The impact of AI on lifecycle processes:
a security and privacy viewpoint

Antonio Kung

CEO Trialog

25 rue du Général Foy 75008 Paris

www.trialog.com

Introduction

- Engineering background
  - Coordinator PRIPARE (pripareproject.eu) 2013-2015
- Privacy standards
  - Rapporteur Impact of AI on privacy (ISO study period)
  - Privacy engineering for system lifecycle processes (ISO/IEC 27550 editor)
  - Privacy guidelines for smart cities (ISO/IEC 27570 editor)
  - Security and privacy guidelines for IoT (ISO/IEC 27030 co-editor)
  - User-centric framework for the handling of PII based on privacy preferences (ISO/IEC 27556 co-editor)
  - Big data – Security and privacy fabric (ISO/IEC 20547-4 contributor)
  - Consumer protection -- Privacy by design for consumer goods and services (ISO 31700 contributor)
- Cybersecurity standards
  - Towards an ITS cybersecurity framework (ITU/SG17/Q13 study)
- IoT standards
  - Interoperability for IoT systems - Part 3: semantic interoperability (ISO/IEC 28123-3 co-editor)
- Others
  - FG-DPM (D4.1 Framework of Security and Privacy in Data Processing Management)
  - European Innovation Platform – Smart Cities and Communities
    - Citizen approach to data: privacy-by-design

Ecosystem oriented standards

IPEN member (ipen.trialog.com)

https://ipen.trialog.com/wiki/Wiki_for_Privacy_Standards

Page Discussion

Wiki for Privacy Standards and Privacy Projects

(Redirected from Wiki for Privacy Standards)

Contents [hide]

1 Objective of this Wiki
2 Content
3 Membership
4 More on IPEN - Internet Privacy Engineering Network
5 Sponsors and Support

Main page
Recent changes
Wiki help
Organisation
Contacts
Standardisation
CEN-CENELEC-ETSI
IEEE
IETF
ISO
ITU
OASIS
OpenId
W3C
National Level
Tools
What links here
Related changes
Upload file
Special pages
Printable version
Permanent link
Page information

Objective of this Wiki [edit]

The objective of this Wiki is to be a tool allowing stakeholders interested in privacy engineering and standardisation to find resources and to id

Content [edit]

Privacy standards	Privacy engineering projects	Reports, Events, Pr
• CEN-CENELEC-ETSI	• APP Pets (ULD project)	• DPIA and PIA guidelines
• IETF Activities	• AN.ON-Next (ULD project)	• Studies
• IEEE standards	• CREDENTIAL (EC project completed)	• OWASP
• ISO/IEC	• DNT Guide	• Business Process Cookboo
• ITU standards	• PARIS (EC project completed)	• Events
• OASIS	• PDP4E (EC project on-going)	• Presentations
• OpenID Foundation	• PRIPARE (EC project completed)	
• W3C Activities	• PRISMACLOUD (EC project completed)	
• National Level Standards	• Privacy framework (NIST project on-going)	
	• Privacypatterns	
	• Signatu	

More info on privacy standards [Expand]
More info on privacy engineering projects. [Expand]
More info on reports, events, presentations [Expand]

https://ipen.trialog.com/wiki/ISO

ISO IEC

Contents [hide]

1 Introduction
2 Some conventions on ISO standards
3 Meetings
4 Standards and Projects
- 4.1 19608 TS Guidance for developing security and privacy functional requirements based on 15408
- 4.2 20547 IS Big data reference architecture - Part 4 - Security and privacy
- 4.3 20889 IS Privacy enhancing de-identification techniques
- 4.4 27018 IS Code of practice for protection of PII in public clouds acting as PII processors
- 4.5 27030 IS Security and Privacy for the Internet of Things
- 4.6 27045 IS Big Data Security and Privacy - Processes
- 4.7 27550 TR Privacy engineering for system lifecycle processes
- 4.8 27551 IS Requirements for attribute-based unlinkable entity authentication
- 4.9 27552 IS Extension to ISO/IEC 27001 privacy management - Requirements
- 4.10 27555 IS Establishing a PII delection concept in organisations
- 4.11 27556 IS User-centric framework for the handling of personally identifiable information (PII) based on privacy preferences
- 4.12 27570 TS Privacy Guidelines for Smart Cities
- 4.13 29100 IS Privacy framework
- 4.14 29101 IS Privacy architecture framework
- 4.15 29134 IS Guidelines for Privacy impact assessment
- 4.16 29151 IS Code of Practice for PII Protection (also a ITU document - ITU-T X.1058)
- 4.17 29184 IS Online privacy notices and consent
- 4.18 29190 IS Privacy capability assessment model
- 4.19 29191 IS Requirements for partially anonymous, partially unlinkable authentication
- 4.20 31700 IS Consumer Protection - Privacy-by-design fo consumer goods and services

5 On-going Study Periods
- 5.1 Privacy consideration in practical workflows (Started in April 2018)
- 5.2 Additional Privacy-Enhancing Data De-identification standards (Started in April 2018)

Outline

- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Governance of AI-based systems
- Conclusions

