Додаток 2. Технічне завдання зі створення моніторингової бази даних Truth Hounds Версія 1.0

1. Передумови

З 2023 року ТН підтримує спеціалізовану **Моніторингову базу даних**, побудовану на архітектурі MySQL + PHP, яка містить понад 85 000 задокументованих інцидентів, які містять ознаки воєнних злочинів у російсько-українській війні.

Система слугувала центральним сховищем для зберігання, класифікації та аналізу інцидентів, але її поточна архітектура обмежує масштабованість, автоматизацію та інтеграцію з сучасними інструментами. Для забезпечення сталої діяльності та підготовки до майбутніх аналітичних потреб, ТН прагне **переписати та модернізувати Моніторингову базу даних**, зберігаючи всі дані та докази, а також забезпечуючи ефективну моніторингову роботу.

2. Мета тендеру

Мета проєкту – створити **безпечну, модульну, автоматизовану платформу**, яка покращить управління даними, інтегрує передові аналітичні інструменти та підтримуватиме постійну документацію міжнародних злочинів.

Обраний підрядник буде відповідальним за **проєктування, розробку, інтеграцію, міграцію** та **довгострокову підтримку і обслуговування** нової системи, працюючи в тісній координації з ТН в рамках **Agile, ітеративного процесу**.

3. Обсяг робіт

Обсяг проєкту охоплює наступні компоненти:

3.1. Основні завдання розробки

- Спроєктувати та впровадити нову структуру бази даних на основі концептуального документа та специфікацій, наданих ТН.

- Розробити веб-інтерфейси користувача (українською та англійською мовами) для введення даних, пошуку, візуалізації та звітності.

- Впровадити основні модулі, включаючи:

    - **Управління інцидентами** – створення та редагування інцидентів, згрупованих під спільними фактичними наративами ("сюжетними лініями").

- **Управління доказами** – ручне та автоматизоване завантаження файлів (фото, скріншоти, відео, PDF) з повним відстеженням метаданих.
- **Модуль осіб** – структуровані записи для жертв, свідків, співучасників та інших ролей, з перехресними зв'язками між інцидентами та валідацією даних.
- **Управління локаціями** – ієрархічна географічна структура (країна → регіон → населений пункт → вулиця → будівля) з редагованими координатами та підтримкою картографування кількох районів.
- **Правова кваліфікація та модуль відповідальності** – багаторівнева класифікація інцидентів та акторів (підрозділи, формування, організації).
- **Інструменти автоматизації** – отримання даних через API, автоматичне створення скріншотів, проставлення часових міток, виявлення дублікатів, підтримка транскрипції та перекладу, допомога з граматикою.
- **Моніторинг наративів** – відстеження та аналіз інформаційних кампаній, пропаганди та дезінформації з можливістю зв'язування з фабулою та категоризації за джерелами, аудиторією та темами.

3.1.1 Управління інцидентами

Модуль управління інцидентами формує основу Моніторингової бази даних Truth Hounds (TH-MD). Він повинен поєднувати встановлену логіку існуючої системи з покращеною структурною гнучкістю та автоматизацією.

Основні принципи

- **Один інцидент = один фактичний запис.** Кожен інцидент представляє окрему подію, верифіковану або спостережену з відкритих або закритих джерел.
- **Групування за сюжетною лінією.** Декілька інцидентів можуть бути пов'язані під єдиним загальним наративом («сюжетна лінія» або «ланцюг справи»), що представляє один фактичний контекст (наприклад, тривалий обстріл або послідовність пов'язаних атак).
- **Відображення пов'язаних епізодів.** Пов'язані інциденти в межах однієї сюжетної лінії повинні відображатися як вертикальний ланцюг зі спільним описом, при цьому кожен інцидент зберігає власні метадані, докази та правову кваліфікацію.

Ключові функції

- **Інтерфейс створення та редагування** зі структурованими полями даних та інтелектуальною валідацією.
- **Розпізнавання приміток монітора** – текст у дужках після «(прим. мон.: …)» розглядається як приховане поле, яке можна фільтрувати або експортувати окремо.
- **Текстові шаблони** для ефективного введення даних, що дозволяють вставляти типові фрази або структури за допомогою гарячих клавіш.
- **Автоматичне проставлення часових міток** для подій створення, модифікації та верифікації.

