

Investigation of Framework Conditions for AI-Driven Population Warning and Preparedness Systems in Digital-Generation Crisis Communication

Mesterséges intelligencia alapú polgári védelmi lakossági tájékoztató és felkészítő rendszer keretfeltételeinek vizsgálata a digitális generációk válságkommunikációjában

Tamás Szalóki
Ludovika University of Public Service
MA student
szalokitt@gmail.com
ORCID: 0009-0008-3325-1144 

Absztrakt:

A tanulmány egy helyi telepítésű (on-premise), mesterséges intelligencia alapú polgári védelmi lakossági tájékoztató és felkészítő rendszer megvalósíthatósági keretfeltételeit vizsgálja, különös tekintettel a digitálisan aktív Y és Z generációk specifikus válságkommunikációs igényeire, valamint az online oktatási modellek tesztkörnyezeti implementációjára. A rendszer illeszkedik a Magyarország Mesterséges Intelligencia Stratégiája 2025–2030 célkitűzéseire, amelyek az AI-alkalmazások közszolgálati és katasztrófavédelmi kiterjesztését prioritásként jelölik meg, ugyanakkor a tudomány és a technológia oly gyorsan fejlődik, hogy a stratégia egyes elemei már a kidolgozásuk pillanatában is folyamatos frissítést igényelnek.

Kulcsszavak: mesterséges intelligencia, polgári védelem, lakossági tájékoztatás, digitális iker, kritikus infrastruktúra védelem

Abstract:

This study examines the feasibility framework conditions of a locally deployed (on-premise), artificial intelligence-based civil protection population information and preparedness system, with particular focus on the specific crisis communication needs of digitally active Generation Y and Z, as well as the test-environment implementation of online educational models. The system aligns with the objectives of Hungary's Artificial Intelligence Strategy 2025–2030, which identifies the extension of AI applications to public services and civil protection as a priority, while acknowledging that science and technology are advancing at such a pace that certain elements of the strategy already require continuous updating at the very moment of their formulation.

Keywords: Artificial intelligence, civil protection, population warning, digital twin, critical infrastructure protection

1. INTRODUCTION

The 21st-century security environment, shaped by climatic extremes, hybrid warfare threats, and the growing vulnerability of critical infrastructures, is fundamentally transforming the task system of civil protection information and preparedness. The Scientific Council of the National Directorate General for Disaster Management (BM OKF) recognized in 2025 as a prize-winning research contribution the study demonstrating that artificial intelligence represents both a challenge and an opportunity for revolutionary change in preparing for natural and man-made disasters, one capable of fostering a higher level of population resilience. Podholiczki (2025) comprehensively documents the civil protection application spectrum of artificial intelligence in the Defence Science journal, noting that while technological potential is outstanding in areas ranging from early warning systems to decision support, significant deficiencies remain in the field of real-time, reliable, and personalized population information [1].

Traditional civil protection population warning relies on unidirectional, static channels: leaflets, posters, television and radio announcements, the Electronic Siren System (MoLaRi), the VÉSZ (Emergency Notification Service) mobile application, and the BM OKF Facebook and website information platforms. This model is fundamentally incompatible with the information consumption patterns of digital generations, particularly Generation Z born between 1995 and 2010. Members of Generation Z prefer real-time, personalized, interactive, and mobile-accessible channels, and favor active participatory forms over passive reception [2].

For Generation Alpha, born after 2010, this digital nativeness runs even deeper: they are the first cohort socialized from birth in a world of smart devices, voice-controlled assistants, and short-form visual content. Among Generation Alpha members, video-based, gamified, and instant-feedback formats constitute the primary reception channels rather than text-based messages, creating corresponding expectations for civil protection preparedness systems. At the same time, civil protection cannot overlook older, digitally less active generations, for whom voice-based, natural-language interaction represents the safest path to information. A further critical consideration is that during emergencies, the human capacity of professional staff and dispatch centres is finite; AI agents managing mass population enquiries do not merely perform a convenience function, but become indispensable complements to human activity, relieving operators from routine tasks.

The system aligns with the objectives of Hungary's Artificial Intelligence Strategy 2025–2030, which identifies the extension of AI applications to public services and civil protection as a priority, and emphasizes the defense applicability of domestic HPC infrastructure, including the Komondor and Levente supercomputers and the Data Treasury concept [3]. Science and technology, however, particularly in generative AI and LLM-based agent systems, are advancing so rapidly that certain elements of the strategy require continuous adaptation even at the moment of their formulation. This dynamic is not a weakness but a characteristic of the field, and justifies a continuous, adaptive research approach.

The methodological basis of the research is an interdisciplinary approach combining investigative procedures from security studies, computer science, and social sciences. Based on the theoretical foundation, a systematic literature review was conducted focusing on the latest international reference projects and trends in agent-based artificial intelligence. Analysis of the regulatory and legislative environment is carried out using a doctrinal method, with particular attention to the integration of the EU AI Act and the NIS2 Directive into the domestic legal order. The technological framework was developed using system modelling methodology for the design of the hybrid RAG-based architecture and digital twin simulations. The target group analysis identifies the information consumption patterns of digital generations, and at the conclusion of the research process, the synthesis of collected data yields the professional and technical recommendations required for the final implementation of the system.

2.1. Current Mapping of EU Civil Protection AI Applications

Numerous relevant reference projects can be identified within the European Union. The DRIVER+ (Driving Innovation in Crisis Management for European Resilience) project, funded by the EU's 7th Framework Programme, established a pan-European test-bed-based crisis management innovation platform within which more than 30 organizations tested innovative solutions. The project's Portfolio of Solutions database documented civil protection innovations and provided a model for the sectoral test-bed concept [4].