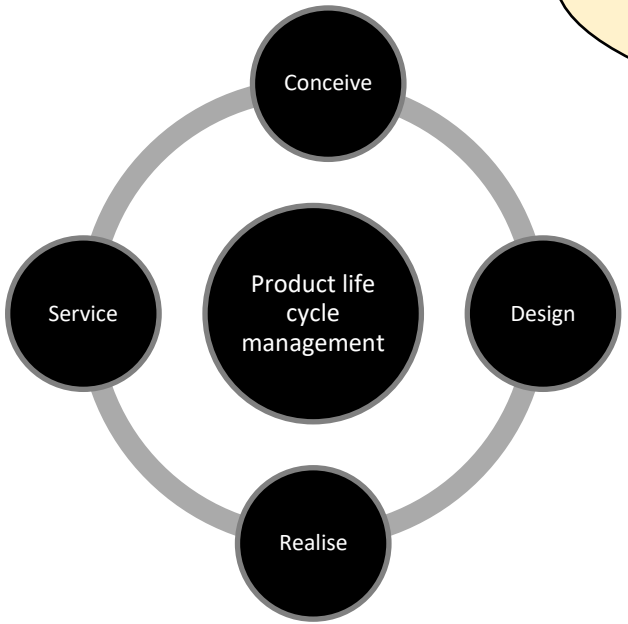
Outline

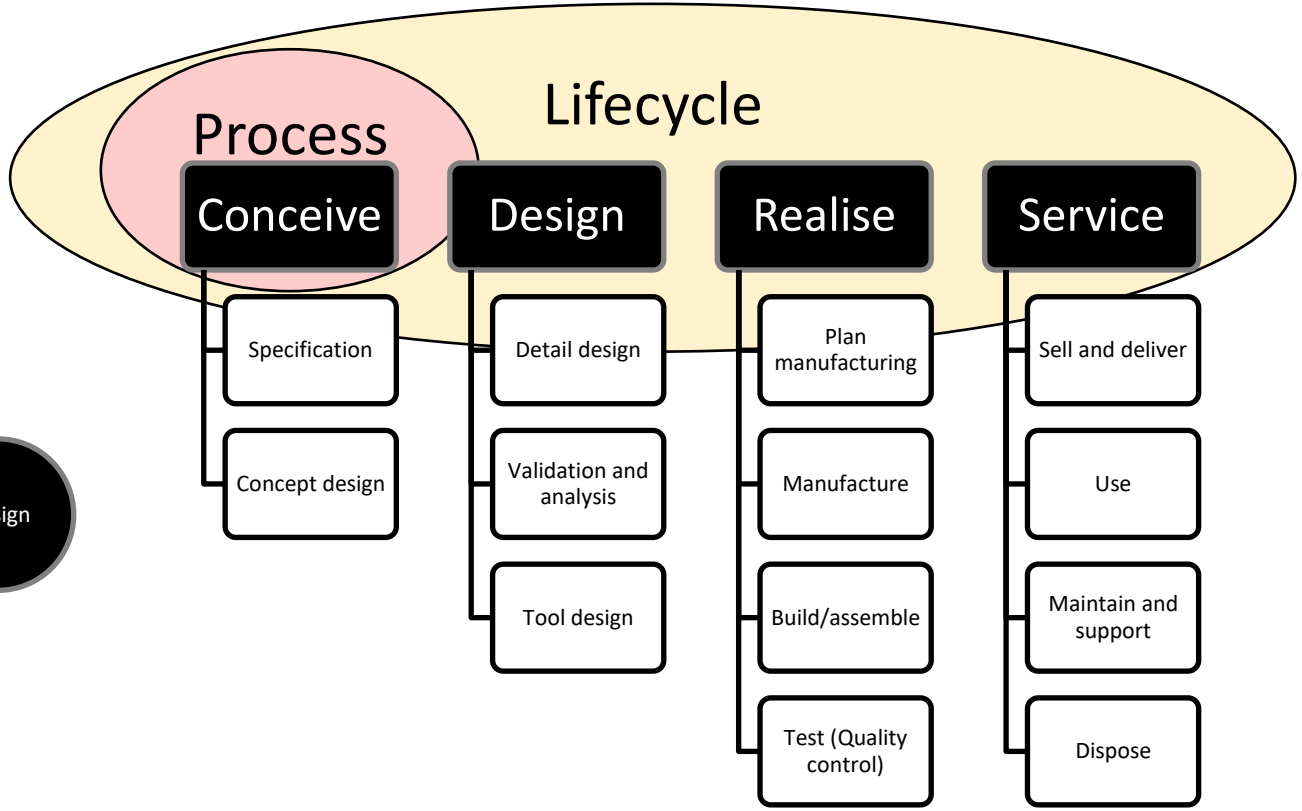
- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

Definitions

- Lifecycle
  - evolution of a system, product, service, project or other human-made entity
  - from conception through retirement
  - [ISO/IEC/IEEE 15288]
- Process
  - set of interrelated or interacting activities
  - use inputs to deliver an intended result
  - [ISO 9000]

Example: Product Lifecycle Management (PLM)





Life Cycle Processes (ISO/IEC/IEEE 15288)

Agreement

- Acquisition
- Supply

Organisational project-enabling

- Life cycle model management
- Infrastructure management
- Portfolio managment
- Human resources management
- Quality management
- **Knowledge management**

Technical management

- Project planning
- Project assessment and control
- **Decision management**
- **Risk management**
- Configuration management
- Information management
- Measurement
- Quality assurance

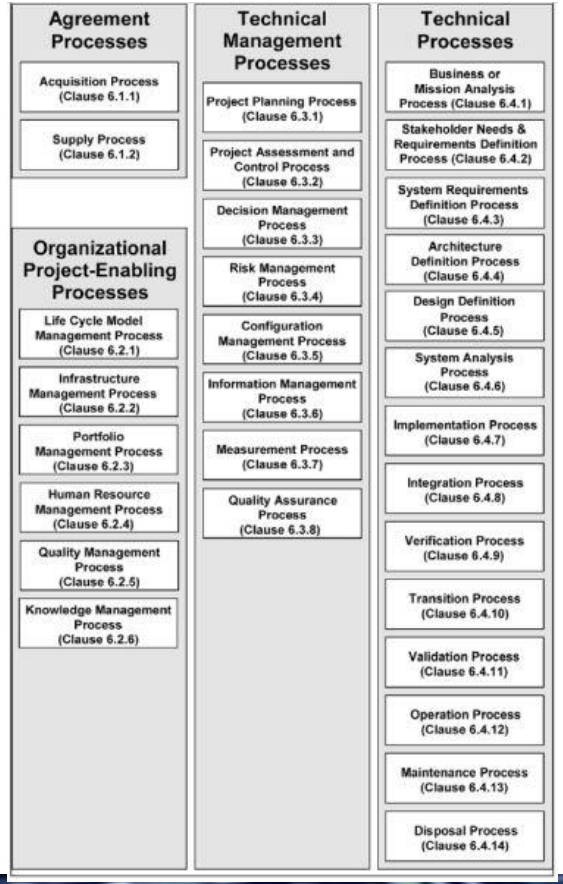
Technical

- Business of mission analysis
- Needs & requirements
- System requirements
- Architecture definition
- Design definition
- System analysis
- Implementation
- Integration
- Verification
- Transition
- Validation
- Operation
- Maintenance
- Disposal

