black hat®
USA 2023

AUGUST 9-10, 2023
BRIEFINGS

What Does an LLM-Powered Threat Intelligence Program Look Like?

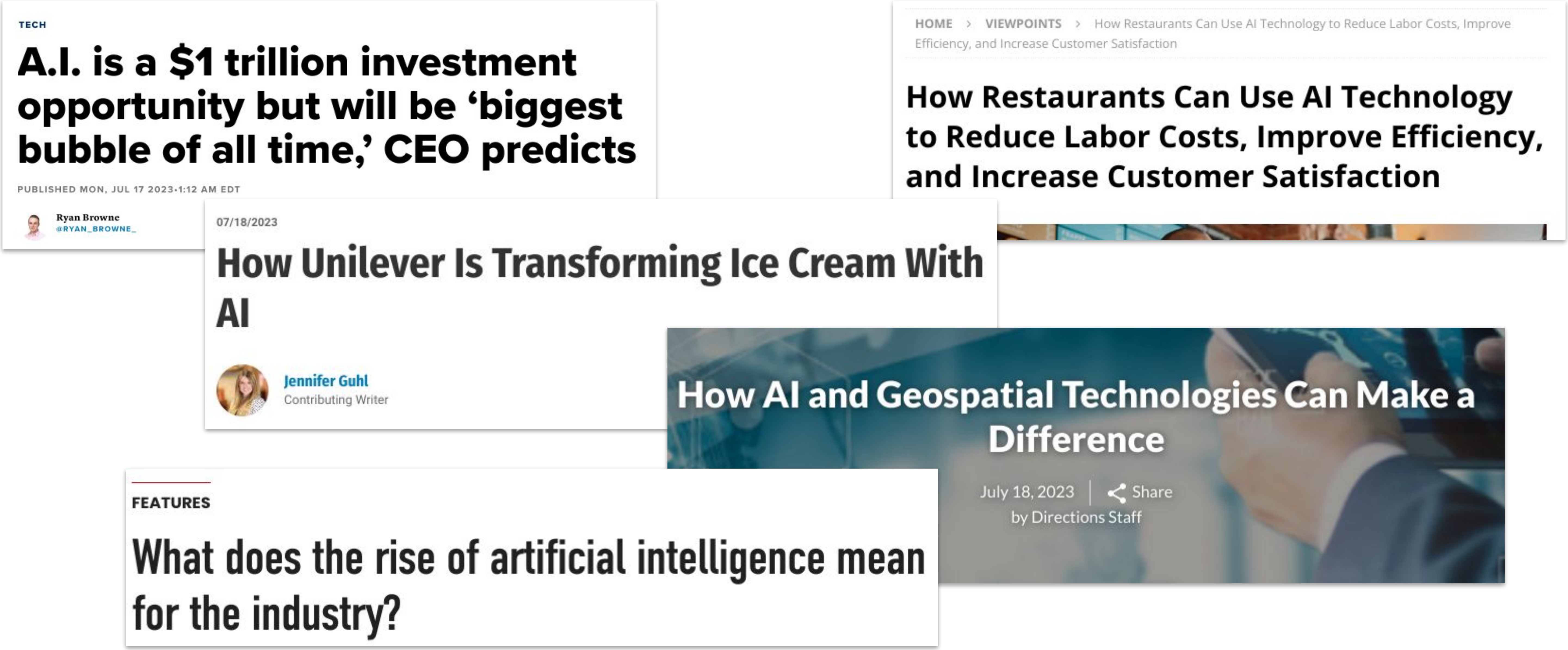
Speakers: Ron Graf & John Miller

black hat USA 2023

Speakers

Ron Graf
Data Scientist

John Miller
Head of Mandiant Intelligence Analysis



TECH

A.I. is a $1 trillion investment opportunity but will be 'biggest bubble of all time,' CEO predicts

PUBLISHED MON, JUL 17 2023·1:12 AM EDT

Ryan Browne
@RYAN_BROWNE_

HOME › VIEWPOINTS › How Restaurants Can Use AI Technology to Reduce Labor Costs, Improve Efficiency, and Increase Customer Satisfaction

How Restaurants Can Use AI Technology to Reduce Labor Costs, Improve Efficiency, and Increase Customer Satisfaction

---

07/18/2023

How Unilever Is Transforming Ice Cream With AI

Jennifer Guhl
Contributing Writer

How AI and Geospatial Technologies Can Make a Difference

July 18, 2023 | Share
by Directions Staff

FEATURES

What does the rise of artificial intelligence mean for the industry?

