow as it becomes easier and cheaper to connect devices across the globe, and thus the volume of data flowing through these platforms increases. The potential to leverage this wealth of information during crisis events, however, remains largely untapped. Natural disasters create a surge in media activity, particularly in social-networking sites, and analyzing this data can improve situational awareness: people



themselves share information about incidents and request rescue via Twitter, Facebook, and Google+. Other public data streams, such as earthquake early-warning systems, open-source-sensor networks, and satellite images, become available and can be made accessible through Application Programming Interfaces (APIs) for use in online applications. This information is complementary to the reports and messages transmitted by conventional emergency-management services. When combined with information from Internet of Things (IoT) networks and predictions from models, such sources constitute a form of Big Data that can be analyzed and reacted to in a timely manner.

The use of Big Data for crisis management is particularly appealing as resource allocation for response operations is seldom sufficient compared to demand, especially in low-income countries with limited financial resources. Allocating the share of available resources, such as rescue workers, medic crews, helicopters, and food, to minimize the overall impact is a difficult task. Probabilistic models can capture uncertainties causing unquantified damage, such as the potential blockade of a road, while taking complementarity of resources into account. An efficient solution can be produced if a reasonable upper bound on the time needed to process the allocation optimization problem is respected. An efficient Big Data architecture is hence required, offering also a data-fusion layer able to merge and understand coarse information flowing from different sources. Various Volunteered Geographic Information and social-media data streams are made accessible through APIs and, together with streaming-sensor data, represent a unique opportunity for real-time crisis-management systems.

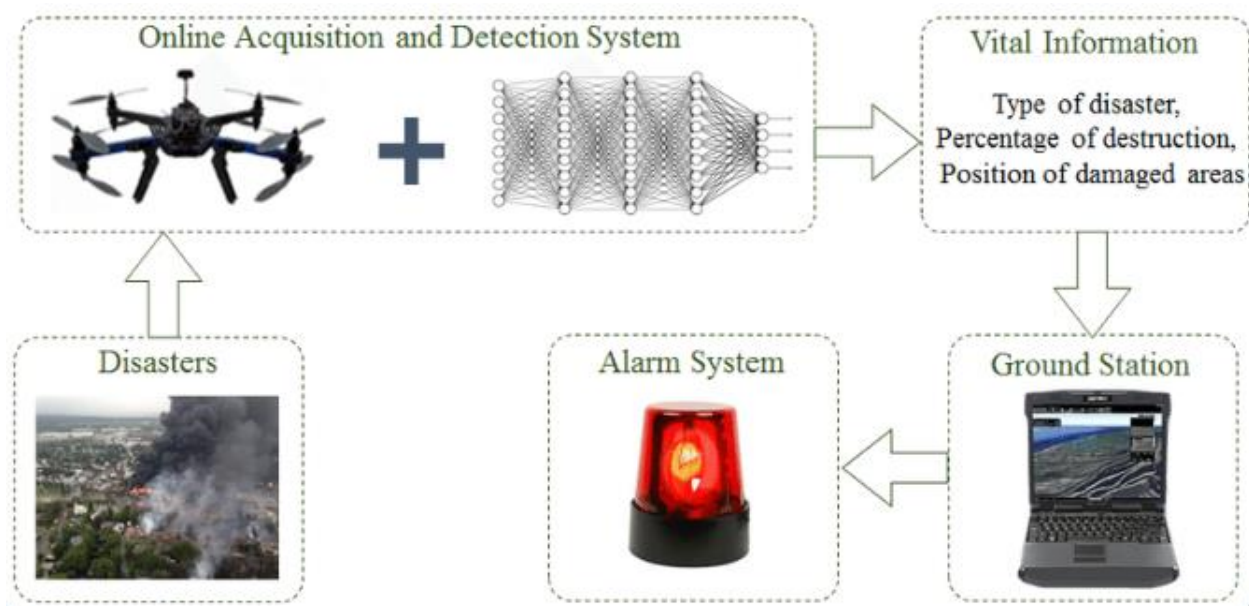


Fig 1: Real time disaster detection system

1.2. Research design

With the increasing frequency and severity of extreme natural hazards, including earthquakes, storms, landslides, and floods, the need for efficient disaster response systems is becoming more pressing. A variety of interconnected technologies now make it possible to use social media, web-based open data for disaster management, sensor networks for hazard monitoring, and crowd-sourced sensor arrays for damage assessment during disaster events. The proposed research studies the requirements and design



principles for an integrated AI-based Big Data system infrastructure that integrates these capabilities into a coherent unit offering real-time operational support for decision-makers. A scenario-driven methodology is employed to develop and validate the system architecture.

Data acquisition, integration, and management are shaping up as the key challenges of a Big Data approach to disaster response, and the research focuses in particular on the distribution of data-gathering responsibilities across a large number of sources. New or differently used data sources are being sought, adaptive query execution strategies are being examined, and integrated solutions combining the features of event-driven, data-streaming, and complex-event-processing paradigms are being developed. Sizing and performance metrics of Big Data infrastructures for disaster response are being defined in conjunction with users to ensure the suitability of the design.

Equation 1: Assign a latency to each stage

Let:

- T_{gen} = data generation latency
- T_{dist} = distribution/communication latency
- T_{proc} = processing/analytics latency
- T_{cons} = consumption latency (visualization/decision/dispatch time budget)

• Because these stages occur sequentially in the loop, total time is:

$$T_{\text{total}} = T_{\text{gen}} + T_{\text{dist}} + T_{\text{proc}} + T_{\text{cons}}$$

2. Background and Foundations

Big Data Architectures for Disaster Management

Architectures for managing large amounts of data related to natural disasters are beginning to emerge. These architectures integrate various data sources and streams, including social media, sensor and data feeds, remote sensing, and other open data sources that can aid in disaster management. Real-time system capabilities are included within these environments, although often at a different scale from large scale batch processing. These emerging systems have been termed `Cloud Computing for Natural Disasters` by the authors. Cloud Computing from the viewpoint of consumers is a Sigma utility model, where both storage and processing can be performed within a Cloud Service Provider enabling exploitation of massive tools and offers fault-tolerance and capacityfulfillment issues for low-cost or non-paying users.

It is with a view to automate the especially critical area of real-time impact assessment systems that these architectures have been proposed. The ability to quickly assess damage or loss as near realtime as possible has a tremendous influence on the timing and effectiveness of any relief response, and failure to do so results in a waste of precious time and resources during actual disaster events.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18



Real-Time Analytics and Streaming Systems

Real-time analytics, harnessed from Open Data, combined with Machine Learning algorithms offer tremendous potential for computing damage in near Real Time after a natural disaster. Recently, a Big Data cloud-based real-time analytics architecture has been proposed for natural disasters. The architecture aggregates available Open Data from various trustworthy data providers and utilizes public-generated content from media platforms such as Twitter and Facebook during an emergency contingent. The architecture supports the federation of multiple Big Data sources, efficiently executes the process of data/time filtering and text classification. Within the system, two examples of non-geographic data analysis are implemented to illustrate how to apply the analytics as part of situational awareness and response support during a crisis.

