

MetaTrace: Web-Based Metadata Forensics Tool

Final Presentation

WE Group 12

April 15, 2025

Problem Statement

- **Importance of Metadata:**

- Critical for verifying file authenticity, ensuring data integrity, and supporting investigations.
- Offers insights into file origins, modifications, and hidden data.

- **Challenges with Existing Tools:**

- Advanced tools like ExifTool require deep technical expertise.
- Lack integrated visualization and intuitive, predictive analytics.
- Limited accessibility for non-technical users, hindering broader adoption.

MetaTrace: A Comprehensive Metadata Analysis Platform

- **User-Friendly Interface:** A web-based tool that simplifies metadata extraction, visualization, and analysis.
- **Broad File Format Support:** Extracts metadata from images, documents, multimedia files, and more.
- **Advanced Analytics:** Predictive insights and anomaly detection.
- **Empowering Users:** Makes metadata accessible for both technical experts and non-technical users to verify authenticity, trace leaks, and investigate cybercrimes.

Key Features

- ① **File Upload:** Supports images, PDFs, and documents.
- ② **Metadata Extraction:** Uses ExifTool for detailed metadata analysis.
- ③ **Predictive Insights:** Detect anomalies in metadata.

Technology Stack:

- **Frontend:** Next.js, Tailwind CSS
- **Backend:** Python with FastAPI and ExifTool integration.
- **Database:** MongoDB
- **ML:** Isolation Forest & Groq LLM

Timeline

- Week 1, Dec 2024 – Research & approval
- Week 2, Dec 2024 – Project planning & task allocation
- Week 3, Dec 2024 – Secondary approval
- Week 4, Dec 2024 – UX design & ExifTool documentation
- Weeks 1–3, Jan 2025 – Frontend & backend development
- Week 4, Jan 2025 – Metadata extraction; integration failed
- Week 2, Feb, 2025 - JWT Authentication & API Handling
- Week 3, Feb, 2025 - Integration
- Week 4, Feb, 2025 - ML Optimization
- Week 1, Mar, 2025 - ML Optimization & Integration

Demo!

`https://meta-trace.vercel.app/`

- **Challenges:**

- Metadata variations across file formats
- Ensuring proper API Handling
- Finding labelled dataset for anomaly detection

Thank you

Thank you for your time and consideration!

WE Group 12