Original Research Article

Big Data & Society
January–June: 1–13
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517221108369
journals.sagepub.com/home/bds
SAGE

When the future meets the past: Can safety and cyber security coexist in modern critical infrastructures?

Ola Michalec[1,2], Sveta Milyaeva[1] and Awais Rashid[1,2]

Abstract

Big data technologies are entering the world of ageing computer systems running critical infrastructures. These innovations promise to afford rapid Internet connectivity, remote operations or predictive maintenance. As legacy critical infrastructures were traditionally disconnected from the Internet, the prospect of their modernisation necessitates an inquiry into cyber security and how it intersects with traditional engineering requirements like safety, reliability or resilience. Looking at how the adoption of big data technologies in critical infrastructures shapes understandings of risk management, we focus on a specific case study from the cyber security governance: the EU Network and Information Systems Security Directive. We argue that the implementation of Network and Information Systems Security Directive is the first step in the integration of safety and security through novel risk management practices. Therefore, it is the move towards legitimising the modernisation of critical infrastructures. But we also show that security risk management practices cannot be directly transplanted from the safety realm, as cyber security is grounded in anticipation of the future adversarial behaviours rather than the history of equipment failure rates. Our analysis offers several postulates for the emerging research agenda on big data in complex engineering systems. Building on the conceptualisations of safety and security grounded in the materialist literature across Science and Technology Studies and Organisational Sociology, we call for a better understanding of the 'making of' technologies, standardisation processes and engineering knowledge in a quest to build safe and secure critical infrastructures.

Keywords

Safety, cyber security, risk, critical infrastructure, materiality, expertise

Introduction

What happens when new tools enter the old world? For decades, inaccessible legacy computing systems have been running critical infrastructures, like power plants, train stations or wastewater facilities. These so-called operational technologies (OTs), traditionally consisted of isolated computers controlling sensors and actuators, often use simple binary logic (e.g. a machine turning on/off depending on a sensed ambient temperature). The proponents of infrastructure modernisation argue that legacy systems are due an upgrade – after all, they remained the same for decades (cf. Schiølin, 2020). Connecting critical infrastructures to the Internet and the world of big data would equip practitioners with the possibility of remote operations, predictive maintenance or real-time monitoring of industrial processes (Brass et al., 2018; Urquhart and McAuley, 2018). Although this paradigm shift offers interesting prospects, it also brings a novel concern, namely cyber security (Thomas et al., 2020). Moreover, the requirement for cyber security cannot be divorced from safety as the consequences of cyber security attacks in critical infrastructure systems move into the material realm (Tanczer et al., 2018). Cyber security attacks on OTs can lead to explosions, collisions and blackouts. This necessitates novel risk management practices which are simultaneously attuned to security and safety.

Although the professional practice of cyber security risk management is novel in critical infrastructures, risk management in other domains has a long-standing tradition of evolving through controversies. What is construed as

[1]The University of Bristol, Bristol, UK
[2]Bristol Cyber Security Research Group, Bristol, UK

Corresponding author:
Ola Michalec, The University of Bristol, Bristol, UK.
Email: ola.michalec@bristol.ac.uk; aleks.michalec@gmail.com

'risky', 'secure' or 'safe' is a matter of debates, testing regimes and evolving standards – in other words, it is socially constructed (Marres and Stark, 2020; Stilgoe, 2021). The same could be said about how practitioners grapple with new computing technologies to arrive at judgements about a novel criterion of cyber security while preserving their traditional goal of safety. Although safety and cyber security concerns have different origins, the current direction in policy and practice is to integrate these two requirements through the harmonisation of regulatory frameworks, product standards and professional training (Kriaa et al., 2015).

The Science and Technology Studies (STS) and Organisational Sociology research on safety and risk in complex systems helps us understand to what extent the security and safety risk management practices could be integrated. Looking at how the adoption of big data technologies in critical infrastructures shapes understandings of risk management, we focus on a specific case study from the UK cyber security governance. We argue that the implementation of the Network and Information Systems Directive (NIS) (European Commission, 2016) commenced the process of the integration between safety and security concerns across critical infrastructure providers. It has, therefore, legitimised the modernisation of OTs. By addressing the traditional engineering requirement of safety, the proponents of big data technologies managed to position the changes as acceptable to critical infrastructure practitioners. However, we also show that the risk management practices in big data-enabled critical infrastructures cannot be directly transplanted from the safety realm, as cyber security is grounded in anticipation of the future adversarial behaviours rather than the history of equipment failures. Therefore, while integration of safety and security is important for the delivery of reliable critical infrastructure services, it cannot be taken for granted. Precisely, this clash in temporalities between legacy infrastructures and big data technologies creates a gap in conceptualisations and methodologies for risk management.

The remainder of this article will proceed as follows. First, we provide an overview of security, safety, and OTs. Here, we synthesise the literature of theoretical 'turns' across security and safety, highlighting the material, cultural and political differences between information technologies and OTs. Next, we introduce the conceptual lens of safety, risk and security as social constructions (Pinch and Bijker, 1984; Barnes, 1993). In the section 'Case study: NIS implementation in the United Kingdom', we contextualise the article by familiarising the reader with the outline of the regulatory landscape. Following that, the section 'Research design' reports on methods used and reflects on the opportunities and challenges of close collaborations between researchers and practitioners.

We present our argument in four parts. First, by establishing that safety-security integration was key for engineers accepting the modernisation agenda. Second, by outlining collective risk management practices that enabled diverse practitioners to collaborate. Third, by highlighting how practitioners borrowed elements from safety culture and incorporated it to security. Fourth, by cautioning that epistemic and material differences between the old world of legacy technologies and novel big data tools pose limits to the future of critical infrastructures modernisation.

Background

Safety versus Security

Safety and security might seem synonymous, however, there are many technical, political and cultural differences which distinguish these two requirements. Broadly, in infrastructure research, safety is concerned with prevention, protection and recovery from *unintentional* accidents, while (cyber)security is interested in dealing with *malicious and deliberate* incidents (Pietre-Cambacedes and Chaudet, 2010). However, even this high-level distinction has been a subject to multiple theoretical developments. Researchers identify four main paradigms in cyber security: (1) fixing and breaking technical objects; (2) erroneous use of computers; (3) malicious political actions by the means of digital tools; (4) social construction of expertise around what is deemed worth protecting (Adams and Sasse, 1999; Dunn-Cavelty, 2018; Klimburg-Witjes and Wentland; 2021; Renaud et al., 2018). In parallel, a number of 'turns' have been recognised in safety research: from safety being the priority goal, trumping efficiency of processes; through accident prevention via designing-in safety into complex systems, to, finally, placing responsibility on the end-user or the operator (Elish, 2019; Norton, 2015). It is worth noting that these paradigms often co-exist over the same timescales, although they tend to reside in different professions and disciplines, without challenging each other's assumptions. Therefore, the premise of 'new tools entering the old world' – managing novel risks from big data in legacy critical infrastructures – provides a unique opportunity to re-consider the established ways of thinking about both safety and cyber security.