How do I navigate this environment when planning for a threat intelligence function?

"How will AI affect next year's human resources needs?"

"What will we be able to deliver that we couldn't before?"

"How much faster will we respond to incidents?"
…and more

Session Roadmap

- Background on AI and CTI
- Framework for components of a CTI program
- Historical CTI reporting through lens of framework
- LLM impacts to components of a CTI program
- LLM implementation considerations
- Takeaways for CTI program planning

What does a threat intelligence program deliver?

Improve security decisions by providing answers to difficult & uncertain questions

- *What are the top threats facing our business?*
- *Which security events are the most malicious?*
- *How can we hunt for undiscovered threats?*
- *Which vulnerabilities should we prioritize?*
- *How can we test our security controls?*

What does a threat intelligence program need to succeed?

Threat Visibility	**Processing Capability**	**Interpretation Capability**
Direct visibility into relevant security data	Convert data to standardized, useful observations	Apply processed data to address key questions for stakeholders
Example: emails detected as malicious	*Example: Identify email attachment is Malware X*	*Example: Answer “Is Malware X the top threat to us?”*

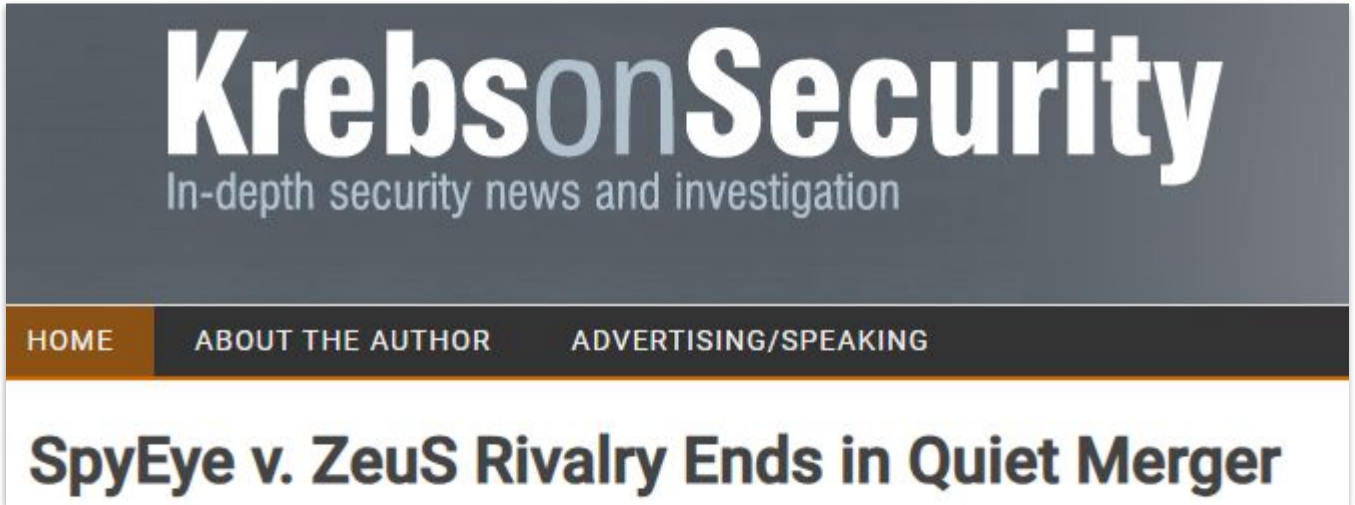
How could a TI function benefit from LLM capabilities?

Threat Visibility	Processing Capability	Interpretation Capability
Direct visibility into relevant security data	Convert data to standardized, useful observations	Apply processed data to address key questions for stakeholders

What are possible capacity challenges, in these terms?

How can LLMs be applied to help?

