

## Project Design Document

**Projekt:** Global Unified Automated Response and Disaster Intelligence Alert Network (GUARDIAN AI)

**Przedmiot:** Zarządzanie projektem informatycznym

**Grupa nr:** 2

**Zespół:**

- Ruslan Zhukotynskyi - kierownik
- Remigiusz Sęk
- Jakub Pawlak

### Spis treści:

<b>1. Lab 2</b>	<b>1</b>
a. Idea	1
<b>2. Lab 3</b>	<b>2</b>
a. Koncepcja technologii	2
b. Lista rzeczy	2
<b>3. Lab 4</b>	<b>2</b>
a. Moduł integracji	3
b. Moduł interpretacji AI	3
c. Moduł sugestii służb	3
d. Moduł sprawdzenia wyposażenia	3
e. Moduł informowania służb	3
f. Moduł informowania obywateli	3
<b>4. Lab 5</b>	<b>3</b>
a. Zadanie 1	3
b. Zadanie 2	5
Opis koncepcji	9
Skład zespołu i role	9

## 1. Lab 2

### a. Idea

W związku z narastającym problemem, jakim jest globalne ocieplenie, ludzkość mierzy się z rosnącymi w siłę klęskami żywiołowymi. Nie jest tajemnicą, że temperatura na naszej planecie z roku na rok rośnie, co w końcowym rozrachunku może doprowadzić do przytłaczającej liczby kataklizmów. Do rozwiązania tego problemu można wykorzystać potencjał danych wytwarzanych przez sensory z różnych części ziemi, mogłaby to być dobra okazja do usprawnienia synchronizacji i integracji między istniejącymi już lokalnymi systemami takimi jak np. satelity czy systemy wykrywania trzęsień ziemi. Być może w ten sposób udałoby się lepiej informować i instruować mieszkańców zagrożonych terenów lub poprawić współpracę między organami państwowymi. **Taki projekt mógłby też pomóc z utworzeniem odpowiednich fortyfikacji przed nadciągającymi klęskami żywiołowymi.**

## 2. Lab 3

### a. Koncepcja technologii

Przeprowadzono wstępną analizę możliwości technologicznych związanych z integracją danych o klęskach żywiołowych na terenie Polski. Ustalono, że optymalnym rozwiązaniem jest stworzenie systemu opartego na architekturze API, który będzie integrował dane w czasie rzeczywistym z różnych źródeł, takich jak IMGW, EUMETSAT, NASA FIRMS, Copernicus czy EMSC. Opracowywany system ma działać jako warstwa pośrednia (fusion API), która będzie interpretować wyniki z zewnętrznych interfejsów, przetwarzać je i udostępniać w zunifikowanym formacie dla innych aplikacji i systemów ostrzegania.

W ramach koncepcji zaprojektowano, aby system analizował dane w czasie rzeczywistym dotyczące m.in. powodzi, pożarów, burz, sztormów oraz trzęsień ziemi, a następnie na ich podstawie generował odpowiednie komunikaty ostrzegawcze. Mechanizm ten może być zintegrowany z lokalnymi kanałami powiadamiania, takimi jak SMS, e-mail, aplikacje mobilne czy systemy miejskie, umożliwiając szybkie i automatyczne informowanie mieszkańców zagrożonych obszarów.

Zastosowanie takiej technologii w przyszłości pozwoliłoby zwiększyć skuteczność działań prewencyjnych i operacyjnych służb ratunkowych, usprawnić wymianę informacji między instytucjami państwowymi oraz poprawić ogólny poziom bezpieczeństwa obywateli.

### b. Lista rzeczy

#### i. API

1. NASA FIRMS (do pożarów)
2. EMSC Real-time Feed (do trzęsień ziemi),
3. Copernicus GFM (Sentinel-1) (do powodzi)
4. EUMETSAT Meteosat RSS (do burz/wichur)
5. Copernicus CDS (ERA5-Land) (do upałów/susz)
6. Copernicus Marine (CMEMS) (do sztormów morskich)

#### ii. Inne

1. 6-10 komputerów stacjonarnych o wysokiej mocy obliczeniowej (z kartami GPU) tak aby było można trenować modele sztucznej inteligencji oraz przetwarzać duże zbiory danych,
2. serwer do obsługi API i integracji danych
3. środowisko testowe,
4. środowisko chmurowe,

5. dostęp światłowodowy do Internetu,
6. system zarządzania bazami danych
7. środowisko programistyczne i analityczne na każdym komputerze + github i ERP (Python, TensorFlow, PyTorch, FastAPI, Docker, Visual Studio Code, GitHub / GitLab, Jira / Trello)

### 3. Lab 4

W ramach III poziomu gotowości technologicznej systemu GUARDIAN-AI zostaną przeprowadzone prace badawcze oraz potwierdzenie słuszności założenia koncepcji.