Security and safety are distinguished by their unique temporalities. A key feature specific to the cyber security field is that it rests on novelty. New threat actors, vulnerabilities and theoretical attacks come to light regularly, often at a pace faster than the creation of regulations (Matthew and Cheshire, 2016). Cyber security risks in critical infrastructures often originate from high-profile malicious activities of organised criminals or state actors, lending itself to the use of political rhetoric, high levels of secrecy and large budget spending as means to protect critical infrastructures from any presumed existential threats (Dunn-Cavelty, 2013). Consequently, security by automation, prediction, and testing becomes especially challenging in such environments due to difficulties in access to data or trusted informants. Meanwhile, in legacy OT systems, safety risk has been traditionally understood probabilistically as a 'failure rate', a frequency with which an engineering component fails when tested, expressed in failures per unit of time. Failure rate figure is deeply rooted in physical properties of the system and a wealth of historical data (Ani et al., 2016). The dynamic characteristic of cyber security contrasting with a static (or, at best, slowly moving) nature of safety implies there are limits to the integration of traditional OT safety paradigms to the context of modern, interconnected critical infrastructures (Slayton and Clark-Ginsberg, 2018).

Operational technologies versus information technologies

Throughout the article, we distinguish between information technologies (computers commonly found in homes and offices) and OTs (computers operating engineering machinery) to understand how these technologies were historically constructed as separate and how they are now poised as integrating with each other. In this claim we follow calls from Kinsley (2014) and Aradau (2010) to pay attention to materiality in computers and infrastructures. Objects of cyber security – sensors, buildings, code – are not passive technologies waiting to be filled with discourses. They are not characterised by 'essential' features separating them from humans, either (Fouad, 2021). Instead, materiality is critical for noticing how objects become 'agents' of social change through practices which are both discursive and material (Aradau, 2010).

Historically, the consequences of security incidents in IT were materially different from OT because these systems were traditionally built for different purposes. While IT professionals are typically concerned with the damage to data, hence lost revenue, customer trust or reputation, OT practitioners are mainly concerned with human safety, equipment damage and continuous supply of 'essential services'. Traditionally, OT systems were designed with physical resilience and safety in mind; cyber security was not a typical requirement due to the practice of 'air-gapping', that is, isolating OT computers from the unsecured networks like the public Internet (Byres, 2013). IT systems, in contrast, are commonly interconnected, which necessitates the need for security and privacy by design and regulation (Michalec et al., 2020). As both IT and OT systems are gaining Internet connectivity and real-time analytics functionalities, they are 'blending' into a single entity. And so are the previously separate concerns for cyber security and safety (Michalec et al., 2021). In short, contemporary 'big data' practices of OT and IT professionals are reconfiguring what critical infrastructures are made of.

The differences between OT and IT were historically not only material but also cultural, such as varying degrees of professionalisation (i.e. typical career routes and education required), or the juxtaposition of safety culture of OT engineers and innovation culture of IT workers (Guldenmund, 2000; Reece and Stahl, 2015; Thekkilakattil and Dodig-Crnkovic, 2015). Infrastructure providers running on OT systems are also organised very differently compared to IT companies – critical infrastructures are often hierarchical and governed through public-private partnerships, while IT companies range from start-ups to monopolies and they are most often private sector entities (Dunn-Cavelty and Suter, 2009; Murray et al., 2017). These distinctions are important as they inform who gets to conceptualise risk, how do they do it and why.

Theoretical framework

How do we know if machines are safe?.

In order to frame this research paper, we used the conceptual lens of social construction of safety, risk and security. This paradigm explains how technological expertise emerges, stabilises, gets contested or widely accepted (Barnes 1993; Pinch and Bijker, 1984). Such research examines *how* different actors arrive at their assessments, rather than examining whether their assessments are true. Examining technological expertise involves an inquiry into situated practices, materials of day-to-day work and debates surrounding technoscientific developments (Collins, 2007; Pinch and Swedberg, 2008; Suchman, 2007). In that vein, we first build our argument by reviewing the literature on social construction of safety and risk before moving to the analysis of our data which focuses on the construction of security.

The long history of safety research in complex systems like aviation (Downer, 2010) nuclear engineering (Wynne, Waterton and Grove-White, 2007; Polleri, 2020; Perrow, 1984) or autonomous vehicles (Haugland, 2020; Stilgoe, 2021) thoroughly documents and analyses the evolution of testing regimes and assurance schemes to minimise risks, recover from incidents and anticipate a range of possible scenarios. How do experts establish whether a complex system is 'safe enough'?

Focusing on controversies, STS scholars have been tracing how debates evolve to establish complex, emerging or high-stakes technologies as 'safe' or 'risky'. For example, industrial manufacturers' framed safety as a matter of feeding more data to proprietary machine learning algorithms in case of autonomous vehicles (Stilgoe, 2018) or raising awareness of machine operators working with robots (Elish, 2019). Meanwhile, Norton (2015) showed how road safety was deprioritised over decades as the automotive industry grew in the United States. These examples illustrate that the research on safety and cannot be limited to laboratory experiments and test beds as categories like 'safety' and 'risk' are inherently riddled with political and organisational contingencies.

Other researchers contributed to the social construction of safety through understanding accidents. Downer (2010), brings attention to what he calls '*epistemic* accidents' – man-made calamities resulting from fundamental limitations of engineering tests and models which, by design, are never perfect representations of the 'real world' conditions. Such events happen if 'scientific or technological assumptions prove to be erroneous, even though there were reasonable and logical reasons to hold the assumptions before the event' (Downer 2010: 752). Epistemic accidents offer valuable lessons for organisations and practitioners by revealing inherent shortcomings of the current engineering and design paradigms. By researching how experts work towards safety in complex engineering systems, we can better see the potential for epistemic failures and build-in practices to learn from them.