The ENGAGE project, implemented under Horizon 2020 funding, developed an AI-based chatbot integrated into the EU-ALERT system to strengthen emergency communication and disaster risk resilience. The system, developed by students at Saxion University of Applied Sciences, demonstrated that modular AI architecture is capable of interpreting EU-ALERT message content and generating personalized responses. The beAWARE project, between 2017 and 2019, improved the quality and speed of information sharing by integrating artificial intelligence into early warning systems for floods and storms [5]. The SecurIT programme's AI Disaster Emergency Com project extended the EU-ALERT system with an upward communication channel in which real-time situation reports from citizens are processed by an AI-based pre-diagnostic system [6].

These initiatives serve as technological forerunners, yet also reveal the research gap that forms one of the foundations of the present study. The majority of earlier systems did not yet employ generative language models and digital twin-based simulations for the personalized automation of population preparedness. It is also noteworthy that the native multilinguality of LLM-based architectures opens a new dimension in European-level interoperability. The application of the examined models enables real-time, language-barrier-free adaptation of civil protection information, which is critically important in managing cross-border emergency situations and in handling the linguistic diversity arising from free movement and tourism within the European Union. The developed system can therefore be interpreted not only as a local solution but as a scalable, inclusive protection tool at the EU level.

2.2. Regulatory and Legislative Framework

The EU Artificial Intelligence Act (EU AI Act, 2024/1689/EU), which entered into force in August 2024, employs a risk-based approach whereby civil protection applications fall into the high-risk category, since they are used for classifying emergency calls or prioritizing emergency response, imposing mandatory requirements for transparency, documentation, human oversight, and robustness. Cybersecurity resilience is underpinned by Directive (EU) 2022/2555 (NIS2), which Hungary transposed into domestic law through Act LXIX of 2024 and Government Decree 418/2024 (XII. 23.), entering into force on 1 January 2025. Kis et al. (2025) [7] note in the Military Engineering journal that the domestic introduction of NIS2 presents numerous challenges, as the scope of organizations covered by the directive expanded approximately twelvefold compared to the previous regulation, posing a significant preparedness burden particularly for newly affected organizations not previously subject to regulation. Vászárhelyi (2024) [9] confirms in the National Security Review that NIS2 and the parallel CER Directive together aim to establish a uniform, high level of resilience, also responding to the lessons of the Russo-Ukrainian war. From the perspective of Regulation (EU) 2016/679 (GDPR), the on-premise solution offers a decisive advantage since personal data does not reach third-party cloud infrastructure. The Sendai Framework for Disaster Risk Reduction 2015–2030 and the EU Civil Protection Mechanism (Decision 2021/836/EU) emphasize the strategic importance of preventive, audience-specific information.

3. HYBRID MATHEMATICAL MODELS UNDERPINNING DEEP LEARNING AND PARALLEL ARCHITECTURE

Every form of artificial intelligence, from the simplest regression model to the most complex large language models, is founded on mathematical structures. This fundamental truth is not merely of theoretical significance: when designing a civil protection AI system, understanding the underlying mathematical models is critically important, as they determine what types of problems the system can solve, how quickly it learns, and under what circumstances it may become unreliable. The models presented in the following subsections are not alternatives to one another but complementary, parallel layers of a hybrid architecture, the purpose of which is the adaptive, self-correcting, and optimized operation of civil protection information [7]. Kirovné Rácz and Scholtz (2025) [8] confirm in the Military Engineering journal that, according to the current state of science, combining classical methods with multiple machine learning and artificial intelligence models, including fuzzy logic and neural networks, yields the best results in civil protection applications.

The scientific foundation of the hybrid approach is provided by the ensemble methods literature. The parallel application of different algorithms systematically produces better performance than any single algorithm, particularly in uncertain, dynamic, and partially known environments such as every real emergency situation [10]. The logic of the proposed four-layer hybrid architecture is that MDP and Q-learning provide the adaptive decision foundation, ACO and PSO form the optimization layer, ANFIS constitutes the structured rule layer, and LLM and RAG form the natural language generation and validation layer (Figure 1).