Intelligence Example #1: SpyZeus



KrebsonSecurity

In-depth security news and investigation

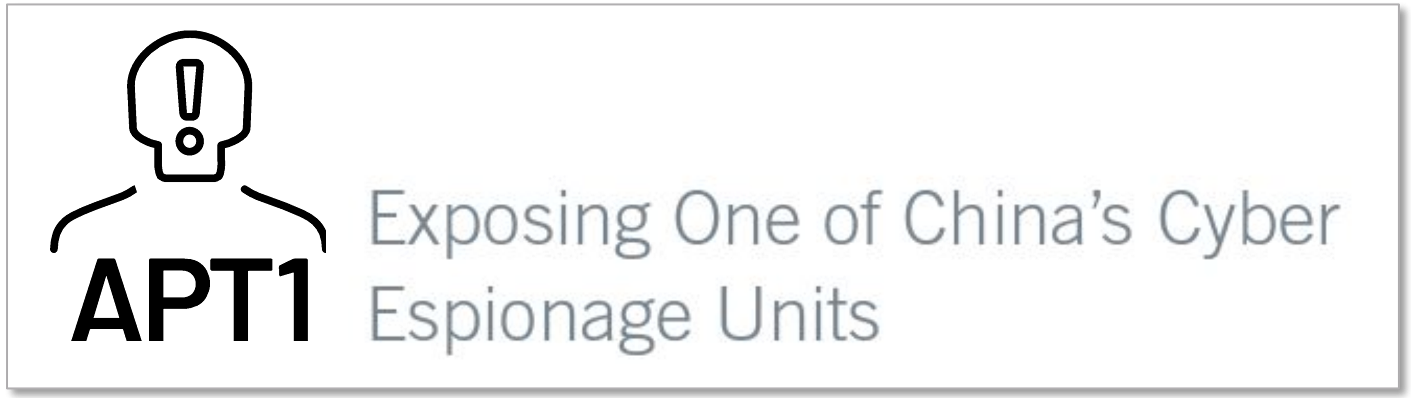
SpyEye v. ZeuS Rivalry Ends in Quiet Merger

Summary	Two prominent malware types reportedly merged into single “product” by underground vendors
Intelligence Type	Qualified assessments on prominent security issue
Sources	Primarily direct/manual “dark web” research
Data Scope	Weeks of activity in specific forums
Length	~3 pages
IOCs	Not in focus

What are my barriers to scaling this deliverable?	
Threat Visibility	- Access to dark web
Processing Capability	- Not a primary requirement
Interpretation Capability	- Capacity to summarize, explain & assess dark web activity

Intelligence Example #2: APT1

Summary	Sophisticated nation-state intrusion operation exposed
Intelligence Type	Detailing of intrusion operations & assessed sponsor
Sources	DFIR data & open-source information on assessed sponsor
Data Scope	"Nearly 150 victims over 7 years" 30+ cited open sources on sponsor
Length	74 pages
IOCs	3,000+


APT1 Exposing One of China's Cyber Espionage Units

What are my barriers to scaling this deliverable?	
Threat Visibility	• Regular DFIR data over extended period • Access to open sources
Processing Capability	• Capacity to develop data points from threat artifacts (e.g. malware) • Chinese-language translation
Interpretation Capability	• Capacity to plan & direct research on sponsorship questions • Capacity to synthesize gathered data on sponsorship questions