Analysis of epistemic accidents matters as it shifts attention from an individual's error to *interactions* between operators, organisational cultures and politics and machines. In doing so, STS scholars go beyond seeing safety accidents conventionally, as failure of individuals and their erroneous use of complex systems (Pinch and Bijker, 1984; Stilgoe, 2018). Here, the notions of error and safety are intimately connected to (the limits to) 'knowability' of complex systems (Downer, 2010; Spinardi, 2019). They feed into safety testing and modelling, and therefore, everyday decisions about risk (Marres and Stark, 2020).

Finally, while the expertise from the safety world cannot be directly transposed to the security context, there is an overlap between these fields. Politically, both safety and security of infrastructures are prioritised by governments as they fundamentally relate to the 'normal' functioning of the society (Agrafiotis et al., 2018; Shove and Trentmann, 2018). However, as certain security incidents and safety accidents cannot be prevented in complex systems (Perrow, 1984), critical infrastructure operators often emphasise resilience and risk management. In practice, this means that the same teams could be made responsible for both safety and security. Even though cyber security incidents and safety accidents require separate root cause analyses, they might manifest as the same consequences (Agrafiotis et al., 2018; Kriaa et al., 2015), thus share commonalities in terms of risk management practices.

Risk management: Between calculation and anticipation. Understanding risk management in critical infrastructures is a multifaceted issue of both qualitative and quantitative nature (Shreeve et al. 2020). Despite the rise of rule-based and probabilistic risk methodologies, for example, attack trees, attribute-based algorithms (Tatam et al., 2021), security risk is 'incalculable' since there are limits of what could be inferred from scientific data (Amoore, 2014: 424). Risk methodologies are 'already political' as they involve combinatorial possibilities whose arrangement has effects on risk scores, and associated countermeasures (Amoore, 2014: 423). In the case of OT cyber security, this means that we cannot simply assume that the risk rises proportionately to the number of Internet-connected devices. In practice, assigning risk scores in a given organisation depends on asset criticality (i.e. how important are devices and datasets), motivations of potential attackers, the available budget, just to name a few (Cherdantseva et al., 2016). Moreover, risk decision makers need to account for issues which are not specific to cyber security, for example, public responsibility for delivery of reliable essential services, business models, insurance, risk appetite, reputation (Henrie, 2015; Nurse et al., 2017; Pieters and Coles-Kemp, 2011). While recent research offers reviews of risk assessment frameworks (Kriaa et al., 2015; Cherdantseva et al., 2016), it leaves a gap for understanding to what extent these frameworks are applied in the real-world context.

Previous risk studies embedded in the critical infrastructure context highlighted that risk assessments are collaborative processes, rather than a matter of following a formal methodology (Frey et al., 2019; Shreeve et al., 2020). In doing so, they challenge the trope of 'security expertise' being solely a technical and individual matter. This shows security expertise as inherently emergent, contextual and subjective. For example, security practitioners might use a variety of reasoning strategies, such as 'risk first' (following governmental risk assessment framework) or 'opportunity first' (identification of investment opportunities before considering risks). Moreover, in practice, people exercise both kinds of reasoning, with vulnerabilities (i.e. weaknesses in computer systems) are thought about most commonly, and assets (i.e. equipment, documents, employees) least often, leading to an over-reliance on vulnerabilities-centred threat assessments (Shreeve et al. 2020). Moreover, risk thinking is 'front-loaded,' as practitioners tend to think about risk in the beginning of the decision-making process, rather than systematically throughout the lifecycle of OT systems (Shreeve et al., 2020). Meanwhile, as threats in cybersecurity evolve over time, risk management ought to be iterative and regularly updated (Ani et al., 2016; Frey et al., 2019).

Risk management in the context of security often draws from a practice called threat modelling to anticipate likely attackers, incident pathways, possible consequences of attacks and best ways to respond to them. The techniques under the umbrella of threat modelling vary; from qualitative expert workshops (Wuyts et al., 2020), through mathematical models based on probabilities (Markov chains, game theory) to graphical representations (in the forms of tables, data flow diagrams and attack trees), with some threat modelling techniques promising full automation and quantification of risks (Tatam et al., 2021).

In practice, when it comes to classifying potential impacts, and evaluating attackers' motivations, threat modelling relies on qualitative expert judgement, usually a small group of domain specialists. However, as cyber security 'spills out' beyond simply protecting computers, there is a call for broadening the scope of threat modelling. Critical social scientists argued for anticipating risks of emerging technologies by including non-experts (Slupska et al., 2021), understanding security in tandem with privacy and surveillance (Kazansky, 2021; Wuyts et al., 2020), and approaching non-human actors (code, hardware, algorithms) as active co-creators of geopolitics (Dwyer, 2021; Fouad, 2021). The strength of such a 'critical threat modelling' approach would then lie in the capacity to imagine and anticipate a wide range of outcomes and curate a space for explicitly normative discussions about living with digital technologies. Including actors outside of cyber security profession allows the multiplicity of futures to become visible, as there is no single objective and optimal choice between security, privacy, risk appetite, resources available, reputation, innovativeness, and many other factors.

Case study: NIS implementation in the United Kingdom

To address how critical infrastructure practitioners conceptualise and practice security risk management, we use the case study of the NIS, as implemented in the United Kingdom (DCMS, 2018). The NIS implementation practices reveal how practitioners from diverse sectors grapple with the modernisation of legacy OT systems. Their NIS compliance practices are balancing acts to build interconnected and secure infrastructures, without compromising on the traditional engineering goals like safety, or reliability of essential services like water, energy or transport.

NIS originated as a high-level supranational directive ratified by the European Parliament in 2016. Since then, it has been transposed to the EU Member States and the United Kingdom as NIS *Regulations* (DCMS, 2018). This move meant that while high-level objectives and international cooperation mechanisms were set by the EU, the scope of what is regulated as well as implementation mechanisms are decided by each state and sector individually. In the United Kingdom, the implementation of NIS follows the principles of 'appropriateness and proportionality' (Michels and Walden, 2018), which necessitates careful deliberation over designation of the operators falling under the purview of regulations, thresholds of incident reporting and maximum penalties. NIS is known as 'principles-based regulation,' meaning that critical infrastructure operators work towards meeting the governmental objectives without specification how to achieve such goals (Michels and Walden, 2018). The government's reasoning behind this move is to avoid 'box ticking' style of compliance and contextualize risk management. In the eyes of the UK's National Cyber Security Centre, 'this encourages innovation and expands the breadth of technologies we can assure' (NCSC, 2021).