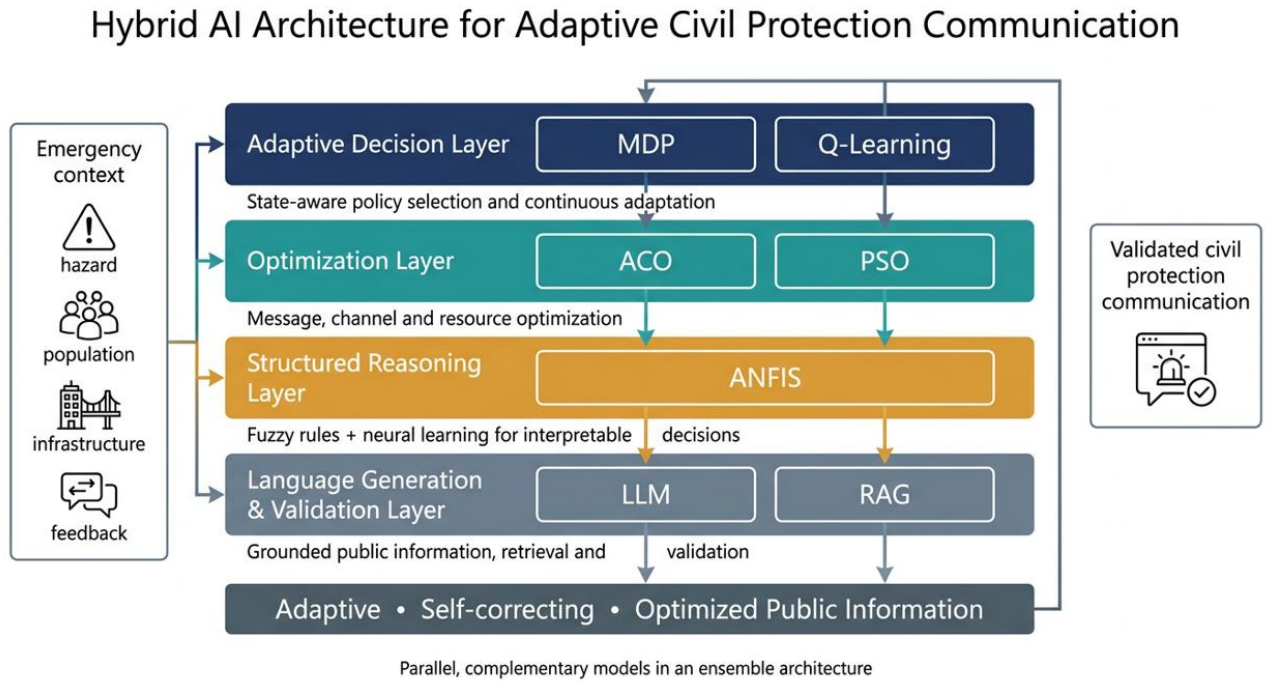


Figure 1: Hybrid AI Architecture for Adaptive Civil Protection Communication. The four-layer ensemble architecture receives emergency context inputs (hazard, population, infrastructure, feedback) on the left and produces validated civil protection communication on the right. The Adaptive Decision Layer (MDP, Q-learning) performs state-aware policy selection; the Optimization Layer (ACO, PSO) handles message, channel and resource optimization; the Structured Reasoning Layer (ANFIS) applies fuzzy rules with neural learning for interpretable decisions; the Language Generation and Validation Layer (LLM, RAG) generates grounded public information with retrieval-based validation. The feedback loop (top arrow) enables continuous system adaptation. Source: Author's own compilation.

3.1. Markov Decision Processes — The Mathematical Framework for Adaptive Decision-Making

The Markov Decision Process (MDP) is the most general mathematical framework for adaptive decision-making, upon which the entire theory of reinforcement learning is built. An MDP is formally defined as a quintuple where S is the set of states, A is the set of possible actions, P is the transition probability function between states, R is the reward function, and γ is the discount factor governing the present value of future rewards. The Markov property of the MDP means that the probability of the next state depends exclusively on the current state and the current action, without requiring knowledge of the full history [11].

$$\text{MDP} = (S, A, P, R, \gamma) \quad (1)$$

where

$$P(s_{t+1} \mid s_t, a_t) = P(s_{t+1} \mid s_t, a_t, s_{t-1}, a_{t-1}, \dots) \quad (2)$$

In the civil protection information system, the MDP state space encodes the user's knowledge level, geolocation, and the current phase of the emergency situation; the action space contains combinations of information content, channel, and message density; and the reward function is definable on the basis of user feedback and knowledge level changes. Alqithami (2025) [12] demonstrated that applying MDP-based reinforcement learning combined with social media data is effective in optimizing resource allocation in emergency situations, directly adaptable to optimizing the prioritization of information channels.

3.2. Q-Learning — The Self-Correcting Solution of the MDP

Q-learning is a model-free reinforcement learning algorithm that learns the optimal policy of an MDP without prior knowledge of the transition probabilities. The algorithm learns a Q value function that estimates the long-term expected reward for each state-action pair. The update rule is summarized below, where α is the learning rate, R is the immediate reward, γ is the discount factor, S' is the next state, and $\max_a Q(S', a)$ is the maximum Q -value achievable in the next state:

$$Q(S, A) \leftarrow Q(S, A) + \alpha [R + \gamma \cdot \max_a Q(S', a) - Q(S, A)] \quad (3)$$

In the civil protection system, this algorithm enables the system to learn from every user interaction and refine the delivery of emergency response protocols. Yu et al. (2021) demonstrated through tabular Q-learning that the algorithm can find the optimal solution in emergency resource allocation tasks up to six times faster than conventional dynamic programming, confirming the method's civil protection applicability. When combined with deep neural networks, the Deep Q-Network approach, it remains applicable in large state spaces [13].

3.3. Particle Swarm Optimization — Optimization in Continuous Parameter Spaces

Particle Swarm Optimization (PSO) is a metaheuristic algorithm inspired by the collective behavior of natural bird flocks, developed by Kennedy and Eberhart in 1995 [14]. PSO searches for the optimum in continuous parameter spaces by having each particle represent a position and velocity vector in the solution space, updated iteratively towards the particle's own best experience and the swarm's global best solution. In the velocity and position update rules, $P_{i,\text{best}}$ denotes the particle's local best position, P_{gbest} the swarm's global best position, and φ_1 and φ_2 the cognitive and social learning coefficients:

$$V_{\text{id}} = V_{\text{id}} + \varphi_1 \cdot \text{rand}_1(P_{i,\text{best}} - X_{\text{id}}) + \varphi_2 \cdot \text{rand}_2(P_{\text{gbest}} - X_{\text{id}}) \quad (4)$$

$$X_{\text{id}} = X_{\text{id}} + V_{\text{id}} \quad (5)$$

In the civil protection system, PSO plays a role in optimising continuous parameters of information content, including message length, visual element ratio, and transmission timing.

