Abbreviation and Term Reference

Abbreviation	Full Term	Description
AI	Artificial Intelligence	Computational techniques that enable machines to perform tasks typically requiring human intelligence, such as reasoning and decision support.
AI Agents	Artificial Intelligence Agents	Autonomous or semi-autonomous AI components designed to perform specific tasks within a larger analytical workflow.
CTI	Cyber Threat Intelligence	The collection, analysis, and dissemination of information about cyber threats to support informed security decisions.
CURES	Cranfield University Research Ethics System	The institutional ethics review body responsible for approving and overseeing research involving human participants.
DCS	Distributed Control System	A control system commonly used in industrial environments to manage complex processes through distributed controllers.
ICS	Industrial Control Systems	Integrated hardware and software systems used to monitor and control industrial processes.
IEC	International Electrotechnical Commission	An international standards organization responsible for developing standards for electrical and industrial systems.
Industry 4.0	Fourth Industrial Revolution	A paradigm describing the integration of digital technologies, automation, and data exchange in industrial environments.
LLM	Large Language Model	A machine learning model trained on large text corpora to generate, analyze, and reason over natural language.

NLP	Natural Language Processing	A field of AI focused on enabling machines to understand, process, and generate human language.
OT	Operational Technology	Hardware and software systems that directly monitor or control physical industrial processes.
PLC	Programmable Logic Controller	A specialized industrial computer used to control machinery and processes in real-time environments.
RAG	Retrieval-Augmented Generation	An AI technique that combines information retrieval with text generation to ground outputs in relevant data sources.
Reasoner Agents	Reasoning Agents	AI agents responsible for analysing, correlating, and drawing inferences from retrieved threat intelligence data.
Retriever Agents	Retrieval Agents	AI agents tasked with retrieving relevant contextual or historical information from a knowledge store.
SCADA	Supervisory Control and Data Acquisition	A system used to monitor and control industrial processes across distributed locations.
SOC	Security Operations Center	A centralized team or facility responsible for monitoring, detecting, and responding to cybersecurity incidents.
Triage Agents	Triage Agents	AI agents that filter, prioritize, and classify incoming data to reduce analyst workload.
Human-in-the-Loop		An approach that explicitly involves human experts in reviewing AI-generated outputs and making final decisions, particularly in safety-critical or high-risk environments.

Study Title: Evaluation of an AI-Enabled Cyber Threat Intelligence (CTI)Framework for Industrial and Critical Infrastructure Environments Research[1]

Purpose: This interview is part of a PhD research study conducted at Cranfield University. The aim of the study is to evaluate a proposed Cyber Threat Intelligence framework that employs Natural Language Processing(NLP) and Large Language Models (LLMs) to support threat intelligence and analysis in Industrial Control Systems (ICS) and Industry 4.0 environments.

Why You Have Been Invited: You are invited to participate because of your professional experience in **[ICS Security / Threat Intelligence / Industrial AI / Industrial Automation/ Control Systems/ Critical Infrastructure/ Digital Transformation]**. Your expertise is valuable in assessing the proposed framework.

Voluntary Participation: Participation is entirely voluntary. You may decline to answer any question or withdraw at any time before submitting the survey. There are no penalties or disadvantages for choosing not to participate. (Participation is voluntary and anonymous.)

Confidentiality & Data Handling:

- No personally identifying information is required.
- Responses will be stored securely and used only for academic research purposes.

Risks & Benefits: There are no known risks associated with participating in this interview. While there may be no direct personal benefit, your input will contribute to research that may support the development of more effective cyber defense methodologies for industrial environments.