Risk assessment is embedded in NIS implementation from the beginning. The implementation procedures in the United Kingdom begin with a self-assessment stage (known as the Cyber Assessment Framework; NCSC, 2019). The Cyber Assessment Framework is the key operational document pertaining to the question of cyber security risk management of critical infrastructures in the United Kingdom. Fourteen principles of the Cyber Assessment Framework are set out as so-called 'Indicators of Good Practice' (NCSC, 2019), or recommended outcomes of security improvements rather than specification *how* to improve cyber security. For the purpose of self-assessments, each of the 14 outcomes is self-assessed in diverse teams comprising of both OT and IT practitioners according to a three-grade scale as either 'fully achieved', 'partially achieved' or 'not achieved'. Following the completion of self-assessments, operators and regulators draw agreements on the improvement plans, and conduct external audits (Shukla et al., 2019; Wallis and Johnson, 2020). Since the successful implementation of cyber security regulations requires collaboration across the IT and OT teams, it makes the cross-cutting issues of safety and security visible (Michalec et al., 2021)

Research design

We conducted a qualitative study of experts managing big data risks to critical infrastructures. Between November 2019 and January 2020, we interviewed 30 practitioners and observed two industry events focused on the implementation of the NIS Regulations. Our interviewees ranged from the critical infrastructure operators, regulators, consultants, lawyers, to OT equipment manufacturers. We aimed to cover a range of sectors (e.g. energy, water, transport) and roles (e.g. technical, managerial, consultancy, regulatory). We conducted semi-structured interviews focusing on historical perspectives on the development of OTs, participants' outlooks on the future of modernisation, interpretations of the Regulations and the issues around communicating security risk across professional boundaries. Questions were tailored to each participant in order to account for differences in sectors and professions. Interviews took place either at the participant's organisation, our institution or via online calls, with the lead author conducting all interviews. All conversations were recorded with the interviewees' consent. No reimbursement was given for participation. Our analysis is complemented by an in-depth reading of the Cyber Assessment Framework (NCSC, 2019), a UK-specific document outlining what the outcomes of 'good' risk management is security look like.

Our approach responds to the calls by de Goede (2020) for increased engagements between empirical research on expert practices and critique. By treating the implementation of cyber security regulations as 'situated practices', we bring our attention to the notion of expertise construction and de-centre policy discourses or legal analysis. By following practitioners and practices, we were able to gain trust of our informants, appreciate the diversity of their expertise, and their disagreements and material artefacts they work with. As a result, long after data collection period finished, the first author of the paper is still collaborating with practitioners; publishing government guidance or giving regular industry talks. The downside of research approaches relying on in-depth engagement with practitioners is the possibility of losing 'critical distance' and getting 'co-opted' by practitioners' agenda (de Goede, 2020). This is especially challenging when working with practitioners whose goals are both normative and open to interpretation, like security and safety. In our case, we navigated that tension by highlighting the plurality and contingency of expertise, rather than promoting a single vision. In terms for further research avenues, this research agenda would benefit from an in-depth investigation of single quantitative threat modelling methodology (e.g. Markov chains); following the datasets, construction of algorithms, modeller's assumptions and how results of risk assessments are translated (or not) into organisational decisions.

Towards modernisation of critical infrastructures

Our research shows that the introduction of security regulations into the world of legacy safety-critical systems prompted harmonisation of these two requirements. In turn, this move legitimised the modernisation of legacy OT environments. However, due to fundamental differences between managing safety and security risks, the modernisation of critical infrastructures cannot be taken for granted.

At first, engineers exhibited resistance to the modernisation agenda: '*our sector is adopting Industrial Internet of Things at a frightening rate, and we'll have little idea as to what it looks like and how to secure it*' (interview with water sector operator 1). The dominating mood was cynicism about big data technologies being introduced to increase manufacturers' profits: '*people see opportunities to deliver a new shiny box, a new system, a new bit of software, a new service. So, that is really, really driving and almost pushing along innovation in the market*' (interview with OT security consultant 1). However, a pivotal moment occurred when safety and security professionals started working together with the regulators to identify how their requirements map onto the Cyber Assessment Framework and create a common benchmark for the whole sector: '*it is like an exam board where you get together and make sure all the markers are assessing against the same criteria. We often went back to the regulator pointing out where NIS did not make sense in our context of OT technologies*' (Interview with water operator 2).

Precisely, that bringing together of diverse experts enabled safety-security integration. Regulators, by listening to the concerns from safety engineers, adjusted the Cyber Assessment Framework guidance, to facilitate digital connectivity in critical infrastructures. This agrees with the overarching justification behind the UK National Cyber Strategy which claims that the ongoing and rapid expansion of digital connectivity is a main driver behind cyber security regulations (Cabinet Office, 2022: 29). Effectively, NIS is the first step to legitimise the modernisation of critical infrastructures:

> *'there are some instances where the best answer would be to innovate legacy. NIS has not ever come up with a recommendation that there should be greater digitisation, but what it did say is: "There are certain expectations, particularly around configuration and software management where it was very hard to deal with a legacy." So, some people found themselves caught in a business case between a technology refresh which, frankly, was overdue anyway, or retaining legacy systems for reliability reasons with negative cyber security implications.'*

(Interview with energy sector working group lead)

However, while bringing the diversity of expertise allowed to advance and integrate risk management practices, there are fundamental differences between safety and security. Therefore, the future of big data in critical infrastructures is still uncertain. In what follows, we will examine practices which enabled that integration as well as highlight epistemic and materials differences between these two requirements.

Hiveminds and other collaborations

Diverse expertise

What makes risk management across security and safety successful? First, it is contingent on the *access to diverse expertise* within an organisation, and how effectively recommendations are communicated to those in charge of decision making, who are usually senior managers without the expertise in security: '*So the security engineers might be quite grumpy because the manager just does not understand their problems. But the engineers also do not understand there is a bigger picture going on here, e.g., that a power station needs to provide an ongoing supply of electricity*' (Interview with engineering consultant 2). Second, cyber security risk assessment requires diverse inputs – apart from traditional technical experts, human factors practitioners are needed to anticipate how workers could be employing workarounds against security measures, so that they could improve the usability of security practices. In the words of our interviewee, a water regulator, *'You can create more risks by going overboard with too stringent and annoying security measures where people try and find work arounds. Water plant operator working with time-critical systems cannot afford 30 seconds delay if they typed their password incorrectly'* (Interview with water regulator no 3). As such, NIS does not only regulate technologies, but also how people use them.