Intelligence Example #3: Carbanak

CARBANAK APT

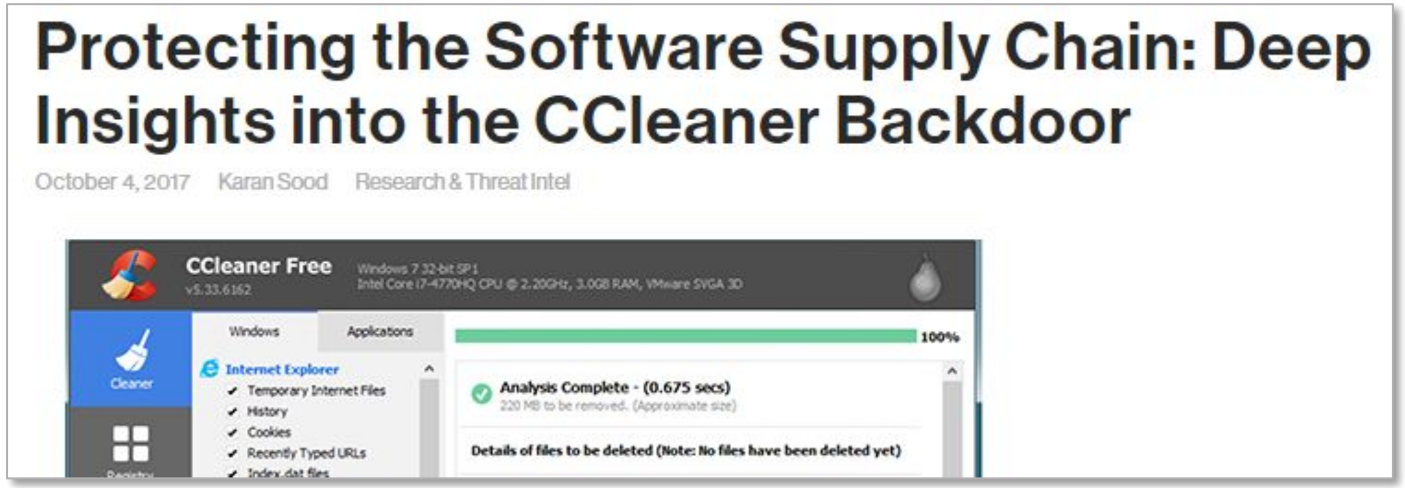
THE GREAT BANK ROBBERY

Version 2.1
February, 2015

Summary	Exposition of criminal intrusion operation into financial institutions yielding $1B in losses
Intelligence Type	Technical detailing of intrusions & connected malware operations
Sources	DFIR data, malware repository information & malicious infrastructure
Data Scope	“100 banking entities impacted,” 2-year scope, multiple control servers
Length	39 pages
IOCs	300+

What are my barriers to scaling this deliverable?	
Threat Visibility	• Extended-period DFIR data • Commercial technical sources
Processing Capability	• Capacity to develop data points from threat artifacts (e.g. malware)
Interpretation Capability	• Capacity to describe intrusion observations

Intelligence Example #4: CCleaner Backdoor



Protecting the Software Supply Chain: Deep Insights into the CCleaner Backdoor

October 4, 2017 Karan Sood Research & Threat Intel

Analysis Complete - (0.675 secs)
220 MB to be removed. (Approximate size)

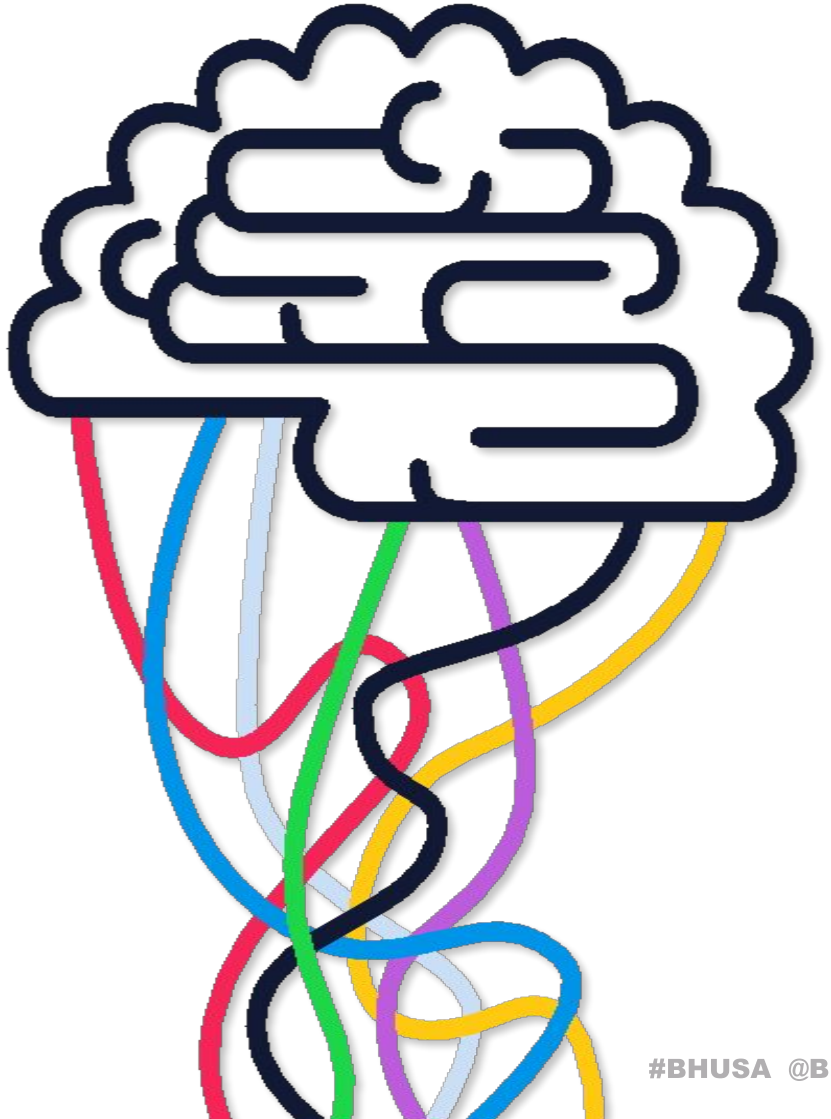
Details of files to be deleted (Note: No files have been deleted yet)