Guan et al. (2024) [15] applied multi-objective particle swarm optimization (MOPSO) to optimize the placement of balloon base stations in disaster communication networks and demonstrated that the algorithm's convergence speed and coverage maximization performance substantially exceeds stochastic alternatives, confirming the method's applicability in emergency communication [15].

3.4. Ant Colony Optimization — Graph-Based Priority Optimization

Ant Colony Optimization (ACO) is a metaheuristic algorithm inspired by the pheromone trail-based collective behavior of natural ant colonies, developed by Marco Dorigo in the early 1990s [16]. ACO yields the best performance in discrete, graph-based optimization tasks, where PSO excels in continuous spaces. In the pheromone update rule, τ_{ij} is the pheromone level of edge (i,j) , ρ is the evaporation coefficient, m is the number of ants, Q is a constant, and L_k is the path length traversed by the k -th ant. In the transition probability formula, α and β are weighting exponents and d_{ij} is the edge length:

$$\tau_{ij} \leftarrow (1-\rho)\tau_{ij} + \sum_{k=1}^m (Q/L_k) \quad p_{ij}^k = (\tau_{ij}^\alpha \cdot (1/d_{ij})^\beta) / \sum_{h \in \text{allowed}} (\tau_{ih}^\alpha \cdot (1/d_{ih})^\beta) \quad (6)$$

In the civil protection system, ACO plays a role in prioritizing information channels and user segments: under constrained capacity, it determines the optimal routing for delivering critical information. Muhammad and Wu (2023) [17] applied the ACO algorithm to plan volcanic eruption evacuation routes, demonstrating that simultaneous consideration of five parameters including distance, speed, obstacles, density, and safe points enables the method to generate optimal, dynamically changing evacuation routes. This result is directly adaptable to the channel prioritizations tasks of the civil protection information system [17].

3.5. ANFIS — Synthesis of Human-Readable Rule Systems and Systematic Learning

ANFIS (Adaptive Neuro-Fuzzy Inference System) integrates a Takagi–Sugeno type fuzzy inference system with a neural network learning mechanism. Brassai (2019) [18] documents in detail in his monograph that the combination of human-readable rule systems and systematic parameter tuning satisfies precisely the dual requirement posed by civil protection information, namely the joint demands for comprehensibility and accuracy. In the Gaussian form of the membership function, c denotes the center and σ the standard deviation. In the Takagi–Sugeno rule system, p , q , and r are the adaptively learned consequence parameters. Parameter tuning minimizes the squared error E using a hybrid algorithm of backpropagation and least squares estimation:

$$\mu_A(x) = \exp(-(x-c)^2/2\sigma^2) \quad (7)$$

$$R^{(i)}: \text{IF } x_1 \in A_1^{(i)} \text{ AND } x_2 \in A_2^{(i)} \text{ THEN } y^{(i)} = p^{(i)}x_1 + q^{(i)}x_2 + r^{(i)} \quad E = (1/2) \cdot \sum_k (y_k - \hat{y}_k)^2 \quad (8)$$

ANFIS integrates expert civil protection IF-THEN rules with data-driven learning, providing explainable decision logic for human oversight. Riano-Briceno et al. (2013) [19] applied ANFIS in an emergency classification task and found that the system produces stable outputs, readily interpretable by human experts, in determining triage categories through the synergy of expert knowledge and data-driven learning, which is an indispensable prerequisite for meeting the transparency requirements of the EU AI Act [20].

3.6. Synthesis of the Four-Layer Hybrid Architecture

The parallel application of the four models forms a coherent, multi-layer hybrid architecture in which each layer operates at its own strength while neighboring layers compensate for weaknesses. The adaptive decision core continuously optimizes the information policy based on user feedback through MDP and Q-learning. In the optimization layer, ACO optimizes discrete channel and segment prioritization in a graph-based framework, while PSO handles continuous content parameters; their parallel execution produces significantly better solutions than either alone [10]. In the structured rule layer, ANFIS encodes expert knowledge as explainable fuzzy rules, providing transparent decision logic for human oversight.

In the generation and validation layer, the LLM generates natural-language content, RAG retrieves context from a controlled knowledge base, and the VOTE-RAG ensemble filters hallucinations through cross-validation. This architecture draws a direct parallel with the general logic of deep learning: just as the stacked layers of deep neural networks learn progressively higher-level representations, the layers of the hybrid architecture realize progressively higher-level decision abstractions, from the rewards of individual interactions through optimized channel selection to personalised natural-language information [6]. Fulfilling the public service AI application objectives defined by Hungary's Artificial Intelligence Strategy 2025–2030 requires precisely such multi-layer, validated, human-supervised systems, beyond mere technological innovation rhetoric.