2.1. Big Data Architectures for Disaster Management

Disaster-management systems have advanced from legacy static designs to become multi-tier Big Data architectures capable of real-time data ingestion and processing. Ultimately, response and recovery depend on decisions made in the command-and-control center and on field operations. Their effectiveness is determined by how well they address the four dimensions of disaster management: prediction, prevention, impact assessment, and resource allocation.

Command-and-control operations have four goals:

- Monitor the disaster situation using sensor networks, social media, and available open data sources.
- Understand the evolution of the disaster and rapidly assess its impact by estimating the affected population and infrastructure damage.
- Proactively manage the deployment of resources in disaster-prone areas.
- Establish a decision-management system that combines all necessary information and provides the command-and-control team with services for decision support, reinforcement learning, and the orchestration of field operations.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

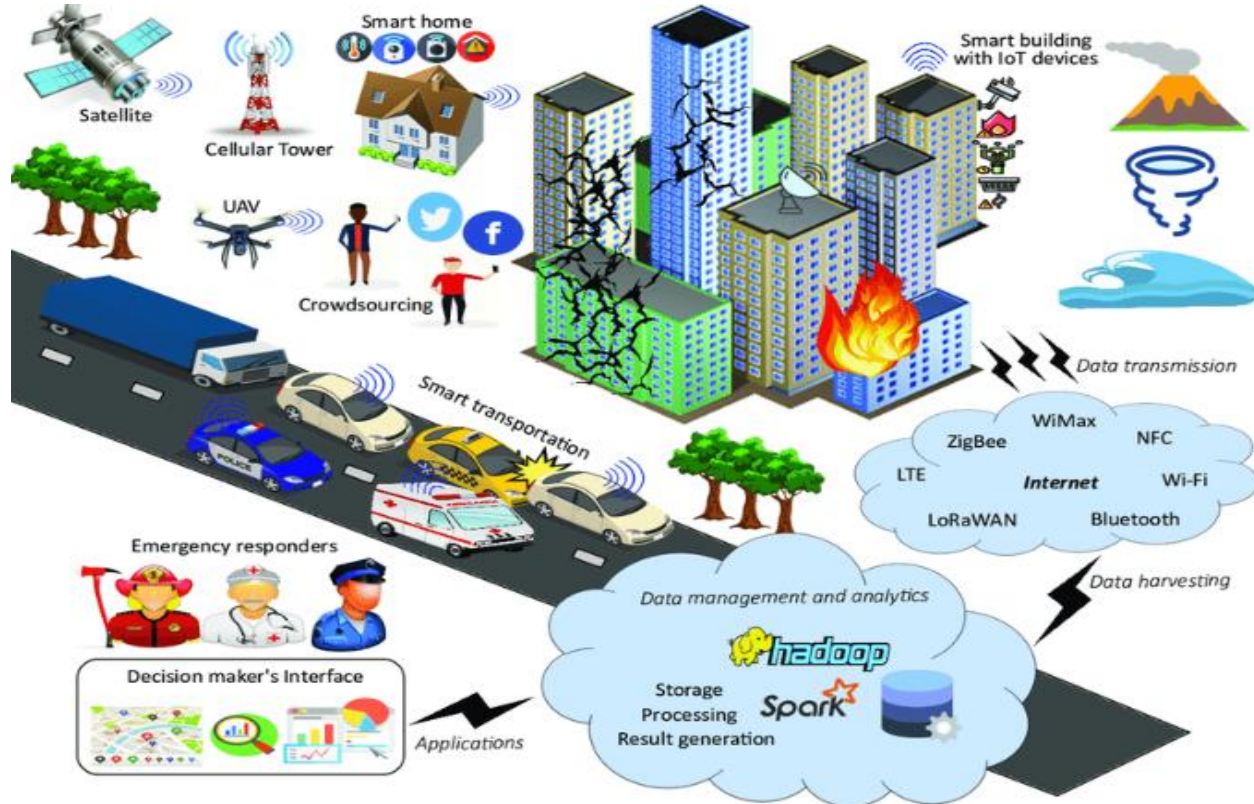


Fig 2: Big Data Architectures for Disaster Management

2.2. Real-Time Analytics and Streaming Systems

Data-driven disaster management relies on the analysis of continuously updated stream and sensor-based data sources. Consequently, real-time data analytics are increasingly required to complement traditional batch processing. Streaming architectures are now widely used to create and apply predictive models for sudden-onset natural disasters in tandem with data supplied by social media. However, early-warning and evaluation systems for slow-onset disasters—such as drought, epidemics, and conflicts—remain predominantly batch oriented. Recent advances in data-management and analytics technologies—particularly for data streaming and sensor integration—suggest the merit of complementing batch-based predictive modeling and impact assessment with real-time change-detection and evaluation methods.

Streaming systems are best suited for processing workloads of low latency, continual arrival patterns, and bounded processing requirements. Alarm-triggered predictive modeling is a natural candidate for processing streaming workloads, using historical training sets to derive prediction models that are subsequently updated as new data arrive. For some models, real-time change detection can be employed to ascertain whether the predictive relationship persists. In this context, both Big Data technology and application should be suitable for latencies of several minutes. Although streaming technologies for sensor-network data have advanced significantly, such data still lack rich deployment among local disaster-management organizations.

Equation 2: Allocation under uncertainty: “delay worsens damage” → utility function



Let $D(d)$ be an increasing function: higher delay \rightarrow higher overall damage.

So:

$$\frac{dD(d)}{dd} \geq 0$$

A simple direct definition is:

$$U(d) = -D(d)$$

3. Data Sources and Integration

Social media is changing how disasters are perceived and understood, enabling a fast and dense distribution of information. However, Twitter, Facebook, and other platforms also produce colossal amounts of data. It has been suggested that social media data stream can augment traditional sources in emergency management, contributing to more timely situational awareness, textual insights on event dynamics, and knowledge on the people affected, their needs, and available resources. Crisis-related posts are becoming an opportune complement to official sources, opening new avenues for research and development. Nevertheless, undertaking the processing of such a flow at a suitable scale requires careful attention since not every post is useful or truthful.

Different types of traditional information infrastructures have been conceived for large-scale disasters, yet they rarely work when needed; they suffer from missing data and inconsistent cadences; and their deployment is labor-intensive on affected territories often populated by first responders. Following the “small is beautiful” paradigm and the advent of low-power wide-area networks, low-cost radio-enabled sensor systems can also be deployed ad-hoc to collect data on specific phenomena, but their data must be integrated with those coming from social media and official open data for enhanced situational awareness.

Equation 3: Make it risk-averse under uncertainty

AI-Based Big Data Systems for R...

Let d be a random variable, then the decision objective is often:

$$\max \mathbb{E}[U(d)]$$

Risk aversion means the utility is **concave** in “benefit” (or negative damage). One common way:

- Define benefit $B(d) = B_{\max} - D(d)$ (larger is better)
- Choose concave $u(\cdot)$

$$U(d) = u(B(d))$$

3.1. Social Media and Open Data Streams

Data sources that are publicly accessible and support a high data rate are critical for disaster response applications. Social media

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18



activity provides timely information about events at the location of deployment, such as earthquakes. Open data from other organizations can provide context about the meteorological, geological, or infrastructural characteristics of the affected area. In many cases, these channels provide information about disaster events ahead of the official government assessments. However, the unstructured, informal nature of social media data makes it relatively challenging to model.

