An AI-based Cognitive Architecture for Augmenting Cybersecurity Analysts

Salvatore Vella, Moussa Noun
Salah Sharieh, Alex Ferworn

Toronto Metropolitan University

Who we are

- Salvatore Vella (Toronto Metropolitan University)
- Moussa Noun (Royal Bank of Canada)
- Salah Sharieh (Toronto Metropolitan University)
- Alex Ferworn (Toronto Metropolitan University)

- Salvatore Vella - Re-retired technology executive - ex-IBM, ex-Royal Bank of Canada; currently a doctoral student at the Toronto Metropolitan University

Agenda

- Goal
- Abstract
- Cybersecurity Analyst Role
- CorpIA (Corporate Intelligence Augmentation) Framework
- Bloom’s Taxonomy
- Results
- Discussion
- Future Work

Goal

How might we use Generative AI as augmentation for Knowledge Workers using a Cybersecurity Analyst as an example?

Toronto Metropolitan University

Abstract

We present a **Generative AI-based cognitive architecture** and an agent specifically developed for the complexities of **Cybersecurity analysis**. White-collar roles, exemplified by Cybersecurity analysts, are multifaceted and rely on **declarative knowledge, procedural understanding, and diverse tools**. The ability to learn and adapt to the nuances of the job is crucial. This paper **introduces CorpIA, a cognitive architecture that provides an agent with knowledge, tools, and the capacity to acquire on-the-job experience**. This system enhances human performance by providing suggested solutions and continuous **mentoring.** Our research demonstrates that the CorpIA agent **can learn from interactions using Bloom's Taxonomy**. We provide the source code for these experiments. **https://github.com/salvella/CorpIA**

Methodology

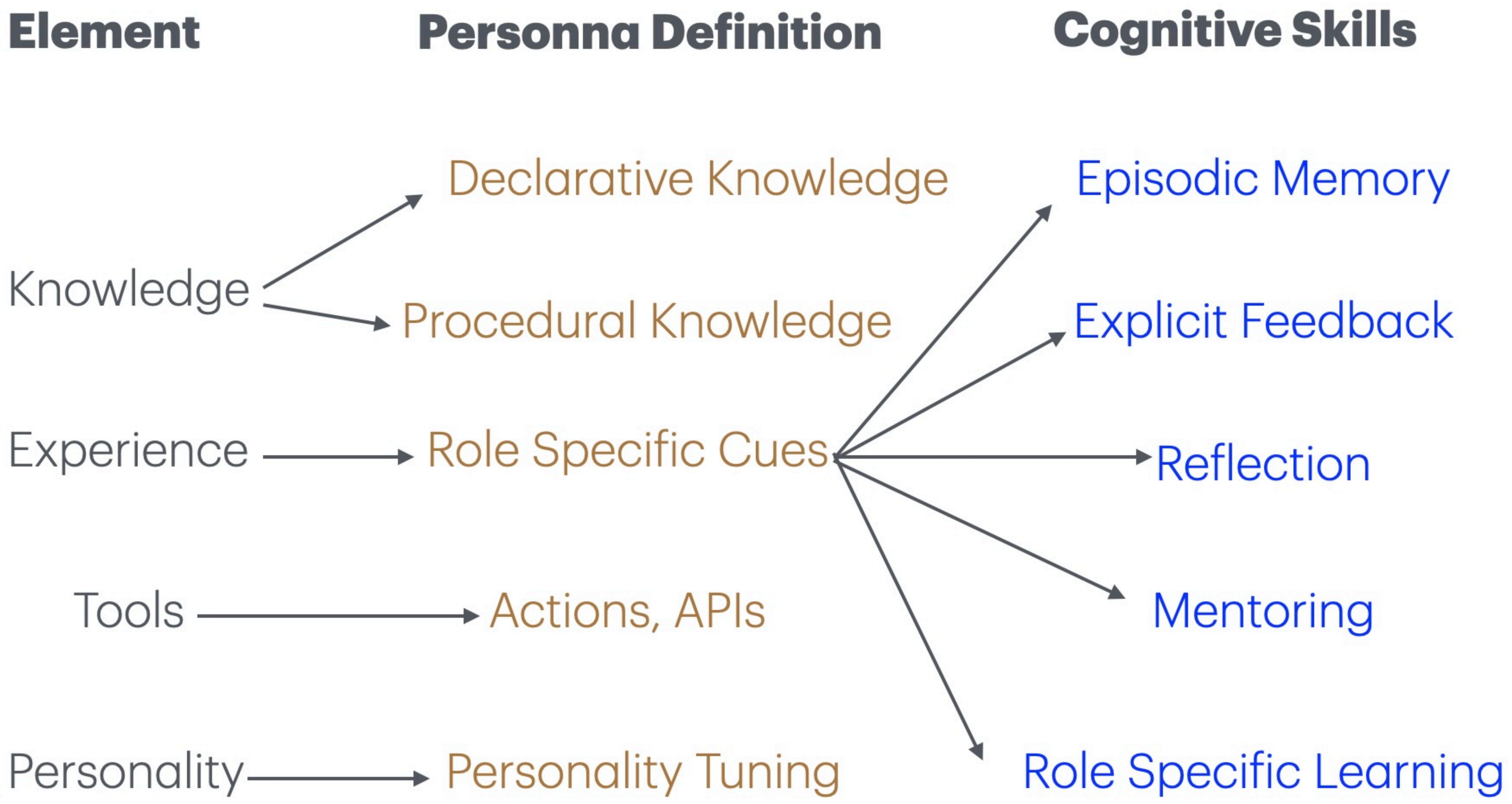
- Define a Cybersecurity Gen AI agent with a set of Knowledge, Experience and Skills
- Create a series of discussions with a fictional client that provides the Gen AI agent with additional information about the client
- Test the learning of the Gen AI agent with a set of questions derived from Bloom’s Taxonomy

