NTNU

Norwegian University of Science and Technology

# Ethics and Security in a Transparent World

NORDTEK 2022 Annual Meeting and Conference

Stephen Wolthusen
9 June 2022

# Digital Twins and Security

Digital Twins are used both for decision support and automated decisions usually in cyber-physical systems

- Domains and depth may vary from precise models of well-defined systems to large-scale environments such as cities
- Consequences of faults and attacks can hence range from physical damage to wrongful policing decisions

A typical flight relying on a modern Rolls-Royce Trent 700 will generate 500MB of data per flight per engine

- This is retained in a Microsoft Azure platform for real-time analysis including via ML techniques
- Incorrect analysis, such as provoked by adversarial ML may e.g. result in deferring maintenance for too long

NTNU

# Cascading Decisions

Manipulation of (automated) decisions, including dependencies and cascades by way of manipulating sensor data or data sets makes it difficult to bound effects of faults and attacks

- Can and should (critical) systems be constructed that rely on integrity we are hard-pressed to guarantee?
- Safety and security deemed acceptable in information technology still lags behind what is acceptable in other disciplines – arguably this has **deteriorated** in the last 30+ years

# Connected Systems: Case Study

The European synchronous power grid is arguably still just about the largest machine on the planet

- Whilst managed by multiple TSO/DSO, these must co-ordinate, and integrate a variety of loads and generators from EVs acting in both roles via renewable generators to conventional bulk power generators
- Green shift in smart grids requires tightly integrated wide area monitoring and control systems and automation
- On January 8, 2021, a fault in a Croatian substation led to the ENTSO-E network islanding to avoid frequency imbalance, causing Europe-wide repercussions
- Severe incidents in Europe are rare (e.g. 2006), but need for intervention is increasing – without deliberate action as seen e.g. 2015-16 in Ukraine

# The Role of Ethics

Ethics in research (and by extension technology) evolves from the role of the profession in society:

- Produce knowledge, safe, secure, functional, and reliable technology
- Meet needs of society without violating rights or causing *disproportionate* harm

**Are we anchoring research ethics sufficiently, embedding professional ethics for future IT professionals adequately?**

# Dilemmata

Strategic Command
Defence Intelligence

In a talk at the 2021 RUSI Strategic Command Conference, General Sir Patrick Sanders (CG UK Strategic Command) stated

"…the one ring to rule them all, is Artificial Intelligence. […] defending against AI capable adversaries without employing AI is an invitation to disaster.

AI will compress decision timeframes from minutes to seconds, expand the scale of attacks, and demand responses that will tax the limits of human cognition. *Human operators will not be able to defend against AI-enabled cyber or disinformation attacks, drone swarms or missile attacks without the assistance of AI enabled machines.*"

# Designing for Failing Better

With apologies to Samuel Beckett – a key task is to “fail better” so as to limit near-inevitable damage

- It is highly hubristic to assume that a determined adversary won’t find an exploitable weakness in a given system or protocol

**Humility and caution**: We also cannot determine how a downstream system or process may depend on the target of evaluation, either now or in the near future

- Boeing’s 737MAX used a **single** angle-of-attack transducer as input to its MCAS system
- Design specifications changed mid-stream from manual flight to automated flight, expanding the function drastically along the way

# Ethical Implications and Imperatives

Researchers may be asked or tempted to conduct potentially harmful work – is refusal possible, or perhaps pointless if more flexible colleagues will step in?

For IT professionals the choices may be even more stark as refusal to work on a project is unlikely to endear with management

**Addressing vulnerabilities caused in part by technology (whether through faults or adversary action) through yet more technical means may be both inevitable yet is likely to create an escalation spiral that only follows an intrinsic logic**

# Transparency and Trust

Despite the enactment of GDPR in the EU/EEA, personal data collection still continues apace and can be used in ways that are not always compatible with the principles of autonomy and transparency

- Security and privacy breaches are likely to have more severe consequences as connectivity and data footprints reach further into the physical domain – consider e-mobility
- Could complacency and ignorance transform into hostility if consequences of breach become intolerable?

# Research Perspectives and Conclusions

One of our contributions lies in the development of formal models for understanding attacks and vulnerabilities in distributed systems such as cyber-physical systems or supply chains:

- At present requiring specialised mathematical skills, but we hope it will eventually enable non-specialists to perform more extensive risk and threat analyses
- The “engineering attitude” of approximating a solution within given parameters may need questioning when security and safety are at stake

# Double Standards?

Medical “research” in Nazi Germany murdered over 15000 victims

Operation Reinhard led to approx. 1.5m murdered victims, enabled by census information and timetabling of trains to death camps across Europe

Deutsche Hollerith (owned by IBM) provided IT equipment and staff