Summary	Analysis of supply-chain compromise of popular software
Intelligence Type	Walkthrough of malicious alterations to software
Sources	Primarily malware samples & sample reverse engineering
Data Scope	3 identified code samples from specific incident
Length	~10 pages
IOCs	20+

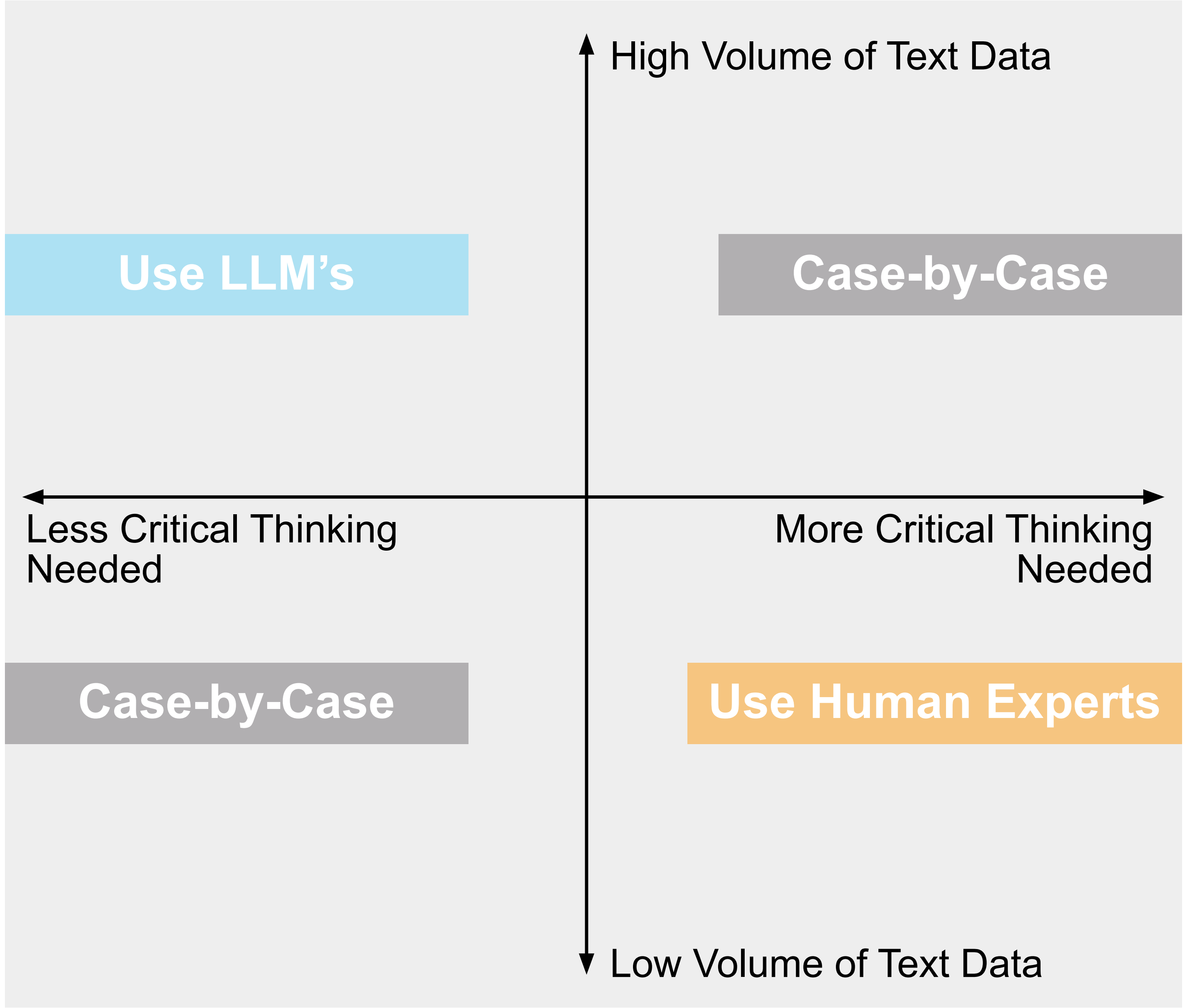
What are my barriers to scaling this deliverable?	
Threat Visibility	Access to malicious code telemetry
Processing Capability	Rapid-turn malware reverse engineering capability
Interpretation Capability	Capacity to rapidly characterize risks from emerging events