Outline

- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

AI to Assist System Lifecycle Processes

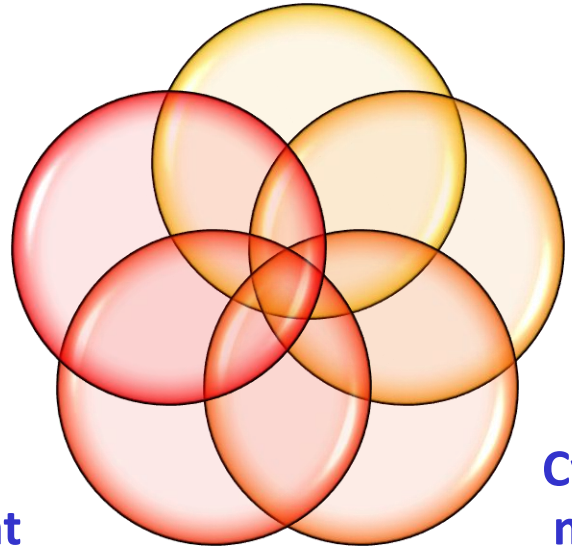


Process	AI support
Agreement	AI-assisted data sharing agreement
Organisational	**AI assisted decision making** **AI assisted knowledge management**
Technical management	**AI assisted risk analysis** AI assisted compliance
Technical process	AI-assisted risk analysis AI-assisted design AI-assisted verification AI assisted operation AI assisted maintenance

Other Lifecycles

Product lifecycle

Privacy management lifecycle



Big data management lifecycle

Risk management lifecycle

Cyber security management lifecycle

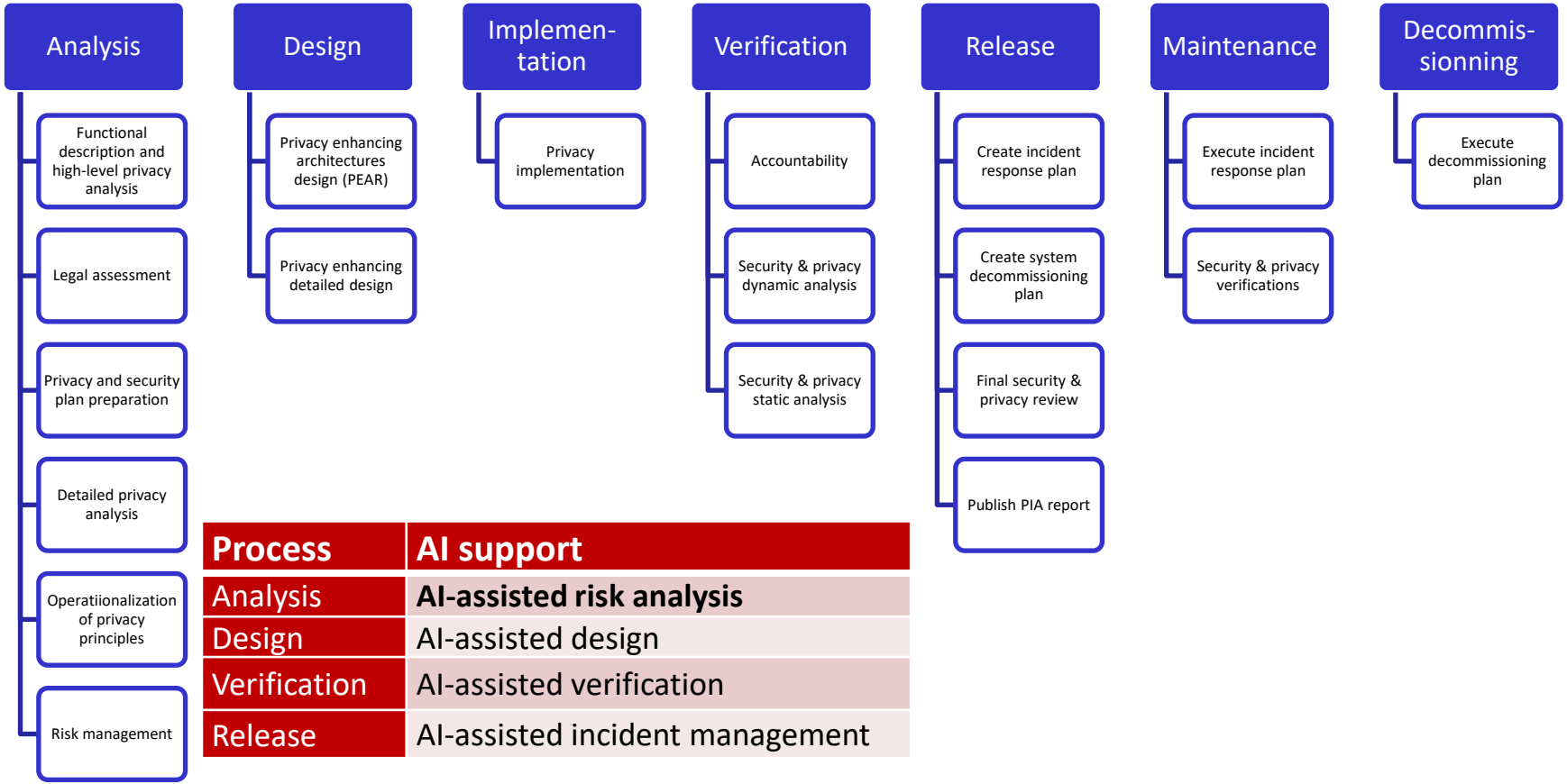
Big Data Management Lifecycle

Process	AI support
Ingestion	Assisted collection
Preprocessing	Assisted cleansing Assisted curation
Analysis	**Machine learning, Deep learning**
Storage	Assisted selection of storage scheme
Destruction / removal	