Trust in collaborations

Security risk assessment is also contingent on trust. In particular, it is trust between IT workers and OT workers filling the Cyber Assessment Framework: '*when I do my risk assessment of systems we rely on, I've got to assume that the guy doing the IT bit has got his IT correctly'* (interview with OT water operator 1). Furthermore, successful risk management happens if security experts manage to establish a trusting relationship with the board members and gain '*buy-in from senior management to invest in cyber'* (Interview with energy regulator 2). Ultimately, senior managers are the budget holders and ought to see how security improvements translate into organisational goals, be it by providing reliable energy supply, or ensuring workplace safety on a train station. While participants acknowledged that connecting security practitioners to board level executives has traditionally been a challenge, they are now gaining techniques for better engagement: '*I would stop talking about the threats, the executives know about the threats. Instead, say how we are looking after the business and its core critical functions*' (interview with a vendor of security products). Ultimately, cyber security risk management is seen in the context of broader risk management, where practitioners across diverse teams are encouraged to reflect: *'How much risk can we tolerate as an organisation? How much do we value our reputation? What is our attitude towards legislation and regulation? It is all interconnected'* (Interview with an IT security consultant). Indeed, in this case, security is a matter of care (Kocksch et al., 2018) where security budgets are considered as a matter of long-term maintenance of whole organisations rather than cutting-edge technological 'solutions'.

Building a 'risk thinking' hivemind

One of the pressing questions for the critical infrastructure practitioners is how NIS could avoid being a tick-box exercise. The UK Government designed the Cyber Assessment Framework as an outcomes-based document to 'discourage compliance thinking' (NCSC, 2019). However, by providing a set of 'good outcomes' rather than policies on *how* to achieve them, the Cyber Assessment Framework received criticisms for '*leaving everything up for negotiation'* (interview with energy regulator 2). On the one hand, outcomes-based regulations are suitable for dynamic contexts, like cyber security, where new risks emerge regularly and there are multiple ways to 'do the right thing'. On the other hand, outcomes-based regulations rely on a baseline level of expertise where practitioners can exercise expert judgement on risk: '*we want people use the* Cyber Assessment Framework *as a sanity check rather than a procedure to follow to the letter to protect their own reputation*' (interview with energy regulator 1). And so, practitioners called for raising the level of expertise across the whole sector, what we call a 'risk thinking hivemind'.

In the eyes of participants, these hiveminds, usually expressed as semi-formal working groups, are better suited for sharing expertise than regulations. In other words, risk management practices preferred by the participants are relational and collaborative, rather than top-down and individualised. As our interviewee put it, '*working groups could be tasked with a creation of sector-specific process standards which would be a collective endeavour rather than an individual activity of 'box ticking*' (interview with water operator 1). An example from an energy sector working group shows that collaborating on risk assessment was easier *before* regulations came into force:

> '*In 2013, we did a UK-wide risk assessment. We anonymised responses from individual companies, we aggregated it and so we could come up with two things. First, where collective gaps and difficulties, so that we could request help from the government departments. Second, we found out there was a difference between the best practice in some and those who were struggling and there we introduced knowledge sharing opportunities. We then allowed the good ones to present their approaches and the others could learn so we got a best practice learning environment*'

(Interview with energy sector working group lead).

The implementation of the Cyber Assessment Framework reveals three crucial aspects pertaining to the social construction of risk management: professional practices as objects of regulations, cyber security mapped to broader organisational goals, and practitioners collaborating to create 'risk thinking hiveminds' that capture risk management practices across their sectors. Just like safety regulations in critical infrastructures (Downer, 2010), NIS regulates trust in professional practices, rather than technologies. Cyber security has been placed in the broader organisational context of safety, usability or reliability. Finally, faced with the novelty of cyber security regulations in the legacy environments, practitioners collaborated to manage the overlapping risks of safety and security. Lacking prescriptive guidance, they created a 'risk-thinking hivemind' to collectively work towards their goals.

Towards harmonisation of safety and security

Let us now turn to how cyber security integrated practices from safety engineering in their work to blend the 'digital' and 'engineering worlds'.

Threats and incidents reporting

In the event of a cyber security incident, operators will have to report it to the regulator and evidence that they took 'appropriate and proportionate' measures to mitigate risks in order to avoid a penalty (NCSC, 2019). However, there is no obligation to report ongoing threats, that is, prospective malicious activities and actors that are yet to hit a computer network. The above caveats resulted in the ongoing debates on defining reporting thresholds for incidents and even distinguishing between a threat and an incident (DCMS, 2021). The dilemma lies in the fluid nature of the above terms. On the one hand, encouraging reporting of the ongoing threats improves the collective intelligence, the aforementioned 'risk thinking hivemind'. On the other, if a threat reported by one organisation turns into an incident in another, both organisations may be receiving fines. This lateral way malware propagates is a well-known phenomenon in interconnected complex environments (Dwyer, 2018) but historically it was not a concern in disconnected critical infrastructures. As a result, these contingencies of threat reporting pose a risk that operators will minimise their reporting all together. The evidence from the critical infrastructure security regulations in the United States shows that fear of fines created a counterproductive environment for information sharing (Clark-Ginsberg and Slayton, 2019).

In order to encourage operators to report on the developing threats, water regulators broadened the reporting scope so that all security incidents and safety accidents, however minor, had to be reported under the same umbrella[1]. This also led to discussions among practitioners to report 'near misses,' threats which did not have a significant impact on their network[2], showing that thresholds of harmful events are a subject to ongoing debates. This move signals that both incidents and accidents are bound to happen and reporting of the ongoing threats (even if not yet materialised as security incidents or safety accidents) will not be stigmatised.

However, this practice is not uniform across all critical infrastructure sectors. Right now, energy regulators do not have the same level of insight. In order to allow further integration of security and safety, regulators advocated for improved capabilities to observe the dynamic nature of threat actors and typical attacks: '*it would be of a real interest to us, but currently this is a voluntary procedure*' (Energy regulator 1). Although 2020 saw numerous attempts of security breaches attempts, none of them were reported to NIS regulators as they did not lead to the loss of supply or power outages; such lenient reporting criteria also raise suspicions in the national news, which questions whether NIS' reporting criteria in the energy sector is fit for purpose (Martin, 2021).

