ASSER INSTITUTE
Centre for International & European Law

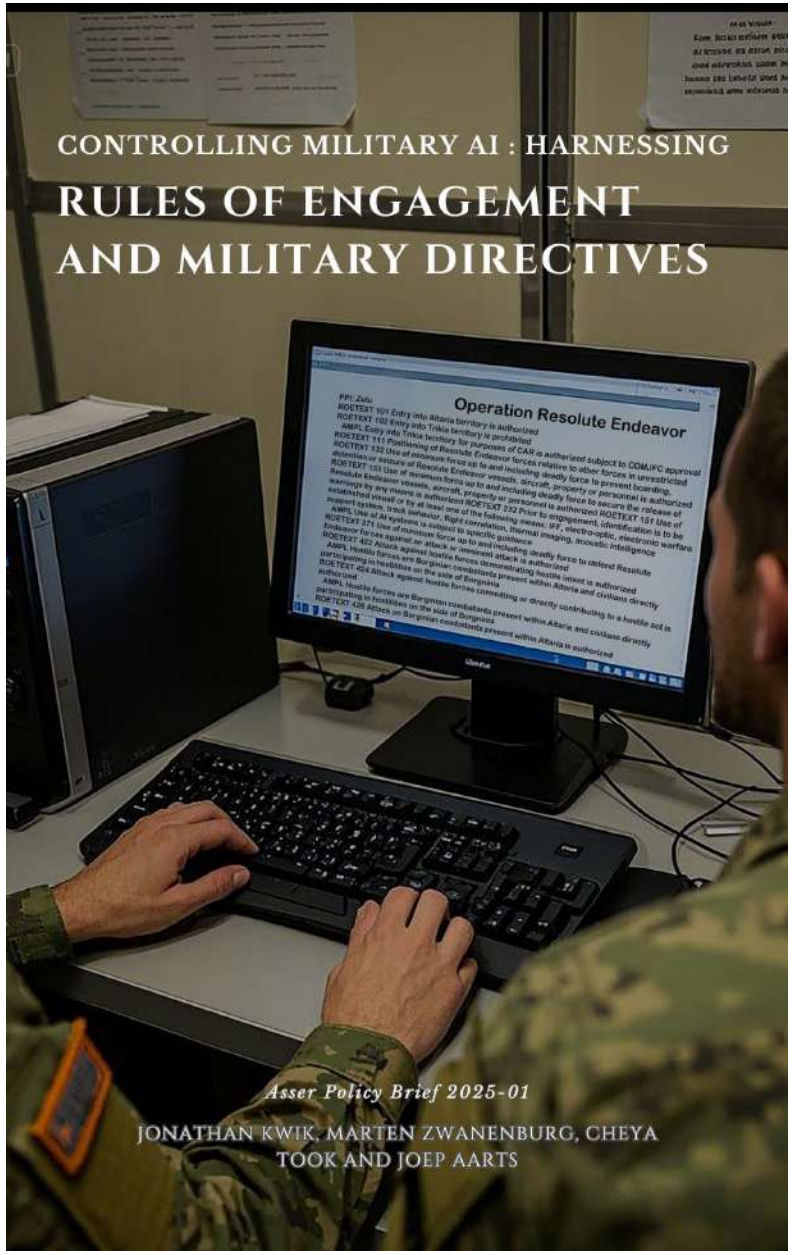
ELSA LAB DEFENCE

Risks to IHL Compliance - A Technical, Operational and Institutional Perspective

Dr Jonathan Kwik

Centre for the Study of Military Law and the Law of Armed Conflict
Brussels, 25 June 2026


DIGITAL YES-MEN
HOW TO DEAL WITH SYCOPHANTIC MILITARY AI?
Global Policy, Special Section Article
JONATHAN KWIK


CONTROLLING MILITARY AI : HARNESSING
RULES OF ENGAGEMENT AND MILITARY DIRECTIVES
Operation Resolute Endeavor
Asser Policy Brief 2025-01
JONATHAN KWIK, MARTEN ZWANENBURG, CHEYA TOOK AND JOEP AARTS