How LLM's Impact Processing and Interpretation

- Exploit data which is often overlooked due to volume
- Toil reduction for analysts
- Better, faster responses to RFI's



Processing Tasks



High Volume of Text Data

Use LLM's

Case-by-Case

Less Critical Thinking Needed

More Critical Thinking Needed

Case-by-Case

Use Human Experts

Low Volume of Text Data

- Exploit data overlooked due to volume
- Prevent waste of scarce experts on trivial tasks
- Many high-impact, low risk LLM-applications here

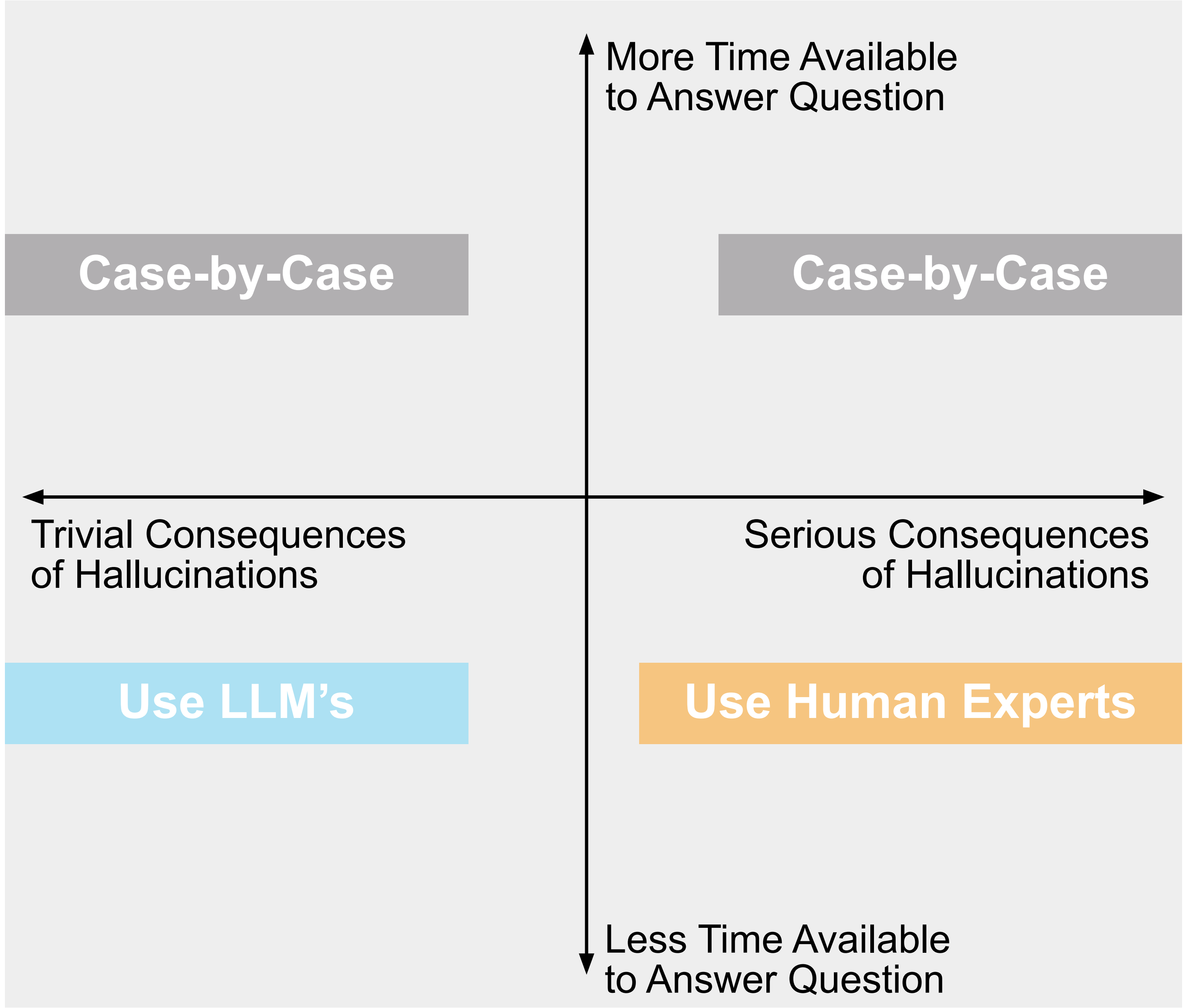
LLM Examples

- Malware RE reports for less complex samples
- Translate log into standardized, meaningful format
- Weekly summary of open source/third-party intel

Human Expert Examples

- Malware RE report for novel, complex samples
- Translate foreign language docs critical for attribution

Interpretation Tasks


More Time Available to Answer Question
Case-by-Case
Case-by-Case
Trivial Consequences of Hallucinations
Serious Consequences of Hallucinations
Use LLM's
Use Human Experts
Less Time Available to Answer Question

- Most workflows should remain human-in-the-loop
- Fewer low risk applications vs. processing

LLM Examples

- SOC triage of a high-priority, time-sensitive alert
- Patch prioritization for low CVSS score vulns
- Prioritizing dark web forum monitoring alerts

Human Expert Examples

- Incident response report writing
- Patch prioritization for high CVSS score vulns
- Assess likelihood of intruders lateral movements

Impact of Hallucinations - Consequential Example

Actor sends invoice lure with malicious PDF to Accounting

SOC analyst alerted, conducts standard response

Automated triage on email & attachment

Results into LLM-powered tool to digestible format for SOC analyst

LLM hallucinates about risks, determines email benign

Impact of Hallucinations - Nuisance Example