- Use CorpIA, our custom-built agent framework for knowledge workers, as the Gen AI agent framework to create the agent, GPT-4o as the underlying large language model

CorpIA (Corporate Intelligence Augmentation) - A Flexible AI-Agent System for Knowledge Workers

- Memory
- Experience
- Skills
- Policies and Procedures

Framework Elements



Element

Personna Definition

Cognitive Skills

Knowledge → Declarative Knowledge

Knowledge → Procedural Knowledge

Experience → Role Specific Cues

Tools → Actions, APIs

Personality → Personality Tuning

Role Specific Cues → Episodic Memory

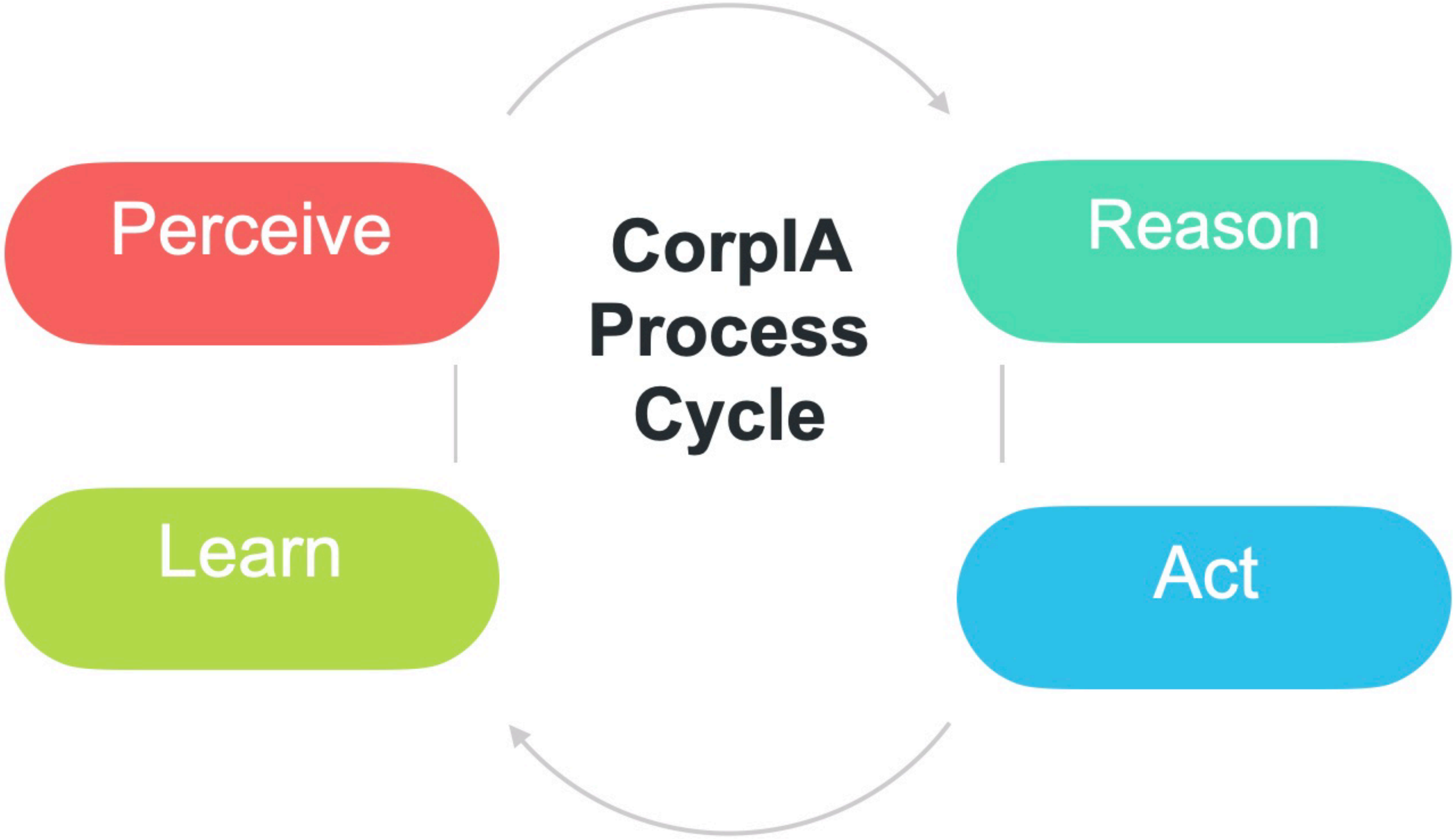
Role Specific Cues → Explicit Feedback

Role Specific Cues → Reflection

Role Specific Cues → Mentoring

Role Specific Cues → Role Specific Learning

Framework Flow



Perceive

CorpIA Process Cycle

Reason

Learn

Act

Cybersecurity Analyst Definition

Role: Cybersecurity Analyst. As a cybersecurity expert with CISSP, OSCP, and CASP+ certifications, your role is to provide clear, concise answers to cybersecurity questions from business users. You will assess risks, identify potential threats to the organization, and recommend appropriate mitigations or remediations. You work in a global enterprise environment that includes Windows, Linux, and Mac systems and is subject to various regulatory and legal requirements. Maintain a professional tone and ensure your explanations are easy to understand.

Knowledge: Possesses deep knowledge of computer security, network protocols, and systems administration. Well-versed in cybersecurity threats, risk analysis techniques, and security standards such as ISO/IEC 27001, GDPR, and NIST frameworks.

Experience: Typically has several years of experience in IT or cybersecurity roles, focusing on threat detection, security assessments, and incident response. Experience often includes conducting vulnerability scans and managing security solutions to protect against threats.

Skills: Proficient in technical skills such as intrusion detection, malware analysis, and the use of SIEM (Security Information and Event Management) tools. Strong analytical skills are crucial, as well as the ability to quickly adapt to new threats. Effective communication skills are also important for explaining technical details to non-technical stakeholders.

