Master thesis : UniForeXT: A Unified Data Model for Cross-Tool Integration in Digital Forensic Triage

Auteur : Robert, Louan
Promoteur(s) : Donnet, Benoît
Faculté : Faculté des Sciences appliquées
Diplôme : Master en sciences informatiques, à finalité spécialisée en "computer systems security"
Année académique : 2025-2026
URI/URL : https://github.com/louanrobert/UniForeXT.git; https://zenodo.org/doi/10.5281/zenodo.20396207; http://hdl.handle.net/2268.2/26068

UniForeXT: A Unified Data Model for Cross-Tool Integration in Digital Forensic Triage

Student	Robert Louan
Section	Computer Science
Academic year	2025–2026
Supervisor	Prof. Benoît Donnet

Abstract

The heterogeneous outputs of specialized forensic tools are a bottleneck in digital forensics and incident response (DFIR) triage. This thesis introduces UniForeXT, a tool-agnostic framework that consolidates diverse forensic artifacts into a unified, searchable knowledge graph, to accelerate triage workflows. UniForeXT comprises three integrated components: (1) a lightweight ontology that models forensic entities (e.g., detections, events, processes, users, hosts, etc.) using controlled vocabularies, (2) a YAML-to-RDF mapping tool driven by configuration and implemented in Java that generates an RDF graph and validates output using SHACL shapes, and (3) a UniForeXT tailored JavaFX/HTML visualization client that offers five complementary views to support analysis from an overview to a drill-down. Evaluation on realistic CERT Belgium datasets shows ingestion of $\sim$500 MB in ~ 31 seconds and full SHACL validation in ~ 2 minutes (both run offline, ahead of analysis). Representative triage queries return in under 300 ms. Future directions include web-based, multi-user visualization, entity reconciliation, RML/YARRLM interoperability, LLM-assisted workflows, and broader user studies. UniForeXT demonstrates that ontology-driven integration, pragmatic mapping, and purposeful visualization can reduce forensic fragmentation and are likely to improve triage efficiency.