AI-ON-AI PERFIDY
Articles of War
SAMUEL WHITE, JONATHAN KWIK, DORA VELENCZEI

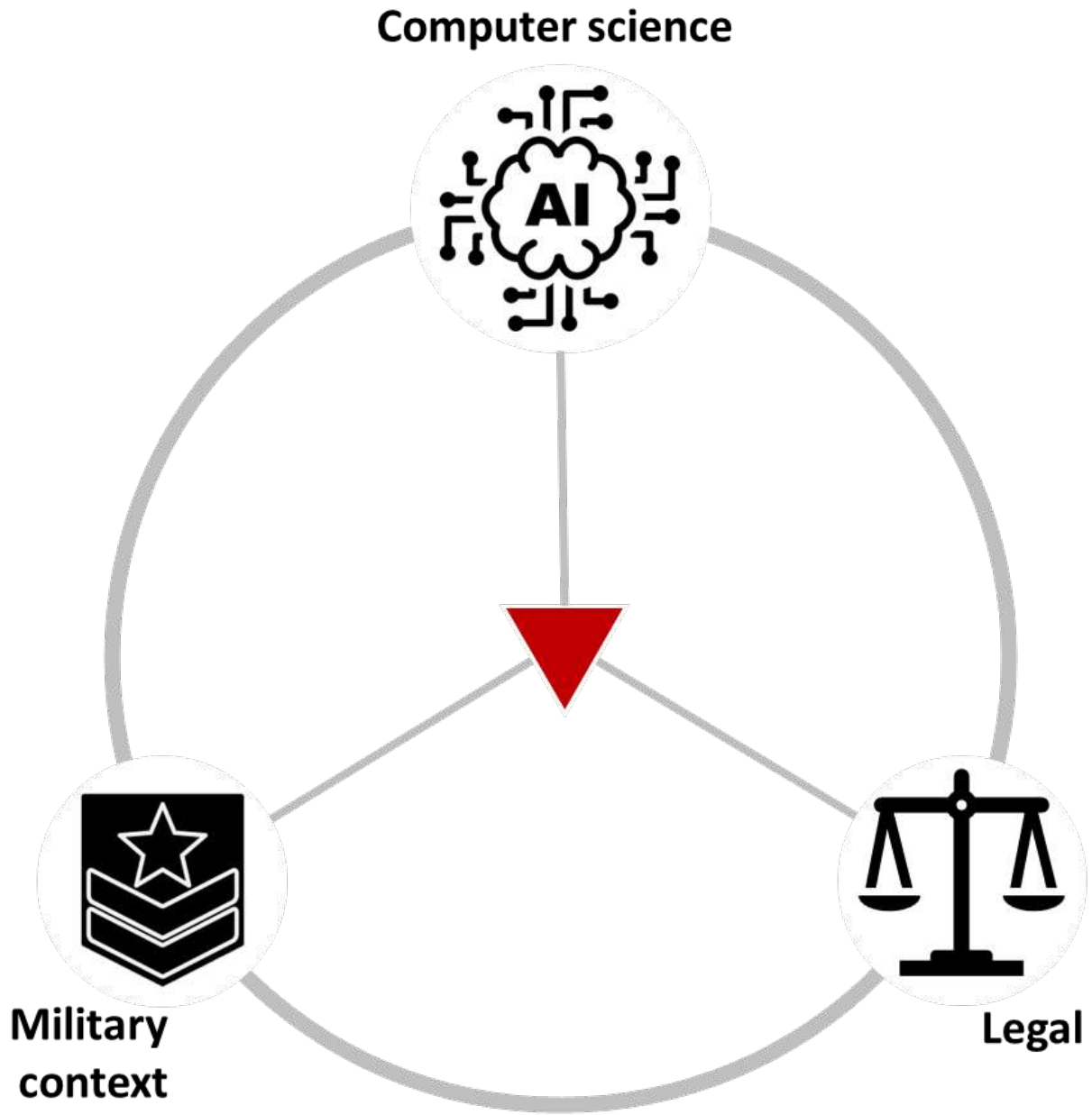
LEGAiD

[Grant award] LEGAiD: A Critical Study of AI Legal Assistants in the Military Domain