4. LITERATURE REVIEW: AI, DIGITAL TWINS, AND LLM AGENTS

4.1. LLMs and AI Agents in Disaster Communication

The most important academic reference for the application of large language models in disaster communication is the study by Liu et al. (2024) [21] published in the journal *Patterns*, which systematically examines the opportunities, challenges, and future directions of LLMs in the field of disaster management. The authors emphasize that there is a fundamental tension between the informational nature of LLMs and the accuracy requirements of emergency communication, which cannot be resolved without structured validation mechanisms. The agent-based approach represents an advance beyond simple chatbot solutions. Xia et al. (2024) [22] presented at the IEEE ETFA conference work integrating an LLM-based multi-agent system into digital twin simulations in which AI agents make decisions in response to state changes of the simulated physical system [22]. The Scene-Learning-Decision (SLD) model (Xie et al., 2025) [23] is the most advanced published framework for the application of LLM agents in disaster management. The model integrates agent-based simulation with an LLM module: the ABM engine generates the dynamics of the emergency situation, the LLM agent receives the semantic input, and makes collective decisions accordingly. Results validated in a 2022 wildfire case study show that the hybrid system performs substantially better than conventional simulations or standalone LLMs in complex, dynamic emergency environments. Al-Shamrani and Hasan (2025) [24] demonstrated in the CrisisGuideGPT framework that agent-based AI systems are capable of generating contextualised guidance in disaster response plans, drawing a direct parallel with the preparedness function of the system proposed in the present study, as illustrated in Figure 2.

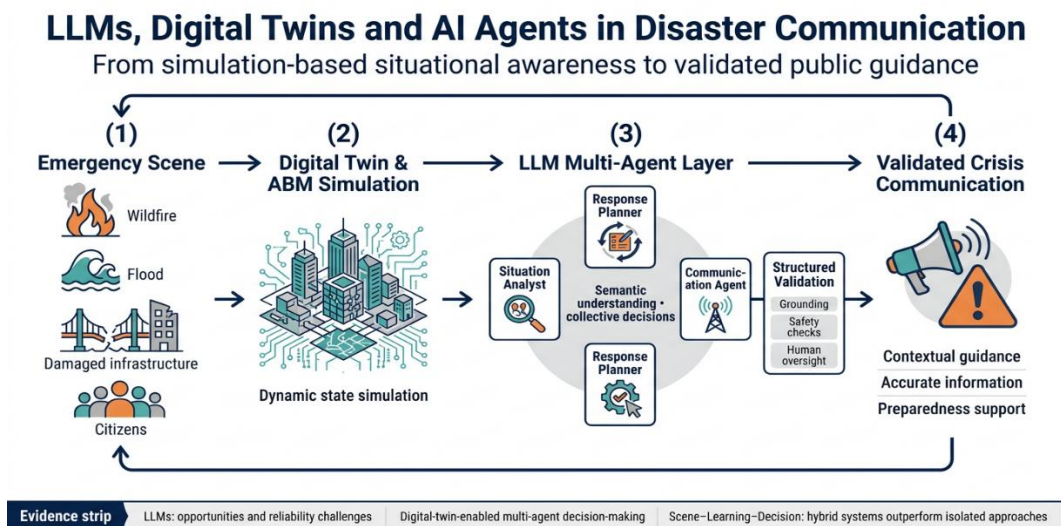


Figure 2: Large Language Models, Digital Twins, and AI Agents in Disaster Communication — from simulation-based situational awareness to validated public guidance. The four-step process proceeds from emergency scene inputs (wildfire, flood, damaged infrastructure, citizens) through dynamic state simulation (digital twin and ABM) to structured validation by a multi-agent LLM layer (Situation Analyst agent, Response Planner agents, Communication Agent), encompassing grounding, safety checks, and human oversight, producing validated crisis communication with contextual guidance, accurate information, and preparedness support. The feedback loop enables continuous system adaptation based on citizen feedback. Source: Author's own compilation.

4.2. Digital Twins and LLMs

The concept of the digital twin originally referred to a digital, real-time replica of a physical system. In the AI era, LLM-based digital twins no longer merely map the states of physical systems but also simulate human behavior, decision-making, and advisory functions [25]. The critical gap identified by Sudanthiran et al. (2025) [26] is that simulation models operate in isolation, unable to dynamically integrate mutually interacting system elements. LLM-based AI agents address precisely this gap: as cognitive architectures they are capable of coordinating various simulations and generating holistic response strategies.

4.3. Challenges of the Hungarian Language

Hungarian as an agglutinative, low-resource language presents particular challenges for LLMs. The data-driven analysis of Harang (2024) [27] demonstrated that the tokenizers of large LLMs process Hungarian text considerably less efficiently than English: the same information in Hungarian requires more tokens, resulting in more expensive, slower, and potentially less accurate processing. From the perspective of confabulation risk, the semantic entropy-based hallucination detection method published by Farquhar et al. (2024) [28] in Nature constitutes a fundamental reference: LLMs hallucinate most frequently when training data on the queried topic is sparse. The VOTE-RAG methodology [29] applies the ensemble voting principle to compare the results of multiple parallel RAG queries, reinforcing only consistent responses.

5. SYSTEM ARCHITECTURE OF THE HYBRID LLM, DIGITAL TWIN, AND AI AGENT MODEL

5.1. Principles of the On-Premise Architecture

The proposed system is based on on-premise operation for four critical reasons. Data sovereignty means that every user interaction and personal data remains on the authority's own infrastructure, meeting the requirements of GDPR and the EU AI Act, as well as the strict data isolation expectations of the NIS2 Directive. Operational autonomy ensures that when internet connectivity is interrupted, civil protection information remains operational in offline mode, which is particularly important in those emergency situations where disruption of communication infrastructure is most likely. For the purposes of auditability and transparency, the system logs every output, fulfilling the mandatory record-keeping obligation under the EU AI Act. Locally running models can be fine-tuned and customized for specific Hungarian civil protection terminology.

