SYSTEMS ENGINEERING RESEARCH CENTER

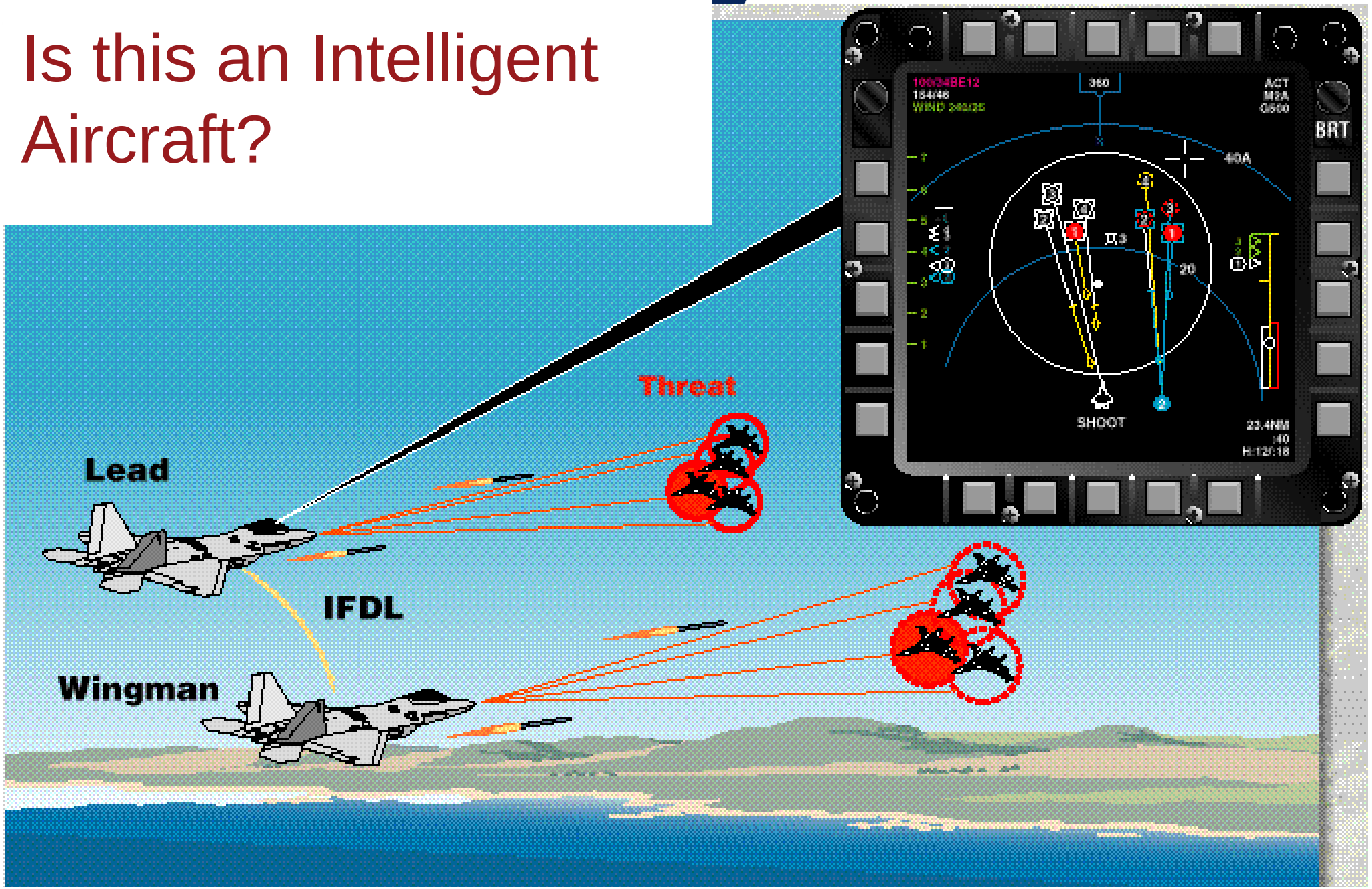
AIRC ACQUISITION INNOVATION RESEARCH CENTER

INCOSE

A Systems Engineering Process to address Trust in Computing Systems and AI

Tom McDermott (Stevens Institute of Technology)

Is this an Intelligent Aircraft?



