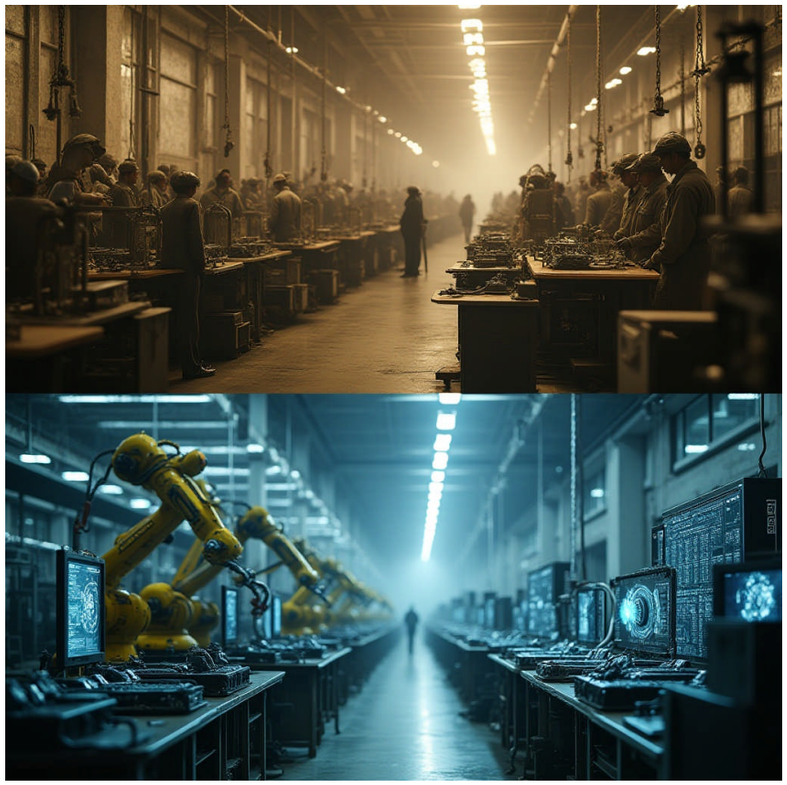
Chair Introduction

LISBON
2025

Welcome to Panel #4:

Cyber Resilience - Pre-AI vs. AI-Enhanced Approaches



Dr. Steve Chan
VT, VTIRL/DE-CAIR,
USA

Chair Introduction

LISBON 2025

Moderator:

- Dr. Steve Chan, VTIRL, VT, USA

Panelists:

- Pr. Res. Masahito Kumazaki, National Institute of Informatics, Japan
- Dr. Damjan Fujs, University of Ljubljana, Slovenia
- Dr. Kathy-Ann Fletcher, Abertay University, UK
- Dr. Gábor György Gulyás, Vitarex Stúdió Ltd., Hungary
- Dr. Steve Chan, VTIRL, VT, USA

Dr. Steve Chan
VT, VTIRL/DE-CAIR, USA

Chair Introduction

LISBON 2025

Pre-AI:

- Cybersecurity & Cyber-Resiliency -> Static & Reactive - Kathy-Ann Fletcher
- Reaction - Dr. Gábor Gulyás
- Transformations in Higher Education - Masahito Kumazaki; Damjan Fujs
- Threat Intelligence -> Human-Curated - Dr. Gábor Gulyás
- Incident Response -> Playbooks - Dr. Gábor Gulyás

AI Enhanced:

- Cybersecurity & Cyber-Resiliency -> Dynamic & Proactive - Kathy-Ann Fletcher
- Cyber/Cyber-Resiliency -> Manipulative AI Content - Damjan Fujs
- Cyber/Cyber-Resiliency -> AI-facilitated threats/exploitation - Masahito Kumazaki; Damjan Fujs
- Prediction - Dr. Gábor Gulyás
- Threat Intelligence -> Machine-Augmented - Dr. Gábor Gulyás
- Incident Response -> Adaptive Defense - Masahito Kumazaki; Dr. Gábor Gulyás
- Bias, Explainability, Trust, etc. - Dr. Gábor Gulyás; Dr. Steve Chan
- AI Reliability/AI Hallucinations/AI Coherence - Masahito Kumazaki; Dr. Steve Chan