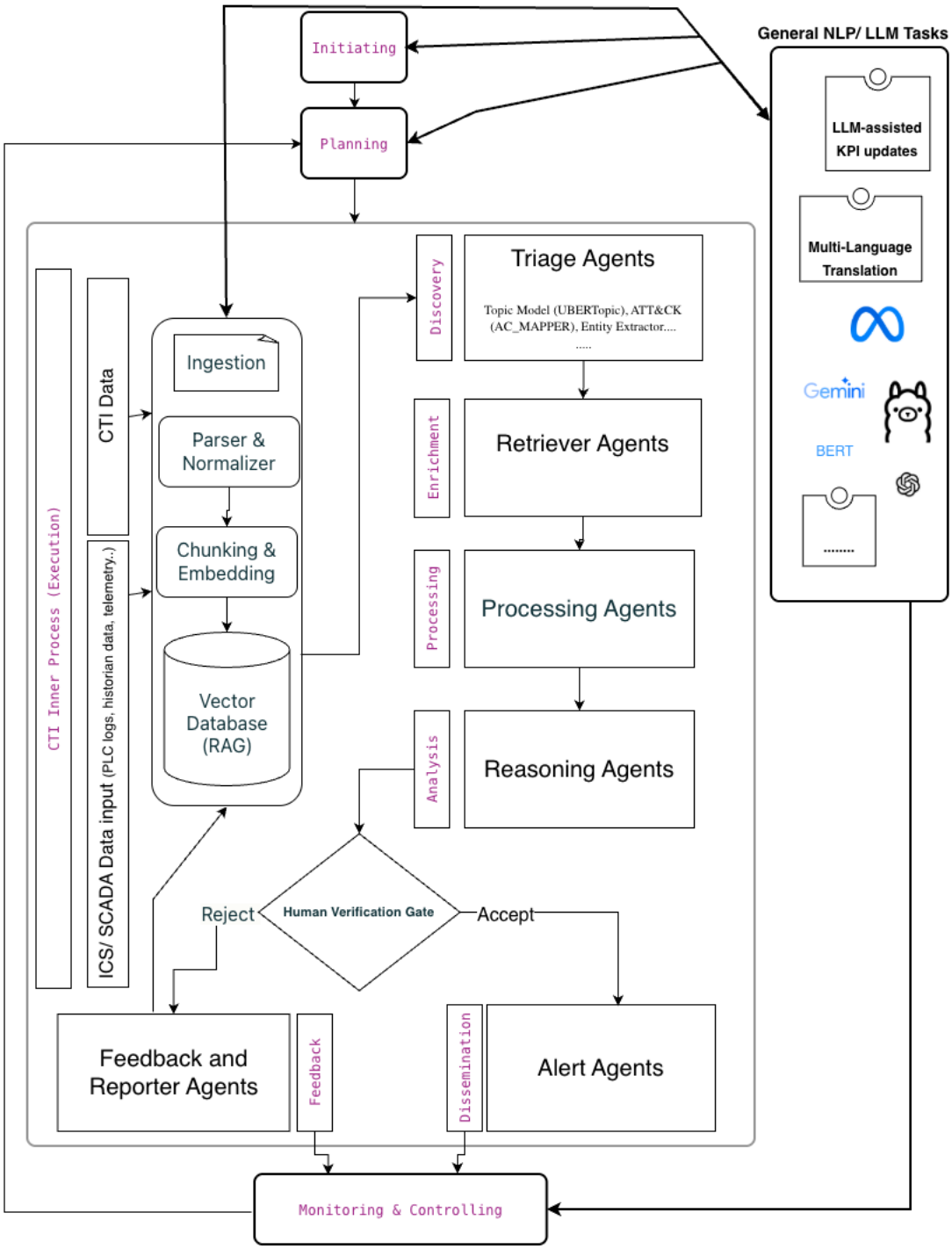
Ethics Approval: This research has been reviewed and approved by the Cranfield University Research Ethics System (CURES), Reference Number: CURES/26101/2025.

Contact for Questions: If you have questions about this study, please contact: Cranfield University Email: majed.albarrak.592@cranfield.ac.uk

[1] See the abbreviation table

Consent Statement: By proceeding with this interview, I confirm that:

I am 18 years of age or older,
I have read and understood the information provided above,
I understand that participation is voluntary and anonymous, and
I agree to participate in this research study.


