WEBINAR RECORDED

PECB

# ISO/IEC 27001 & ISO/IEC 42001: The Twin Pillars of AI Governance and Cybersecurity

MARCH 27 | 03:00-04:00 PM CET

**Anthony English**
VP Security / CISO

**Deepinder Chhabra**
Head of GRC Professional Services (EMEA)

*#GlobalLeadingVoices*

PECB

# Disclaimer

I would like to clarify that the views and opinions expressed in this presentation are my own and do not reflect the views, positions, or opinions of my employer, any organizations I am associated with, or any third-party entities.

I am speaking in my personal capacity, and the information presented here is for informational and educational purposes only. It should not be construed as legal, professional, or business advice.

Deepinder Chhabra

# Agenda

- Overview of ISO/IEC 27001 vs ISO/IEC 42001
- ISMS vs AIMS
- Key considerations for 42001
- Integration of ISMS & AIMS
- Risk Management and Compliance
- Developing AI SecOps

# ISO 27001 vs. ISO 42001

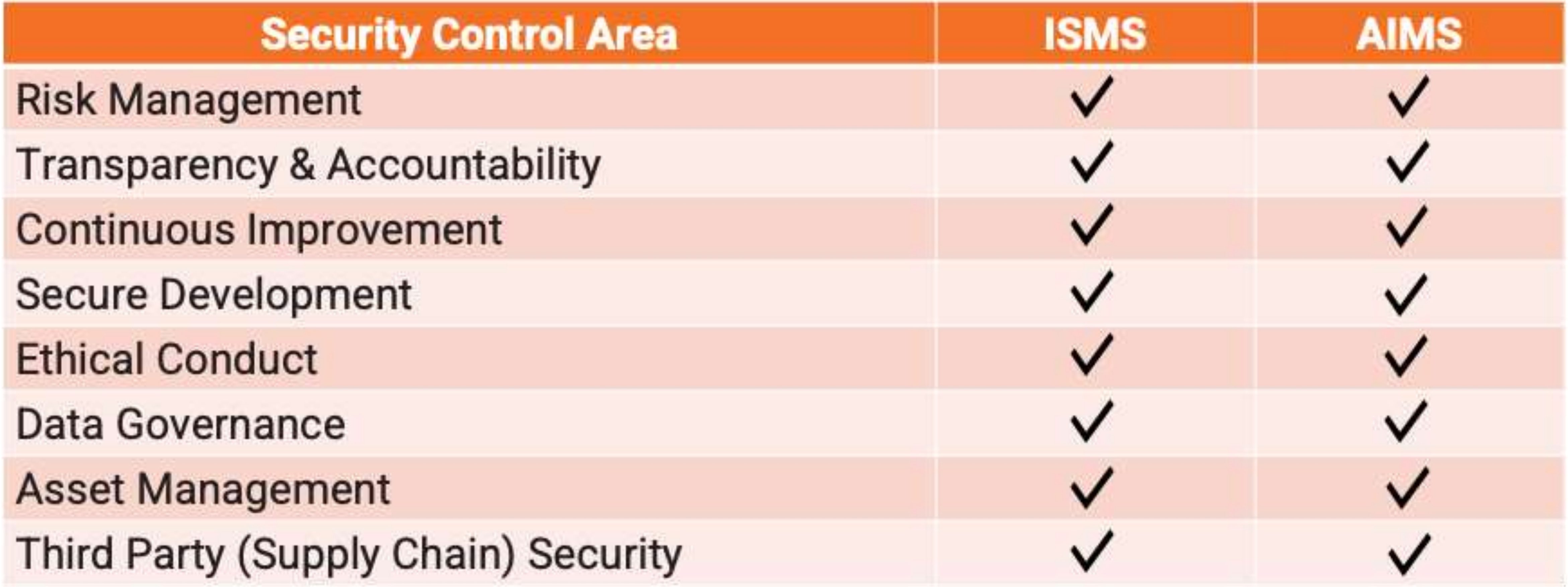
## What is the relationship?

**ISO/IEC 27001** defines the requirements to create, implement, and maintain an **Information Security Management system (ISMS)**.

**ISO/IEC 42001** defines the requirements to design, build, and continually improving an **Artificial Intelligence Management System (AIMS)**.