The unstructured nature of social media messages makes it necessary to establish a workflow for monitoring and processing such observations. Keywords related to emergency assessment, such as “fire,” “earthquake,” “flood,” and so on, usually help to filter the data streams. Associated languages can be predefined for specific regions. While data from any online source could be leveraged for impact assessment, it may not always be necessary. Specialized data, such as that from the Global Disaster Alert and Coordination System's (GDACS's) automatic earthquake impact estimation tool, would clearly provide more precise information than social media. However, relying solely on official estimates may lead to delays. Because social media messages are often unstructured and rapidly generated, it is important to design a clear workflow to monitor and process such information effectively. One common approach is to filter incoming data streams using specific keywords related to emergencies, such as “fire,” “earthquake,” or “flood.” These keywords help identify relevant posts quickly, allowing analysts to focus on messages that may indicate a developing crisis. In addition, the languages associated with particular regions can be predefined so that messages in those languages are prioritized for analysis. Although data from various online sources can be used for impact assessment, it is not always necessary to collect information from every available platform. In many cases, specialized systems such as the Global Disaster Alert and Coordination System's (GDACS) automatic earthquake impact estimation tool can provide more accurate and structured data. Nevertheless, depending only on official estimates can sometimes cause delays in response, whereas social media can offer immediate, real-time observations from affected populations. Therefore, combining social media monitoring with reliable official data sources can help create a faster and more comprehensive disaster assessment process.

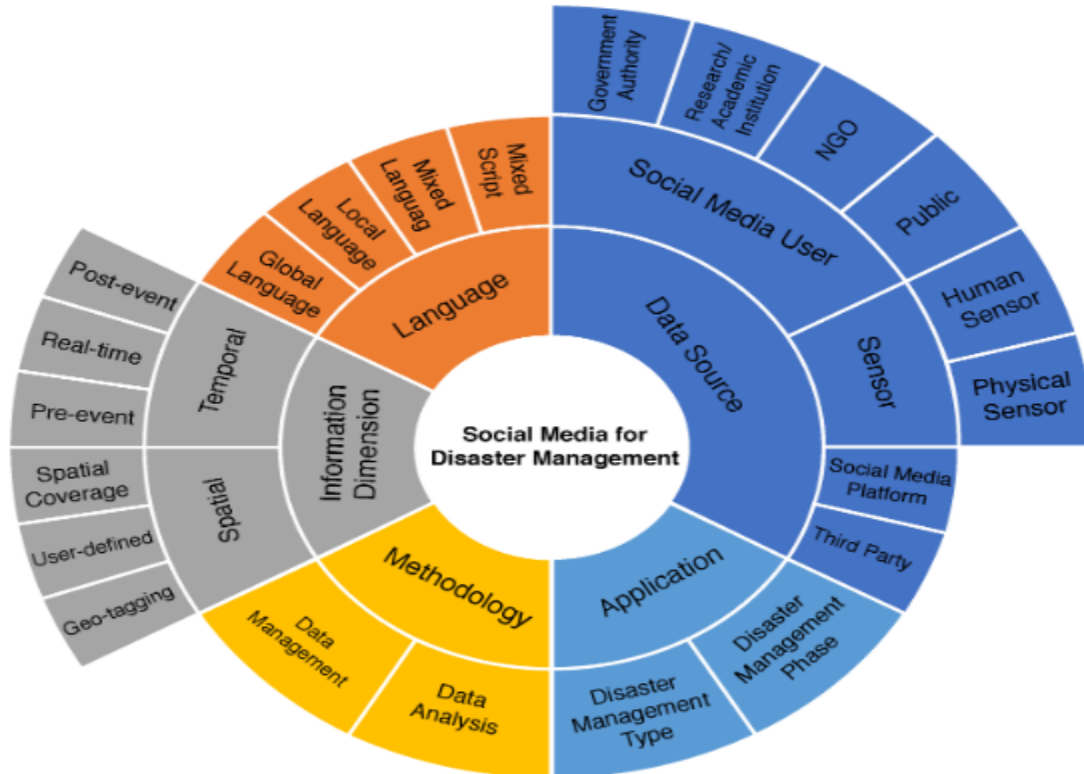


Fig 3: Social Media Data in Disaster Management

3.2. Sensor Networks and IoT Deployments

The advent of the Internet of Things (IoT) and the significant investments worldwide in sensor networks for a variety of applications spanning environmental monitoring, physical security, infrastructure condition assessment, and more combined with the recent tremendous growth of event-sensitive video and audio scanning networks offer a wealth of new real-time data sources available for disaster response and resource management tasks. These sources attract interest beyond their primary intended use (e.g., for hazard monitoring or security applications) and are often motivated by the desire to visually record any disaster impact in order to assist the public and relief agencies with disaster information, assessments, and resource acquisition. Their coordination is key to true value generation, for example allowing users to merge these heterogeneous data streams with real-time social media, damage-impact models, and the decision-support models mentioned previously in order to produce timely and actionable information.

Real-time social media analytics and open-data fusion mechanisms will leverage these new data sources in order to monitor developing situations and to conduct field-operations fertilization and resource-allocation optimization processes. The latest enhancement in real-time social media fusion capability extends the spatial surveillance provided by these camera- and audio-based networks and supports the near-real-time identification and visualization of unwanted and crisis events and their underlying conditions in specific urban and rural areas. Audio event-detection models based on Deep Learning techniques have recently been extended to jointly enable the automatic detection of specific sounds and the location of their sources, even in the presence of multiple overlapping sound events. Advanced fusion techniques, including deep visual-field-based video fusion, now provide



the capability to assemble the views of available monitoring camera networks in order to generate a global perspective of recorded events in the monitored urban and rural locations.

4. Modeling for Real-Time Decision Support

A variety of applies models may be integrated into a real-time disaster-management system's architecture to enhance decision support in impacted communities. Predictive models estimate potential human impact and infrastructure damage with actionable reliability during an evolving disaster. Stochastic optimization problems accommodate the inherent uncertainty in resource requests and priority levels derived from estimated damage.

Whether used independently or together, predictive models and resource-allocation modules can be applied to community-disaster-management systems capable of operating in a fully automated mode. In such a setup, a region exposed to a natural hazard is monitored, and, as soon as a disaster occurs, models run in real time. Corrections are provided to disaster-response organizations, and resources are allocated automatically. Clearly, the automatic mode is only a secondary feature. An increasing degree of automation is intended to enable organizations to comply with legislation requiring compliance with NIMS at the appropriate level.