19 April 2026 Admin 06.44 Share

Compliance Risk Framework for Evaluating the Legal Acceptability of Military AI Systems

Jonathan Kwik and Marta Bo

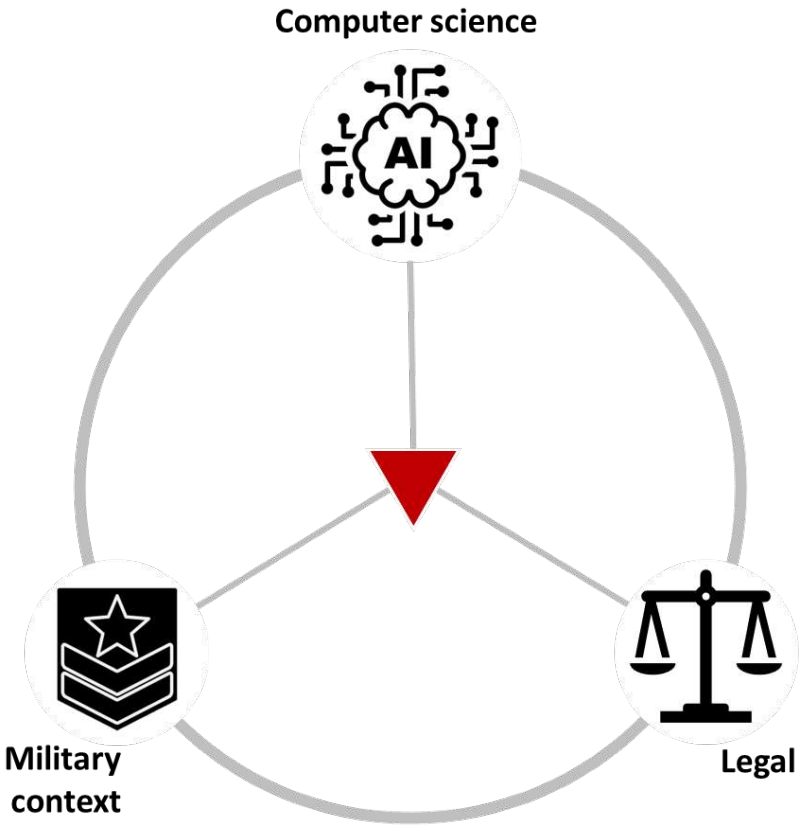

Computer science
AI
Military context
Legal

Developer

Acquisition & evaluation

Commander

Operator



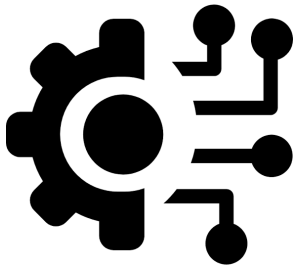
Problem 1: Abstract norms

Abstract norms for technically-complicated systems → What technical/operational/ institutional factors to consider? → Maladaptive reviews, overlooking factors, poor decisions, bad communication

tension

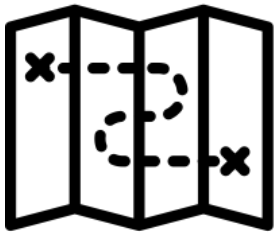
Abstract principles, many lack legal expertise → Difficulties in translating IHL into concrete specifications → Poor investments, difficulty planning, IHL-noncompliant features/specifications

Problem 2: Context-dependence



Performance metrics
AI model training process
Data used
Interface design
Updating, modifications

Objective
Type of operation
Operator conditions
Doctrine, procedure
Institutional culture



Type of environment
Rate of change
Presence of civilians
Interconnectivity
Adversary activity

Problem 2: Context-dependence

Statement:
If an AI-system cannot distinguish, it is unlawful

Counterpoint:
BUT IF… it is used in a desert?

Developer

What factors might become requirements set by your customer?

Acquisition & evaluation

What factors should you evaluate?
How do you communicate this to your developers?
What limitations should you set for your operators?

Commander

Operator

What factors should you take into consideration when making an operational decision?