#### a. Moduł integracji

System zbiera dane ze wszystkich wybranych źródeł API w sposób niezależny siebie, następnie normalizuje otrzymane dane do jednolitej postaci w celu prostszej interpretacji. Zbiera te dane i wysyła do innego modułu w znormalizowanej formie.

#### b. Moduł interpretacji AI

Po otrzymaniu znormalizowanych danych z modułu integracji, AI interpretuje wyniki i tworzy wytyczne dla odpowiednich służb odnośnie zalecanych poczynąń oraz wyposażenia. W ramach tego etapu opracowany zostanie wstępny model AI, który zostanie wytrenowany z wykorzystaniem przygotowanych zbiorów danych symulacyjnych oraz dostępnych archiwalnych danych środowiskowych pochodzących z publicznych źródeł. Model ten będzie miał za zadanie poprawnie analizować przekazane dane wejściowe, identyfikować wzorce charakterystyczne dla poszczególnych zjawisk oraz generować komunikaty i zalecenia operacyjne dla służb.

#### c. Moduł sugestii służb

Na podstawie interpretacji AI oceniane jest ryzyko katastrofy i szacowane są potrzebne zasoby do rozwiązania sytuacji. Gdy zasoby są już oszacowane poszukiwane są odpowiednie służby w podanej ilości, następnie jest sprawdzane wyposażenie danych służb, jeżeli spełniają kryteria to dane są przekazywane do modułu wysyłki, jeżeli nie to szuka się dalej aż zostaną spełnione wymogi AI.

#### d. Moduł sprawdzenia wyposażenia

Każda służba musi prowadzić monitoring stanu wyposażenia oraz zasobów ludzkich dla informacji ogólnej na podstawie której można szybko przesyłać informacje zwrotne do systemu bez opóźnień w czasie rzeczywistym.

#### e. Moduł informowania służb

Zalecenia otrzymane przez AI zostają rozesłane do odpowiednich służb, z informacjami takimi jak zalecane wyposażenie, miejsce zjawiska, zalecanymi działaniami oraz siłą kataklizmu.

## **f. Moduł informowania obywateli**

Rozesłane zostaną na numery telefonów obywateli zagrożonego obszaru odpowiednie informacje. Numery telefonów i adresy urzędów docelowych są wybierane na podstawie geolokalizacji oraz danych administracyjnych z rejestrów mieszkańców lub lokalnych baz danych.

## **4. Lab 5**

### **a. Zadanie 1**

- i. Moduł integracji - Pobrano dane z różnych API z losowo wybranych 50 miejsc z Polski. Moduł integracji przetwarzał początkowo dane bez dodatkowych poprawek, co ujawniło błędy w normalizacji, takie jak różne jednostki miar temperatury, niejednolite znaczniki czasu, brakujące pola oraz różny sposób określania lokalizacji – w niektórych przypadkach obszar katastrofy był podany jako współrzędne punktowe, a w innych jako szerszy region administracyjny.

Po wprowadzeniu funkcji korekcji i standaryzacji danych, w tym konwersji jednostek, ujednolicenia formatu daty, walidacji pól obowiązkowych oraz normalizacji sposobu zapisu lokalizacji, moduł poprawnie znormalizował wszystkie rekordy i przesyła je dalej do modułu interpretacji AI.

- ii. Moduł interpretacji AI - Przygotowano zestaw kontrolny 200 przypadków symulacyjnych obejmujących pożary, powódzie, burze i trzęsienia ziemi z wykorzystaniem danych archiwalnych wykonano 100 serii testów weryfikujących poprawność analizy danych wejściowych, identyfikacji wzorów zjawisk oraz generowania komunikatów operacyjnych.

Model wykazał odchylenia w rozpoznawaniu intensywności burz - błędnie klasyfikował 15% przypadków. Ustalono, że na wynik ma wpływ niedostateczna reprezentacja danych o burzach w zbiorze treningowym. Uzupełniono zbiór o 50 dodatkowych przypadków i przeprowadzono trening na nowo. Po ponownych testach dokładność klasyfikacji wzrosła do 94% a generowane zalecenia były zgodne z procedurami operacyjnymi służb.

- iii. Moduł sugestii służb - Przygotowano zestaw kontrolny 80 scenariuszy katastroficznych z różnymi poziomami ryzyka. Wykonano 80 serii testów weryfikujących poprawność oszacowania zasobów, doboru służb według kryteriów oraz walidacji wyposażenia.

Moduł wykazał opóźnienia w wyszukiwaniu służb dla scenariuszy wymagających więcej niż 5 typów jednostek jednocześnie. Ustalono, że na wynik ma wpływ nieoptymalne zapytanie do bazy danych służb. Zoptymalizowano algorytm wyszukiwania i ponownie wykonano testy. Czas doboru służb skrócił się o 60%, a wszystkie scenariusze zostały obsłużone zgodnie z wymogami AI.