Predictive models that assess human impact and infrastructure damage rely on traditional impact models to estimate fatalities and injuries. Additionally, Open Street Map (OSM) data, newly available during a disaster, are combined with the core building inventory to quantify the impact (building loss or damage) on the primary road network and other important classes of buildings (hospitals, police stations, and fire brigades). The predicted degree of damage to buildings is used as input for HydroGeoSphere—an integrated surface-subsurface hydrological model—to assess the potential effect of flooding on those areas for which no models are provided.

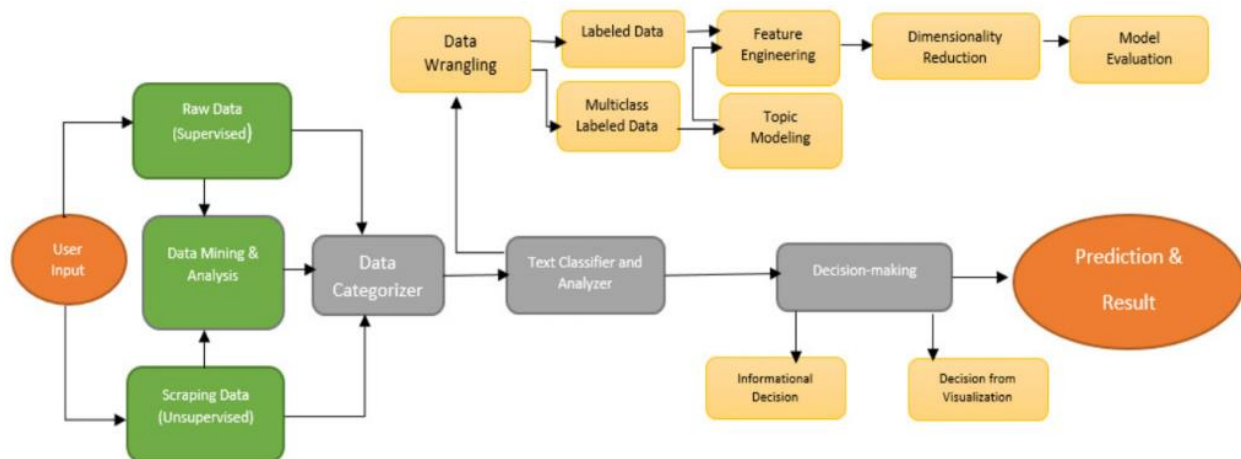


Fig 4: Real-Time AI-Based Informational Decision-Making Support System

4.1. Predictive Modeling for Impact Assessment

Full-scale disasters such as hurricanes and flood require real-time action on geographical and operational parameters.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18



Communication and transportation resources such as airports need to be momentarily shut-down for incoming supplies, while other resources like hospitals need temporary increase of capacity to take care of the victims. Typically it is the disaster-science communities who used to be the end users of the scientific models. Consequently, the scales of the models have not been designed for real-time action. Over the years many modeling platforms are developed, which can predict the impact on different dimensions, but the synchronicity and coherence among these models has not been clearly defined. Many of these models also have not generated data in real-time manner. Operations managers need to know the change in operations capacity in terms of logistics, shelter, food, command and control, communication systems — both terrestrial and satellite, hospital facilities, and transportation networks, etc. in real-time. Some models provide estimates but the model scales or time steps may not be correct for emergency decision-making. Community-based participatory simulation systems such as Synergia provide prediction tool kits but they have open-ended scales.

During Disasters, Impact on Natural Resources is quite important. Systemic impact evaluation tools such as MCDM provide predictive services by generating big data bases on natural-resource front. Again, the cool-and-hot map of the Natural Vulnerability and Exposure are present for the examined area, and the maps can be statistically knowledgeable for real-time prediction. These two areas of impact assessment are quite critical and require updated information.

Equation 4: Helicopter heuristic as a decision inequality

Let:

- $t_g(i)$ = ground travel time to site i
- $t_a(i)$ = air travel time to site i
- $\Delta t(i) = t_g(i) - t_a(i)$ = time saved
- τ = user-defined threshold

Decision rule:

$$\text{Use helicopter for } i \Leftrightarrow \Delta t(i) > \tau$$

4.2. Allocation Optimization under Uncertainty

Minimizing estimated field operations costs typically involves minimizing the use of expensive resources such as helicopters. However, this goal should not come at excessive delays that worsen the disaster's general impact. Consequently, the use of helicopters can be preferred when the expected increase in operations costs is small compared to the expected improvements in the area's overall damage. A heuristic can be defined to recommend helicopter use when the estimated time saved in reaching a site due to air transport exceeds a user-defined threshold. Broadly categorized, field operations are often executed by emergency teams. Actions involving teams of a specific type (e.g. search-and-rescue teams) are generally executed first; the order of execution among actions involving teams of different types is commonly defined based on local damage estimates; finally, actions of the same type can often be performed in parallel. Such prioritization and ordering can therefore be easily integrated in the allocation modeling step.

The absence of full knowledge regarding disaster impact naturally translates into uncertainty in the field operations allocation decision. Aiming to recommend allocations that minimize expected allocation costs, the model incorporates a risk-averse behavior quantified through a utility function defined in terms of the impact on overall damage caused by delays in performing



the operations. The utility function depends on an estimate of the relationship between delays and overall damage, which can be established based on historical data. The decision on whether a particular decision should be performed can follow similar principles. If the allocation of field operations is based on damage estimates dx_1, \dots, dx_R , representing their values in case of complete knowledge, the decision on operation i can take place if and only if the allocation costs (the resources actually used to perform operation i) belong to a defined price interval associated with dx_j and dx_k , where j and k correspond to the operations previously defined as first and second priorities, respectively.

5. System Architecture and Infrastructure

Architecture and infrastructure decisions are critical to the success of the approach. The real-time response requirements impose strict latency restrictions on the system, dictating suitable technology choices and deployment configurations. Notably, geospatial data that are generated and consumed by field assessments become painfully out of date in rapidly turning situations, making the time to ingest, fuse, and utilize location-targeted information a crucial aspect of the design. In addition to the analytic frameworks, the system architecture and infrastructure must also support data coherence and fusion, given the utilization of disparate common-user data sources that are independently produced and updated.

Duality is a key property of real-time data-stream fusion processes, giving rise to efficient computing strategies. Efficient resource utilization within a distributed computing environment is realized through workload partitioning that explicitly decomposes the total response time into the constituent latencies of data generation, distribution, processing, and consumption. This development highlights the time required for data generation in distributed environments, which depends on the generation and processing latencies at the source nodes and the traversing time over the communication links. Appropriately, analytic approaches that are conducted at the source nodes usually impose minimal time overhead on the total budget.