Analyst Prompt to LLM:

What malware families are associated with APT32?

LLM Response:

APT32, also known as OceanLotus or APT-C-00, has been associated with several malware families throughout its cyber espionage campaigns. Some of the notable malware families linked to APT32 include:

❌ 1. Cobalt Kitty: A custom remote access trojan (RAT) used by APT32 for initial compromise and establishing a foothold in the target network.

2. KOMPROGO: A backdoor RAT employed by APT32 to maintain persistence and execute commands on compromised systems.

❌ 3. Kasperagent: A modular malware framework utilized by APT32 for reconnaissance, data exfiltration, and lateral movement within the victim network.

4. BEACON: A backdoor malware used by APT32 to gain unauthorized access to systems, collect information, and execute commands remotely.

5. RemyRat: A remote access trojan used by APT32 to maintain persistence and conduct surveillance on compromised systems.

Beyond Interpretation

- Certain automated actions could be enabled by LLM interpretation capabilities
- Will require thorough vetting to mitigate the impacts of hallucinations
- Examples:
    - Change firewall rules/network configurations
    - Patch vulnerabilities
    - Take system offline until a human clears it to bring it back online
    - Force user password change

Impacts of Hallucinations

- Contemplate: which functions can & can’t tolerate fabricated information?

Unacceptable - Hallucinate vulnerability details that change patch prioritization

Acceptable - Misinterpret benign log entries not being reviewed by human anyways

- Prioritize grounding model outputs in factuality should be prioritized
  - Knowledge graphs & other sources of truth provide options for grounding

“Black Hat Sound Bytes”

What are takeaways for threat intelligence functions: AI integration & capability planning?