Provide an **accessible, common frame of reference**

on factors relevant for legal compliance, and provide a

shared vocabulary and understanding

of complicated technical and operational concepts.

Norms

- What are the high-level norms that govern conduct?

Variables

- What legally-relevant factors can be identified?

Evaluation

- How to you evaluate them?

Relationships

- How do they relate to other variables?

D

distinction

C

civilian harm

E

epistemic,
process-based

ENVIRONMENT

- Civilian Presence
- Changing Environment
- Environment Complexity
- Adversarial Threat

TARGETS & ID

- False Positive
- False Negative
- 0.8 Confidence Threshold
- Target Proxy
- Bias
- Civilian Detection
- Status Fluidity

AWARENESS

- Contextual Awareness
- Feedback Loop
- Understand-ability
- Technical Knowledge

SAFETY

- Input Data
- System Security
- System Safety

EFFECTS

- Intended & Likely Effects
- Possible & In-direct Effects
- Unpredict-able Effects

INTERACTION

- Intervention Ability
- Out-of-Scope Use
- Modification

DEVELOPMENT

- Training Data
- Engagement w/ Developer
- Repurposing

Modification

RISK

RISK

RISK

D

Causes **performance drop**

- Reduces reliability in **distinguishing targets** from protected persons/objects
- Violates **Principle of Distinction**

C

Causes **performance drop**

- Reduces reliability in **identifying collateral concerns**
- Violates **Principle of Proportionality**

E

Becomes **difficult to foresee** when major behaviour/performance shift is induced

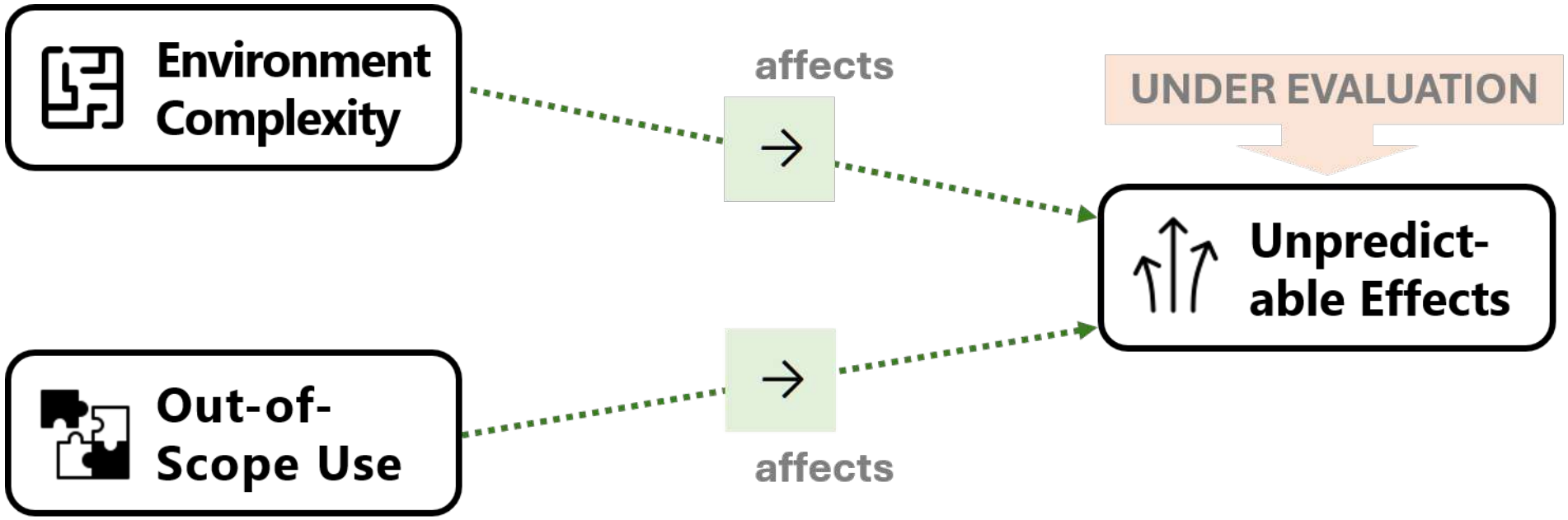
- **Prevents reasoned & informed decisions** from being made
- Violates **Precautionary Principle**