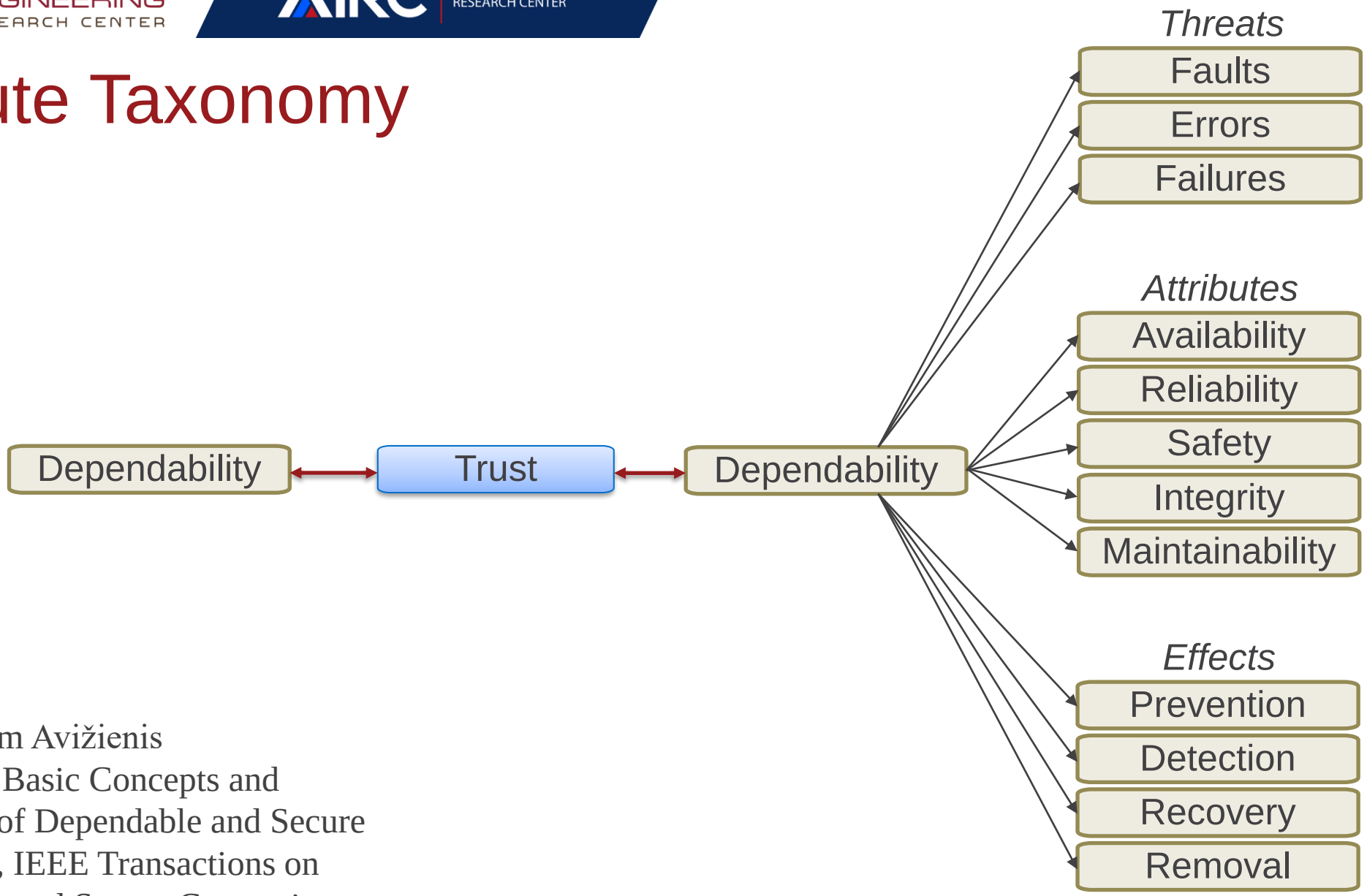
Contextual Trust

Engineered Trust

Concept of Trust in computing systems

- In computing systems, there is a defined relationship between dependability and trust. This relationship is defined by the dependence of one system on another, and the acceptance that the other system is also **dependable**. [Avižienis, et al, 2004].
- This dependence can be either human/machine or machine/machine.

