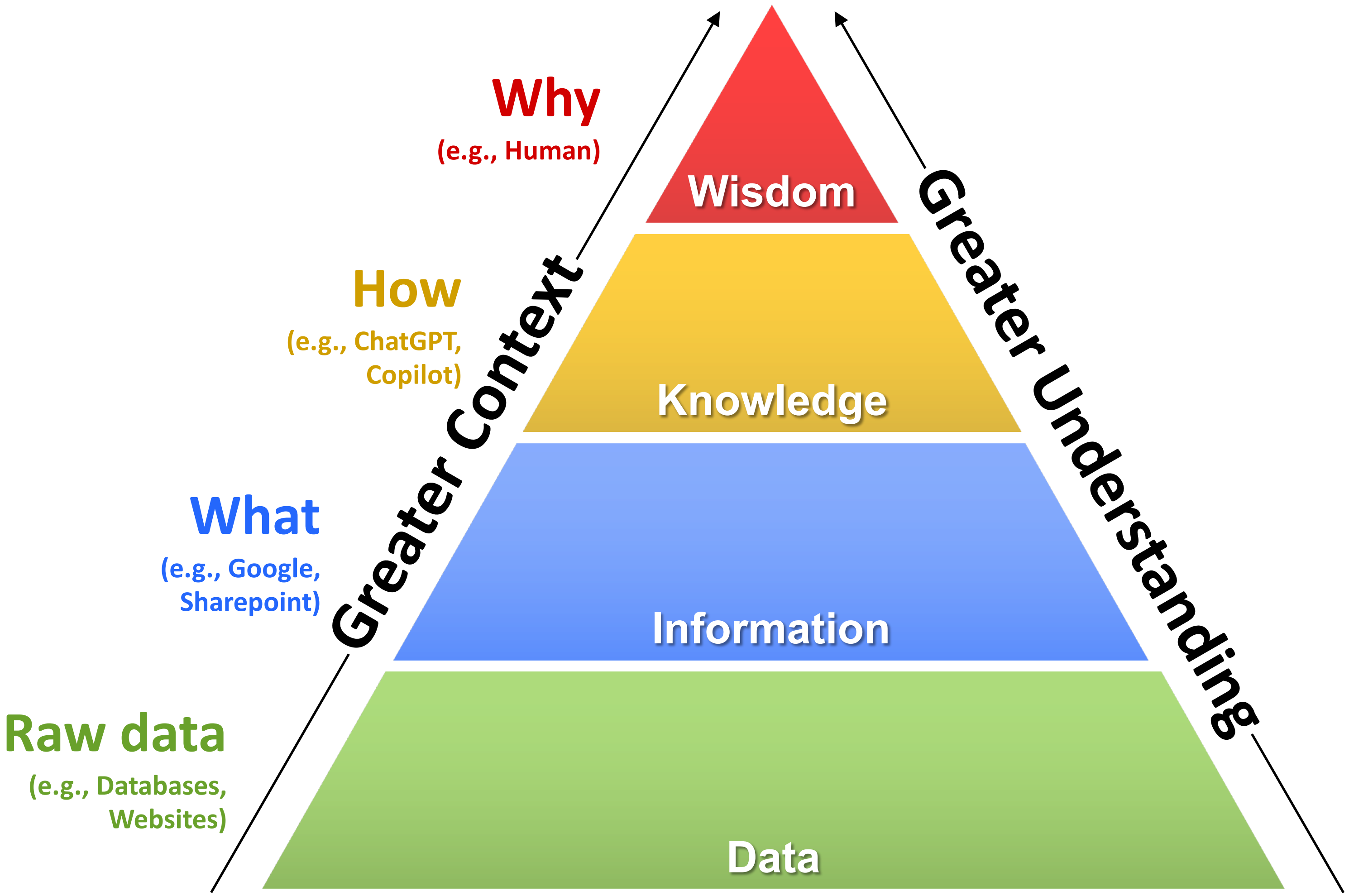
The Limits of Data Security in the Age of LLMs