Maintenance contracts

Deciding on the ownership of cyber security risks proved very challenging: '*when you start looking at the scope of NIS, which is one of the first things you do, you ask yourself, what do you really depend on? Very complicated, and no one person owns it, and you look at all the independences and as you get further away from the core*' (interview with engineering consultant). In particular, it is the international nature of internet services (e.g. cloud providers), which highlights the difficulty with drawing a clear boundary around cyber security risks (and, indeed, the scope of NIS itself!): '*a whole chunk of security is now outsourced to the Cloud provider overseas, so critical infrastructure operators lose control over it*' (IT security vendor).

Yet again, well-established practices from safety engineering could come to rescue, with maintenance contract between third-party suppliers and operators recommended as ways to uphold good standards of security over time: '*long term improvement is a matter of maintenance contracts. So that is important also, is that if you are buying an expensive piece of equipment you want to have it supported for a long time, otherwise you do not have a business case to use that supplier*' (interview with a rail engineer).

While borrowing professional practices from the safety culture might help engineers with understanding of cyber security, the complexities around global supply chains and the scope of NIS remain. In an example from one of the critical infrastructure sectors (Wallis and Johnson, 2020), for data centres located outside of the United Kingdom, the NIS regulators cannot oversee their security measures. However, critical infrastructure operators are still legally obliged to arrive at bilateral contracts with data centre providers to meet the requirements of NIS. The requirement for security remains but less so the clarity about who validates the process (Wallis and Johnson, 2020).

To conclude, by borrowing established practices and terms from the safety culture context, NIS practitioners were able to make cyber security more familiar to critical infrastructure engineers. Encouraging broader incident reporting and establishing maintenance contracts opened new discussions highlighting the complexity of cyber security in interconnected, big data environments like cloud.

Dissonant harmonies: The limits to integration of safety and security

Despite the opportunities of safety-security harmonisation as expressed through professional practices, this section argues that there are fundamental epistemic and material differences between legacy OT environments and big data practices.

Prescriptive thinking

First, let us return to collective risk assessments we identified earlier in the analysis. The creation of 'risk thinking hiveminds' which consolidate security knowledge across the sector could be complicated by the tendency to work in a prescriptive manger common in safety engineering. One of the water regulators appeals:

> 'G*etting companies used to doing risk assessment rather than compliance is key. Our safety framework was completely prescriptive: a list of measures that you must have on water plants –, e.g., you must have this kind of lock fitted in this kind of way by one of these companies. Which companies love because they can cost it up and go along to Ofwat*[3] *and say, 'We need exactly this much money to do this much work over this number of years.'* (Interview with water regulator 1)

Risk thinking at the intersection of safety and security would necessitate encompassing novel big data practices, that is, back-ups for real-time environments, asset inventories for equipment operating automated processes, anomaly detection on segregated computer networks (NCSC, 2021). These practices are not familiar to safety engineers who typically work with legacy systems where computer networks were not traditionally monitored, backed-up or segregated. Moreover, no single risk management framework covers all recommended risk management practices with '*various countries having their own standards. So, it is horses for courses and some of the best solutions I have seen are basically taking a blend of several of the standards*' (interview with OT security consultant). But, to what extent are engineers willing to let go of prescriptive thinking and, instead, start blending various frameworks or anticipating futures? A shift to 'risk thinking' culture would necessitate a major change in the 'epistemic culture' in safety engineering, to borrow a term from Knorr-Cetina (1999). Epistemic culture refers to an established way of accessing, validating and advancing knowledge in a given expert community (Knorr-Cetina, 1999). Nonetheless, changing culture established over many decades is a mammoth task beyond the scope of a single regulatory initiative.

Secrecy restricts learning. Our second point relates to the secrecy challenges with accessing data required to differentiate between security and safety. Is an anomaly in the system due to an error or a hacker? Did the blackout result from a storm or a cyber security attack?

Earlier in our analysis, we examined integrated reporting of security and safety events to harmonise these two requirements. However, integrated reporting of security incidents and safety accidents yields limited lessons for the operators if they cannot learn what caused a harmful event. Currently, the lack of separate root cause analysis limits further integration of safety and security paradigms in engineering:

> '*There is a lot of reporting that does not tend to think it is cyber security but actually that could feed into the cyber risk picture that we need to bring into the mix, we should be asking: was it safety incident, security, or something else? What did we learn from it? It all needs to be put into the pot. It is like you are telling people, "Something bad has happened". They need to know: "Well, actually, what can I do about it?." And I think there needs to be more done about turning those incidents into lessons for best practice'* (interview with engineering consultant).

A parallel gap resides in the practice of threat modelling in critical infrastructures. The lack of historical data on security attacks in OT environments poses challenges to the modelling of future threats: '*when you're looking for a record of past incidents to take to your senior management and all you can show them is a brief declassified document with barely any information, they can say, "Well is that all you have got? If there are hardly any incidents, maybe we should not be spending more money and effort?"* (Interview with engineering consultant 3).

Why is it so difficult to obtain data? The access to information on threat actors and past incidents is highly limited due to the sensitivity of this topic. Complex procedures around data classification, information exchanges and even day-to-day interactions give rise to secrecy as a dominating practice in social interactions. For example, some of our participants were unable to have their cameras on during the interviews due to their work incorporating both offensive and defensive security (i.e. they were simultaneously hackers and defenders). Such restrictive norms around communication raise a possibility of 'epistemic accidents' (or, rather 'incidents', if we are concerned with intentional and malicious nature of security attacks), events highlighting the limits to established practices across engineering and computing (Downer, 2010). A telling example would be a cyber security incident hitting an underprepared organisation that incorrectly extrapolated the rarity of cyber security attacks based on scant declassified data. In such case, a cyber security attack would be a consequence of poor communication and mistrust across critical infrastructure organisations.

Logics of risk assessment

The final point of contention relates to the very logics of risk assessment across safety and security. While the probability of safety failures is well grounded in historical records and components testing (Michalec et al., 2021), security incidents in the OT space are a function of anticipating malicious behaviours and relying on sparse historical data, which does not lend itself easily to the logics of probabilistic prediction. Considering active adversarial actions from highly skilled actors like organised criminals or state-sponsored hackers brings a contentious dimension to the practice of risk management. In practice, it means that engineers will have to conduct an explicitly political and normative analysis and they are not necessarily ready to acknowledge this: *"if state sponsored hackers bring a power station down, then we have to react. But that is difficult, because then you are definitely into politics. We are a non-political, non-government organisation, we only do what we can"* (Interview with Incident Response Director).