Cybersecurity Lifecycle (ISO/IEC 27101 – NIST)

Process	AI support
Identify	**AI assisted risk analysis**
Protect	Pattern recognition for the design of security and privacy controls
Detect	**Anomaly detection** - off-line analysis - on-line detection
Respond	**Assisting and training operators**
Recover	**Autonomous decision taking?**

Privacy Management Lifecycle (PRIPARE)



Process	AI support
Analysis	**AI-assisted risk analysis**
Design	AI-assisted design
Verification	AI-assisted verification
Release	AI-assisted incident management

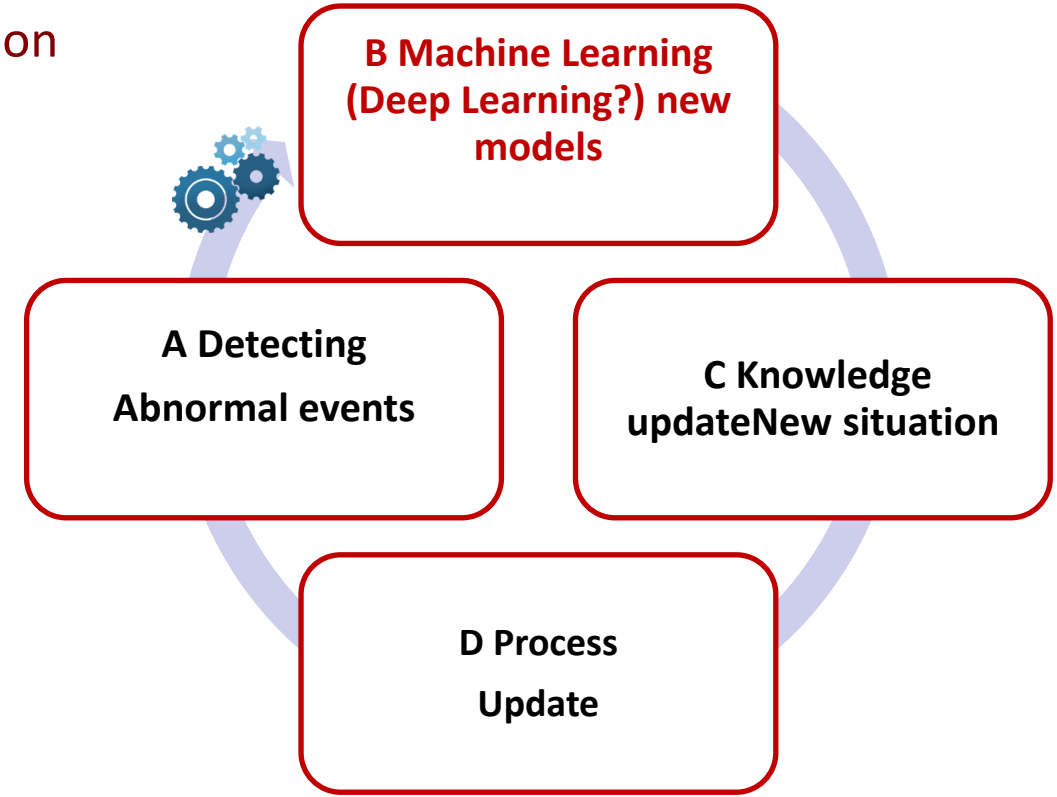
Information security risk management (ISO/IEC 27005)

- Establish context
  - Risk criteria
  - Scope and boundaries
- Risk Identification
  - Assets
  - Threats
  - Existing controls
  - Vulnerabilities
  - Consequences
- Risk analysis
  - Assessing consequences
  - Assessing incident likelihoods
  - Determining level of risks
- Risk evaluation
  - Prioritization
- Risk treatment
  - Risk modification
  - Risk retention
  - Risk avoidance
  - Risk sharing
- Risk monitoring and review
  - Risk factors
  - Risk management process