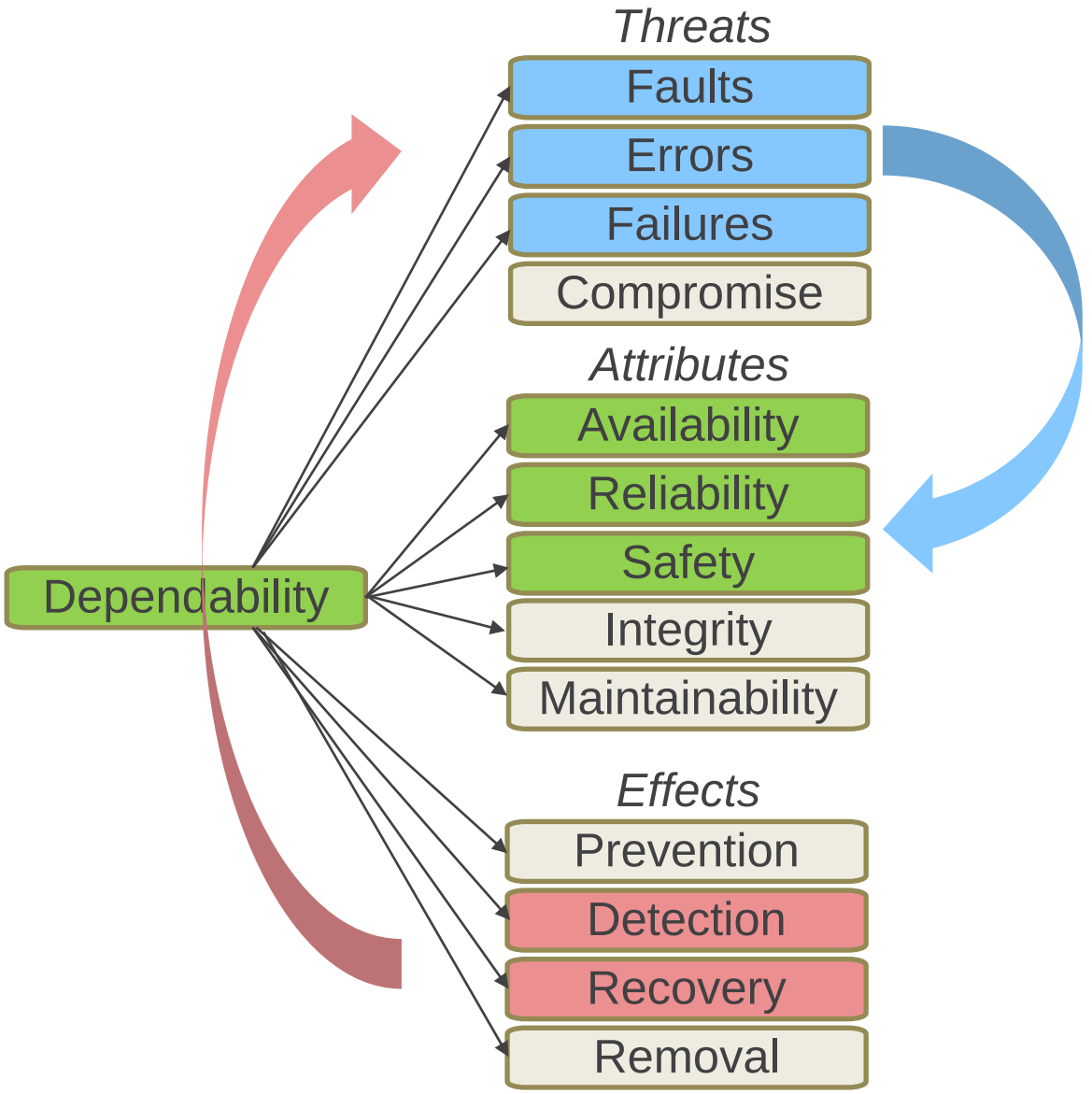
Attribute Taxonomy



adapted from Avižienis et al, 2004. Basic Concepts and Taxonomy of Dependable and Secure Computing, IEEE Transactions on Dependable and Secure Computing