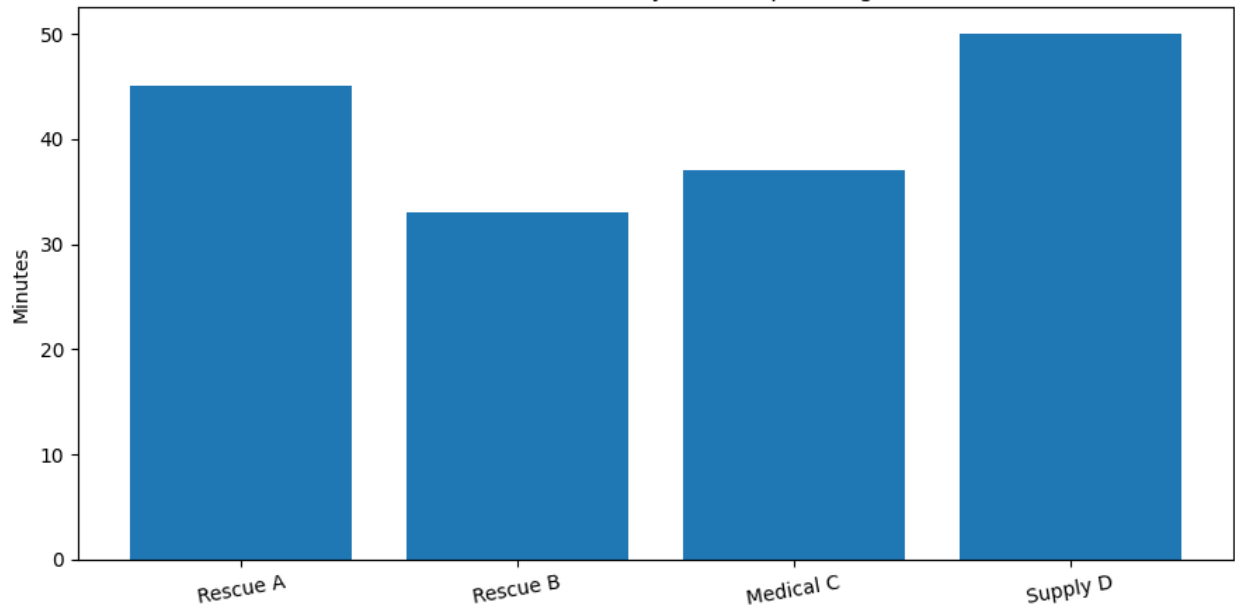
5.1. Distributed Computing for Latency-Sensitive Workloads

The volume and velocity of data generated during disasters necessitate the use of distributed cloud architectures so that advanced AI capabilities can be deployed without affecting system responsiveness. Frameworks such as Apache Spark, Apache Flink, or Storm enable the dynamic, scalable execution of data-intensive processes on clusters of computers in the cloud. An executive dashboard can employ predictive models to inform high-level decision-making, but reaction and mitigation actions require low latency, as relevant information often deprecates within minutes. Danger warning, evacuation, and damage inspection planning require predictable response times and delivery of optimally chosen resources to field operations in the right place and in time. Such operational research problems lead to dynamic optimization formulations that can also be visualized using simple spatial graphics.

Spark Streaming implements mini-batch sampling, where the constantly accumulating data stream is periodically assembled into batches, allowing workloads to take full advantage of the advanced capabilities of cloud infrastructures. Real-time management of disaster situations poses highly demanding and complicated tasks that often combine streaming with batch processes, but still make it possible to leverage the analytical capabilities of cutting-edge Big Data structures without compromising their inherent characteristics: speed, streaming capability, and the presence of domain-experts who can use the information before it can be assimilated by conventional decision-support systems. Smart city infrastructures and Internet of Things deployments can exploit the availability of resources such as high volume and high velocity data streams to help improve the lives of their citizens.



Illustrative time saved by air transport vs ground



5.2. Data Fusion and Coherence Mechanisms

Real-time data generated through social media, remote sensing, proactive dissemination campaigns and on-site sensor deployments presents unique advantages for disaster management but also formidable challenges. In a single disaster, data may originate from hundreds of thousands of independent sources, reporting on the same topics concurrently and often referring to the same spatial-temporal domains. While the availability of such information can be harnessed by real-time decision-management systems to support impact assessment and resource allocation, it can also produce a massive spike in the required data-processing capacity and create annoying noise in the final outcome.

To address these challenges, complex-event-detection and data-reduction methods can be approached in a more coordinated manner. Three levels of coherence can be defined for disaster-oriented big-data infrastructures: spatial coherence, which guarantees that information originating from the same geographical area is pre-processed together in order to enable coordinated fusion (for instance, sensor readings monitored by crowdsourcing applications); temporal coherence, which imposes the same condition on the time dimension, making it possible to add temporal aggregation procedures to the regular data preparation and enrichment cycle; and semantic coherence, facilitating the integration of pieces of information that use different sensors but still refer to similar phenomena, permitting functionally-oriented aggregations. Such mechanisms can significantly reduce both the amount of information to be processed and the processing latency.

6. Decision-Management Frameworks

Data fusion, predictive modeling, resource allocation, and impact assessment all represent key components of a coherent system for disaster response. Yet, as a real-world application demonstrated during the 2013 Colorado floods, these elements must also be

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

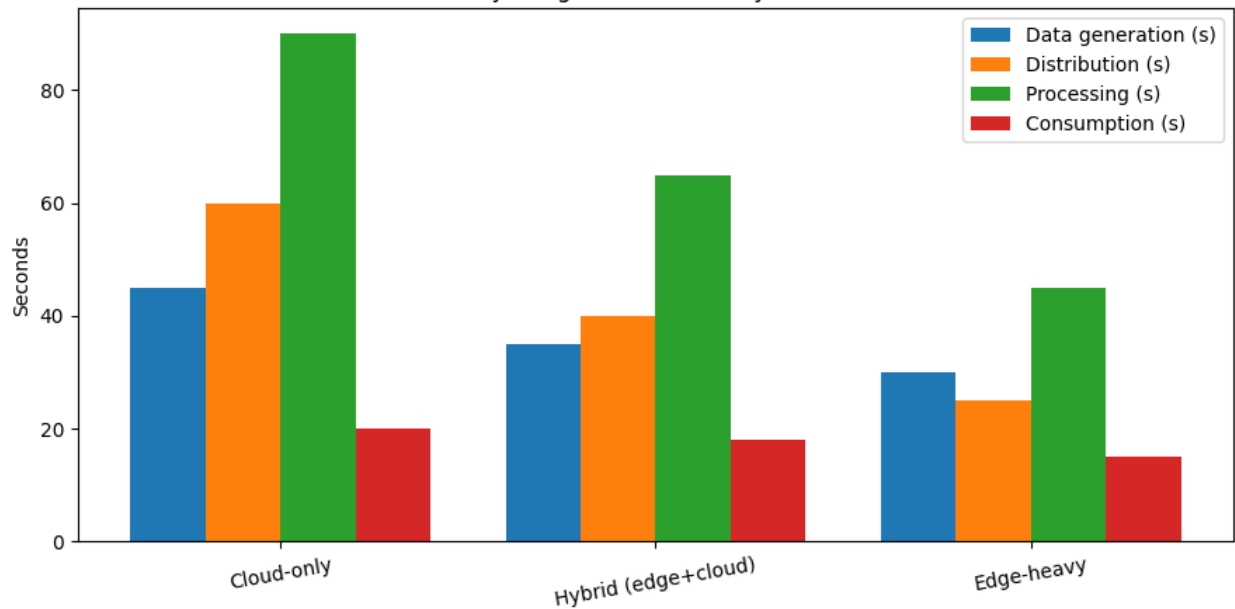
PUBLISHED: JUNE 18



assembled into a decision-management framework capable of supporting simulated-annealing-style operations at multiple layers of a disaster-management command and control hierarchy. Such a framework encompasses a user interface for the control center as well as higher-order support for orchestrating the different missions undertaken by on-the-ground operations. An effective decision-management framework should close the loop around the predictive modeling and resource-allocation functions, feeding their results back into the control-center interface.