- iv. Moduł sprawdzenia wyposażenia - Przygotowano zestaw kontrolny 30 jednostek służb ratunkowych zróżnicowanych pod względem stanu wyposażenia, liczby personelu oraz dostępności pojazdów. Wykonano 150 serii testów weryfikujących kompletność danych, poprawność mapowania pól oraz aktualność raportów przesyłanych w czasie rzeczywistym. W trakcie testów zidentyfikowano niespójność danych pomiędzy jednostkami. W 7 przypadkach raporty zawierały nieaktualne informacje o liczbie dostępnych pojazdów lub błędne oznaczenia typów sprzętu. Ustalono, że przyczyną był brak walidacji schematu danych oraz różnice w formacie raportowania pomiędzy jednostkami.

Wprowadzono jednolity schemat raportowania, ujednolicono nazewnictwo pól oraz dodano automatyczną walidację kompletności rekordów. Po ponownym wykonaniu 150 serii testów wszystkie dane były aktualizowane w czasie poniżej 10 sekund, a spójność i kompletność rekordów osiągnęły poziom 98%.

- v. Moduł informowania służb - Przygotowano zestaw kontrolny 60 komunikatów dla różnych typów służb i poziomów zagrożenia. Wykonano 60 serii testów weryfikujących kompletność przekazanych informacji (wyposażenie, lokalizacja, działania, intensywność) oraz czas dostarczenia.

Moduł wykazał brakujące dane o sile kataklizmu w 20% komunikatów dla zdarzeń o średnim ryzyku. Ustalono, że na wynik ma wpływ nieprawidłowe mapowanie danych z modułu interpretacji AI. Poprawiono schemat integracji danych i ponownie wykonano testy. Wszystkie komunikaty zawierały kompletne informacje i były dostarczane w czasie do 5 sekund.

- vi. Moduł informowania obywateli - Przygotowano zestaw kontrolny 500 numerów testowych rozlokowanych w 10 różnych strefach geograficznych. Wykonano 100 serii testów weryfikujących poprawność selekcji numerów na podstawie geolokalizacji, treść komunikatów oraz pokrycie zagrożonego obszaru.

Moduł wykazał pominięcie 12% numerów w strefach granicznych między dwoma obszarami administracyjnymi. Ustalono, że na wynik ma wpływ precyzyjny algorytm określenia przynależności do strefy. Zmodyfikowano logikę geolokalizacji z uwzględnieniem buforów granicznych i ponownie wykonano testy. Pokrycie obszaru wzrosło do 99.5% a wszystkie komunikaty zostały dostarczone w ciągu 30 sekund.

## **b. Zadanie 2**

- i. Ryzyko 1 - Istnieje ryzyko otrzymania błędnych lub niepełnych danych z jednego z zewnętrznych API. Może się to zdarzyć w wyniku awarii serwisu lub chwilowego braku połączenia. W takiej sytuacji system GUARDIAN-AI może błędnie ocenić sytuację zagrożenia lub pominąć niektóre zdarzenia, co zmniejsza skuteczność jego działania.

Rozwiązanie: Aby ograniczyć to ryzyko, system zostanie wyposażony w mechanizm automatycznej weryfikacji danych pochodzących z różnych źródeł. Jeśli jedno API przestanie działać lub zwróci dane o podejrzanym strukturze, moduł integracji automatycznie wykorzysta alternatywne źródła lub dane historyczne. Dodatkowo przewidziana jest możliwość automatycznego raportowania błędów do zespołu technicznego oraz bezpośredniego kontaktu z dostawcami danych, aby możliwe było szybkie usunięcie przyczyny problemu.

- ii. Ryzyko 2 - Opóźnienia w czasie rzeczywistym przy synchronizacji danych z wielu źródeł API.

Rozwiązanie: Zastosowanie asynchronicznego pobierania danych z kolejkami priorytetowymi i mechanizmem cache'owania dla danych o niskiej zmienności.

- iii. Ryzyko 3 - Istnieje ryzyko, że dane dotyczące wyposażenia i zasobów służb ratowniczych w module sprawdzenia wyposażenia mogą być nieaktualne. Może to prowadzić do sytuacji, w której system przydzieli jednostkę, która w danym momencie nie dysponuje wymaganym sprzętem lub personelem, co obniża skuteczność akcji ratowniczej.

Rozwiązanie: W celu wyeliminowania tego ryzyka zostanie wdrożony system tzw. heartbeat monitoring, który będzie cyklicznie sprawdzał i potwierdzał aktualność danych przekazywanych przez poszczególne jednostki. Dodatkowo przed przydzieleniem zadania system automatycznie przeprowadzi zapytanie weryfikujące stan sprzętu i gotowość operacyjną służb, aby mieć pewność, że informacje są aktualne i wiarygodne.