5.2. The Three-Layer Digital Twin and AI Agent Architecture

The core of the system consists of a three-layer hybrid architecture integrating digital twin simulation, LLM-based AI agents, and the RAG knowledge base. The first layer is the digital twin simulation layer: simulation models based on real Hungarian civil protection scenarios including fire, flood, chemical accident, and power grid failure, which dynamically alter the state of the scenario in response to user decisions. The second layer is the LLM-based AI agent layer, in which the agent receives the state changes of the simulation as semantic input and generates scenario-specific, action-oriented advice, similarly to the logic of the SLD model and the IVSR architecture [23]. The third layer is the RAG knowledge base: domestic and international official BM OKF documents, regulations, and protocols indexed in a vector database, from which the system retrieves relevant documents for every query and passes them as context to the LLM. The multi-model ensemble layer (on the VOTE-RAG principle) ensures that only responses generated consistently by different models reach the output, providing an extra safeguard against hallucination and fulfilling the human oversight requirement of Article 14 of the EU AI Act.

5.3. Preparedness Module Structure and Multi-Channel Access

The preparedness modules of the system consist of five functional units. Interactive training simulations employ branching decision trees for emergency scenarios in which the LLM agent dynamically responds to and evaluates user decisions, providing feedback on correct and incorrect courses of action. The real-time assistance function enables users to ask natural-language questions in Hungarian, to which the RAG system generates structured responses based on BM OKF protocols. The digital rewards and gamification element, through completed simulations, level progressions, and certificates, increases the reception efficiency of Generation Z, empirically confirmed by the research of Kiss and Kozma (2024) [1]. The community preparedness meter generates an anonymized, aggregated readiness level indicator for the local environment, encouraging neighborhood-level preparedness. The voice-based inclusive agent reaches older generations through Voice-over-IP technology in the form of direct telephone calls: using LLM-based speech-to-text and text-to-speech engines, it provides information in natural Hungarian, simulating a human dispatcher's conversation, and by performing basic triage tasks relieves overburdened dispatch centers during crises.

6. OPERATING ENVIRONMENTS AND VOLUNTEER-BASED PILOT PROGRAMME

6.1. Operating Environment Framework and Justification

The system lifecycle takes place across three clearly delineated environments. The development environment is where prompt design fine-tuning, RAG knowledge base vectorization, and the initial alignment of the multi-model ensemble architecture take place. The pilot or test environment enables controlled, systematic testing of the system with external volunteer participation: the online, closed test environment allows measurement of hallucination rate, response reliability, and user acceptance without errors endangering real emergency operations, confirmed by the EU DRIVER+ project experience as the most reliable validation method [4]. The production environment is the final, scaled state of the system, configured in accordance with the highest security class under the NIS2 Directive, with high availability, cyber-resilient monitoring, and a human oversight interface.

6.2. Volunteer Recruitment and Selection Methodology

The pilot programme operates on a voluntary basis: participation is free, without time commitment obligations, and accessible to all. Planned recruitment channels include official communications from BM OKF and Disaster Management Directorates, students of the Ludovika University of Public Service and the University of Pécs, as well as the Hungarian Red Cross and Hungarian Rescue Foundation. The target volunteer profile consists of smartphone-active users over 18 years of age representing the Y and Z generations.

The question of involving minors under 18 concerns a highly regulated legal area, the framework of which is jointly determined by Article 8 of the GDPR and domestic legislation. Under the GDPR, the processing of personal data in the context of information society services offered directly to children is only lawful if the child has reached the age of 16. For children who have not yet reached the age of 16, the processing of personal data is lawful only if consent has been given by the legal representative exercising parental authority, and the data controller is obliged to conduct a data protection impact assessment, as the processing of data of minors constitutes a high-risk data processing activity. Persons between the ages of 14 and 16 may provide their own consent statement together with a legal representative's consent, while for minors under 14, the legal representative's consent alone is sufficient. Accordingly, for minor volunteer participants in the pilot programme, written parental or legal representative consent, child-comprehensible information about the activity, and a data protection impact assessment carried out by the data controller are required. Correspondingly, the primary target group of the pilot programme consists of volunteer adults over 18; the involvement of 14 to 18 year-olds is feasible only with legal representative consent and the involvement of a school or youth protection organization.

6.3. Performance Indicators and Evaluation Framework

The pilot testing measures four groups of performance indicators. Technical performance indicators include hallucination rate against a reference response set, response consistency score as VOTE-RAG consensus indicator, response time in milliseconds, and offline operational capability. In measuring preparedness effectiveness, the difference in knowledge level between pre- and post-tests, the proportion of correct decisions in simulations, and procedural memory retention, i.e. the result of a repeated test 30 days later, are examined. User acceptance is assessable through SUS (System Usability Scale) score, Net Promoter Score (NPS), and abandonment rates measured from event logs. Hungarian language performance indicators encompass terminological accuracy evaluated by civil protection experts and token efficiency. The volunteer-based emergency preparedness simulation of George Mason University confirmed that the gamification approach significantly increases engagement levels [30].

7. CHALLENGES AND FUTURE DIRECTIONS

7.1. Hungarian LLM and Digital Twin Integration Challenges

The most significant technical challenge is the absence of a reliable, low-hallucination-rate Hungarian-language LLM. The assembly of the fine-tuning dataset required for digital twin simulation, built from BM OKF document corpora, domestic emergency case studies, and legislation, is a time-consuming and resource-intensive process. A critical open question regarding the civil protection application of LLM agents is validatability from a policy-making perspective: Gurcan (2025) [31] notes that agent-based simulations are promising but have not yet gained admission to official state decision-support systems and currently function primarily as academic demonstrations. This observation is parallel to the volunteer-based pilot validation proposed in the present study, which is intended to establish the necessary scientific evidence base.