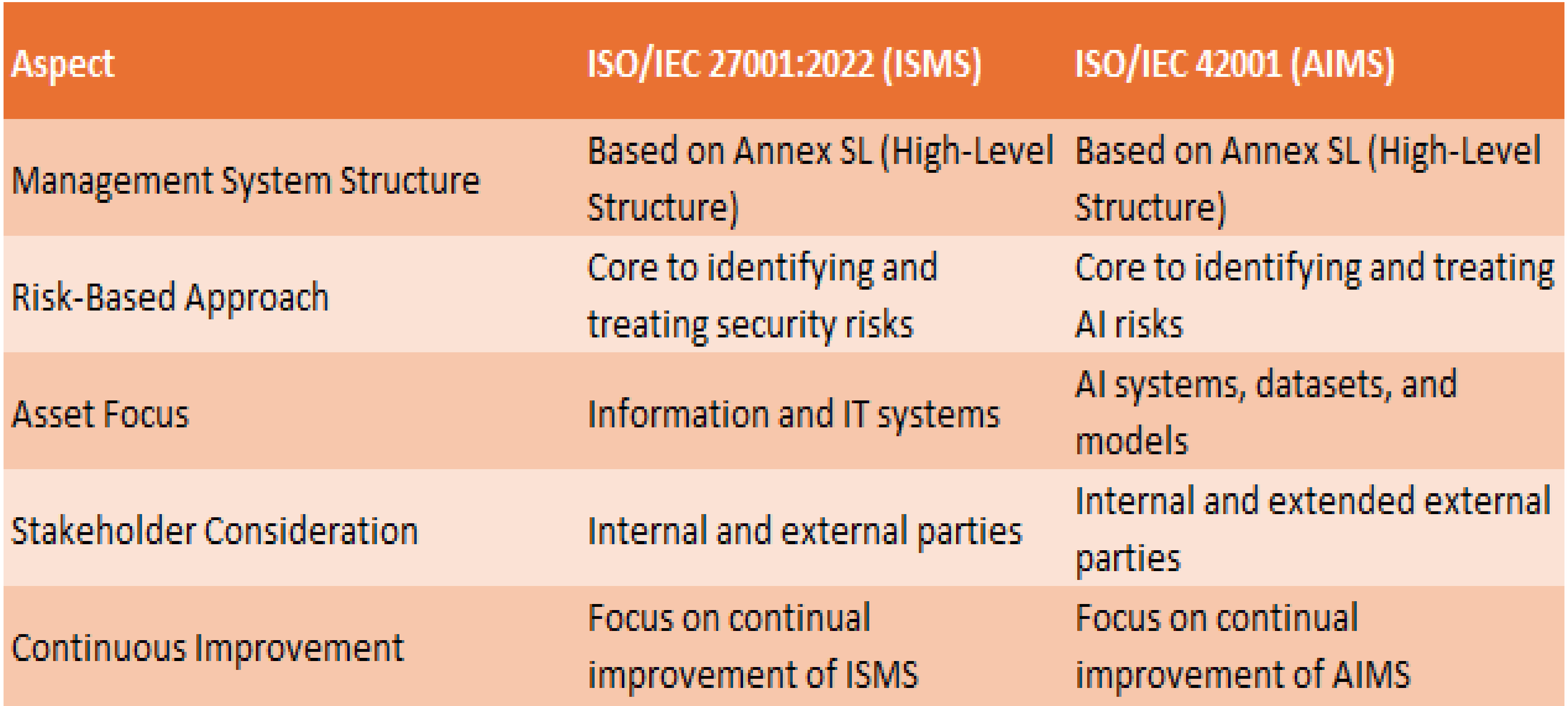
# ISMS vs AIMS

## Interrelated Management Systems



| Security Control Area | ISMS | AIMS |
|---|---|---|
| Risk Management | ✓ | ✓ |
| Transparency & Accountability | ✓ | ✓ |
| Continuous Improvement | ✓ | ✓ |
| Secure Development | ✓ | ✓ |
| Ethical Conduct | ✓ | ✓ |
| Data Governance | ✓ | ✓ |
| Asset Management | ✓ | ✓ |
| Third Party (Supply Chain) Security | ✓ | ✓ |

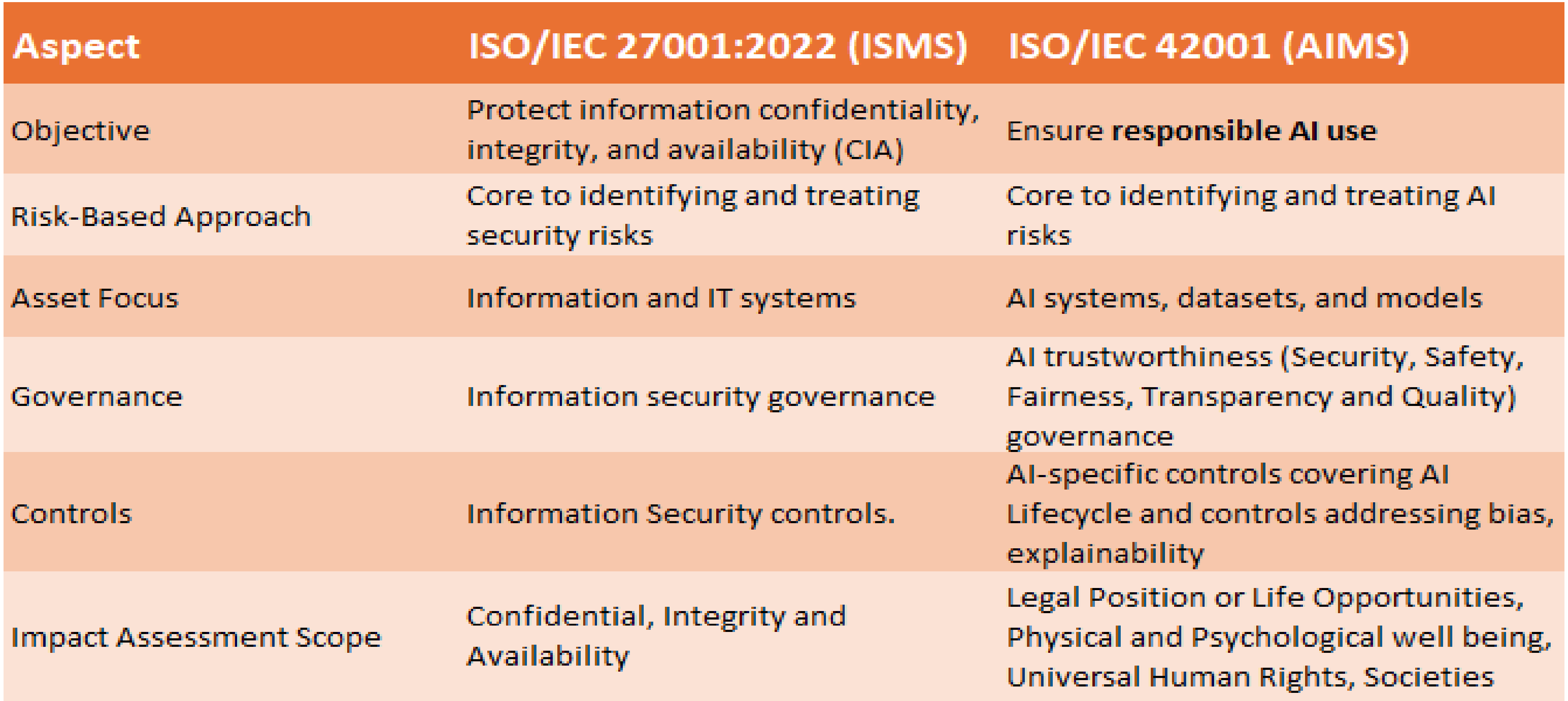
# Key Similarities



| Aspect | ISO/IEC 27001:2022 (ISMS) | ISO/IEC 42001 (AIMS) |
|---|---|---|
| Management System Structure | Based on Annex SL (High-Level Structure) | Based on Annex SL (High-Level Structure) |
| Risk-Based Approach | Core to identifying and treating security risks | Core to identifying and treating AI risks |
| Asset Focus | Information and IT systems | AI systems, datasets, and models |
| Stakeholder Consideration | Internal and external parties | Internal and extended external parties |
| Continuous Improvement | Focus on continual improvement of ISMS | Focus on continual improvement of AIMS |