Tools: Experienced with tools such as firewalls, antivirus software, intrusion detection systems (IDS), and encryption technologies. Familiar with cybersecurity platforms like Splunk, IBM QRadar, or Palo Alto Networks products for monitoring and responding to security incidents.

Personality: Exhibits a detail-oriented, vigilant, and analytical personality. Must be proactive in staying updated on the latest security trends and threats. Strong problem-solving skills are essential, as is the ability to remain calm and focused under pressure during security breaches or attacks.

CorpIA - Some Key Parameters

Declarative Memory - Facts

```
"has_declarative_memory": true,
"declarative_memory_file": “knowledge/CSRB_Log4j.pdf”,
```

Procedural Memory - How to do things

```
"has_procedural_memory": true,
"procedural_memory_file": “knowledge/NIST.pdf”,
```

Active Listening Based on Cues

```
"has_role_memostore": true,
"cues": ["security threats”, "ABC Bank”],
```

List of Possible Team Mate Agents

```
"helpful_agents": [“Lawyer", "IT Specialist"],
```

Cybersecurity Analyst Mentoring
Mentoring Advice to help on the job learning.

As a mentor in cybersecurity, it's essential to reinforce the importance of structured and thorough assessment processes when dealing with vulnerabilities like Log4j (CVE-2021-44228). Here are some key points and advice that can guide you in executing the role of a Cybersecurity Analyst effectively:

Emphasize a Collaborative Mindset
- Team Involvement: Engaging with application development teams during your inventory process is crucial. Ensure that there is clear communication to understand the context of Log4j usage within specific applications. Encourage them to be proactive in security discussions.
- Cross-Functional Collaboration: Foster relationships with other departments, such as compliance and operations, who can provide insights on regulatory requirements or operational impacts related to vulnerabilities.

Continuous Learning and Adaptability
- Stay Updated: Cybersecurity is a fast-evolving field. Regularly allocate time for learning about new threats and vulnerabilities. Subscribe to reputable cybersecurity blogs, attend webinars, and participate in industry conferences to stay informed.
- Hands-On Practice: Engage in lab environments where you can test and understand different vulnerabilities and exploits, enhancing your practical knowledge.

Bloom’s Taxonomy

Bloom's Taxonomy is a hierarchical model used for classification of educational learning objectives into levels of complexity and specificity.

Remember - Recognizing, Recalling

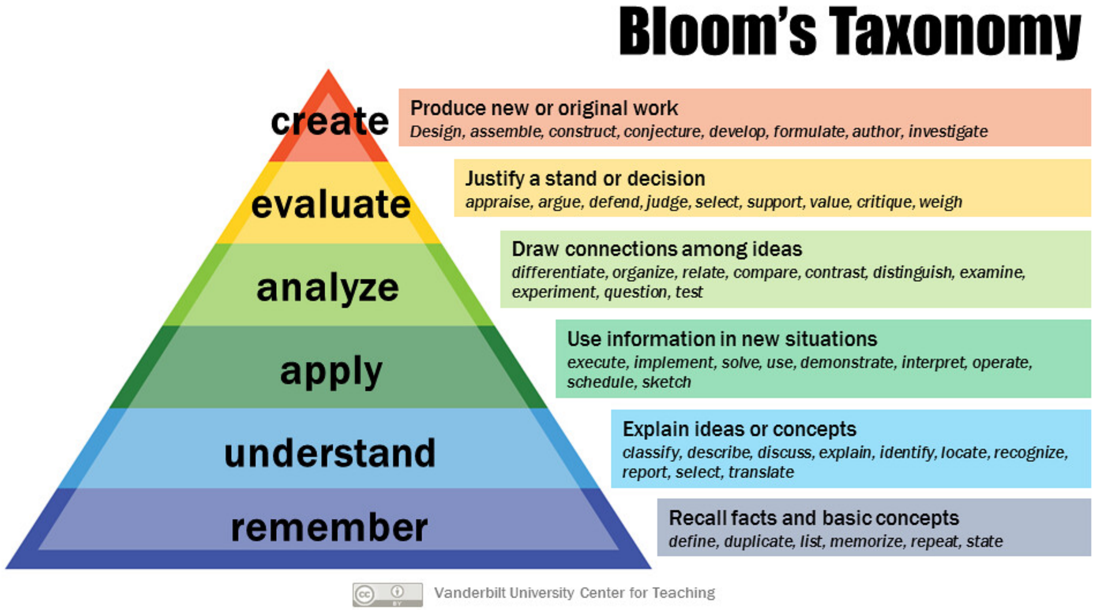
Understand - Interpreting, Exemplifying, Classifying, Summarizing, Inferring, Comparing, Explaining