Initiating
Planning
General NLP/ LLM Tasks
LLM-assisted KPI updates
Multi-Language Translation
Gemini
BERT
.....
CTI Inner Process (Execution)
CTI Data
ICS/ SCADA Data input (PLC logs, historian data, telemetry..)
Ingestion
Parser & Normalizer
Chunking & Embedding
Vector Database (RAG)
Discovery
Triage Agents
Topic Model (UBERTopic), ATT&CK (AC_MAPPER), Entity Extractor....
.....
Enrichment
Retriever Agents
Processing
Processing Agents
Analysis
Reasoning Agents
Reject
Human Verification Gate
Accept
Feedback and Reporter Agents
Feedback
Dissemination
Alert Agents
Monitoring & Controlling

Framework description:

The proposed framework is an **AI-enabled Cyber Threat Intelligence (CTI) framework** designed to support security operations in **Industrial Control Systems (ICS) and Industry 4.0 environments**. It addresses key limitations of current industrial CTI practices, including fragmented data sources, analyst overload, delayed intelligence generation, and insufficient contextualization of threats within operational technology (OT).

The framework adopts a **decision-support approach**, integrating Natural Language Processing (NLP), Large Language Models (LLMs), and Retrieval-Augmented Generation (RAG) into the CTI lifecycle while preserving human oversight. It is intended to augment analyst capabilities rather than automate security decisions or responses.

1. The architecture begins with a **multi-source ingestion layer** that processes both structured and unstructured data, including security logs, threat feeds, incident reports, and technical documentation such as PDFs and vendor advisories. NLP techniques are used to normalize and transform this data into machine-readable representations. Processed information is stored in a **centralized vector-based knowledge store**, which serves as a contextual memory. RAG mechanisms retrieve relevant historical and domain-specific evidence before generating analytical outputs, improving relevance and reducing unsupported inferences.
2. Analysis is performed using a **chain-of-agents architecture** that mirrors real-world CTI workflows. Specialized agents handle triage and prioritization, contextual retrieval, analytical reasoning, and summarization or translation of results for different stakeholders. This modular design supports scalability and transparency.

3. Given the safety-critical nature of industrial environments, the framework includes an explicit **Human-in-the-Loop verification gate**. Analysts review AI-generated outputs and may accept, modify, or reject them. Rejected outputs feed into a feedback loop to support iterative system refinement while maintaining analyst authority and trust.
4. A **governance layer** embeds the framework within the CTI lifecycle, supporting intelligence requirement definition, monitoring, auditability, and risk management. The framework is conceptual and intended for **expert evaluation**, with the following interview questions designed to assess its operational realism, technical feasibility, and strategic value in real-world industrial settings

The Questions (Structured for Design Science Validation)

Part A: Operational Validity (The "Process")

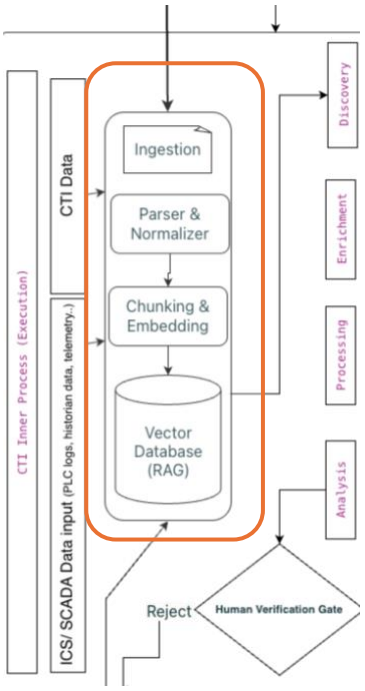
Focus: *Does the workflow match reality?*

Q1. The "Integration" Logic:

The framework introduces a Vector Database (RAG) layer (see diagram, middle-left) to store and retrieve contextual data before any analysis happens.

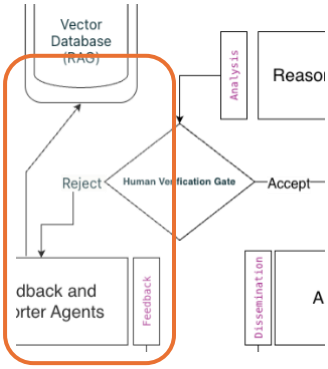
In real OT environments, do you think it is realistic to store and use very different data (such as PLC logs and security reports) in one central system to improve understanding, or would this introduce integration issues?