Created by Deepinder Chhabra

# Key Differences

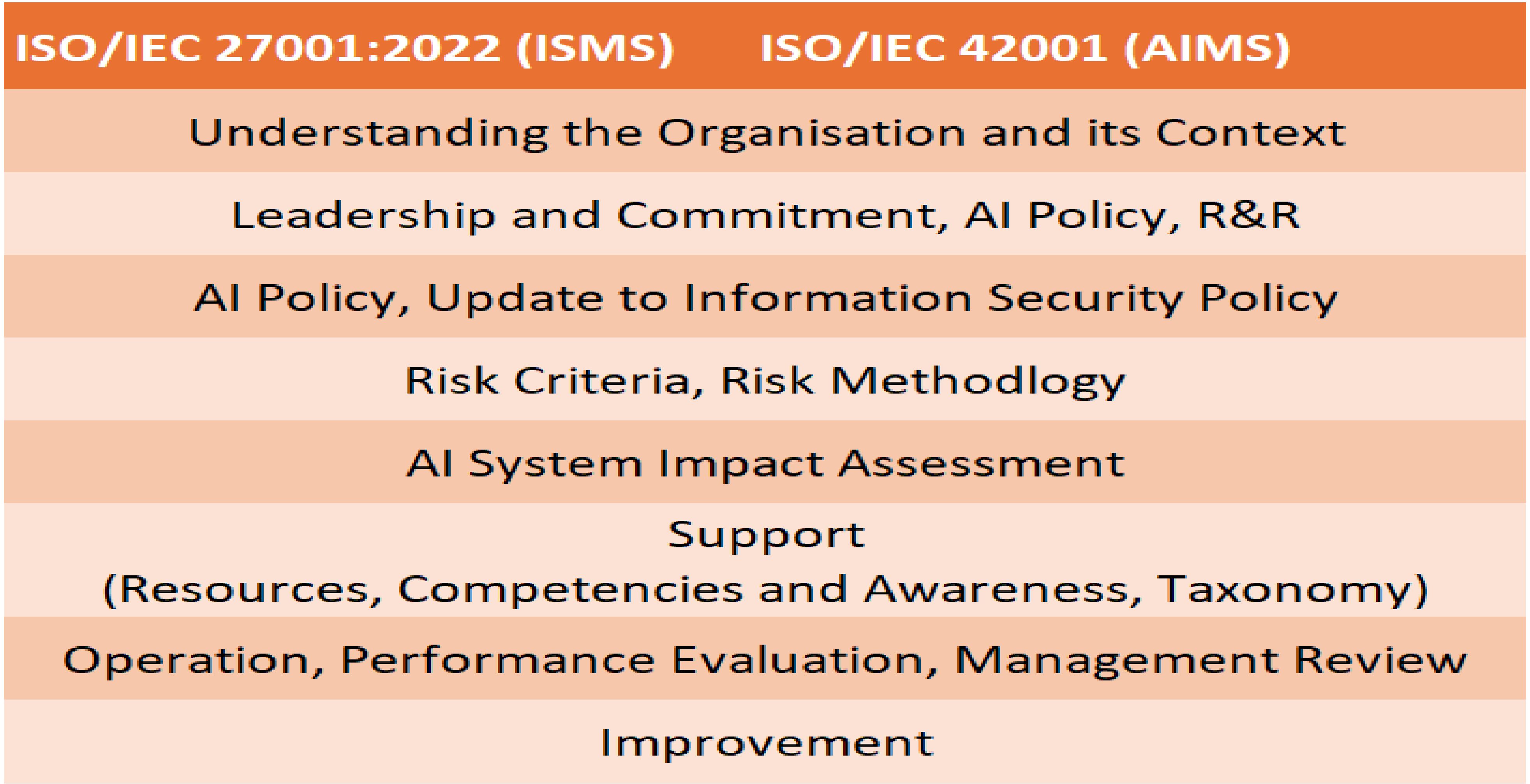


| Aspect | ISO/IEC 27001:2022 (ISMS) | ISO/IEC 42001 (AIMS) |
|---|---|---|
| Objective | Protect information confidentiality, integrity, and availability (CIA) | Ensure **responsible AI use** |
| Risk-Based Approach | Core to identifying and treating security risks | Core to identifying and treating AI risks |
| Asset Focus | Information and IT systems | AI systems, datasets, and models |
| Governance | Information security governance | AI trustworthiness (Security, Safety, Fairness, Transparency and Quality) governance |
| Controls | Information Security controls. | AI-specific controls covering AI Lifecycle and controls addressing bias, explainability |
| Impact Assessment Scope | Confidential, Integrity and Availability | Legal Position or Life Opportunities, Physical and Psychological well being, Universal Human Rights, Societies |