Dr. Steve Chan
VT, VTIRL/DE-CAIR, USA

Panelist Position

LISBON 2025

Masahito Kumazaki
National Institute of Informatics
Japan

- **Effective use of AI:**
  - Education: Creating educational materials, applying to problem-solving,
  - Incident response: Automated resolution via playbooks, detecting zero-day and advanced attacks
- **AI causes various problem**
  - Exploitation by attackers: Enabling advanced attacks
    ->More sophisticated phishing, rapid growth in attack techniques
  - AI Reliability: Correctness of AI Responses
    ->Scope of AI Use and Mission-Criticality
- **future with AI adoption**
  - Defenders and attackers: rapid growth through AI utilization
    ->*Can the general user grow at the same pace?*
    - AI utilization in education and personal security technologies

Panelist Position

LISBON 2025

- Cyber-resilience (important for personal, organizational and citizenship holistic well-being). My research focuses on resilience re: digital marketing effectiveness often exposing all previously named stakeholders to harm.
- Ensuring that marketing operations are secure, trusted and uninterrupted and stakeholders are protected is/should be a fundamental responsibility of any organization seeking to practice digital marketing.
- Farming Cybersecurity off to third parties without taking the steps to learn even the most basic steps in cyber security is negligent in the face of highly publicized breaches of businesses that have happened and threatens the resilience of the business and its stakeholders.
- Resilience protest customer trust and brand reputation – protecting from legal consequences, ensures recovery time while minimizing downtime, revenue loss and keeping the organization viable, is ethical. Be proactive and build a culture of awareness and preparedness

Kathy-Ann Fletcher, Abertay University

Panelist Position

LISBON 2025

- Pre-AI
  - Reactive Security Posture
  - Limited Threat Intelligence
  - Siloed Systems
  - Static Risk Assessments
- Implications for Digital Marketing
  - Manual Data Protection
  - Delayed breach detection
  - Static Compliance Checks
  - Limited personalisation due to risk

- AI Enhanced
  - Proactive and Predictive Defence
  - Enhanced Threat Intelligence
  - Integrated Eco-systems
  - Dynamic Risk Management
- Implications for Digital marketing
  - Automation
  - Data Encryption
  - Real-time anomaly detection in customer behaviour
  - Continuous compliance monitoring
  - Secure, AI Driven Personalisation at scale

Kathy-Ann Fletcher, Abertay University

Panelist Position

LISBON 2025

Damjan Fujs

UNIVERSITY OF LJUBLJANA | FRI Faculty of Computer and Information Science

- **Cyber Resilience - Pre-AI vs. AI-Enhanced Approaches**
  - **Transformations in Higher Education**
    - **Complex tasks -> simple tasks -> complex tasks (because of LLMs)**
    - **Understanding security (software user + software developer)**
  - **Do not forget older adults**
    - **They are more vulnerable to cyber threats (phishing, scams, misinformation)**
    - **Need for accessible training (plain language, hands-on workshops, …)**
  - **Do not forget young people (pre-school, school)**
    - **Critical thinking (spotting misinformation, scams, manipulative AI content)**
    - **Encourage responsible use of AI (chatbots, learning apps, generative tools)**
  - **Questions**
    - **What will be the long-term impact of an LLM in software engineering on cybersecurity?**
    - **What approach should be applied for users?**
    - **What approach should be applied for developers?**

Panelist Position

LISBON 2025

Dr. Gábor Gulyás, PhD
Vitarex Studio Ltd.
Managing Partner

- AI may bring lots of benefits to the table:
  - Core shift: from reaction to prediction
  - Redundancy vs. self-healing
  - Threat intelligence: human-curated vs. machine-augmented
  - Incident response: playbooks vs. adaptive defense
- But also has its risks and problems
  - AI as Double-Edged Sword (defenses vs. attacks)
  - What about adversarial attacks?
  - Bias and false-positive paradox
  - Explainability and trust issues

Panelist Position

LISBON 2025

Dr. Steve Chan
VT, VTIRL/DE-CAIR, USA

- **Mitigating against AI hallucinations/AI incoherence:**
  How suitable is human oversight/validation for addressing AI hallucinations? Putting aside certain “high-stakes”/mission-critical applications (e.g., law, healthcare, etc.), does the Human-in-the-Loop (HITL) approach scale sufficiently well or at speed?