- **Перехресні посилання** на докази, локації та осіб.
- **Відстеження статусу** (редагується, первинний варіант, виправлений, потребує перевірки, перевірений) з кольоровим відображенням.

3.1.2 Управління доказами

Модуль управління доказами забезпечує надійне зберігання, збереження та зв'язування цифрових матеріалів, що підтверджують кожен інцидент.

Основні функції

- **Ручне та автоматизоване завантаження** скріншотів, відео, фото, PDF та посилань.
- **Автоматичне створення скріншотів** при вставці посилання (headless браузер або через API).
- **Захоплення метаданих** – часові мітки, URL-адреси джерел, геолокація (якщо доступна) та EXIF-дані.
- **Виявлення дублікатів** за допомогою хешування файлів та порівняння посилань.
- **Розділення типів доказів:** імпортовані через API або додані вручну.
- **Дотримання ланцюга зберігання:** всі модифікації логуються з посиланням на користувача та час.
- **Попередній перегляд мініатюр та версіонування** для кожного медіа-запису.

3.1.3 Модуль осіб

Цей модуль забезпечує структуроване записування та відстеження зв'язків усіх осіб, пов'язаних з інцидентами.

Ключові вимоги

- **Унікальний ідентифікатор особи**, що призначається автоматично.
- **Детальні атрибути:** ім'я, псевдоніми, стать, дата народження або діапазон, адреса (пов'язана з ієрархією локацій) та категорія ролі.
- **Категоризація:** жертва, свідок, загиблий,співучасник.
- **Міжінцидентні зв'язки:** та сама особа може з'являтися в кількох інцидентах; система повинна виділяти незв'язані дублікати.
- **Фільтрація та візуалізація:** фільтр осіб за роллю, типом інциденту або відповідальним підрозділом; показ усіх інцидентів, пов'язаних з однією особою.

3.1.4 Управління локаціями

Ключові функції

- **Ієрархічна структура:** Країна → Регіон → Район → Громада → Старостат → Населений пункт → Район/Мікрорайон → Вулиця → Будівля.
- **Підтримка кількох районів:** окремі локації можуть належати до адміністративних меж, що перекриваються.
- **Динамічне редагування:** авторизовані користувачі можуть створювати або редагувати локації безпосередньо у формах інцидентів.
- **Точне відображення координат:** клацання по координатах відображає точну точку на карті, а не приблизну область.
- **Інтеграція з інструментами картографування** (OpenStreetMap/Leaflet) та можливість майбутнього розширення ГІС (схема, сумісна з ArcGIS).

3.1.5 Модуль правової кваліфікації та відповідальності акторів

Основні концепції

- **Множинні кваліфікації на інцидент** (наприклад, воєнний злочин + напад на цивільний об'єкт).
- **Ієрархія акторів:** державні та недержавні актори → підрозділи або формування → окремі командири або члени.
- **Зв'язування відповідальності:** кожен актор може бути пов'язаний з інцидентами або особами за роллю та датою.
- **Фільтрація та аналітика:** можливість генерувати списки або карти інцидентів за актором або підрозділом.
- **Довідник акторів:** постійно оновлювана таксономія, що редагується адміністраторами.

3.1.6 Інструменти автоматизації та ШІ

Ключові автоматизації

- **Отримання даних через API** з попередньо визначених відкритих джерел (Telegram, YouTube, Facebook, веб-сайти).
- **Автоматична транскрипція** аудіо/відео (Whisper, Deepgram тощо) з редагованим текстовим виводом.
- **Автоматичний переклад** (DeepL, Google Translate API) з ручним постредагуванням.
- **Мовна допомога** через інтегровану перевірку граматики (LanguageTool або Grammarly API).

- **Набір часових міток:** часова мітка джерела, архівна часова мітка, часова мітка верифікації та часова мітка ручного редагування.