Process	AI support
Risk Analysis	**AI assisted risk analysis**

Continuous improvement

- Example: Cybersecurity Situation Awareness Learning

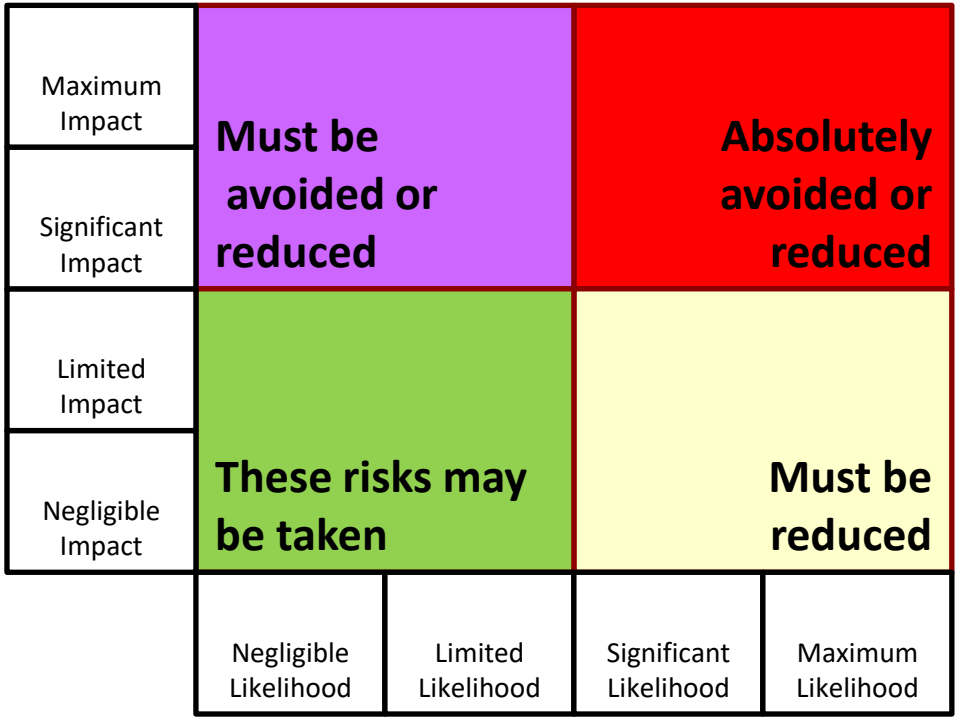


Outline

- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

AI Assisted Risk Analysis Using Risk Maps

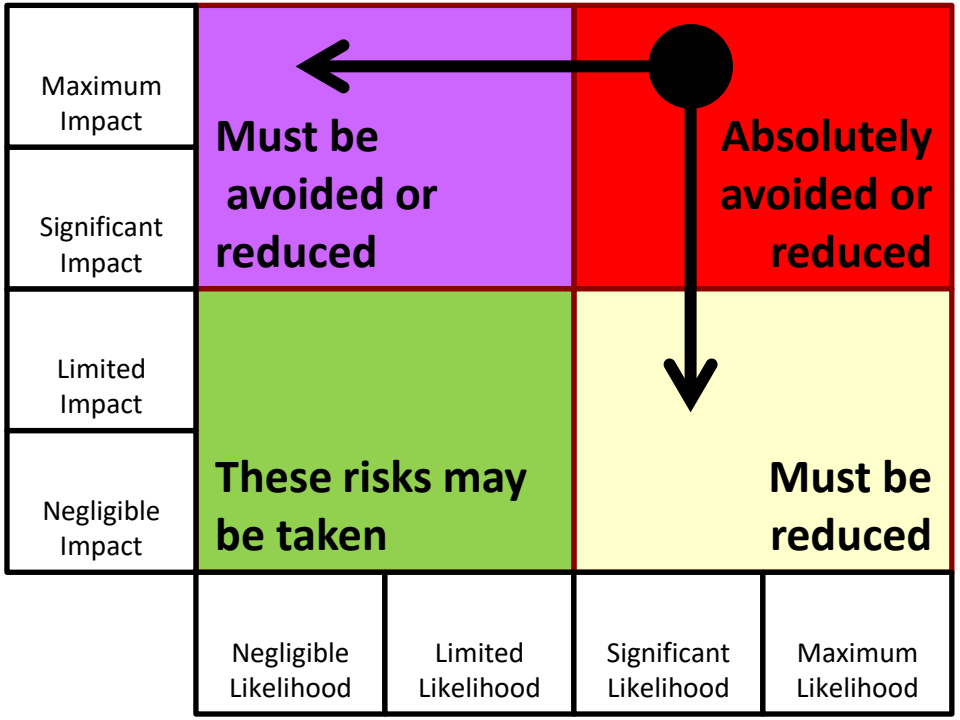
- Security and privacy threat/breach risk level:
  - Likelihood
  - Impact
- Many versions of risk maps
  - More levels
  - Different ways of calculating. Exemples
    - NIST privacy engineering
    - ETSI TVRA
  - This map is from CNIL guidelines



AI to Assist Risk Analysis

- Assistance to avoid attacks (reduce likelihood of threats)

- Assistance to breaches (reduce severity of impact)