Created by Deepinder Chhabra

# Integrating ISMS & AIMS

# Integrated Management Systems



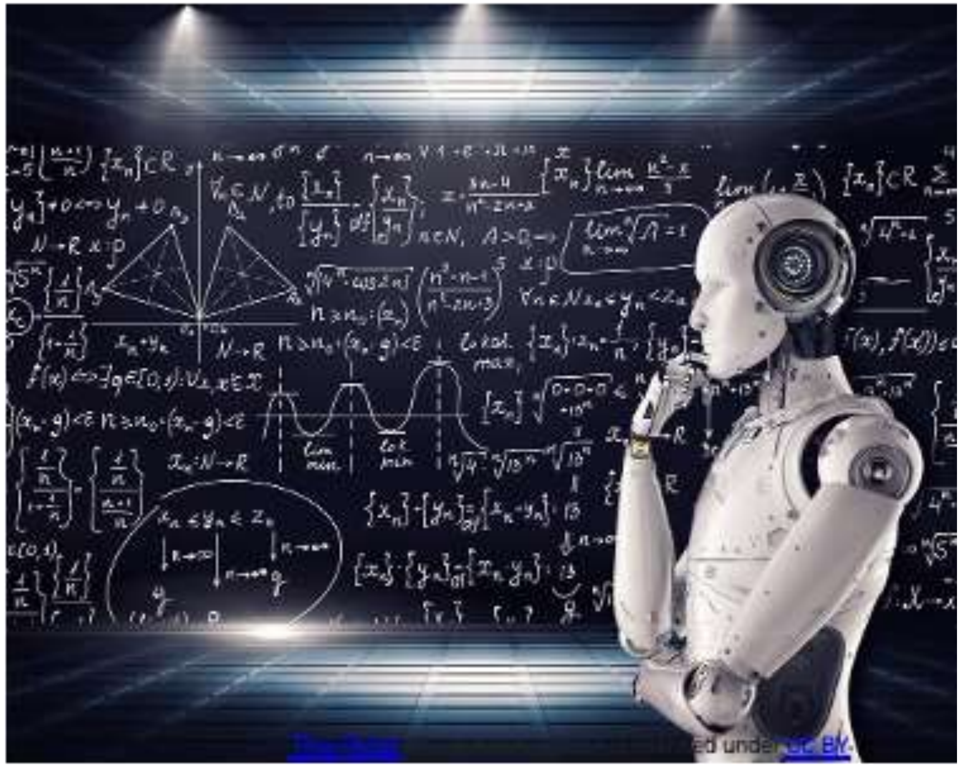
<table>
<tr><th>ISO/IEC 27001:2022 (ISMS)</th><th>ISO/IEC 42001 (AIMS)</th></tr>
<tr><td colspan="2">Understanding the Organisation and its Context</td></tr>
<tr><td colspan="2">Leadership and Commitment, AI Policy, R&R</td></tr>
<tr><td colspan="2">AI Policy, Update to Information Security Policy</td></tr>
<tr><td colspan="2">Risk Criteria, Risk Methodlogy</td></tr>
<tr><td colspan="2">AI System Impact Assessment</td></tr>
<tr><td colspan="2">Support<br>(Resources, Competencies and Awareness, Taxonomy)</td></tr>
<tr><td colspan="2">Operation, Performance Evaluation, Management Review</td></tr>
<tr><td colspan="2">Improvement</td></tr>
</table>

Created by Deepinder Chhabra

# ISO 42001 Key Considerations

***AI Considerations***

- AI risks
  - Data bias
  - Algorithmic fairness
  - AI security vulnerabilities
- AI controls integration
  - Using AI in development
  - Integrating AI into your solutions
  - Using AI for work tasks
  - Supply chain and AI



# Addressing Artificial Intelligence Risks

## AI Content in your ISMS

Build an **AI Policy** that includes:

- Safe use of AI technologies (e.g., use only approved and licensed LLM's, do not share confidential data or names in prompts, etc.)
- Uses for AI in your organization (e.g., building marketing content, assisting with coding, etc.)
- AI tools currently approved for use in your organization

Update your **SDLC** (Software/Secure Development Policy):

- Add testing and design tips for bias
- Add testing and design tips for algorithmic fairness
- Add guidance on protecting PII or other sensitive data in your product

# Addressing Artificial Intelligence Risks

## AI Content in your ISMS

Update your **Supply Chain or Third-Party** Security:

- When risk assessing a new or existing Supplier or third-party partner, ask:
  - Do you have or use AI in the services or product you will be supplying us?
  - If yes, how do you risk assess and secure your use of AI?
  - If yes, does your AI utilize other external AI's?
  - If yes, what happens to prompt data?

Build a dedicated **AI Data Privacy Policy** or Update your existing Privacy Policy:

- Add language that outlines how you handle, store, and process prompt data

# Addressing Artificial Intelligence Risks

## AI Content in your ISMS

Update your **Risk Management Policy** and Process:

- Your risk assessment process should now include:
  - Assessing AI specific risks for AI tools or service usage
    - The new AI tool or service should be assessed for AI risks
    - Any identified AI risks should be in your risk register

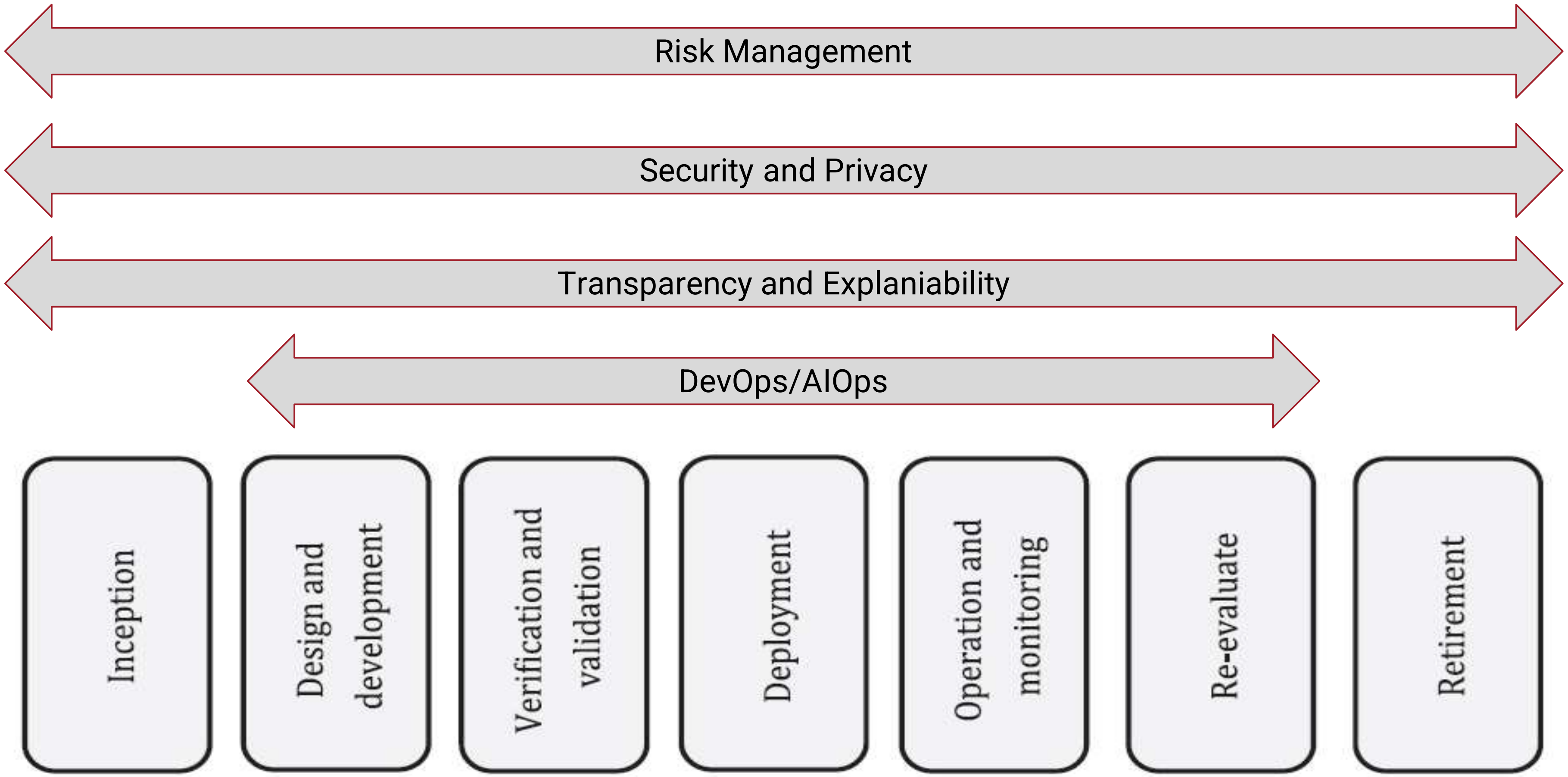
Update your **Asset Inventory** to include AI tools and associated data:

- Prompt data may be retained in a database (e.g., Mongo)
- Prompt data may contain PII or other confidential data

PECB

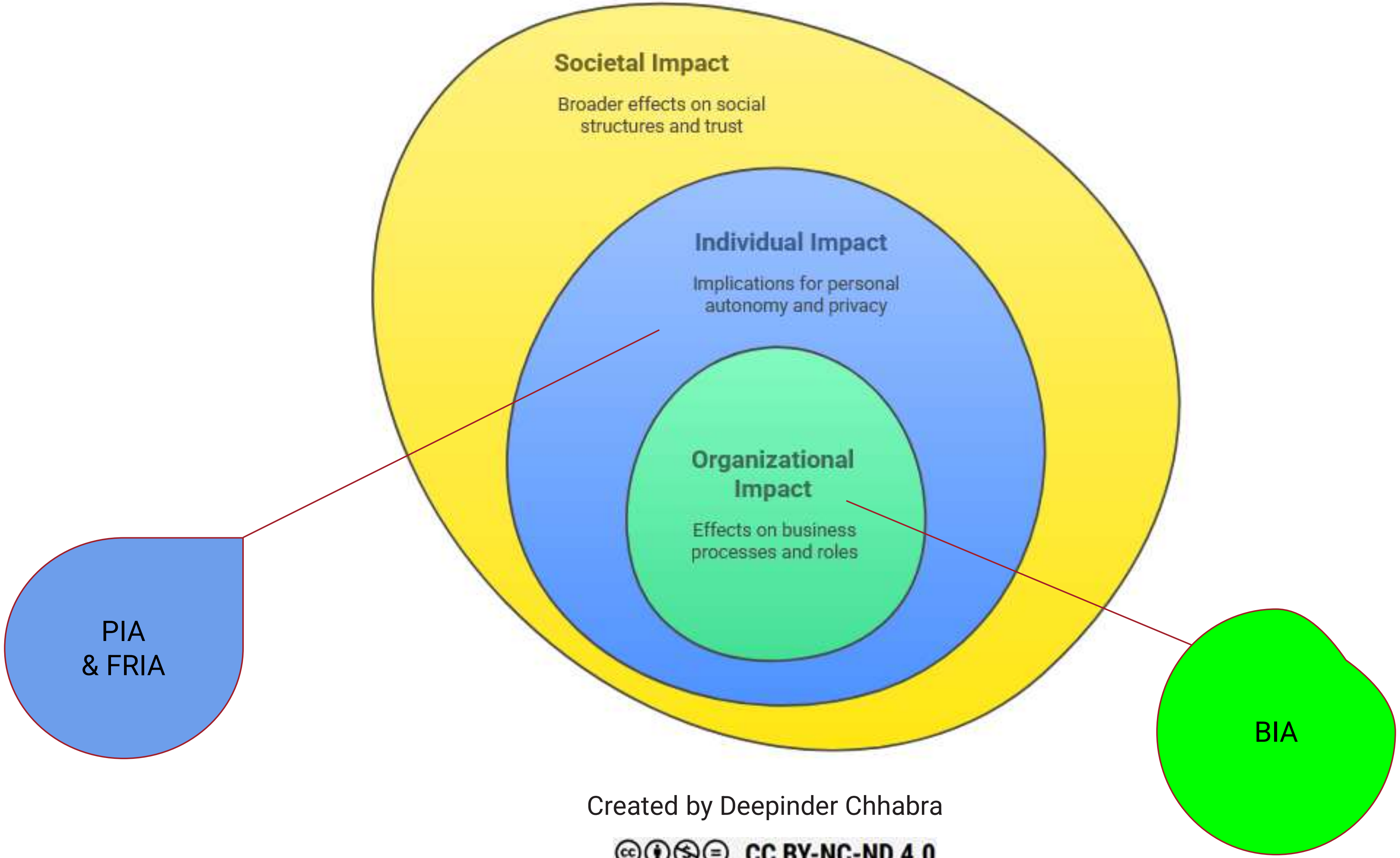
# Risk Management & Compliance

# Risk Management and AI Lifecycle



Adapted from ISO 22989:2022 by Deepinder Chhabra

# AI Impact Assessment



Societal Impact

Broader effects on social structures and trust

Individual Impact

Implications for personal autonomy and privacy

Organizational Impact

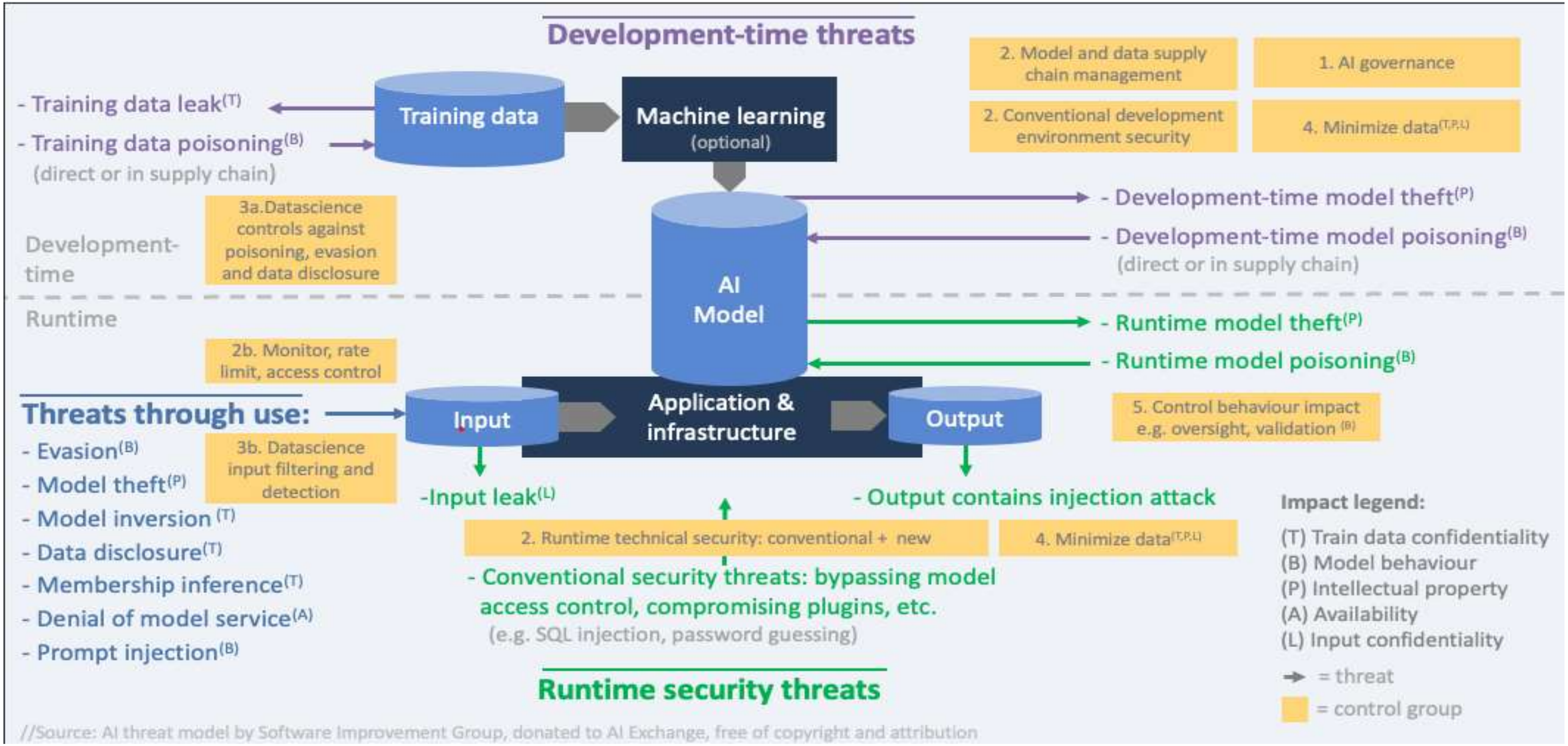
Effects on business processes and roles

PIA & FRIA

BIA

Created by Deepinder Chhabra

# AI Threats and Controls



Development-time threats

- Training data leak(T)
- Training data poisoning(B)
  (direct or in supply chain)

Training data

Machine learning
(optional)

2. Model and data supply chain management

1. AI governance

2. Conventional development environment security

4. Minimize data(T,P,L)

3a.Datascience controls against poisoning, evasion and data disclosure

Development-time

AI Model

- Development-time model theft(P)
- Development-time model poisoning(B)