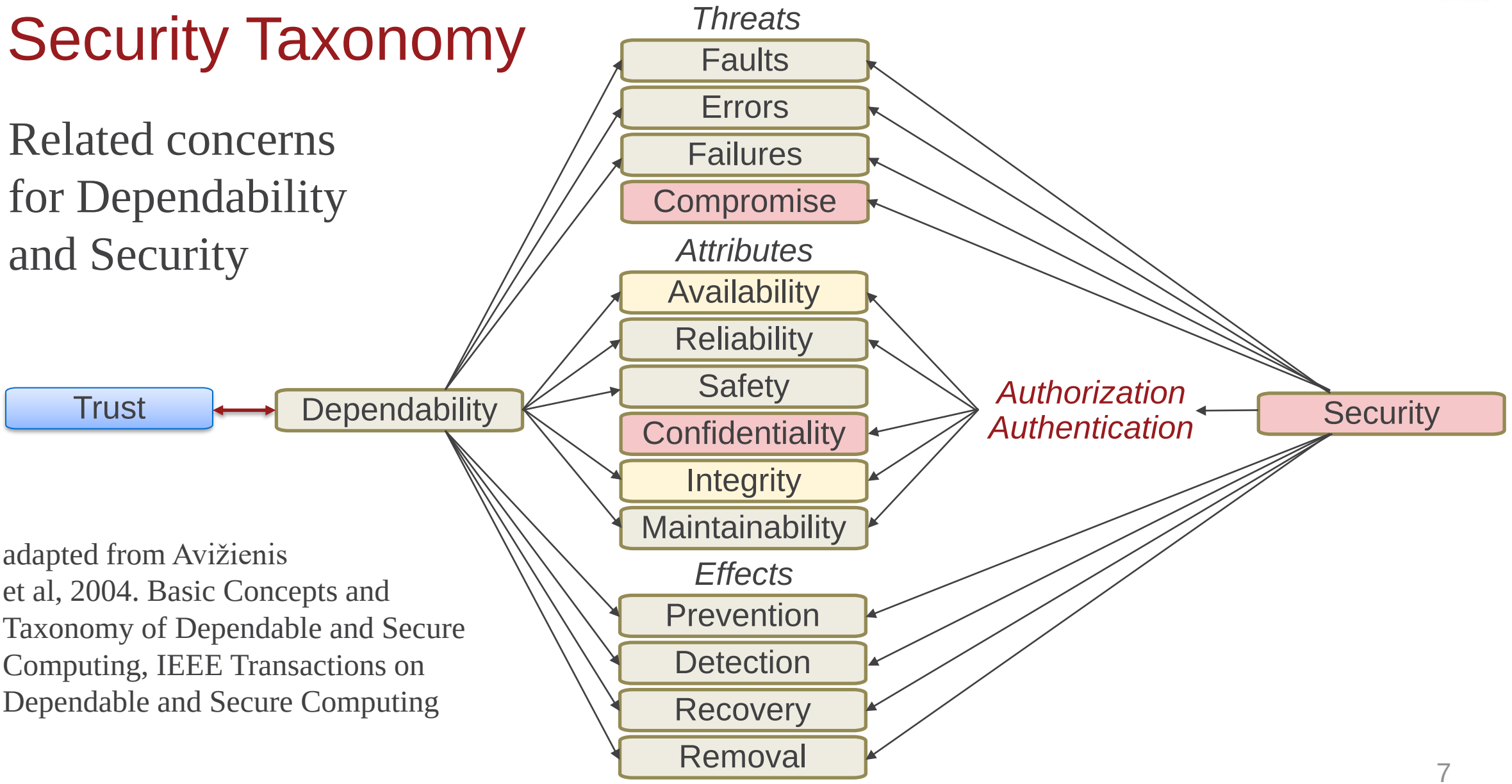
Taxonomy

Ability to *detect* and *recover* from *faults*, *errors*, and *failures* that disrupt elements or control functions to maintain levels of performance or to maintain *reliability*, *safety*, or *availability*