In order to escape being drawn into politics, industry actors propose machine learning as a data-driven, objective means of risk assessment (Dragos, 2019). However, other than *any* risk analysis being far from objective due to aforementioned 'incalculability of risk' (Amoore, 2014), the practice of automated anomaly detection using machine learning in particular is seen as contentious due to the low diversity of modelling data used to train machine learning algorithms: '*There's not enough randomness in the datasets themselves to say the type of algorithms they use were going to have perfect detection rates'* (interview with energy regulator 3). Consequently, practitioners are '*not afraid to lose their jobs*' as '*although the networks are evolving and there is more information, we will always have a human operator checking the anomalies*' (interview with rail engineer).

Overall, despite the attempts to integrate safety with security, the paradox is that big data computing and legacy engineering environment belong to different and incompatible worlds. The dissonance is expressed in the following three forms: (1) epistemic culture: risk versus prescription; (2) secrecy restricting collective learning; (3) different logics of risk assessment. The logics of anticipation and connectivity favoured in the big data environments do not fit easily into the prescriptive and siloed world of OT engineering, leading to the situation in which the modernisation of critical infrastructures will continue to pose challenge and cannot be taken for granted.

Concluding thoughts

Can new tools be useful, or work at all, for the world that has not been designed and built to accommodate them? Can critical infrastructures, with their paramount concern about safety, adjust to the new reality brought about by instant connectivity and big data? Can safety and security coexist? Modernisation of legacy systems with big data technologies brings about the need to reconsider traditional paradigms in both engineering and computing in order to successfully integrate them. Tracing the attempts to harmonise diverse computing and engineering requirements, we draw from the case study of NIS Regulations. NIS Regulations bring attention to the management of risks to critical infrastructures. While previous research on critical infrastructure risk management accounts for the variety and sophistication of risk assessment methods (Kriaa et al., 2015; Cherdantseva et al., 2016) as well as the topical coverage of various frameworks and standards (Topping et al., 2021), we brought attention to the social construction of risk, safety and security. In other words, what happens when traditional safety practices meet novel big data practices.

We argue that the introduction of security regulations into the world of legacy safety-critical systems prompted harmonisation of these two requirements. Integration of safety and security was afforded thanks to collective risk management practices: (1) conducting risk assessment in diverse team; (2) mapping cyber security onto organisational goals with senior stakeholders and (3) practitioners collaborating to create 'risk thinking hiveminds' capturing good practices across their sectors. Next, we also show that the implementation of NIS created opportunities to borrow established terms and practices from safety engineering and incorporate them into security procedures. In doing so, NIS serves as a vehicle that enables incorporating cyber security in the existing engineering professions, organisational structures, and maintenance contracts with third party suppliers. On the other hand, however, there are major epistemic and material differences between safety and security domains, such as prescriptive attitudes to risk in safety engineering standards, or secrecy restricting cyber security information sharing. Ultimately, the NIS Regulations exposed a tension between two vastly different logics of risk assessment across security and safety: future-grounded and explicitly normative anticipation versus past-based probabilistic prediction.

The implementation of NIS is the first step in the integration of safety and cyber security; therefore, it is the move towards legitimising the modernisation of critical infrastructures with big data. But we also show that the cyber security risk management practices cannot be directly transplanted from the safety realm, as cyber security is grounded in anticipation of the future adversarial behaviours rather than the history of equipment failure rates. While the harmonisation of safety and security standards and organisational practices is important for the delivery of reliable critical infrastructure services, this process cannot be taken for granted and, consequently, we call for a better understanding of the making of technologies, standardisation processes and engineering knowledge in a quest to build safe and secure modern critical infrastructures. Despite epistemic accidents and incubation over a long period of un-reliability and controversy, we learn a lot from the histories of safety and engineering paradigms.

Acknowledgements

We thank our participants for sharing their experiences and inviting us to their professional events. We would also like to thank Dr John Downer, Lars Gjesvik as well as two anonymous reviewers for their insightful comments on early versions on the paper. Finally, we would like to extend our gratitude to the European Cyber Security Seminar community, where we had a chance to present a seminar based on this research project.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Cyber Security Centre. “Power^2– Understanding disruptive powers of IoT in the energy sector” funded by the PETRAS National Centre of Excellence (part of UKRI, number EP/S035362/1).

ORCID iDs

Ola Michalec https://orcid.org/0000-0003-3807-0197
Sveta Milyaeva https://orcid.org/0000-0002-0156-5359
Awais Rashid https://orcid.org/0000-0002-0109-1341

Notes

1. Workshop with water suppliers, November 2019, Leeds.
2. Workshop for critical infrastructure operators, Oct 2019, London.
3. Ofwat is the economic regulator for the water sector in England and Wales, setting maximum investment budgets and water pricing.

References

Adams A and Sasse MA (1999) Users are not the enemy. *Communications of the ACM* 42(12):40–46.

Agrafiotis I, Nurse JC, Goldsmith M, et al. (2018) A taxonomy of cyber-harms: defining the impacts of cyber-attacks and understanding how they propagate. *Journal of Cybersecurity* 4(1): 1–15. doi:10.1093/CYBSEC/TYY006.

Amoore L (2014) Security and the incalculable:. *Security Dialogue* 2014;45(5):423–439. doi:10.1177/0967010614539719

Ani UPD and He H. (Mary) and Tiwari A (2016) Review of cybersecurity issues in industrial critical infrastructure: Manufacturing in perspective. *Journal of Cyber Security Technology* 1(1): 32–74. doi:10.1080/23742917.2016.1252211.

Aradau C (2010) Security that matters: critical infrastructure and objects of protection. *Security Dialogue* 41(5): 491–514.

Barnes TJ (1993). Whatever happened to the philosophy of science? *Environment and Planning A* 25(3): 301–304.

Brass I, Tanczer M, Carr M, et al. Blackstock (2018) Standardising a moving target: the development and evolution of IoT security standards. IET Conference Publications, 2018(CP740). doi:10.1049/CP.2018.0024.

Byres E (2013) The air gap: SCADA’s enduring security myth: attempting to use isolation as a security strategy for critical systems is unrealistic in an increasingly connected world. *Communications of the ACM* 56(8): 29–31.