7.2. Social and Legal Obstacles

One key challenge in gaining acceptance for the system is building institutional trust in AI-generated content. Particular emphasis must be placed on the fact that AI agents do not replace decision-makers but extend their capacity for action: by automating mass, low-complexity enquiries, human specialists can focus on genuine life-saving interventions. The statutory audit under Act XXIII of 2023 and prior consultation with the National Authority for Data Protection and Freedom of Information (NAIH) are indispensable prerequisites for the lawful introduction of the system. The regulatory framework of Hungary's Artificial Intelligence Strategy 2025–2030, together with Act LXXV of 2025 on the domestic implementation of the AI Act, establishes the legal basis for institutional acceptance [3].

7.3. Offline Operation and Security

Offline operational capability is not merely a convenience function but a fundamental security requirement: the system's most critical applications, such as first aid protocols, evacuation routes, and hazardous material identification, are needed precisely when disruption of communication infrastructure is most likely. Edge computing and locally stored, quantized models ensure that the system performs basic functions without internet connectivity.

7.4. International Adaptation Opportunities

The proposed framework is not limited to Hungarian application. The modular architecture of the system allows it to serve as a foundation for similar civil protection systems adapted to other low-resource European languages, including Czech, Slovak, Slovenian, and Croatian.

Within the framework of the Visegrad Cooperation and the EU Civil Protection Mechanism, the system could also be realised as a multilateral development project, facilitating the interoperable management of cross-border emergency situations.

8. SUMMARY AND CONCLUSIONS

Based on the results of the study, it can be established that within the European Union there already exists a mature institutional culture for the development and testing of AI-based civil protection systems, represented by the DRIVER+, ENGAGE, beAWARE, and SecurIT projects, providing a solid methodological foundation for domestic developments [4]. The digital twin-based approach integrated with LLM agents proved to be the most promising technological framework, since through situational orchestration it simultaneously enables real-time scenario adaptation and the development of procedural competence, which traditional, static, and unidirectional information models are unable to provide [23, 26]. The study has demonstrated that the synthesis of mathematical optimization models including MDP, Q-learning, PSO, and ACO with generative AI creates a robust framework that simultaneously addresses technical precision and the human factor. The agent-based approach, complemented by voice-based communication, is capable of eliminating the information gap between digital generations and older cohorts, while offering an operational solution to the shortage of emergency human resources. The system aligns with the objectives of Hungary's Artificial Intelligence Strategy 2025–2030, yet the rapid advancement of science demands continuous adaptation of both technological solutions and regulatory frameworks [3, 9]. The necessary but not sufficient condition for the successful implementation of the system is technological innovation. The transparency requirements of the EU AI Act for high-risk systems, the cyber resilience provisions of the NIS2 Directive, the strict legal framework under the GDPR for the involvement of minors, and the professional approval of BM OKF together form the basis of the system's legitimacy. This complex framework, combining technological, legal, and social aspects, goes beyond an innovative population information tool: it can become one of the fundamental pillars in the creation of modern civil protection.

REFERENCES

- [1] BM OKF Scientific Council for Disaster Management (2025). New perspectives for population preparedness using artificial intelligence [Prize-winning study summary]. Ministry of Interior, National Directorate General for Disaster Management. [Online]. Available: <https://www.katasztrofavedelem.hu/application/uploads/documents/2026-02/86850.pdf> (2026.09.10.)
- [2] Kiss, G. & Kozma, G. (2024). Communication habits of digital generations in the workplace. Jel-Kép: Kommunikáció, Közvélemény, Média, 2024(1). [Online]. Available: https://www.jelkep.eu/pdf/2024/1/JelKep_2024_1__Kiss_Kozma_Georgina.pdf (2026.09.10.)
- [3] Petruska, F. (2025). Connection areas between artificial intelligence and national defence, with particular attention to regulation. Defence-Security Regulatory and Governance Studies, 2025(10). [Online]. Available: https://hhk.uni-nke.hu/document/hhk-uni-nke-hu/VBSZK_MT_2025_10 (2026.09.10.)
- [4] DRIVER+ Consortium (2019). DRIVER+ project in a nutshell. European Commission, 7th Framework Programme. [Online]. Available: <https://www.driver-project.eu/driver-project/> (2026.09.10.)
- [5] EC Research and Innovation (2021). The ultimate platform to handle extreme weather events [beAWARE project]. European Commission. [Online]. Available: <https://projects.research-and-innovation.ec.europa.eu/en/projects/success-stories/all/ultimate-platform-handle-extreme-weather-events> (2026.09.10.)