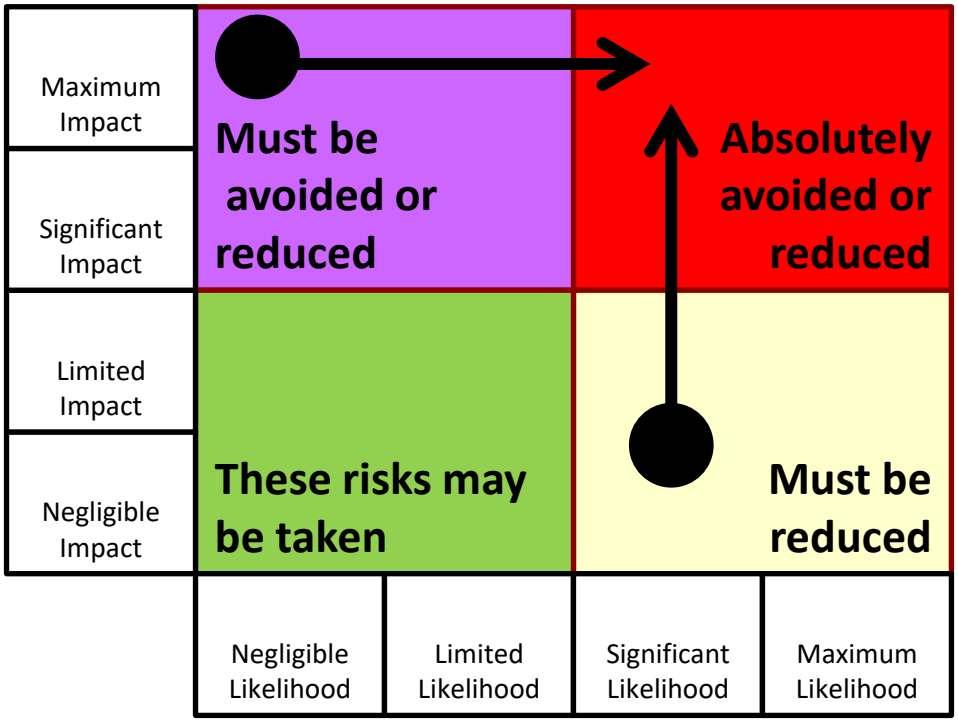
Outline

- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

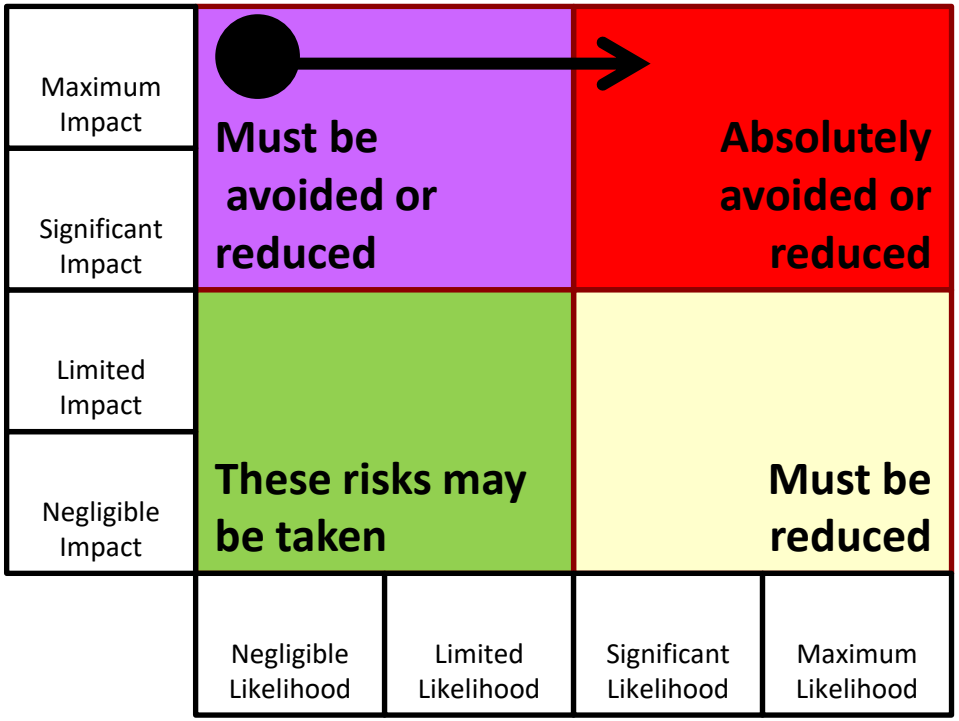
AI to Break Cybersecurity

- security incident / privacy breach is more likely to occur

- Security incident / privacy breach has more impact



AI to break cybersecurity



(TVRA) Threat Vulnerability Risk Analysis		
Attack factor		**Malicious AI assistance**
Time	<= 1 day <= 1 week <= 1 month <= 3 months <= 6 months > 6 months	AI attack creation assistant
Expertise	Layman Proficient Expert	
Knowledge	Public Restricted Sensitive Critical	AI based learning of vulnerabilities
Opportunity	Unnecessary Easy Moderate Difficult Nont	AI based creation of opportunities
Equipment	Standard Specialised Bespoke	Lower cost
Asset Impact	Low Medium High	AI analysis of impact
Intensity	Single intensity Moderate intensity High intensity	AI based swarm attack

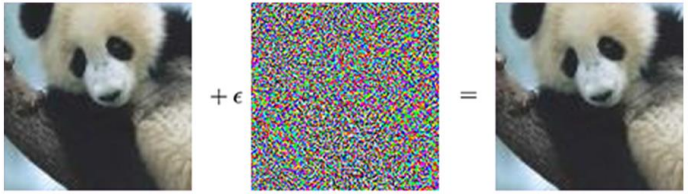
Malicious AI: Enhancing threats / New threats

- Expansion of existing threats
  - Expanding phishing
  - Increasing willingness to carry out attacks
    - increasing anonymity and increasing psychological distance
  - Robotics progress
- Introduction of new threats
  - Mimicking voice
  - New AI capabilities imply new threats
    - Autonomous cars VS image of a stop sign changed
    - Swarm of autonomous systems VS attack on a server to control the swarm

Data Poisoning

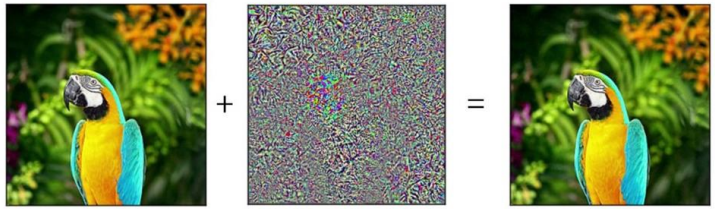
Courtesy Ivo Emanuilov (KUL – citip – Imec)

Adversarial examples and data poisoning..



"panda"
57.7% confidence

"gibbon"
99.3% confidence



X
97.3% macaw

sign($\nabla_X J(\theta, X, Y)$)

X + ϵ.sign($\nabla_X J(\theta, X, Y)$)
88.9% bookcase

3

KU LEUVEN CITIP CENTRE FOR IT & IP LAW

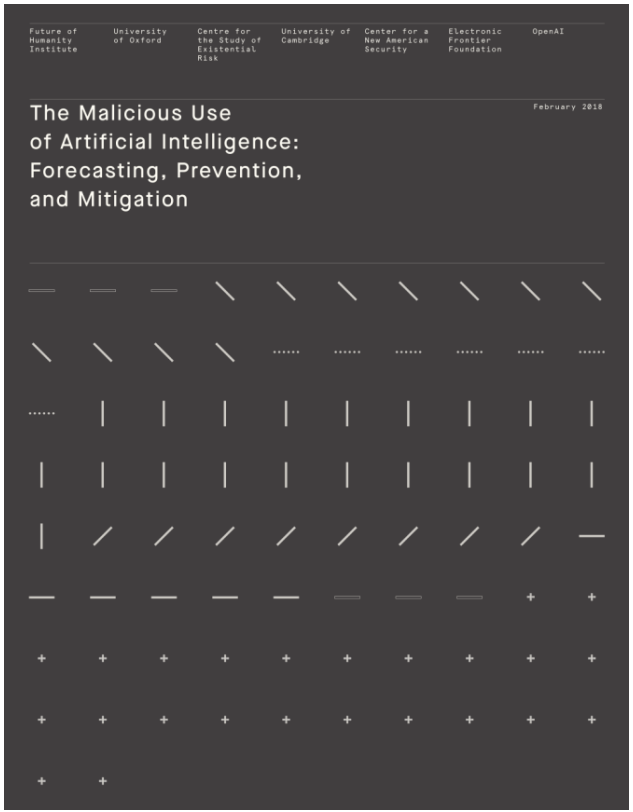
KU LEUVEN CITIP CENTRE FOR IT & IP LAW imec

Poisoned AI: towards data liability?