Apply - Executing, Implementing

Analyze - Differentiating, Organizing, Attributing

Evaluate - Checking, Critiquing

Create - Generating, Planning, Producing



Produce new or original work
Design, assemble, construct, conjecture, develop, formulate, author, investigate

Justify a stand or decision
appraise, argue, defend, judge, select, support, value, critique, weigh

Draw connections among ideas
differentiate, organize, relate, compare, contrast, distinguish, examine, experiment, question, test

Use information in new situations
execute, implement, solve, use, demonstrate, interpret, operate, schedule, sketch

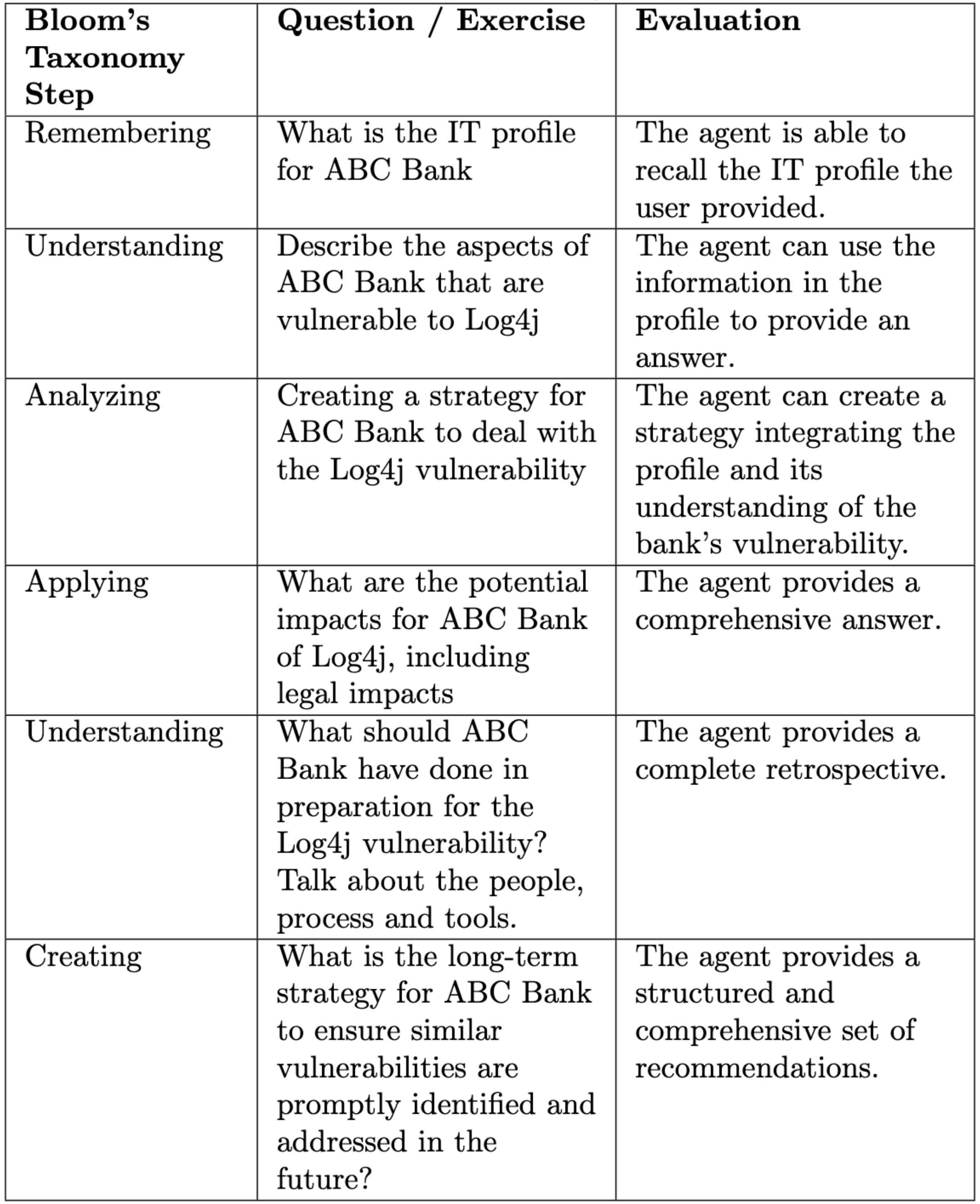
Explain ideas or concepts
classify, describe, discuss, explain, identify, locate, recognize, report, select, translate