  (direct or in supply chain)

Runtime

2b. Monitor, rate limit, access control

- Runtime model theft(P)
- Runtime model poisoning(B)

Threats through use:

- Evasion(B)
- Model theft(P)
- Model inversion (T)
- Data disclosure(T)
- Membership inference(T)
- Denial of model service(A)
- Prompt injection(B)

3b. Datascience input filtering and detection

Input

Application & infrastructure

Output

5. Control behaviour impact e.g. oversight, validation (B)

-Input leak(L)

- Output contains injection attack

2. Runtime technical security: conventional + new

4. Minimize data(T,P,L)

- Conventional security threats: bypassing model access control, compromising plugins, etc.
  (e.g. SQL injection, password guessing)

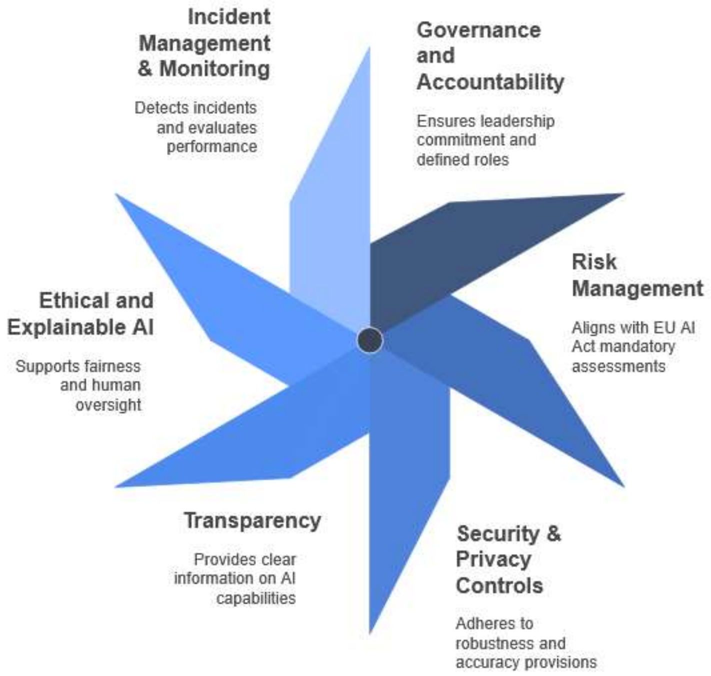
Runtime security threats

Impact legend:
(T) Train data confidentiality
(B) Model behaviour
(P) Intellectual property
(A) Availability
(L) Input confidentiality
➔ = threat
🟨 = control group

//Source: AI threat model by Software Improvement Group, donated to AI Exchange, free of copyright and attribution