Wondrous developments in computerized technology have fundamentally changed the nature of disasters and their consequences. During the 2013 floods in Colorado, for example, the possibility of fatal retreat across the Boulder Canyon was reverse engineered, largely through social-media and satellite data. Rescues considered impossible the day before had occurred by the next morning, aided by crowdsourcing and new approaches in command and control. Capable systems were developed, after the event, to automatically incorporate these lessons into future allocations—a near miracle with respect to past floods. Nonetheless, synchronization between data from the field and the command center remains problematic, indicating that the same data within flows should become the object of on-the-fly fusion for the purpose of secondary, higher-order processes.

Illustrative latency budget breakdown by architecture scenario



6.1. Command and Control Interfaces

A decision-management framework for disaster-management supports officials and first responders leading operations before, during, and after disruptive events. These manage a range of diagnostics, resource supplies, impact estimates, and predictive warnings derived from public-safety incidents, digital volunteer reports, local governmental datasets, on-board time-stamped media, seismological systems, and weather forecasts. Surveillance-camera feeds from local businesses can be included, together with aircraft imagery, and IoT data sparsely covering larger areas. Responses will follow repeated activation and after-tremor scenarios set up by authorities. Prioritized resources will serve mission scenarios gradually saturating the available maritime capacity.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18



Integrating different data sources by means of data fusion techniques is always a challenge, particularly when addressing problems characterized by high uncertainty. Recent research has focused on the recursive distributed and decentralized fusion of data from spatially distributed and possibly heterogeneous sources. Such techniques were developed for the real-time predictive fusion of data from different sensing modalities, working on the recursive storage and fusion of incoming data when new observations become available. These smart data-fusion approaches fully tested on data from different sensing modalities are relatively easy to apply in a disaster-management context.

Equation 5: Priority-based execution + “price interval” rule (formal statement)

Step-by-step formalization

Step 1 (Define operations and damage estimates):

Let operations be $i \in \{1, \dots, R\}$.

Let dx_i be the estimated damage/criticality for operation i .

Step 2 (Define priority ordering):

Let j = index of highest priority operation,

k = index of second priority operation.

(These come from sorting dx_i descending.)

Step 3 (Define allocation cost):

Let C_i be the realized allocation cost of operation i (resources used, monetary cost, etc.).

Step 4 (Price interval rule):

Let $[P(dx_j), P(dx_k)]$ be a cost interval determined by the top priorities.

Then the paper’s rule “if and only if allocation costs belong to a defined price interval...” becomes:

$$\text{Execute } i \Leftrightarrow C_i \in [P(dx_j), P(dx_k)]$$

6.2. Workflow Orchestration for Field Operations

Support for manipulating supporting information in a command-and-control diagram leads to controls for such a design. The emergency management decision-support system enables connections of operations for field teams to online decision-management devices. For incidents categorized as too consequential for tasking by command and-control diagrams alone, authorities establish or delegate responsibilities for high-level management of those operations.

Emergency management command and-control diagrams provide the basis for the standard tasks of a supervisory-control unit, linking incident particulars such as mission, logistical and material requirements, resources needed, and anticipated outcomes. Command-and-control diagrams covering conventional mission responsibilities support a top-down direction mechanism; additionally available symbols suit incident-specific task types, such as execution of prescription-and-mop-up operations on predesignated sites or executing the strategy in semi-structured incidents. In these latter cases, operational-planning officers are authorized to orchestrate all involved assets rather than supervise control-functions allocation on a case-by-case basis, minimizing response time and maximizing morale and collaboration among team members by placing command in familiar hands.



7. Conclusion

AI-based Big Data Systems for Real-Time Disaster Response and Resource Allocation. Use an academic, objective tone with clear, evidence-based statements reflecting the topic.

Timely, accurate information is critical for effective emergency response and management systems during natural disasters. Consequently, Big Data technologies are increasingly deployed for integrating heterogeneous data-streams into unified, low-latency infrastructures. These datasets, employed for crisis management decision support and command-and-control tasks, cover the full cycle of pre-crisis analysis, real-time situational awareness, and coordination of the response and recovery of government agencies and non-governmental organizations.

This work surveyed the available solutions and ongoing research in the context of AI-based Big Data systems, focusing particularly on the real-time operational phase of disaster management. Novel concepts, models, and architectures were proposed for enabling predictive disaster impact assessment and optimizing the online allocation of emergency resources while considering uncertainty in demand estimation. The proposed systems rely on cloud-based distributed Computing for supporting latency-sensitive workloads, integrating data from any number of heterogeneous social-media and sensor-network information streams, and providing coherent situational-awareness services to decision-management frameworks.

Operation	Damage estimate dx	Ground time (min)	Air time (min)	Ground cost (\$k)
Rescue A	0.9	70	25	12
Rescue B	0.7	55	22	10
Medical C	0.8	65	28	11
Supply D	0.5	80	30	9

Table : Toy allocation + helicopter heuristic (illustrative)

7.1. Future Directions

Research on AI and big-data systems for disaster-response applications is still in its early stages. Available approaches are generic and can be adapted to specific use cases, such as floods and earthquakes, only during later phases of a project. Advanced systems are essential, especially for critical components relying on real-time data collection and analysis. A big-data infrastructure dedicated to disaster-response management is needed to deliver valuable real-time information on predictive impact assessment and resource-commitment optimization.

The long-term aim is to develop an integrated platform that offers all the services, models, and predictive analytics required to support decision makers before, during, and after a disaster. Future research activities will include the formal specification and definition of a detailed architecture able to fulfill all requirements, including a workflow for orchestrating the deployment of decision-management systems used by remote commanders to support field operations. Activities will also involve the development of dedicated components, such as a command-and-control system for local smart-city management during disasters



and a generic interface to remotely deploy real-time prediction and resource-allocation models on newly established clusters in the vicinity of an event.