  Do more automated approaches, such as: (1) Guardrails/System Prompts (GSP), (2) Fine-Tuned Models (FTM) for specialized domains, and (3) Real-time Retrieval Augmented Generation (RAG), show promise for mitigation or do they have inherent “glass ceilings?”

  What are some alternative approaches for addressing the AI hallucination/AI incoherence issue?

- **Threshold Tolerances for AI hallucinations/AI incoherence:**
  At what point do the reliability and trustworthiness of AI Systems (AIS), such as Conversational AI Agents, dissipate?

  At certain levels of AI hallucinations/AI coherence, should the involved AIS be restricted from use in mission-critical applications?

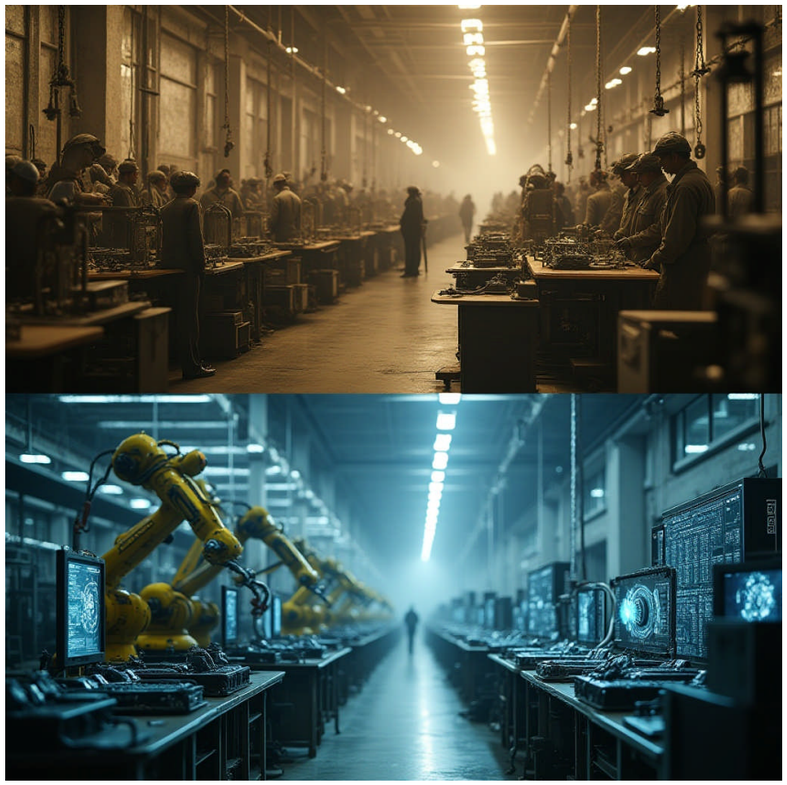
  What is the current state of counterpoising (e.g., conversational flow/fluency versus validity, between/among gaps, non-generation versus pattern—based generation)?

Discussion!

LISBON
2025

for Panel #4:

Cyber Resilience - Pre-AI vs. AI-Enhanced Approaches



Dr. Steve Chan
VT, VTIRL/DE-CAIR,
USA

Closing Remarks

LISBON
2025

Dr. Steve Chan
VT, VTIRL/DE-CAIR, USA

Thank you for attending Panel #4

Cyber Resilience - Pre-AI vs. AI-Enhanced Approaches,

and have a wonderful rest of the conference here at:

The Second International Conference on AI-based Systems and Services (AISyS 2025)