# Compliance

## EU AI Act/ Local AI Regulations



**Incident Management & Monitoring**

Detects incidents and evaluates performance

**Governance and Accountability**

Ensures leadership commitment and defined roles

**Ethical and Explainable AI**

Supports fairness and human oversight

**Risk Management**

Aligns with EU AI Act mandatory assessments

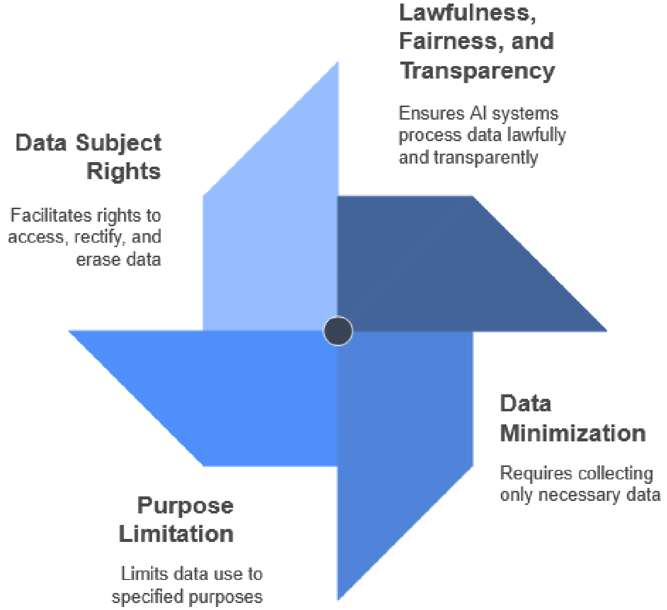
**Transparency**

Provides clear information on AI capabilities

**Security & Privacy Controls**

Adheres to robustness and accuracy provisions

## EU GDPR/ Privacy Laws



**Lawfulness, Fairness, and Transparency**

Ensures AI systems process data lawfully and transparently

**Data Subject Rights**

Facilitates rights to access, rectify, and erase data

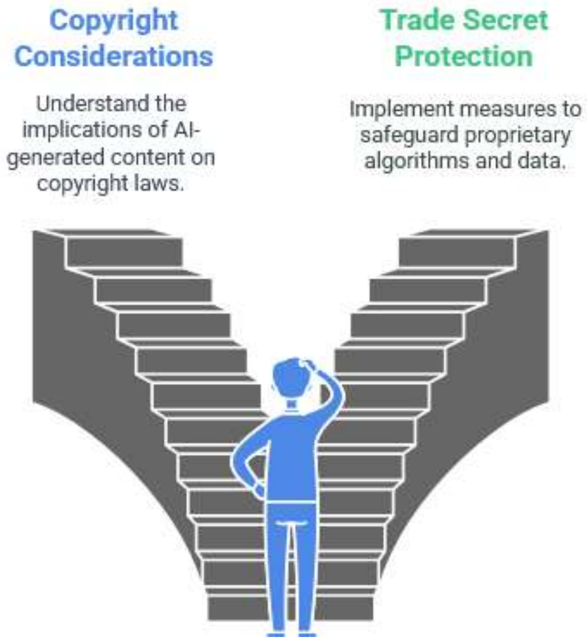
**Data Minimization**

Requires collecting only necessary data

**Purpose Limitation**

Limits data use to specified purposes

## Copyright & Trade Secret Protection



**Copyright Considerations**

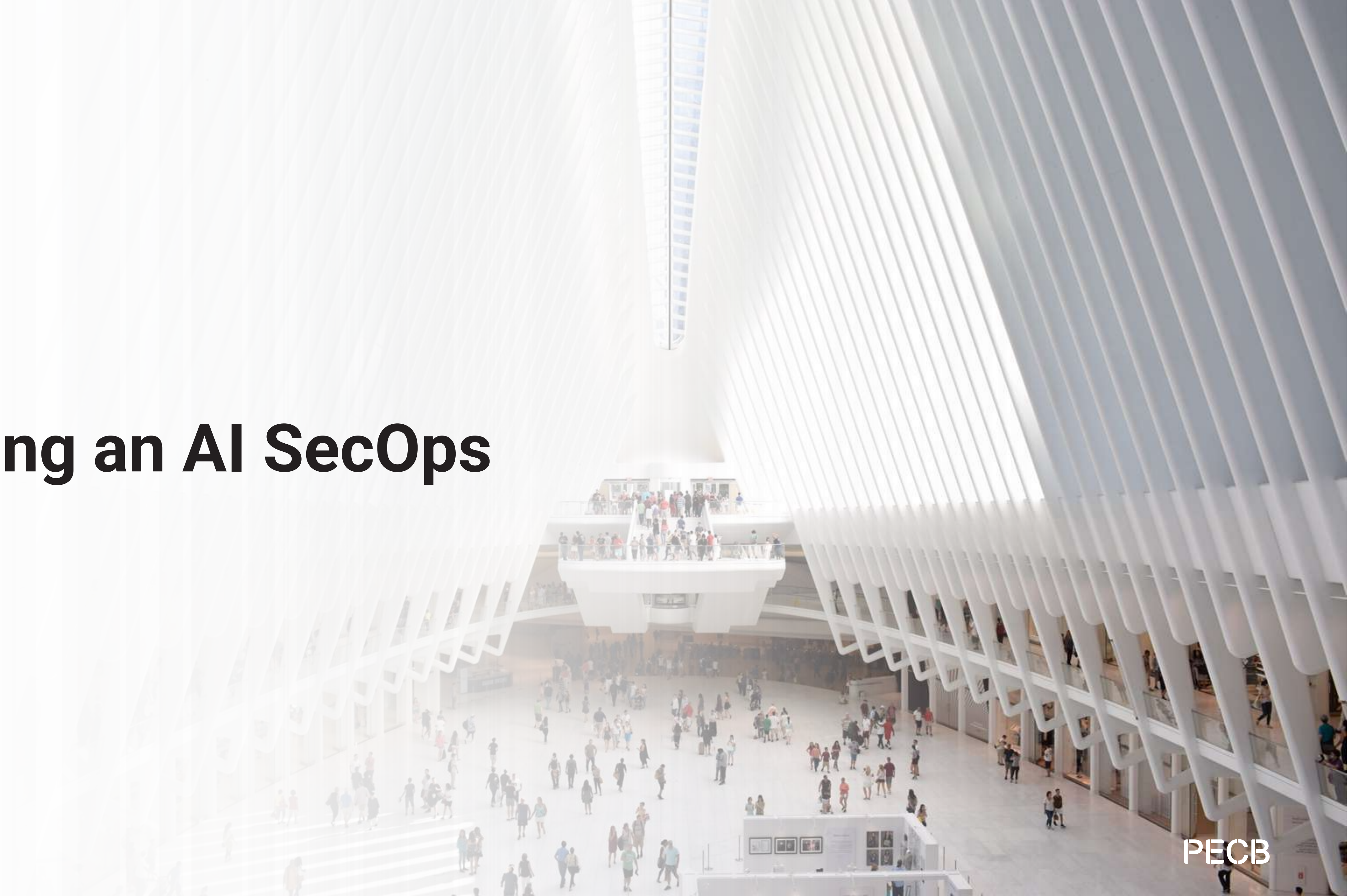
Understand the implications of AI-generated content on copyright laws.