- [6] SecurIT (2023). AI Disaster Emergency Com. SecurIT Project, European Commission. [Online]. Available: <https://securit-project.eu/funded-projects/ai-disaster-emergency-com/> (2026.09.10.)
- [7] Kis, M., Bódi, A. & Számadó, R. (2025). The process and risks of domestic implementation of NIS2. *Hadmérnök [Military Engineering]*, 19(3), 165–182. [Online]. Available: <https://doi.org/10.32567/hm.2024.3.10> (2026.09.10.)
- [8] Kirovne Rác, R. & Scholtz, E. (2025). Application of artificial intelligence and machine learning algorithms in defence against hydrological disasters. *Hadmérnök [Military Engineering]*, 20(4), 125–147. [Online]. Available: <https://doi.org/10.32567/hm.2025.4.8> (2026.09.10.)
- [9] Vászárhelyi, Ö. (2024). The future of Hungary's cybersecurity in the light of the EU NIS2 Directive. *Nemzetbiztonsági Szemle [National Security Review]*, 12(3), 18–35. [Online]. Available: <https://doi.org/10.32561/nsz.2024.3.2> (2026.09.10.)
- [10] Guo, S. et al. (2023). Ensemble Reinforcement Learning: A Survey. arXiv:2303.02618. [Online]. Available: <https://arxiv.org/abs/2303.02618> (2026.09.10.)
- [11] Sutton, R.S. & Barto, A.G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press. ISBN: 978-0262039246. [Online]. Available: <https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf> (2026.09.10.)
- [12] Alqithami, S. (2025). Leveraging Public Sentiment for Resource Coordination in Disaster Response: A Multiagent Framework. arXiv:2505.03979. [Online]. Available: <https://arxiv.org/pdf/2505.03979> (2026.09.10.)
- [13] Mnih, V. et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518, 529–533. [Online]. Available: <https://doi.org/10.1038/nature14236> (2026.09.10.)
- [14] Kennedy, J. & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, 4, 1942–1948. [Online]. Available: <https://doi.org/10.1109/ICNN.1995.488968> (2026.09.10.)
- [15] Guan, M., Feng, Z., Jiang, S. & Zhou, W. (2024). Tethered Balloon Cluster Deployments and Optimization for Emergency Communication Networks. *Entropy*, 26(12), 1071. [Online]. Available: <https://doi.org/10.3390/e26121071> (2026.09.10.)
- [16] Dorigo, M. & Stützle, T. (2004). Ant Colony Optimization. MIT Press. ISBN: 978-0262042192. [Online]. Available: <https://mitpress.mit.edu/9780262042192/ant-colony-optimization/> (2026.09.10.)
- [17] Muhammad, A. & Wu, T. (2023). Examining the evacuation routes of the sister village program by using the ant colony optimization algorithm. *Open Geosciences*, 15(1), 20220512. [Online]. Available: <https://doi.org/10.1515/geo-2022-0512> (2026.09.10.)
- [18] Brassai, S.T. (2019). Neurális hálózatok és fuzzy logika [Neural Networks and Fuzzy Logic]. Kolozsvár: Scientia Kiadó. ISBN: 978-606-975-007-1
- [19] Riano-Briceno, G. et al. (2013). Comparison of ANFIS and ANN model to categorize patients in the emergency department. *PMC/NCBI*. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3776083/> (2026.09.10.)
- [20] ScienceDirect (2025). Advancements in data-driven evolving fuzzy and neuro-fuzzy control: A comprehensive survey. *Engineering Applications of Artificial Intelligence*. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0952197625000739> (2026.09.10.)
- [21] Liu, Y., Zhang, K., Li, Y. et al. (2024). Towards trustworthy large language models for disaster management. *Patterns*, 5(6), 100940. [Online]. Available: <https://doi.org/10.1016/j.patter.2024.100940> (2026.09.10.)
- [22] Xia, Y., Dittler, D., Jazdi, N., Chen, H. & Weyrich, M. (2024). LLM experiments with simulation: Large language model multi-agent system for simulation model parametrization in digital twins. *IEEE ETFA 2024*. [Online]. Available: <https://doi.org/10.1109/ETFA61700.2024.10710961> (2026.09.10.)

- [23] Xie, Z. et al. (2025). Towards Intelligent Emergency Management: A Scenario-Learning-Decision Framework Enabled by Large Language Models. *Mathematics*, MDPI. [Online]. Available: <https://doi.org/10.3390/math13010000> (2026.09.10.)
- [24] Al-Shamrani, S. & Hasan, R. (2025). CrisisGuideGPT: An Agentic-AI for Crisis Response Planning. 2025 IEEE World Forum on Public Safety Technology (WF-PST), 200–205. [Online]. Available: <https://doi.org/10.1109/WF-PST65083.2025.00037> (2026.09.10.)
- [25] Yang, L. et al. (2025). Leveraging Large Language Models for Enhanced Digital Twin Modeling. arXiv:2503.02167. [Online]. Available: <https://arxiv.org/abs/2503.02167> (2026.09.10.)
- [26] Sudanthiran, A. et al. (2025). Disaster Management in the Era of Agentic AI Systems: A Vision for the Future. arXiv:2510.16034. [Online]. Available: <https://arxiv.org/abs/2510.16034> (2026.09.10.)
- [27] Harang, P. (2024). How efficient are LLM tokenisers on Hungarian text? Medium. [Online]. Available: <https://medium.com/@harangpeter/mennyire-hat%C3%A9kony-az-llm-ek-tokeniz%C3%A1l%C3%B3ja-magyar-sz%C3%B6vegen-ce8481e29f93> (2026.09.10.)
- [28] Farquhar, S., Kossen, J., Kuhn, L. & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, 630, 625–630. [Online]. Available: <https://doi.org/10.1038/s41586-024-07421-0> (2026.09.10.)
- [29] Xie, Z. & Sun, Z. (2025). Mitigating Hallucination in RAG via Ensemble Voting [VOTE-RAG]. arXiv:2603.27253. [Online]. Available: <https://arxiv.org/pdf/2603.27253> (2026.09.10.)
- [30] George Mason University (2024). Transforming emergency response training through AI and interactive games. [Online]. Available: <https://www.gmu.edu/news/2024-12/transforming-emergency-response-training-through-ai-and-interactive-games> (2026.09.10.)
- [31] Gurcan, O. (2025). LLM-augmented agent-based modelling for social simulations: Challenges and opportunities. arXiv:2405.06700. [Online]. Available: <https://arxiv.org/abs/2405.06700> (2026.09.10.)