Feedback Loop

Reporting structures	Comprehensive	Relatively comprehensive	Partial	Minimal	None
AND					
Reporting frequency	Constant	Frequent	Often	Rare	Never
AND					
Reporting promptness	Instant	Short delay	Moderate delay	Long delay	Significantly delayed

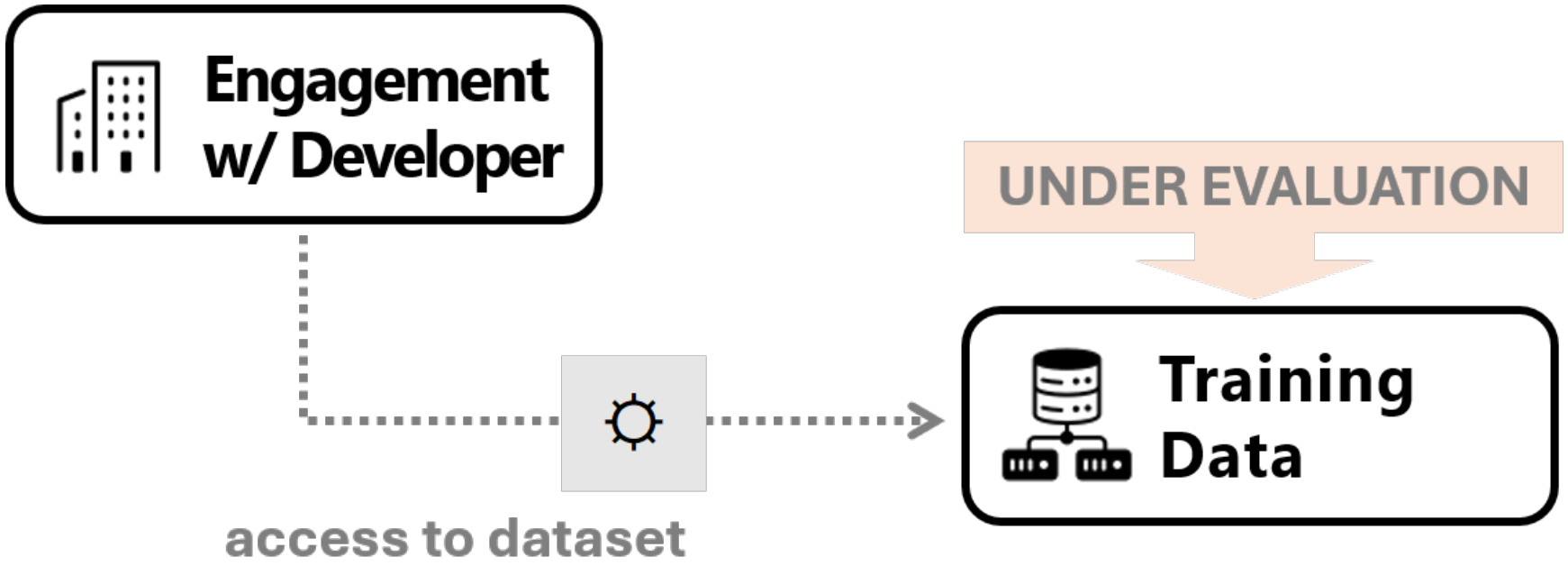
Relationships 1

Correlation & causation



Relationships 2

Necessary to be able to evaluate



Relationships 3

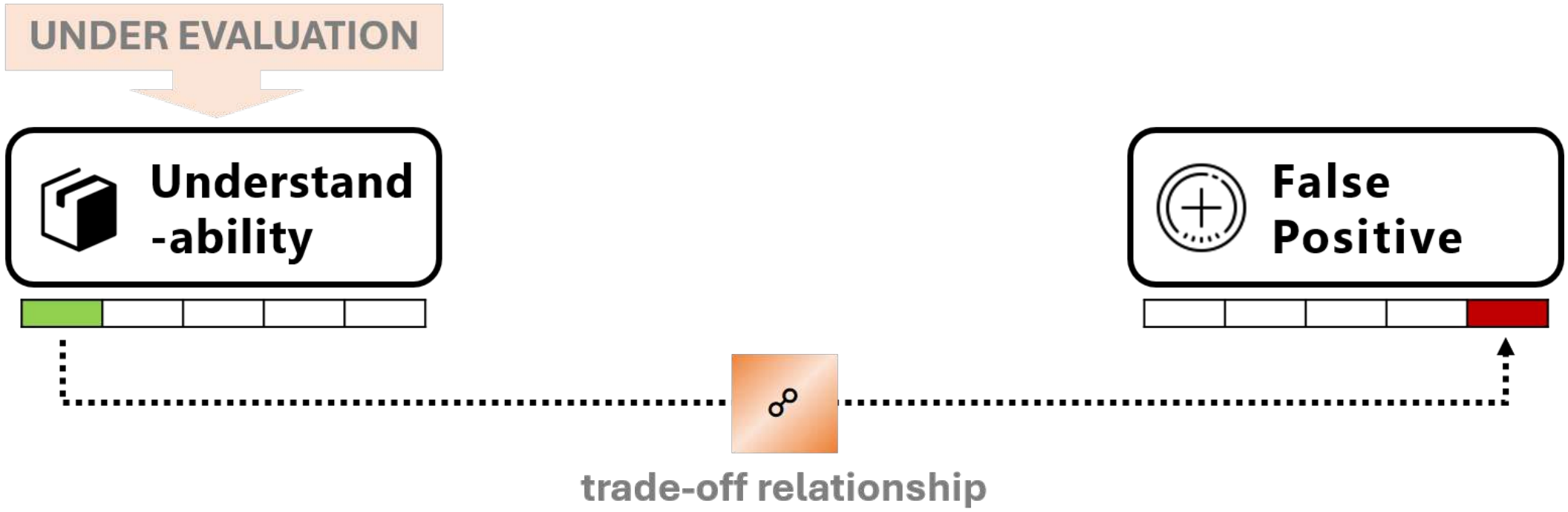
Dangerous combinations

Relationships 4

Critically dangerous combinations