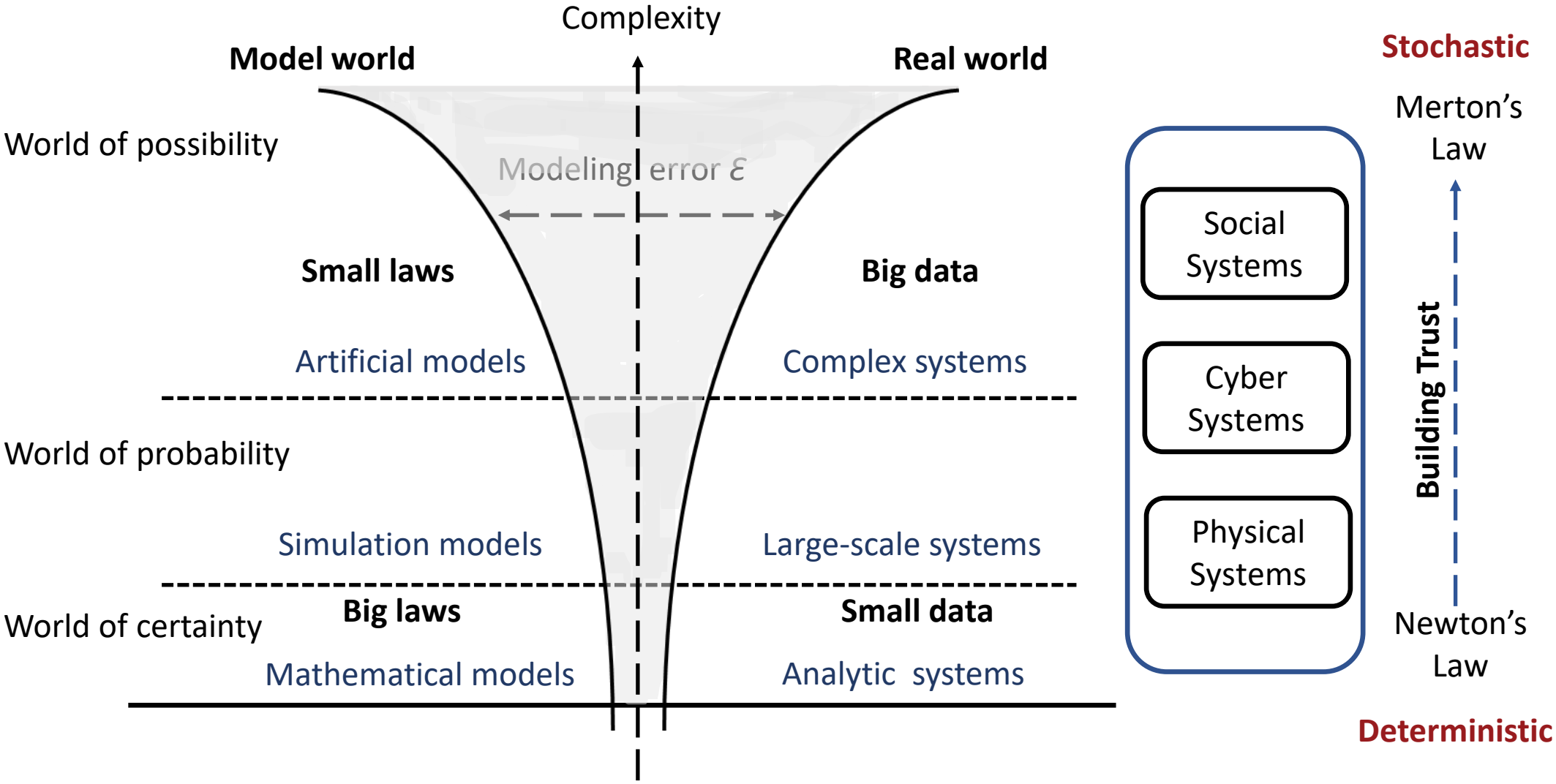
Security Taxonomy

Related concerns for Dependability and Security



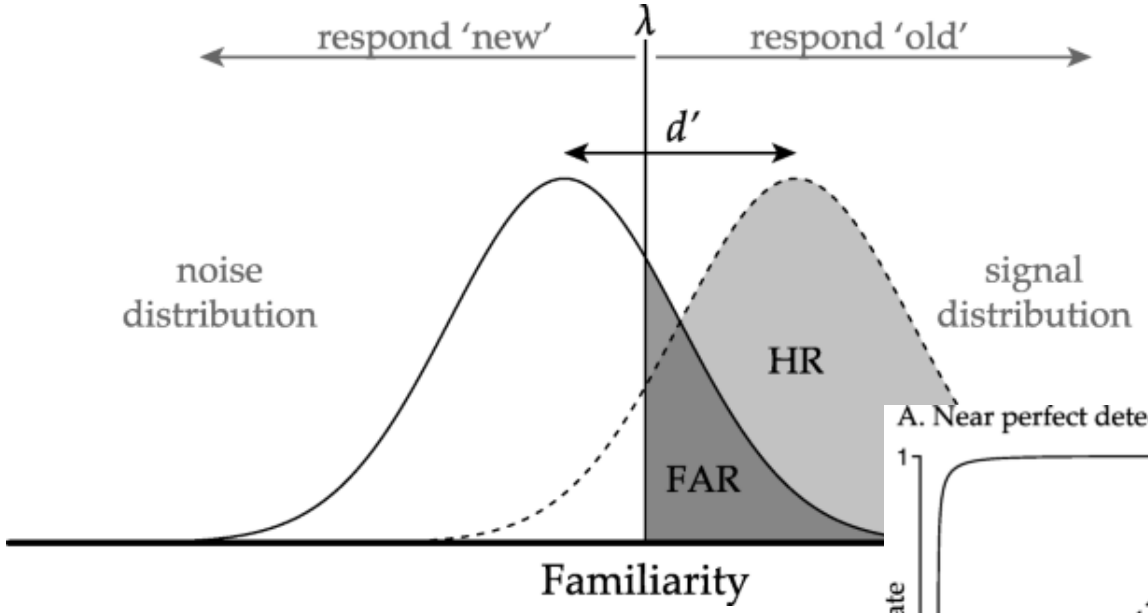
adapted from Avižienis et al, 2004. Basic Concepts and Taxonomy of Dependable and Secure Computing, IEEE Transactions on Dependable and Secure Computing