8. References

- [1] Vardhan Kumar Bandi, V. D. (2024). Automated Feature Engineering Systems in Large-Scale Healthcare Data Environments. *Journal of Neonatal Surgery*, 13(1), 2127-2141.
- [2] Guntupalli, R. (2024). Enhancing Cloud Security with AI: A Deep Learning Approach to Identify and Prevent Cyberattacks in Multi-Tenant Environments. Available at SSRN 5329132.
- [3] Anderson, T. (Ed.). (2008). *The theory and practice of online learning* (2nd ed.). AU Press.
- [4] Challa, K. (2024). Enhancing credit risk assessment using AI and big data in modern finance. *American Data Science Journal for Advanced Computations (ADSJAC)* ISSN, 3067-4166.
- [5] Anikina, Z., & Yakimenko, O. (2015). EdTech and digital learning environments in higher education: Trends and outcomes. *Procedia - Social and Behavioral Sciences*, 206, 432-436.
- [6] Inala, R. Revolutionizing Customer Master Data in Insurance Technology Platforms: An AI and MDM Architecture Perspective.
- [7] Baker, R. S. J. d. (2010). Data mining for education. In B. McGaw, E. Baker, & P. Peterson (Eds.), *International encyclopedia of education* (3rd ed., pp. 112-118). Elsevier.
- [8] Paleti, S. (2024). Transforming Financial Risk Management with AI and Data Engineering in the Modern Banking Sector. *American Journal of Analytics and Artificial Intelligence (ajaai)* with ISSN 3067-283X, 2(1).
- [9] Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
- [10] Kummari, D. N., & Burugulla, J. K. R. (2023). Decision Support Systems for Government Auditing: The Role of AI in Ensuring Transparency and Compliance. *International Journal of Finance (IJFIN)-ABDC Journal Quality List*, 36(6), 493-532.
- [11] Bichsel, J. (2013). *The state of e-learning in higher education: An eye toward growth and increased access*. EDUCAUSE Center for Applied Research.
- [12] Sheelam, G. K., & Koppolu, H. K. R. (2024). From Transistors to Intelligence: Semiconductor Architectures Empowering Agentic AI in 5G and Beyond. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 4518-4537.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



- [13] Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.
- [14] Meda, R. (2024). Agentic AI in Multi-Tiered Paint Supply Chains: A Case Study on Efficiency and Responsiveness. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 3994-4015.
- [15] Brinton, C. G., Rill, R., Ha, S., Chiang, M., Smith, R., & Ju, W. (2014). Individualization for education at scale: Improving online learning via automated student modeling. *IEEE Transactions on Learning Technologies*, 7(4), 347–363.
- [16] Singireddy, S. (2024). Applying deep learning to mobile home and flood insurance risk evaluation. *American Advanced Journal for Emerging Disciplinaries (AAJED)* ISSN, 3067-4190.
- [17] Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, 26(1), 293–331.
- [18] A Scalable Web Platform for AI-Augmented Software Deployment in Automotive Edge Devices via Cloud Services. (2024).
- [19] Campos, P., Desmarais, M., & S. d. Baker, R. (2011). Mining constraint-based tutor data for personalization. *Educational Data Mining*, 3(1), 13–26.
- [20] Garapati, R. S. (2023). Optimizing Energy Consumption in Smart Build-ings Through Web-Integrated AI and Cloud-Driven Control Systems.
- [21] Chi, M. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105.
- [22] Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. McGraw-Hill.
- [23] Mashetty, S., Challa, S. R., ADUSUPALLI, B., Singireddy, J., & Paleti, S. (2024). Intelligent Technologies for Modern Financial Ecosystems: Transforming Housing Finance, Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions. *Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions* (December 12, 2024).
- [24] Davuluri, P. S. L. N. (2024). AI-Driven Data Governance Frameworks for Automated Regulatory Reporting and Audit Readiness. *Metallurgical and Materials Engineering*, 30(4), 996–1010. Retrieved from <https://metall-mater-eng.com/index.php/home/article/view/1936>
- [25] Conole, G., & Alevizou, P. (2010). A literature review of the use of web 2.0 tools in higher education. *Higher Education Academy*.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



- [26] Emerging Role of Agentic AI in Designing Autonomous Data Products for Retirement and Group Insurance Platforms. (2024). *MSW Management Journal*, 34(2), 1464-1474.
- [27] Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue. In *Proceedings of the 6th International Conference on Learning Analytics & Knowledge* (pp. 89–98). ACM.
- [28] Dron, J. (2014). *Teaching crowds: Learning and social media*. AU Press.
- [29] Inala, R. AI-Powered Investment Decision Support Systems: Building Smart Data Products with Embedded Governance Controls
- [30] Ellis, R. A., & Goodyear, P. (2010). Students' experiences of e-learning in higher education: The ecology of sustainable innovation. Routledge.
- [31] Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5–6), 304–317.
- [32] Nandan, B. P. (2024). Semiconductor Process Innovation: Leveraging Big Data for Real-Time Decision-Making. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 4038-4053.
- [33] Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64–71.
- [34] Aitha, A. R. (2024). Generative AI-Powered Fraud Detection in Workers' Compensation: A DevOps-Based Multi-Cloud Architecture Leveraging, Deep Learning, and Explainable AI. *Deep Learning, and Explainable AI* (July 26, 2024).
- [35] Gobert, J. D., Baker, R. S., & Wixon, M. (2015). Operationalizing affective constructs in educational software. In *Proceedings of EDM 2015* (pp. 1–8). International Educational Data Mining Society.
- [36] Graesser, A. C., Conley, M., & Olney, A. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (pp. 451–473). American Psychological Association.
- [37] Guntupalli, R. (2023). AI-Driven Threat Detection and Mitigation in Cloud Infrastructure: Enhancing Security through Machine Learning and Anomaly Detection. Available at SSRN 5329158.
- [38] Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(3), 42–57.
- [39] Meda, R. (2023). *Intelligent Infrastructure for Real-Time Inventory and Logistics in Retail Supply Chains*. Educational Administration: Theory and Practice.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



- [40] Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. In Proceedings of the 2nd European MOOC Stakeholder Summit (pp. 58–65).
- [41] Guntupalli, R. (2024). AI-Powered Infrastructure Management in Cloud Computing: Automating Security Compliance and Performance Monitoring. Available at SSRN 5329147.
- [42] Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.
- [43] Hwang, G.-J., Chu, H.-C., Lin, Y.-S., & Tsai, C.-C. (2011). A knowledge acquisition approach to developing mindtools for organizing and sharing differentiating knowledge in a ubiquitous learning environment. *Computers & Education*, 57(1), 1368–1377.
- [44] Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68, 1961–1990.
- [45] Ionita, A., & Dede, C. (2016). Personalized learning and learning analytics in higher education. *Journal of Learning Analytics*, 3(1), 1–12.
- [46] Varri, D. B. S. (2023). Advanced Threat Intelligence Modeling for Proactive Cyber Defense Systems. Available at SSRN 5774926.
- [47] Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (pp. 170–179). ACM.
- [48] Kolla, S. H. (2024). RETRIEVAL-AUGMENTED GENERATION WITH SMALL LLMS FOR KNOWLEDGE-DRIVEN DECISION AUTOMATION IN ENTERPRISE SERVICE PLATFORMS. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 15(3), 476–486
- [49] Koper, R. (2014). Conditions for effective smart learning environments. *Smart Learning Environments*, 1(1), 1–17.
- [50] Bandi, V. D. V. K. (2024). Intelligent Data Platforms For Personalized Retail Analytics At Scale. *Metallurgical and Materials Engineering*, 30 (4), 1011–1027.
- [51] Kuh, G. D. (2009). The national survey of student engagement: Conceptual and empirical foundations. *New Directions for Institutional Research*, 141, 5–20.
- [52] Kwon, K., & Choi, S. (2020). Effects of AI-based recommender systems on learning engagement and achievement in higher education. *Computers in Human Behavior*, 110, 106385.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