Takeaway: Intentionally codify human expertise

Engineer systems so experts provide feedback to models with no added burden

- Meet the experts where they are
- Reinforcement learning with human feedback (RLHF) should be prioritized as a component of any attempt to power a CTI workflow with an LLM

Invest in generalist, well-rounded experts

- Skill sets tied to specifically to tooling will become less valuable over time
- Breadth of expertise and strong critical thinking skills will become more valuable over time

Takeaways:

Use private models for sensitive use-cases

LLM integration with 3rd-party services (plug-ins/tools) critical

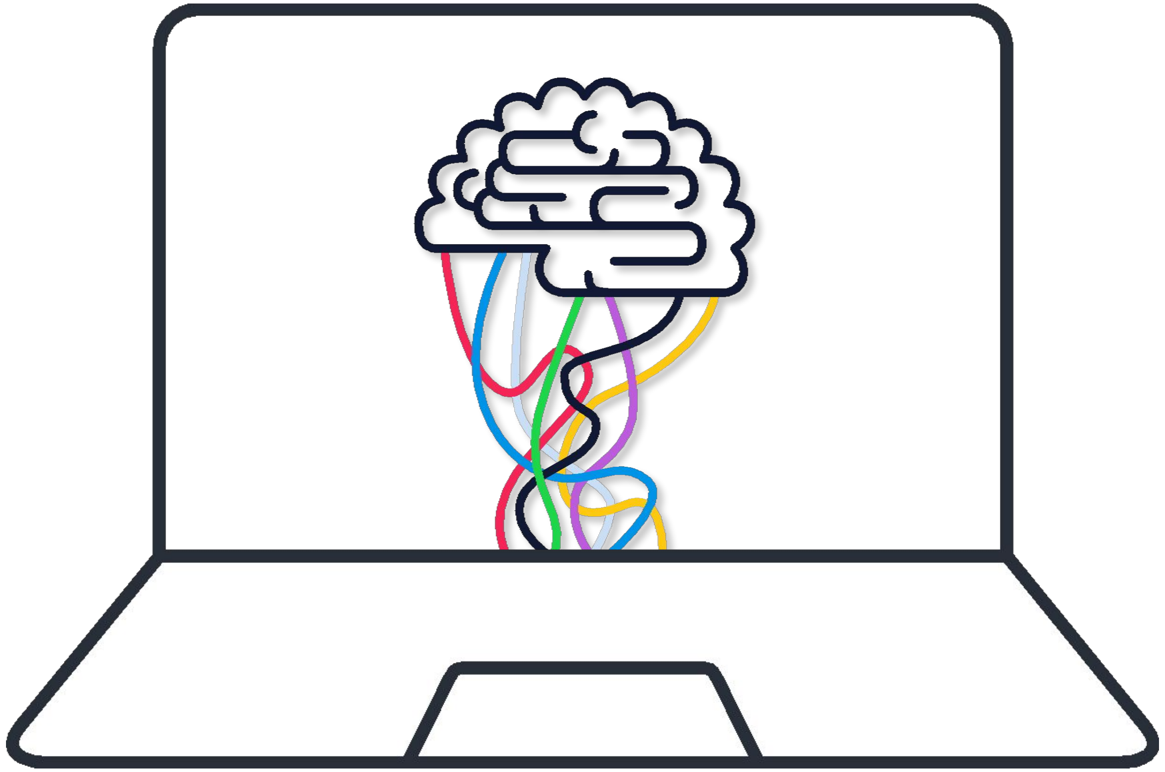
Differentiator: Pre-train / fine-tune on high-quality domain-specific datasets

- *Internally harvested data*
- *Third-party data sources*
- *Acquisition considerations*

Think beyond text - lots of expertise encoded in slide decks as images/diagrams

Takeaway:

Plan for:
continuing baseline of CTI
expertise + expertise ***in LLM behavior***



Takeaway:

Plan for higher ROI from *processing & interpretation* investments

More ***intelligence outcomes delivered*** per ***resource invested***

- Threats assessed
- Vulnerabilities prioritized
- Events actioned
- …

- Incremental hires
- Tools
- Data sources
- …

Takeaway:

Plan for higher ROI from ***processing & interpretation*** investments

More intelligence outcomes delivered per resource invested

Resource reduction vs. **greater ROI**

> *"The workforce gap is not going unnoticed by cybersecurity workers – nearly 70% feel their organization does not have enough cybersecurity staff to be effective."*
>
> *ISC^2, 2022*

Thank you!

Resources

Intelligence content examples

SpyZeus: https://krebsonsecurity.com/2010/10/spyeye-v-zeus-rivalry-ends-in-quiet-merger/

APT1: https://www.mandiant.com/sites/default/files/2021-09/mandiant-apt1-report.pdf

Carbanak:
https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08064518/Carbanak_APT_eng.pdf

CCleaner: https://www.crowdstrike.com/blog/protecting-software-supply-chain-deep-insights-ccleaner-backdoor/