Sounil Yu

Founder/Chief AI Safety Officer, Knostic

🐦 @sounilyu

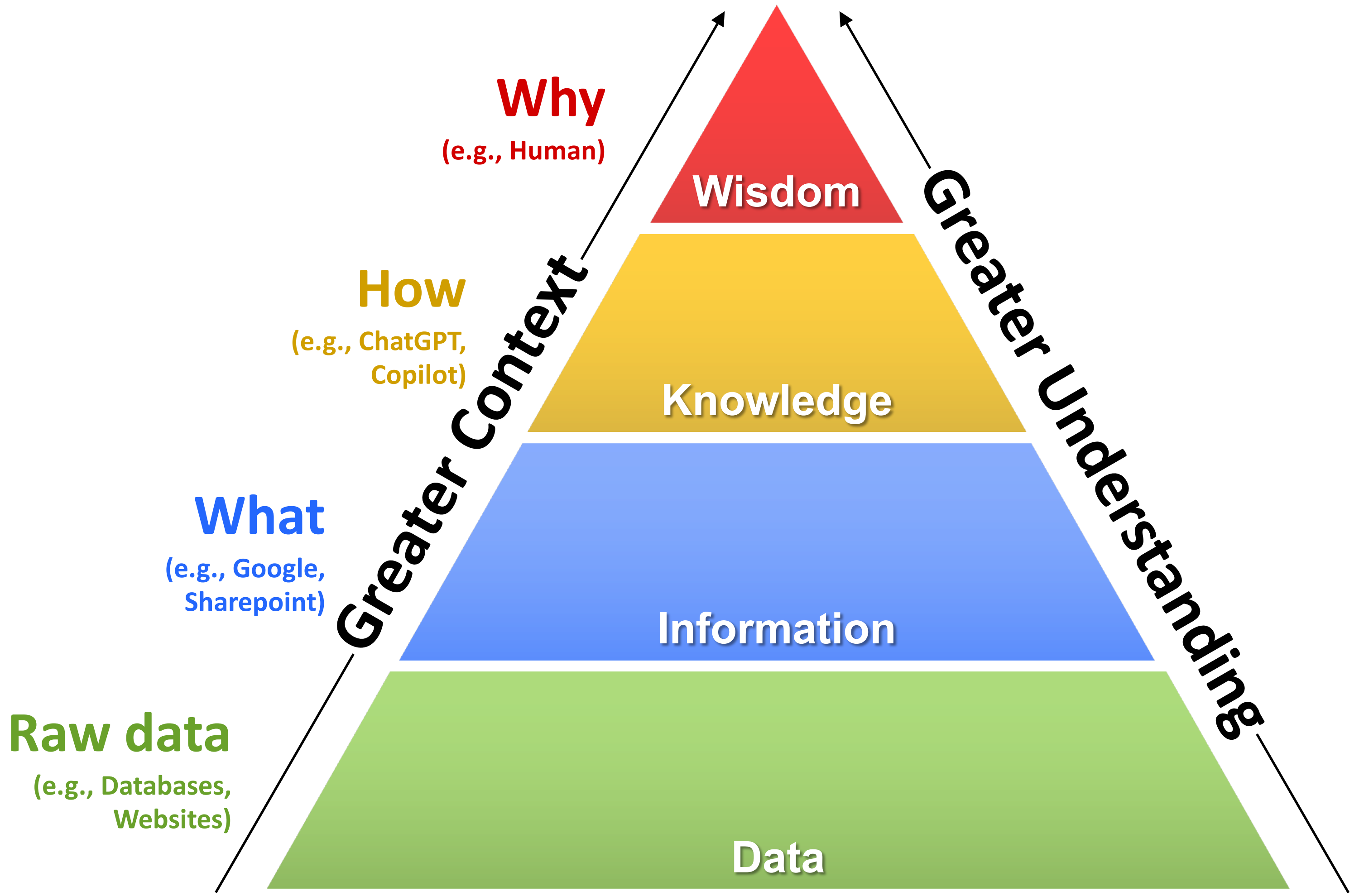
The DIKW Pyramid



KNOSTIC

@sounilyu

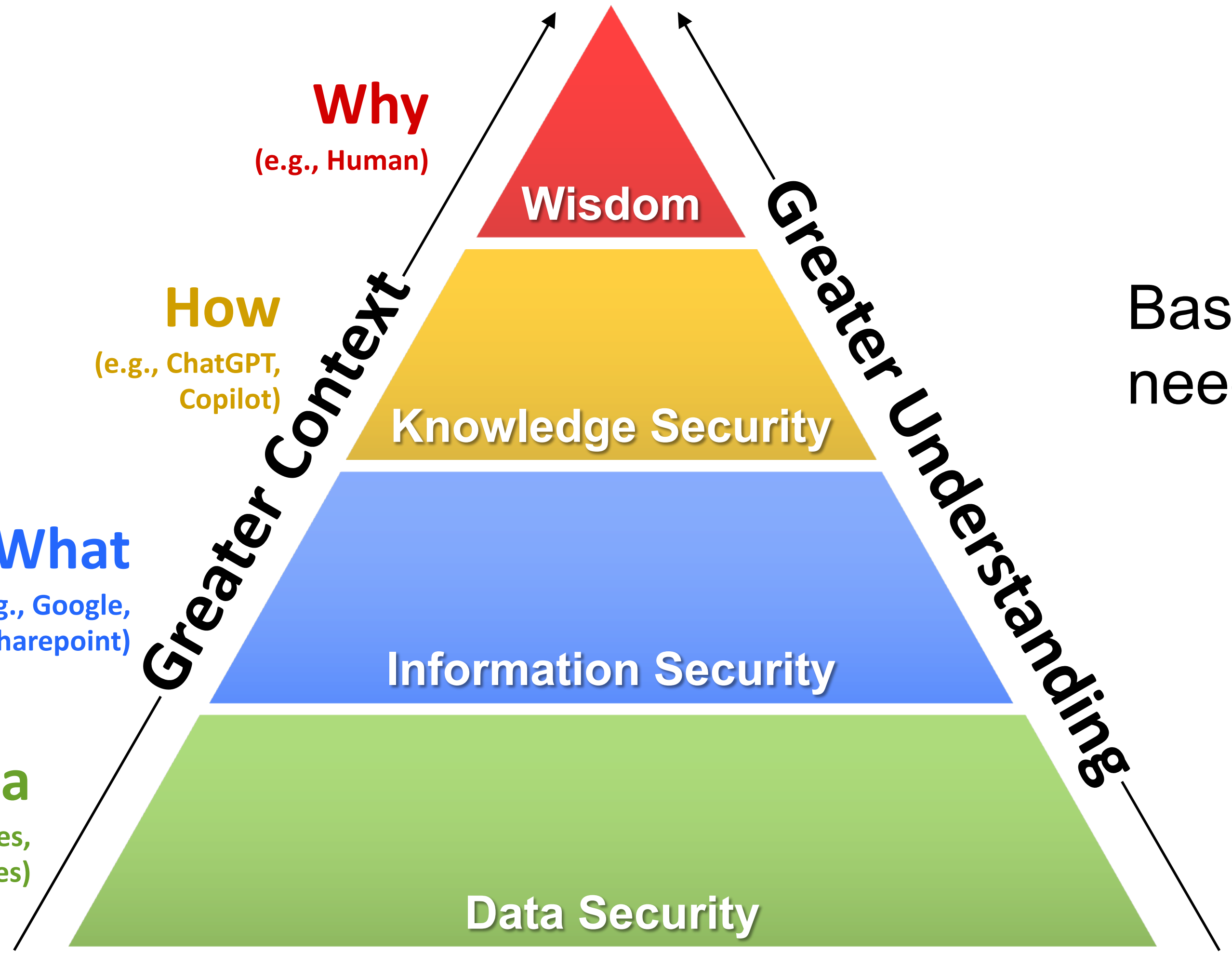
The Knowledge Economy



ChatGPT has unlocked the knowledge economy that will power the AI era

- Knowledge Engineering
- Knowledge Lakes
- Knowledge Pipelines
- Knowledge Quality
- Knowledge Security
- Knowledge Governance
- Knowledge Search
- Knowledge Sharing
- Knowledge-based Decision Support
- Knowledge Classification
- Knowledge Provenance
- Knowledge Management
- Knowledge Privacy

What is Knowledge Security?