- [53] Lahari Pandiri, "AI-Powered Fraud Detection Systems in Professional and Contractors Insurance Claims," *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJREEICE)*, DOI 10.17148/IJREEICE.2024.121206.
- [54] Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 31–40.
- [55] Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- [56] Aitha, A. R. (2023). CloudBased Microservices Architecture for Seamless Insurance Policy Administration. *International Journal of Finance (IJFIN)-ABDC Journal Quality List*, 36(6), 607-632.
- [57] Mayer, R. E. (2009). *Multimedia learning (2nd ed.)*. Cambridge University Press.
- [58] Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
- [59] Milligan, C., & Littlejohn, A. (2017). Why study on a MOOC? The motives of students and professionals. *International Review of Research in Open and Distributed Learning*, 18(2), 92–102.
- [60] Reddy Segireddy, A. (2024). Federated Cloud Approaches for Multi-Regional Payment Messaging Systems. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 15(2), 442–450. <https://doi.org/10.61841/turcomat.v15i2.15464>.
- [61] Mor, Y., & Craft, B. (2012). Learning design: Reflections upon the current landscape. *Research in Learning Technology*, 20, 19196.
- [62] Nelson, K. J., Creagh, T. A., & Clarke, J. A. (2012). Engagement and retention in the first year: Strategies for support. *Higher Education Research & Development*, 31(2), 217–229.
- [63] Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510–1529.
- [64] Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438–450.
- [65] Singireddy, J. (2024). Ai-enhanced tax preparation and filing: Automating complex regulatory compliance. *European Data Science Journal (EDSJ) p-ISSN*, 3050-9572.
- [66] Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review. *Educational Technology & Society*, 17(4), 49–64.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



- [67] Segireddy, A. R. (2024). Machine Learning-Driven Anomaly Detection in CI/CD Pipelines for Financial Applications. *Journal of Computational Analysis and Applications*, 33(8).
- [68] Picciano, A. G. (2012). The evolution of big data and learning analytics in education. *Journal of Asynchronous Learning Networks*, 16(3), 9–20.
- [69] Nagubandi, A. R. (2023). Advanced Multi-Agent AI Systems for Autonomous Reconciliation Across Enterprise Multi-Counterparty Derivatives, Collateral, and Accounting Platforms. *International Journal of Finance (IJFIN)-ABDC Journal Quality List*, 36(6), 653-674.
- [70] Reigeluth, C. M., & Carr-Chellman, A. A. (Eds.). (2009). *Instructional-design theories and models: Building a common knowledge base (Vol. 3)*. Routledge.
- [71] Kaulwar, P. K. (2024). Agentic Tax Intelligence: Designing Autonomous AI Advisors for Real-Time Tax Consulting and Compliance. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 2757-2775.
- [72] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355.
- [73] Rose, C. P., & Siemens, G. (2014). Shared tasks for learning analytics. *Journal of Learning Analytics*, 1(1), 1–5.
- [74] Singireddy, S. (2024). Predictive Modeling for Auto Insurance Risk Assessment Using Machine Learning Algorithms. *European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN*, 3050-9734.
- [75] Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.
- [76] Solomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education*, 23, 1–24.
- [77] Challa, K. (2024). Neural Networks in Inclusive Financial Systems: Generative AI for Bridging the Gap Between Technology and Socioeconomic Equity. *MSW Management Journal*, 34(2), 749-763.
- [79] Simpson, O. (2012). *Supporting students in online, open and distance learning (2nd ed.)*. Routledge.
- [80] Kushvanth Chowdary Nagabhyru. (2023). Accelerating Digital Transformation with AI Driven Data Engineering: Industry Case Studies from Cloud and IoT Domains. *Educational Administration: Theory and Practice*, 29(4), 5898–5910. <https://doi.org/10.53555/kuey.v29i4.10932>.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

ISSN: 3067-1140



[81] Singireddy, S. (2024). Leveraging Artificial Intelligence and Agentic AI Models for Personalized Risk Assessment and Policy Customization in the Modern Insurance Industry: A Case Study on Customer-Centric Service Innovations. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 2532-2545.

[82] Yu, M., Yang, C., & Li, Y. (2018). Big data in natural disaster management: A review. *Geosciences*, 8(5), 165.

[83] Nandan, B. P., & Chitta, S. S. (2023). Machine Learning Driven Metrology and Defect Detection in Extreme Ultraviolet (EUV) Lithography: A Paradigm Shift in Semiconductor Manufacturing. *Educational Administration: Theory and Practice*, 29(4), 4555-4568.

[84] Chamola, V., Hassija, V., Gupta, V., & Guizani, M. (2020). A comprehensive review of the use of machine learning in disaster and pandemic management. *IEEE Access*, 8, 187316–187332.

[85] Singireddy, J. (2024). AI-Driven Payroll Systems: Ensuring Compliance and Reducing Human Error. *American Data Science Journal for Advanced Computations (ADSJAC)* ISSN, 3067-4166.

[86] Xu, C., Xu, X., & Li, Y. (2021). Scalable machine learning techniques for big data analytics in disaster management. *Journal of Big Data*, 8(1), 1–18.

[87] Nandan, B. P. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. *International Journal of Medical Toxicology and Legal Medicine*, 27(5), 759-772.

[88] Banerjee, S., et al. (2020). Scalable frameworks for predictive analytics in disaster management. *ACM Computing Surveys*, 53(2), 1–37.

[89] Challa, K. (2024). Artificial Intelligence and Generative Neural Systems: Creating Smarter Customer Support Models for Digital Financial Services. *Journal of Computational Analysis & Applications*, 33(8).

[90] Wani, M. A., & Jabin, S. (2020). Predictive analytics for disaster management using machine learning. *International Journal of Disaster Risk Reduction*, 48, 101583.

[91] Sriram, H. K., Challa, S. R., Challa, K., & ADUSUPALLI, B. (2024). Strategic Financial Growth: Strengthening Investment Management, Secure Transactions, and Risk Protection in the Digital Era. *Secure Transactions, and Risk Protection in the Digital Era* (November 10, 2024).

[92] Gupta, R., et al. (2020). Real-time flood prediction using machine learning and satellite data. *Remote Sensing*, 12(11), 1842.

[93] Singireddy, S. (2024). The Integration of AI and Machine Learning in Transforming Underwriting and Risk Assessment Across Personal and Commercial Insurance Lines. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 3966-3991.

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 02 ISSUE: 02



RECEIVED: APRIL 11

REVISED: MAY 06

ACCEPTED: MAY 26

PUBLISHED: JUNE 18

[94] Linardos, V., et al. (2022). Machine learning in disaster management: Recent developments in methods and applications. *Big Data and Cognitive Computing*, 4(2), 20.

[95] Nagabhyru, K. C. (2024). Data Engineering in the Age of Large Language Models: Transforming Data Access, Curation, and Enterprise Interpretation. *Computer Fraud and Security*. [78] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30–40.