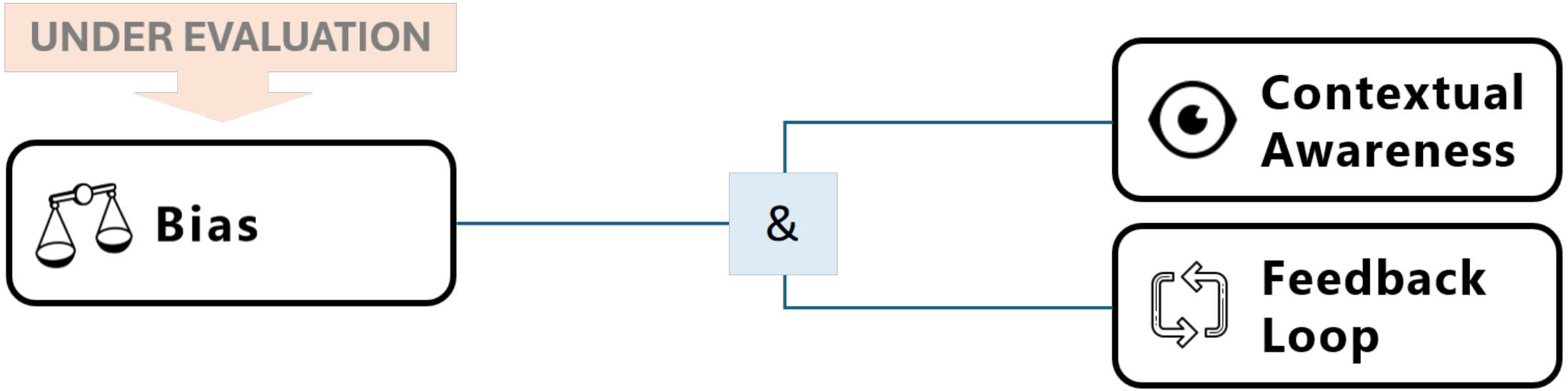
Relationships 5

Trade-offs



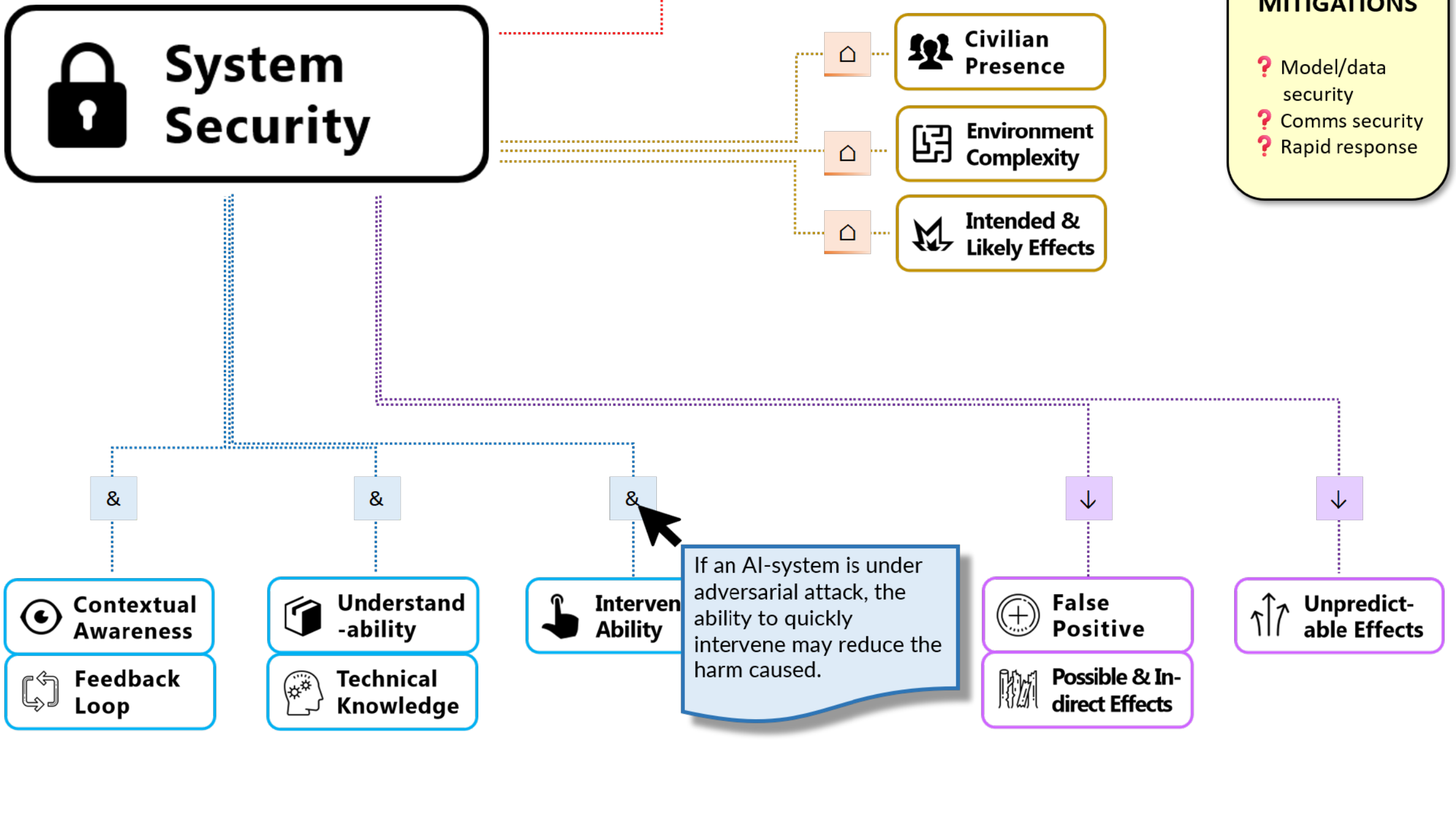
Relationships 6

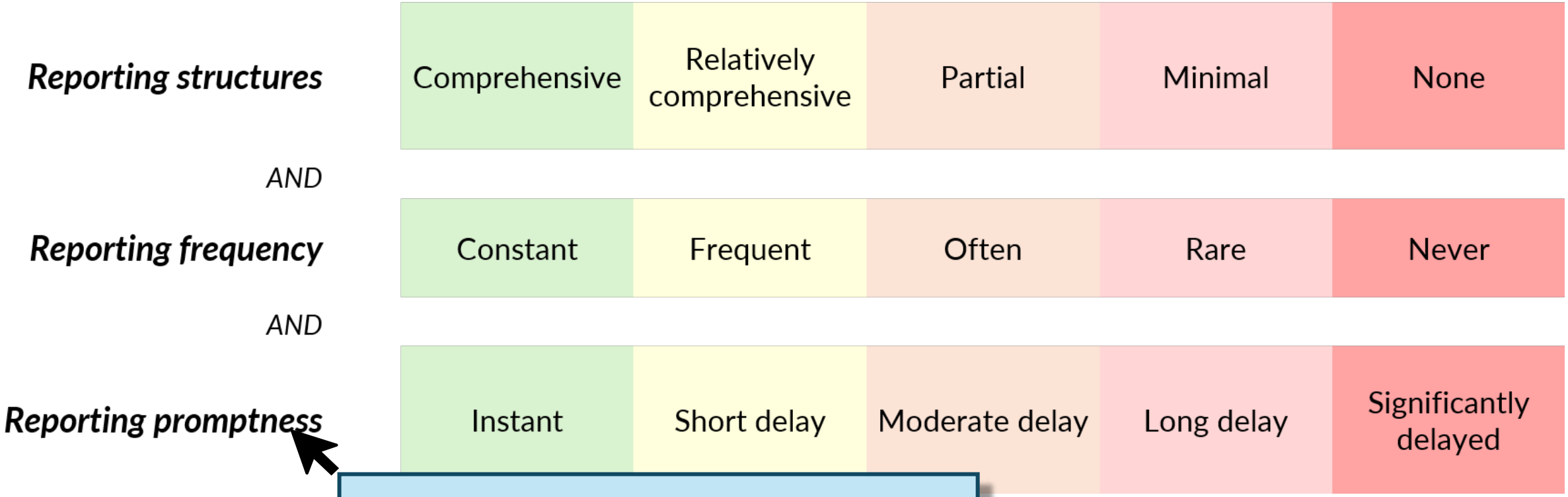
Complementary variables



You can also access via

extranet.asser.nl/elsalab/SEC.aspx


System Security
Civilian Presence
Environment Complexity
Intended & Likely Effects
MITIGATIONS
Model/data security
Comms security
Rapid response
&
&
&
Contextual Awareness
Feedback Loop
Understand -ability
Technical Knowledge
Interven Ability
If an AI-system is under adversarial attack, the ability to quickly intervene may reduce the harm caused.
False Positive
Possible & In-direct Effects
Unpredict-able Effects


Reporting structures
Comprehensive
Relatively comprehensive
Partial
Minimal
None
AND
Reporting frequency
Constant
Frequent
Often
Rare
Never
AND
Reporting promptness
Instant
Short delay
Moderate delay
Long delay
Significantly delayed
How significant the delay is between the impact and users being able to notice this impact. May also depend on the technical structure of the reporting process. See ¶4.

Discussion & feedback

- Any **questions** about the toolkit?
- Any feedback you can provide?
  - **Additional features/functionalities** you think might assist in your work
  - What could **improve the usefulness** of the toolkit further?
- Would **MOD be potentially interested** in using this?
  - Launch open access: Q4 2026
  - How can we facilitate this?

Thank you!

Dr Jonathan KWIK
ལྷ་རུ་མ་ཊཱ་ཡ་ཀྭི་ཀ་ཡ་ཀྭི། 郭漢崇

https://jonathankwik.com/

j.h.c.kwik@gmail.com