Ivo Emanuilov

LICT Workshop Autonomous Systems, 31 May 2018

Malicious AI



Miles Brundage[1] Shahar Avin[2] Jack Clark[3] Helen Toner[4] Peter Eckersley[5] Ben Garfinkel[6] Allan Dafoe[7] Paul Scharre[8] Thomas Zeitzoff[9] Bobby Filar[10] Hyrum Anderson[11] Heather Roff[12] Gregory C. Allen[13] Jacob Steinhardt[14] Carrick Flynn[15] Seán Ó hÉigeartaigh[16] Simon Beard[17] Haydn Belfield[18] Sebastian Farquhar[19] Clare Lyle[20] Rebecca Crootof[21] Owain Evans[22] Michael Page[23] Joanna Bryson [24] Roman Yampolskiy[25] Dario Amodei[26]

1 Corresponding author miles.brundage@philosophy.ox.ac.uk Future of Humanity Institute, University of Oxford; Arizona State University

2 Corresponding author, sa478@cam.ac.uk Centre for the Study of Existential Risk, University of Cambridge

3 OpenAI

4 Open Philanthropy Project

5 Electronic Frontier Foundation

6 Future of Humanity Institute, University of Oxford

7 Future of Humanity Institute, University of Oxford; Yale University

8 Center for a New American Security

9 American University

10 Endgame

11 Endgame

12 University of Oxford/ Arizona State University/New America Foundation

13 Center for a New American Security

14 Stanford University

15 Future of Humanity Institute, University of Oxford

16 Centre for the Study of Existential Risk and Centre for the Future of Intelligence, University of Cambridge

17 Centre for the Study of Existential Risk, University of Cambridge

18 Centre for the Study of Existential Risk, University of Cambridge

19 Future of Humanity Institute, University of Oxford

20 Future of Humanity Institute, University of Oxford

21 Information Society Project, Yale University

22 Future of Humanity Institute, University of Oxford

23 OpenAI

24 University of Bath

25 University of Louisville

26 OpenAI

Authors are listed in order of contribution

Design Direction by Sankalp Bhatnagar and Talia Cotton

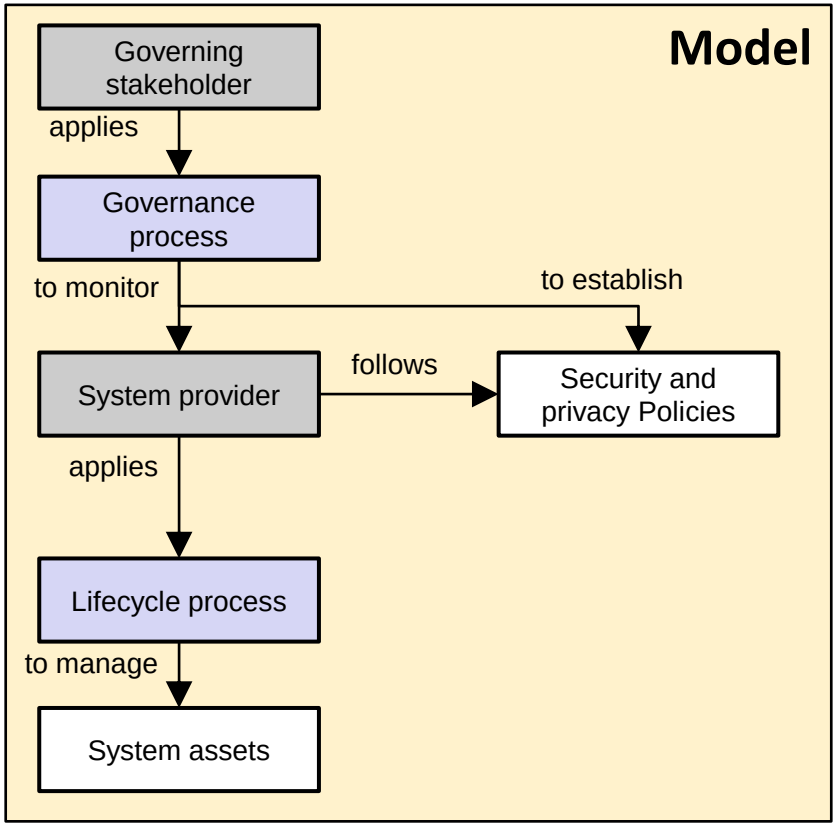
Outline

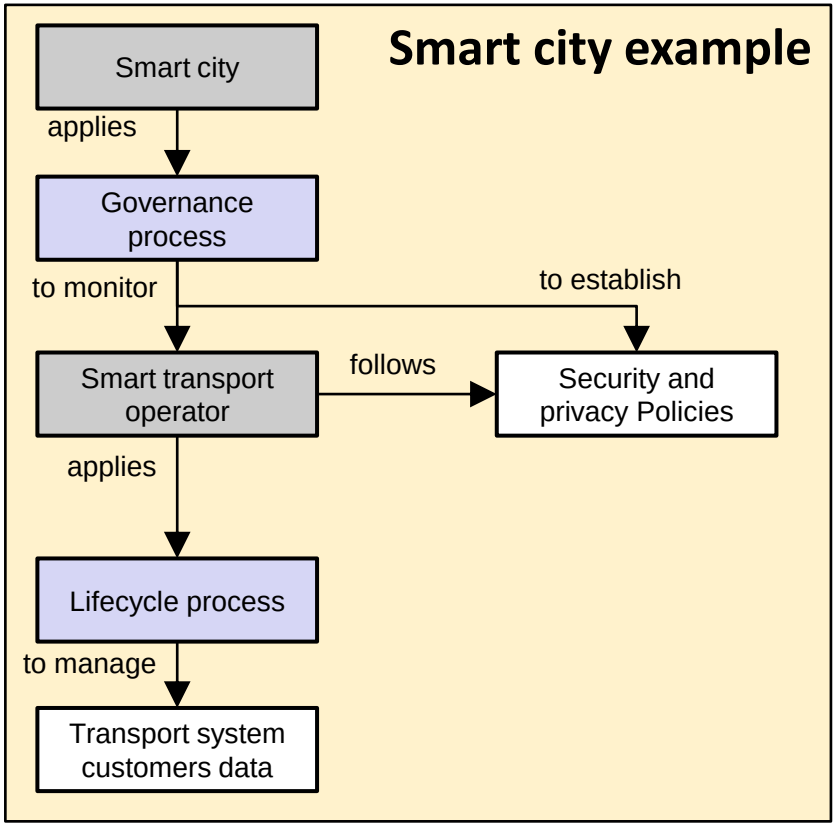
- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