Усі автоматизації повинні включати опції перевизначення користувачем та повне логування аудиту.

3.1.7 Моніторинг наративів

Модуль моніторингу наративів призначений для відстеження та аналізу інформаційних кампаній, пропаганди та дезінформації. На відміну від модуля інцидентів, який фіксує порушення з ознаками злочину, цей модуль зосереджений на моніторингу публічних повідомлень та наративів без обов'язкової правової кваліфікації.

Основні принципи

- **Структурована одиниця моніторингу** – кожен запис наративу є окремою одиницею спостереження з унікальним ідентифікаційним номером.
- **Зв'язок з фабулою** – наративи можуть бути опціонально пов'язані з існуючими сюжетними лініями («фабулою») для кращого контекстуального зв'язку.
- **Фіксація первинного контенту** – збереження оригінального тексту публікацій для подальшого аналізу.

Ключові функції

- **Унікальна нумерація:** кожному наративу автоматично присвоюється унікальний ідентифікаційний номер для відстеження.
- **Дата публікації:** обов'язкове поле для фіксації часу оприлюднення повідомлення з підтримкою точної дати.
- **Зв'язок з фабулою:** опціональне поле для прив’язки наративу до існуючої сюжетної лінії.
- **Повідомлення:** текстовий блок для збереження скопійованого тексту посту або публікації з джерела.
- **Опис наративу:** структурований опис, схожий на опис фабули, що узагальнює ключові тези та меседжі.
- **Файли доказів:** можливість прикріплення скріншотів, відео, фото, а також інформацію з датою, часом публікації (заповнювати за замовчуванням із метаданих, якщо такі наявні), посилання на неї та інших підтверджуючих матеріалів.
- **Тип джерела:** дропдаун-список для категоризації джерела (соціальні мережі, державні ЗМІ, незалежні ЗМІ, офіційні заяви, Telegram-канали тощо).
- **Аудиторія:** дропдаун-список для визначення цільової аудиторії наративу (російська, українська, міжнародна).
- **Теми та ключові слова:** використання тієї ж системи ключових слів, що і в модулі інцидентів, для категоризації тематики.

- **Автор або цитовані особи:** посилання на записи в модулі осіб для відстеження авторів повідомлень або згаданих персон.
- **Категорія:** можливість присвоєння однієї або кількох категорій (легітимізація, заперечення злочинів тощо).

Аналітичні можливості

- **Витяги та звіти:** генерація структурованих витягів на основі відфільтрованих наративів з можливістю групування за темами, датами, джерелами або категоріями.
- **Розширений фільтр:** фільтрація за всіма полями метаданих, включаючи дату публікації, тип джерела, аудиторію, категорії та зв'язок з фабулою.
- **Експорт у набір документів:** можливість експорту відібраних наративів у форматах XLSX, CSV, PDF з налаштуванням структури експорту.
- **Інтеграція з експортом епізодів:** при експорті епізодів інцидентів передбачена опція включення пов'язаних наративів для комплексного аналізу інформаційного контексту.

3.1.8 Контроль якості та робочий процес

Ключові правила робочого процесу

- **Система статусів:** чернетка → попередній → переглянутий → перевірений.
- **Валідація супервайзером:** записи про те, хто перевірив і коли.
- **Виявлення дублікатів:** попередження про повторювані URL, докази або ідентичні описи.
- **Закладки** лише для користувача, які можна фільтрувати для особистого робочого процесу.
- **Замітки для внутрішнього користування:** коментарі або позначки, видимі для конкретних ролей.

3.1.9 Звітність та експорт

Можливості

- **Генерація власних витягів** на основі інцидентів, сюжетних ліній або осіб.
- **Гнучкі формати експорту:** XLSX, CSV, PDF та інтеграція з Google Docs.
- **Консолідований експорт наративів:** об'єднання всіх епізодів під однією сюжетною лінією для зв'язної звітності.
- **Статистична звітність:** агрегація за регіоном, актором, типом події та періодом часу.

3.2. Міграція та синхронізація даних

- Експортувати, очистити та імпортувати всі дані з існуючої системи MySQL + PHP.