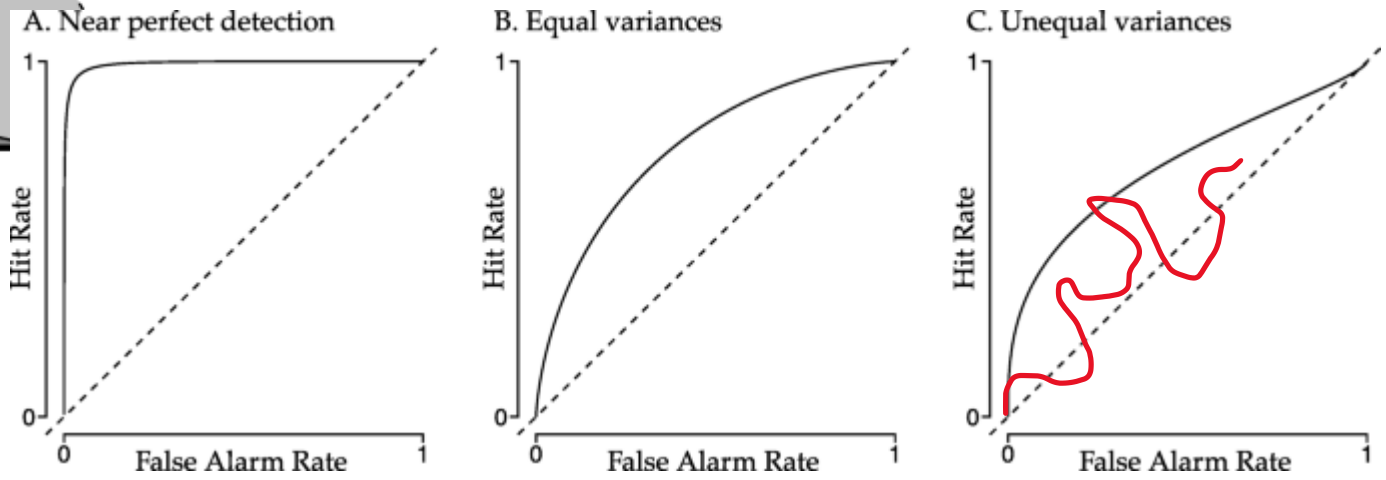
Complexity vs. Intelligence: the cognitive gap



From: Fei-Yue Wang, et al. Parallel intelligence: toward lifelong and eternal developmental AI and learning in cyber-physical-social spaces