Recall facts and basic concepts
define, duplicate, list, memorize, repeat, state

Vanderbilt University Center for Teaching

Results

- Bloom’s Taxonomy questions used to evaluate the learning abilities of the AI Agent
- AI Agent is able to integrate new knowledge, analyze and classify the knowledge, apply it to a customer and situation



Bloom’s Taxonomy Step	**Question / Exercise**	**Evaluation**
Remembering	What is the IT profile for ABC Bank	The agent is able to recall the IT profile the user provided.
Understanding	Describe the aspects of ABC Bank that are vulnerable to Log4j	The agent can use the information in the profile to provide an answer.
Analyzing	Creating a strategy for ABC Bank to deal with the Log4j vulnerability	The agent can create a strategy integrating the profile and its understanding of the bank’s vulnerability.
Applying	What are the potential impacts for ABC Bank of Log4j, including legal impacts	The agent provides a comprehensive answer.
Understanding	What should ABC Bank have done in preparation for the Log4j vulnerability? Talk about the people, process and tools.	The agent provides a complete retrospective.
Creating	What is the long-term strategy for ABC Bank to ensure similar vulnerabilities are promptly identified and addressed in the future?	The agent provides a structured and comprehensive set of recommendations.

Toronto Metropolitan University

Discussion

- AI Agents are good listeners based on pre-defined cues
- This listening capability is useful to integrate On The Job Knowledge which is typical of white collar work
- Using Bloom's Taxonomy as a way to measure, AI Agents can remember, understand, apply, analyze, evaluate and create novel content
- Integration of AI Agents and knowledge professionals is a critical topic - processes, skills. Changes needed beyond introduction of a new tool

Follow On and Ongoing Work

- Parameter based role definition to be able to easily define other white collar roles - Github repo shows several
- Feedback from practicing Cybersecurity Analysts - pending ERB (Ethics Review Board) approvals
- Evaluation and tuning of the Mentoring Feedback - pending ERB approvals

Follow On and Ongoing Work

- Workflows
- Policy and Procedures
  - Paper out for review on this topic - called Policy and Procedure as Code
  - Use LLM to take Procedure, clean it up for inconsistencies, and rewrite in a simplified version of BPEL (Business Process Execution Language) for execution
- Enable Human and AI Interaction

More Future Work

- Benchmarks
  - Compare to humans
  - Define a knowledge worker and be able to do the work of the knowledge worker
- AI Agent Store - Can imagine these don’t have to be all from one place or internal

Thank You!

Toronto Metropolitan University