Why
(e.g., Human)

How
(e.g., ChatGPT, Copilot)

What
(e.g., Google, Sharepoint)

Raw data
(e.g., Databases, Websites)

Based on need-to-know

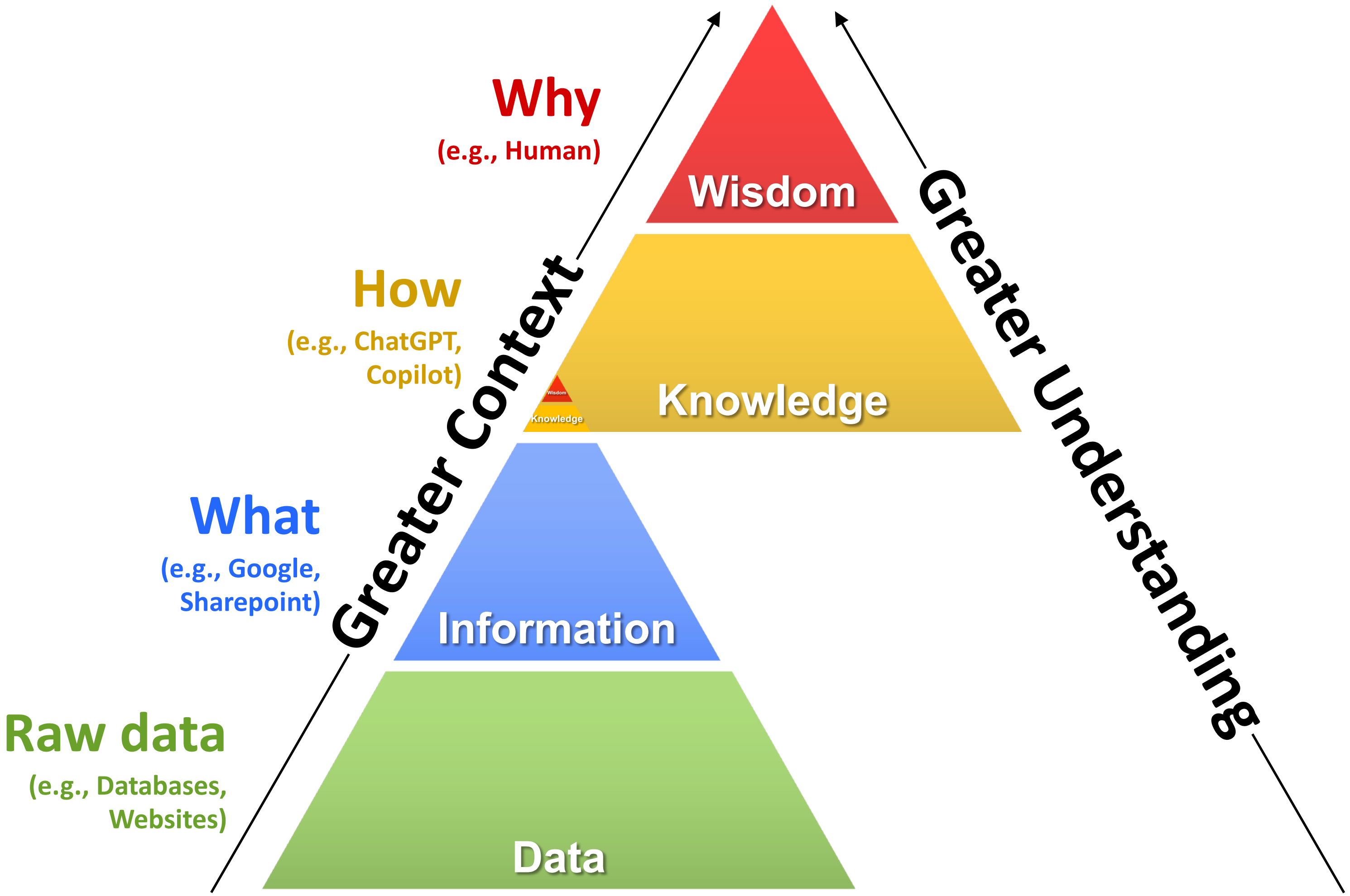
Based on patterns / classification / tags

Based on specific files, datastores, sites, repos

KNOSTIC

@sounilyu

Knowledge Fragments With Data and Information Security



Why

(e.g., Human)

Wisdom

How

(e.g., ChatGPT, Copilot)

Wisdom

Knowledge

Knowledge

What

(e.g., Google, Sharepoint)

Information

Raw data

(e.g., Databases, Websites)

