Panel on **Dependability and Security Challenges in the face of 21st Century Threats and Trends: Industry and Academic Perspectives**

Moderator: Tom Anderson

Rapporteur: Felicita Di Giandomenico

**Panellist and topics covered**

**Lorenzo Strigini -** Pervasive human-computer mingling and complex socio-technical systems

**Paulo Carvalho -** Dependability Challenges in Personal Health Solutions

**John Meyer –** Dependability Assurance Challenges posed by 21st Century Trends

**Wilfried Steiner –** Industry Perspective on Complexity-driven Challenges

**Jay Lala -** Cyber Resiliency and Survivability in Aerospace & Defense Domains

# Panellists Position Statements

- **Lorenzo:** computers are more pervasive and ***more deeply interleaved with human*** functions
  - A number of potential problems in socio-technical systems generating, , e.g., misunderstanding, or dependence or degrading ability to operate autonomously
    - Open to malicious exploitation
  - **Challenges**:
    - Competences of designers of software supports
    - Redefinition of system boundaries for which each designer/vendor/regulatory agency is responsible
    - overtrust in users while undermining their ability to oversee automation
    - Interdisciplinary understanding is crucial
    - emergent effects from pervasiveness may differ from all the above
- **Paulo**: Need of a paradigm shift in health provision from acute illness to chronic illness and personalized/precision medicine
  - **Challenges**
    - Assessment of data → need of standards for deidentification, use and protection of data
    - Certification of AI-based algorithms → different models; issues of interpretability/explainability; reliability/performance
    - Need of a Regulatory Framework for data AI/ML based software for medical devices - initiatives on-going (e.g., FDA draft)

- **John**: Focus on dependability assurance
  - increasing complexity of operating environments for highly dependable and secure systems (e.g., intelligent autonomous vehicles)
  - the notion of ”correct/failed” service provided by such AI-based “support” systems is elusive/nonexisting
  - **To address the challenges**
    - Accurate definition of environment models, to incorporate in an integrated methodologies framework for assuring targeted dependability requirements
    - Definition of novel evaluation measures for AI systems in isolation
    - For AI-enabled systems determine means of inferring possible AI contributions to system failures

- **Wilfried**: Focus on current and future applications of dependable and secure cyber-physical systems
  - Characterizing aspects: autonomy, collaborative behaviour, ..
  - Require a mix of technologies that have their own dependability and security issues (ML, swarm intelligence, blockchain, over-the-air updates, quantum computing...)
  - **Challenges**:
    - difficulties in hierarchical decomposition of system components
    - required use of non-certified COTS and new ways to assess their quality
    - shift from diversity in components implementation to diversity arguments based on diverse usage patterns

- **Jay:**
- The big effort conducted by the research community in the 20$^{th}$ century had allowed to reach very good results wrt accidental faults
- But with the 21th century
  - emerging threats and trends challenge dependability, mainly due to
  - massive interconnectedness of systems which created a large cyber attack surface area
- **Challenges**
  - Cyber resiliency and cyber survivability methods to cope with malicious attacks
    - operate through attacks without human intervention → DARPA OASIS system
  - innovative solutions are needed for cyber survivability, seen as a policy for all critical systems

# Some highlights from the discussion

- AI related
  - Data for training and training process: are they part of the system? (positive and negative opinions)
  - Precise specification of the AI-based system vs restricting the precise specification to unsafe behaviour only
  - Use of AI to enhance (traditional) means for dependability (e.g., testing, detection) → problem of false positives
  - Role of AI: support vs replacing human operation
  - Guidelines for explainability/interpretability → in medicine, certification relies on statistical power
  - Data to use for training: all data vs only relevant data → in medicine, tendency is to use all data
- Safety-critical systems/complexity
  - Monolithic vs distributed fault tolerant architecture → failure of individual components still need to be accounted for
  - Management of emerging behaviours → is it part of the decomposition process?
- Education
  - Educate students to explain the results → change the evaluation criteria rather than forbid use of chatGPT
  - Teach fundamental disciplines in courses: statistics, maths, physics, modeling → but how to attract the interest of students in these subjects?