- Зберегти всі зв'язки між записами (інциденти, докази, особи та правові класифікації).
- Забезпечити верифікацію цілісності через журнали та хешування.
- Впровадити механізм **двосторонньої синхронізації** з поточною базою даних, що дозволить паралельне використання обох систем під час переходу.
- Запланувати повну міграцію та вимкнення старої системи до кінця проєкту.

3.3. Безпека та контроль доступу

- Рольовий контроль доступу з рівневими дозволами (адміністратори, монітори, співробітники, партнери, користувачі з обмеженим доступом).
- Двофакторна автентифікація (ім'я користувача/пароль + TOTP).
- Журналювання та аудиторський слід активності користувачів та редагувань даних.
- Всі дані шифруються під час передачі та зберігання.
- Відповідність GDPR та європейським стандартам конфіденційності.

3.4. Візуалізація та аналітика

- Інтерактивне геопросторове картографування (OpenStreetMap або Leaflet) з точним відображенням координат.
- Візуалізація часової шкали інцидентів та подій.
- Картографування зв'язків між сутностями (Neo4j або еквівалентний графовий движок).
- Аналітичні панелі та експорт звітів (Chart.js, Plotly або подібні).

3.5. Інтерфейс користувача та досвід користування (UI/UX)

- Чистий та інтуїтивний дизайн, оптимізований для нетехнічних користувачів.
- Багатомовна підтримка (українська та англійська як мінімум).
- Розширений пошук та фільтрація за всіма метаданими.
- Функції модерації контенту (фільтри розмиття для графічного контенту).
- Кольорове кодування статусів для етапів робочого процесу (редагування, первинний варіант, потребує перевірки, виправлений, перевірений).
- Приватні примітки користувачів та відстеження завдань з можливістю фільтрації.

3.6. Документація, передача та підтримка

- Надати технічну документацію (архітектура, API, модель даних, посібник з встановлення).
- Надати посібники користувача для моніторів та аналітиків.
- Провести навчальні сесії для команди Truth Hounds.
- Запропонувати **довгострокову технічну підтримку**, включаючи:
    - Постійні оновлення та виправлення помилок.
    - Підтримка сумісності з оновленими браузерами, ОС та залежностями.
    - Оптимізація продуктивності та патчі безпеки.
    - Опціональна розробка функцій на основі змінних потреб ТН.

- Гарантований час відповіді та рівні обслуговування як частина SLA (Угода про рівень обслуговування).

4. Процес розробки

Проєкт повинен дотримуватися **методології Agile**, з акцентом на гнучкість, співпрацю та ітеративну доставку.

4.1. Вимоги до процесу

- Розробка відбуватиметься в **спринтах** (кожні два тижні або частіше за готовності) з регулярними циклами зворотного зв'язку.
- ТН надаватиме постійні коментарі та тестування протягом усієї розробки.
- Заохочуються ранні прототипи та мінімально життєздатні версії для огляду та тестування.
- Документація та оновлення беклогу необхідні в кінці кожного спринту.

4.2. Співпраця

- Щотижневі зустрічі щодо прогресу з менеджером проєкту ТН.
- Спільна платформа для відстеження проблем та документації (наприклад, GitLab, Jira або еквівалент).
- Весь вихідний код та документація залишаються власністю Truth Hounds.

5. Результати та віхи

Віха	**Опис**
M1	Проєктування архітектури та прототип з мінімальним зразком міграції
M2	Функціонування основних модулів (інциденти, особи, локації, докази)
M3	Інструменти автоматизації, візуалізація, встановлення синхронізації
M4	Двостороння синхронізація активна; внутрішня міграція та тестування QA
M5	Повна міграція даних, фінальне тестування, документація та вимкнення старої БД
M6	Ініційована фаза довгострокової підтримки (згідно SLA)

6. Конфіденційність

Всі матеріали, дані та інформація, надані або створені підрядником в рамках цього проєкту, є конфіденційними та залишаються власністю Truth Hounds. Підрядник не повинен розкривати або повторно використовувати будь-які такі матеріали без попередньої письмової згоди.