AI based applications

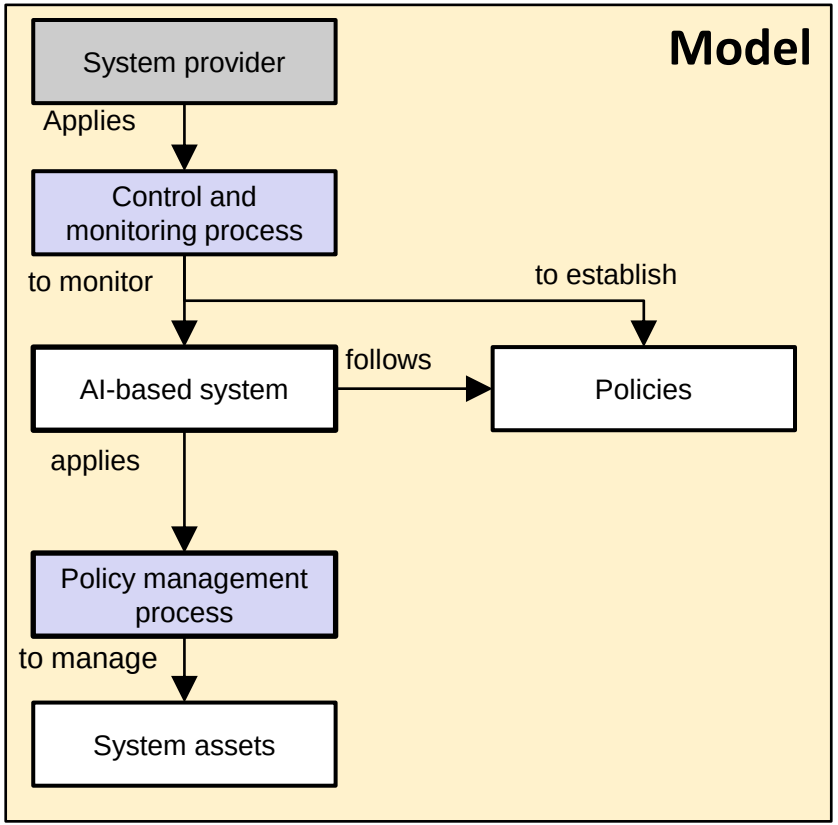
- Automatic speech recognition, machine translation, spam filters, and search engines
- Autonomous cars, Robots for elderly people, Autonomous drones
  - Controlled systems

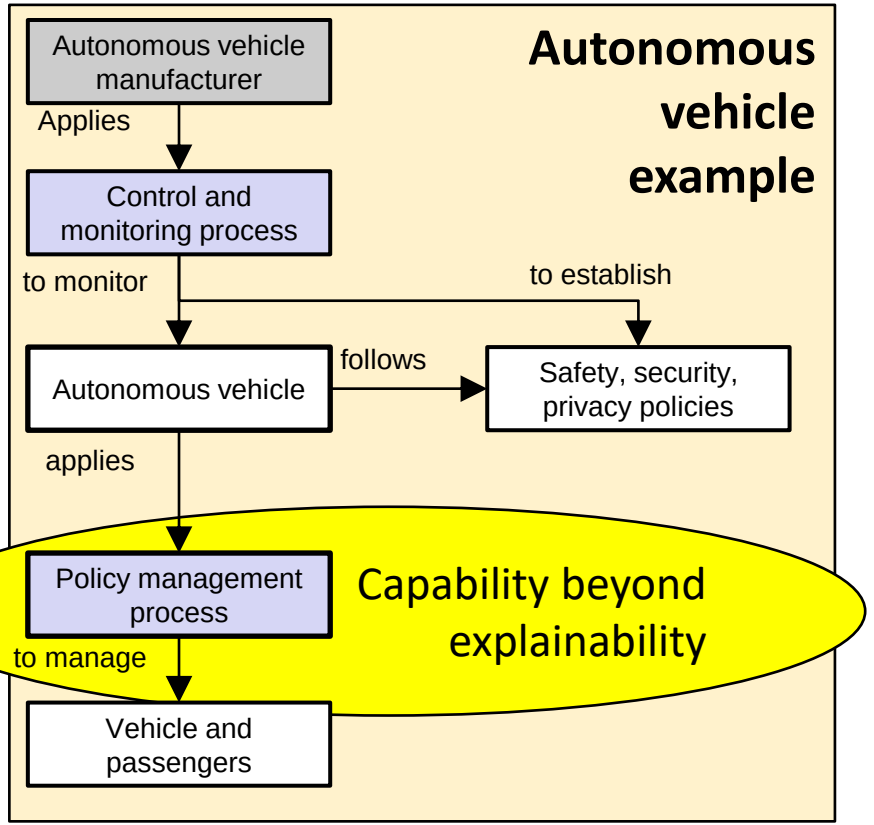
Security and Privacy Governance Model





Security and Privacy Governance Model for AI?





Outline

- Lifecycle processes
- AI assistance for lifecycle processes
- AI assistance for security and privacy risk analysis
- AI for malicious AI
- Security and privacy governance of AI-based systems
- Conclusions

Conclusions

- AI will improve lifecycle processes
- AI will improve security and privacy risk management
- Malicious AI will increase security and privacy risks
- Security and Privacy Governance Model for AI?

www.trialog.com

Questions?