**Trade Secret Protection**

Implement measures to safeguard proprietary algorithms and data.

Created by Deepinder Chhabra

# Building an AI SecOps



# AI Security Operations (SecOps)

## Two types

**AI *in* SecOps**
The use of AI tools or AI enabled tools to help enhance Security Operations teams through expedited threat analysis, big data (e.g., log files) processing, and AI suggested remediations.

**SecOps for AI**
The use of security tools to protect AI tools and/or to analyze AI for security weaknesses (e.g., [https://github.com/meta-llama/PurpleLlama](https://github.com/meta-llama/PurpleLlama) ).

# AI Security Operations (SecOps)

- SecOps for AI

AI coding assist tools can be helpful, but they can also introduce:

- Inaccurate code
- Insecure code
- Irrelevant code

The SDLC must include checks on AI code assist generated source code.

# AI Security Operations (SecOps)

## SecOps for AI

**AI elements in your source code** or application can be vulnerable to:

- Data bias
- Algorithmic fairness risks
- Security vulnerabilities
- Hallucinations

The SDLC must include security tests and security protections against these risks.

# AI Security Operations (SecOps)

## SecOps for AI

**AI tools used in your organization for daily workload** can introduce risks such as:

- Confidential or PII data leak
- Multiple AI tools for the same task
- Inaccurate data from AI
- Third Party (Supply Chain) risks

Your organization must vet and manage AI tool usage by all staff.

# Additional Material

# AI SecOps Tools

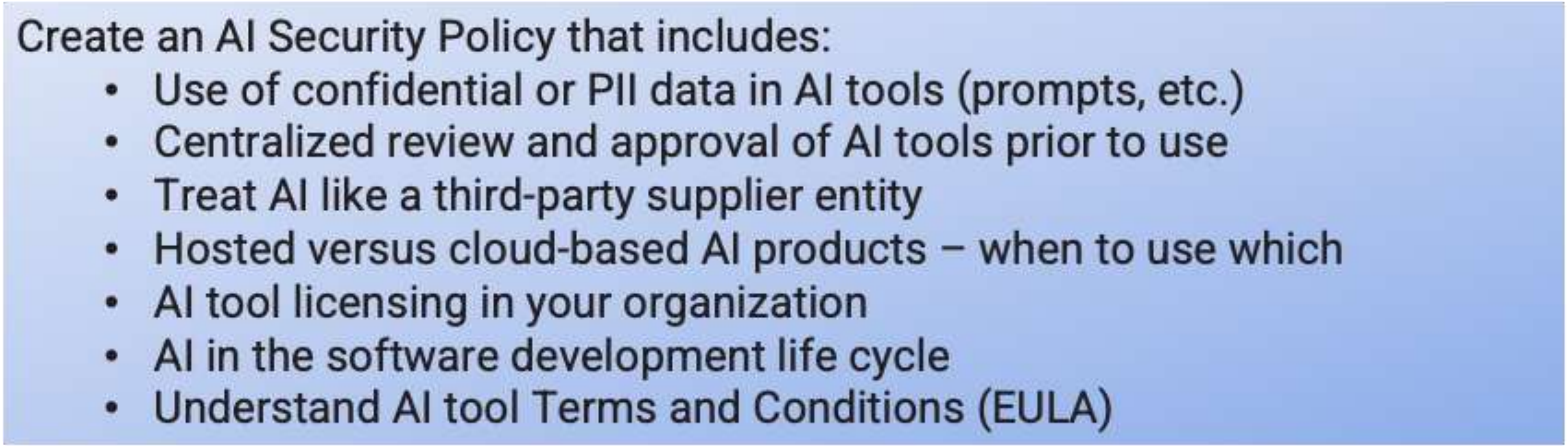
AI Protection Tools:
https://github.com/meta-llama/PurpleLlama

LLM Firewall, Analysis tool, AI Agent tool:
https://Arthur.ai

Zero Trust for AI, LLM runtime security, Automated red teaming:
https://protectai.com

Open source LLM vulnerability scanner:
https://github.com/NVIDIA/garak

# AI Policy



Create an AI Security Policy that includes:

- Use of confidential or PII data in AI tools (prompts, etc.)
- Centralized review and approval of AI tools prior to use
- Treat AI like a third-party supplier entity
- Hosted versus cloud-based AI products – when to use which
- AI tool licensing in your organization
- AI in the software development life cycle
- Understand AI tool Terms and Conditions (EULA)

# AI Security and Legal/Privacy Concerns

**Security**

- Safety vs Security
- Biases/Misinformation
- Data Leaks
- Injections/Breakouts
- Data Storage
- Dependency
- Employee Concerns

**Legal / Privacy**

- Privacy
    - Training Data
    - Prompt Data
- Due Diligence (copyright/trademark)
- Intellectual Property

# AI SecOps Threat Areas

- Supply Chain Vulnerability
- AI model life cycle
- AI Governance
- Creating and Using Trusted AI
- Machine Learning Attacks (e.g., evasion, poisoning, abuse attacks, privacy attacks)

# Some Sources

OWASP
https://owasp.org/www-project-ai-security-and-privacy-guide/

NIST
https://www.nist.gov/itl/ai-risk-management-framework

EU
https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

ISO
https://pecb.com/en/education-and-certification-for-individuals/iso-iec-42001

# Some Sources

NIST AI Cyberattacks
https://www.nist.gov/news-events/news/2024/01/nist-identifies-types-cyberattacks-manipulate-behavior-ai-systems

AI Infrastructure.org
https://ai-infrastructure.org/understanding-types-of-ai-attacks/

# Q&A

# THANK YOU

✉ asenglish@hotmail.com

✉ deepsc_uk@yahoo.co.uk

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

in https://www.linkedin.com/in/deepinder-singh-chhabra-mba-cciso-0656122/