Answer :



Q2. The "Human-in-the-Loop" Gate:

We have placed a "Human Verification Gate" (diamond shape) that explicitly routes rejected AI findings back to the Feedback loop for retraining.



Is it realistic to expect SOC analysts to explain why they reject AI outputs, or would this feedback need to be automated?

Answer :

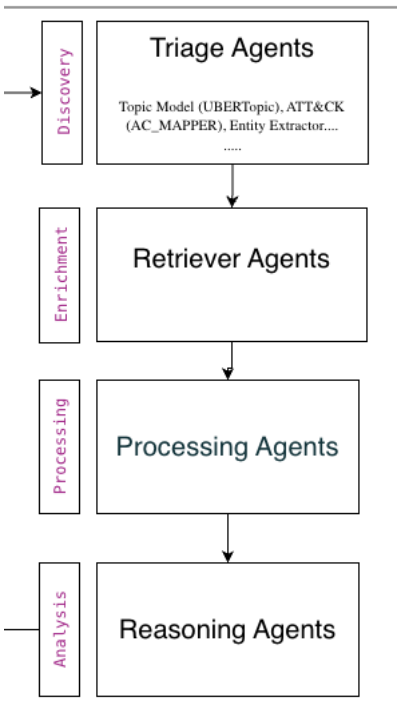
Part B: Technical Feasibility (The "Agents")

Focus: *Can the technology actually do this?*

Q3. The Agent Chain (Triage → Retriever → Reasoner):
Instead of a single "AI Model," we propose a Chain of Agents where a "Triage Agent" filters data before a "Reasoning Agent" analyzes it.

Do you think breaking analysis into multiple steps or agents would improve the quality of threat analysis, or do you prefer a single system that tries to do everything at once?

Answer :



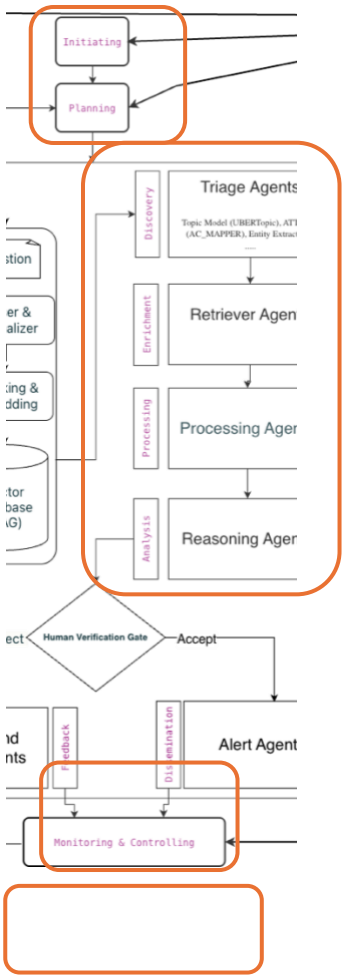
Q4. Latency vs. Depth:
The framework handles complex tasks such as reasoning and Translation (e.g., from Arabic to English) during the Analysis Phase.

In an Industry 4.0 environment, is it acceptable for threat intelligence to take longer if it provides better insight, or does it need to be near real-time to be useful?

Answer :

Part C: Strategic Impact (The "Value")

***Focus**: Is it worth building?*



Q5. The "Governance" Layer:

The framework wraps the technical agents in a Project Management loop (Initiating → Planning → Monitoring).

Do you see value in clearly defining the rules and scope for AI agents (for example, what they are allowed to analyse), or should the framework focus only on the technical side?

Answer :

Q6. The "Fatal Flaw" (The most important question):

From your experience, what would be the primary practical challenge to plan for when introducing this framework in a real manufacturing plant? (e.g., Data Privacy, Trust in AI, Integration complexity, Regulatory compliance?)

Answer :

Thank you for your time and expertise. If you are interested, I would be happy to share the final "refined" framework and the summary of expert findings once the study is complete.

Best regards,

Majed Albarrak
Cranfield University