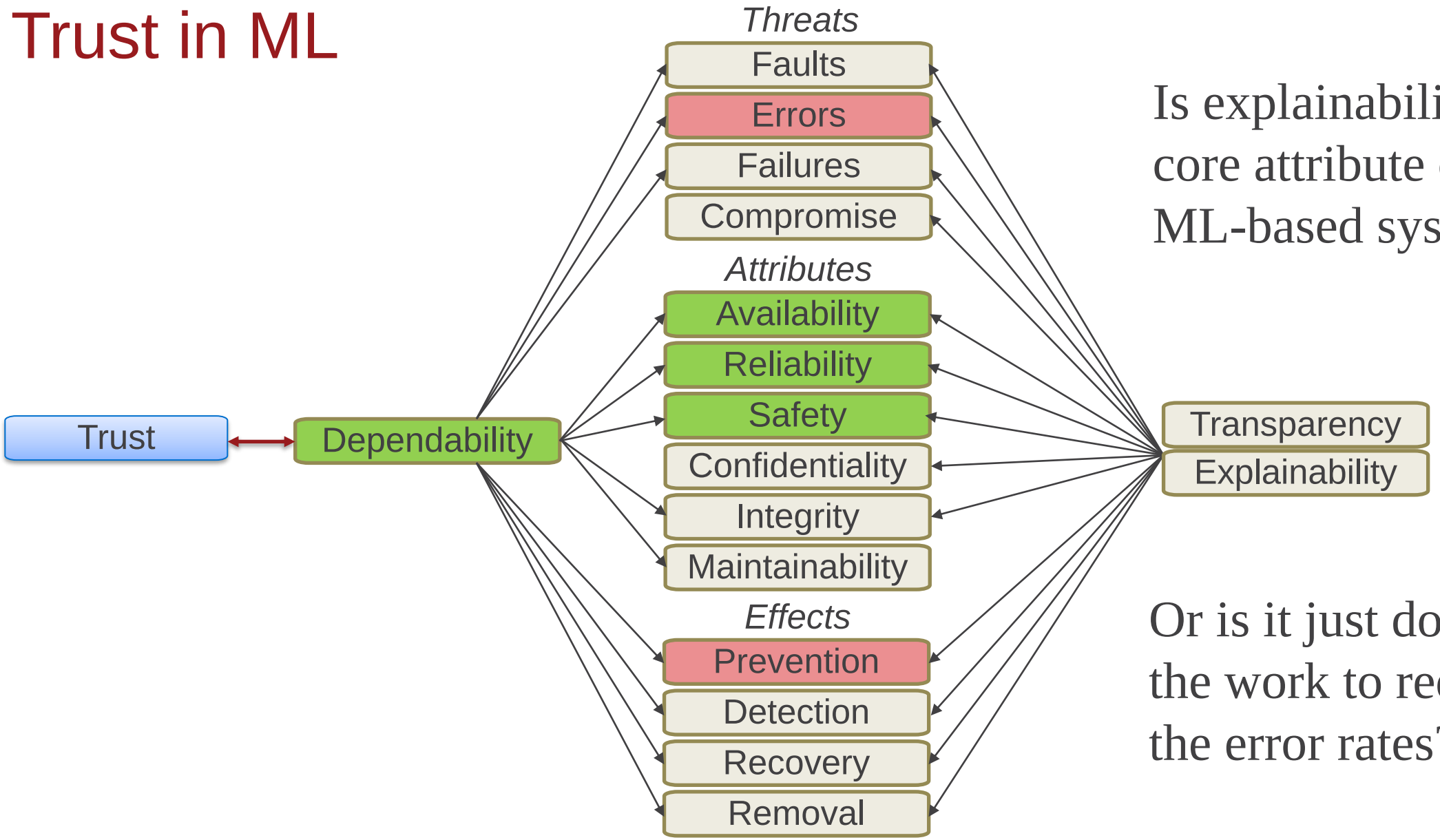
Characterizing error





Figures from: Selker, R., van den Bergh, D., Criss, A.H. *et al.* Parsimonious estimation of signal detection models from confidence ratings. *Behav Res* **51,** 1953–1967 (2019).

Trust in ML



Is explainability a core attribute of ML-based systems?