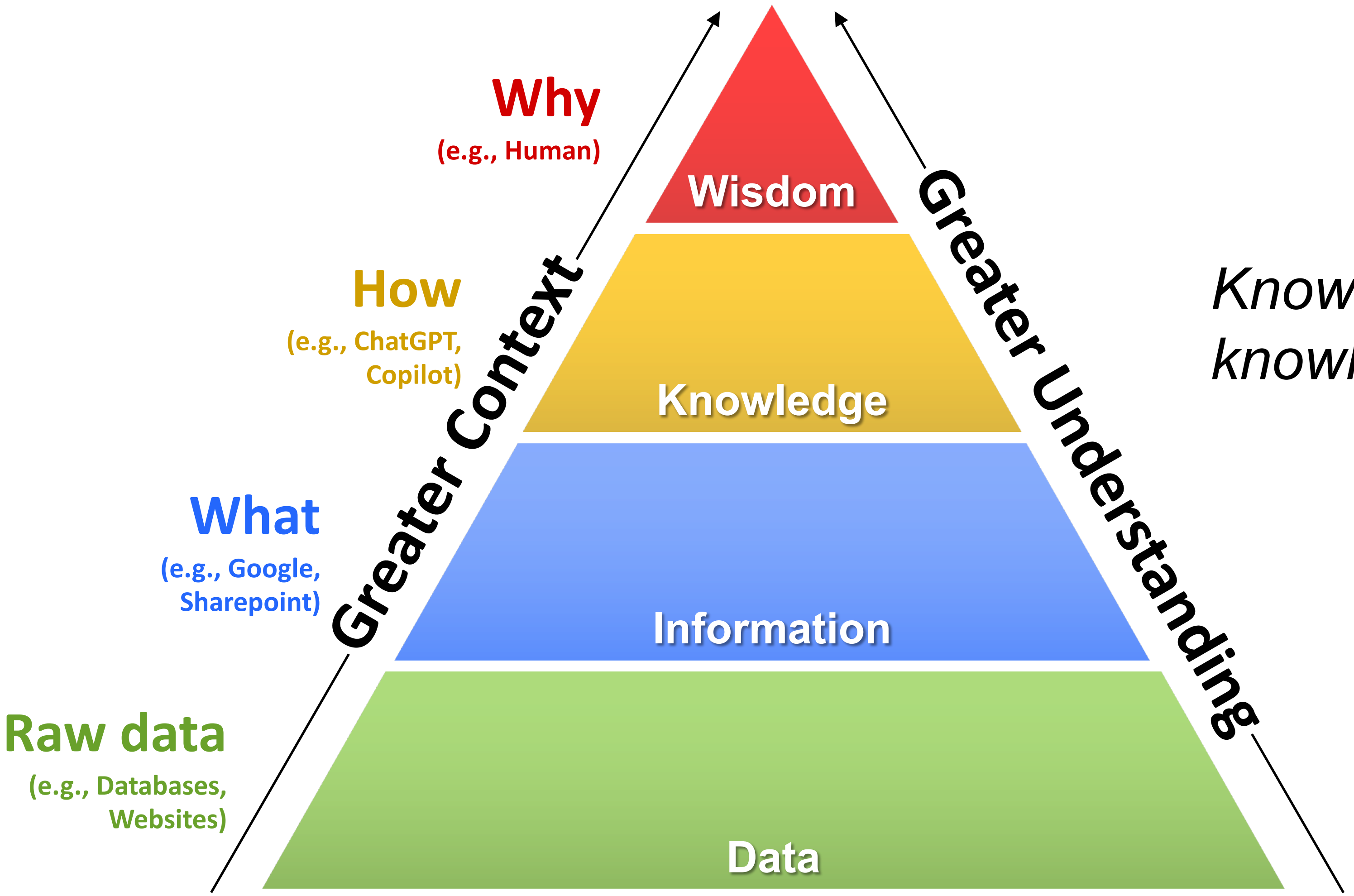
Data

Greater Context

Greater Understanding

Data and information-centric controls will fail in the age of LLMs because of the tradeoff between harm and utility

Conclusion



Knowledge-centric capabilities require *knowledge*-centric controls

Anything else will be suboptimal

Questions?

@sounilyu

sounil@knostic.ai

https://www.linkedin.com/in/sounil

https://www.slideshare.net/sounilyu/presentations