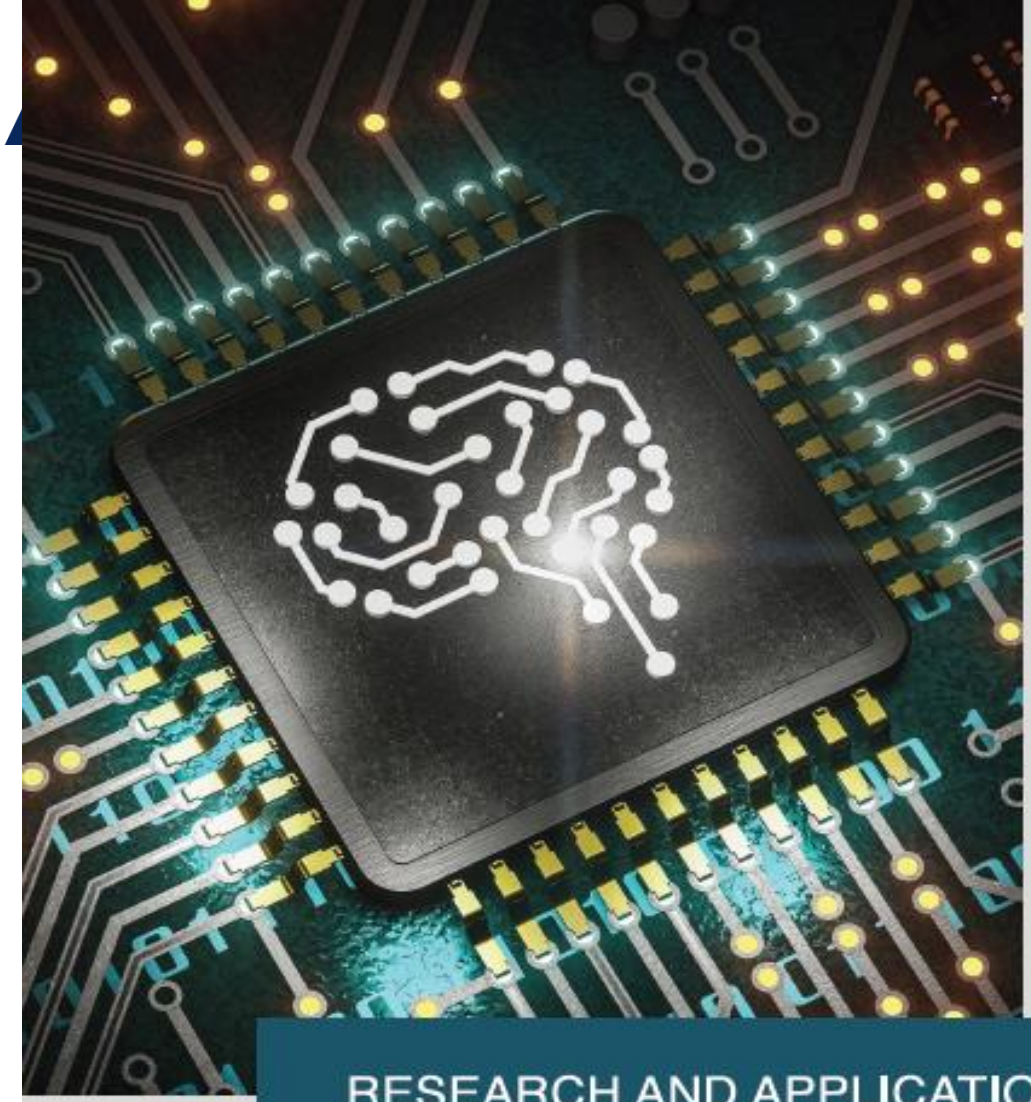
Or is it just doing the work to reduce the error rates?

Maturity of Facial Recognition Systems

Facial recognition error rates in testing conducted by the National Institute of Standards and Technology (NIST):

- In 2014, the leading algorithm had an error rate of 4.1%
- In 2018, the leading algorithm had an error rate of 0.5%
- In 2020, the leading algorithm had an error rate of 0.08%

SYSTEMS ENGINEERING RESEARCH CENTER



SAVE THE DATE

21&22 September 2022

SYSTEMS ENGINEERING RESEARCH CENTER · U.S. ARMY · DEVCOM

RESEARCH AND APPLICATION WORKSHOP

AI4SE & SE4AI

Participation limited to U.S. Citizens only

Registration and call for submissions forthcoming