Cabinet Office (2022) National cyber strategy. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1053023/national-cyber-strategy-amend.pdf (accessed 13 June 2022).

Cherdantseva Y, Burnap P, Blyth A, et al. (2016) A review of cyber security risk assessment methods for SCADA systems. *Computers & Security* 56: 1–27.

Clark-Ginsberg A and Slayton R (2019) Regulating risks within complex sociotechnical systems: evidence from critical infrastructure cybersecurity standards. *Science and Public Policy* 46(3): 339–346.

Collins H (2007). The uses of sociology of science for scientists and educators. *Science & Education* 16: 217–230.

de Goede M (2020) Engagement all the way down. *Critical Studies on Security* 8(2): 101–115.

Department for Digital, Culture, Media and Sport – DCMS (2018) The NIS regulations. Available at: https://www.gov.uk/government/collections/nis-directive-and-nis-regulations-2018

Department for Digital, Culture, Media and Sport – DCMS (2021) Government response to the call for views on amending the security of network and information systems regulations. Policy paper. Available at: https://www.gov.uk/government/publications/government-response-on-amending-the-nis-regulations/government-response-to-the-call-for-views-on-amending-the-security-of-network-and-information-systems-regulations

Downer J (2010) Trust and technology: the social foundations of aviation regulation. *The British Journal of Sociology* 61(1): 83–106.

Dragos (2019) Key Considerations for Selecting an Industrial Cybersecurity Solution for Asset Identification, Threat Detection, and Response. Report. 2019. Available at: https://www.dragos.com/wp-content/uploads/Key-Considerations-Industrial-Cybersecurity-Solution.pdf

Dunn Cavelty M (2013) From cyber-bombs to political fallout: threat representations with an impact in the cyber-security discourse. *International Studies Review* 15(1): 105–122.

Dunn-Cavelty M and Suter M (2009) Public–private partnerships are no silver bullet: an expanded governance model for critical infrastructure protection. *International Journal of Critical Infrastructure Protection* 2(4): 179–187.

Dunn-Cavelty M (2018) Cybersecurity research meets science and technology studies. *Politics and Governance* 6(2): 22–30.

Dwyer AC (2018) The NHS cyber-attack: A look at the complex environmental conditions of WannaCry. *RAD Magazine*, 44.

Dwyer AC (2021) Cybersecurity’s grammars: a more-than-human geopolitics of computation. *Area* 00: 1– 8. doi:10.1111/area.12728

Elish MC (2019) Moral crumple zones: cautionary tales in human-robot interaction (pre-print). *Engaging Science, Technology, and Society (pre-print)* 6: 1–29.

European Commission (2016) NIS Directive. Available at: https://digital-strategy.ec.europa.eu/en/policies/nis-directive

Fouad NS (2021) The non-anthropocentric informational agents: codes, software, and the logic of emergence in cybersecurity. *Review of International Studies* 1–20.

Frey S, Rashid P, Anthonysamy M, et al. (2019) The good, the bad and the ugly: a study of security decisions in a cyber-physical systems game. *IEEE Transactions on Software Engineering* 45(5): 521–536. doi:10.1109/TSE.2017.2782813.

Guldenmund FW (2000) The nature of safety culture: a review of theory and research. *Safety Science* 34(1–3): 215–257.

Haugland BT (2020) Changing oil: self-driving vehicles and the Norwegian state. *Humanities and Social Sciences Communications* 7(1): 1–10. doi:10.1057/s41599-020-00667-9.

Henrie M (2015) Cyber security risk management in the SCADA critical infrastructure environment. 25(2): 38–45. doi:10.1080/10429247.2013.11431973.

Kinsley S (2014). The matter of ‘virtual’geographies. *Progress in Human Geography* 38(3): 364–384.

Klimburg-Witjes N and Wentland A (2021) Hacking humans? Social engineering and the construction of the “deficient user” in cybersecurity discourses. *Science, Technology, & Human Values* 46(6): 1316–1339.

Kazansky B (2021) ‘It depends on your threat model’: the anticipatory dimensions of resistance to data-driven surveillance. *Big Data and Society* 8(1): 1–12. doi:10.1177/2053951720985557

Kocksch L, Korn M, Poller A, et al. (2018) Caring for IT security: accountabilities, moralities, and oscillations in IT security practices. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW): 1–20. doi:10.1145/3274361.

Knorr- Cetina K (1999) *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.

Kriaa S, Pietre-Cambacedes L, Bouissou M, et al. (2015) A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety* 139: 156–178.

Marres N and Stark D (2020) Put to the test: for a new sociology of testing. *The British Journal of Sociology* 71(3): 423–443.

Martin A (2021) UK Cyber security law forcing energy companies to report hacks has led to no reports, despite numerous hacks. *Sky News*. Available at: https://news.sky.com/story/uk-cyber-security-law-forcing-energy-companies-to-report-hacks-has-led-to-no-reports-despite-numerous-hacks-12254296

Matthew A and Cheshire C (2016) Trust and community in the practice of network security. Preprint. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2756244

Michalec OA, Van Der Linden D, Milyaeva S, et al. (2020). Industry responses to the European directive on security of network and information systems (NIS): understanding policy implementation practices across critical infrastructures. In Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020) (pp. 301 – 3317). USENIX, Virtual.

Michalec O, Milyaeva S and Rashid A (2021) Reconfiguring governance: how cyber security regulations are reconfiguring water governance. *Regulation & Governance* 1–18.

Michels JD and Walden I (2018) How safe is safe enough? Improving cybersecurity in Europe’s critical infrastructure under the NIS directive. Queen Mary School of Law Legal Studies Research Paper No. 291/2018, Available at SSRN: https://ssrn.com/abstract=3297470

Murray G, Johnstone MN and Valli C (2017) The convergence of IT and OT in critical infrastructure. *The Proceedings of 15th Australian Information Security Management Conference*. 5–6 December, 2017, Perth: Edith Cowan University, 149–155. doi:10.4225/75/5a84f7b595b4e.

National Cyber Security Centre (2019) Cyber assessment framework guidance. Available at: https://www.ncsc.gov.uk/collection/caf

National Cyber Security Centre (2021) Technology assurance. Guidance. Available at: https://www.ncsc.gov.uk/collection/technology-assurance/future-technology-assurance/whitepaper-developing-a-new-approach-to-assurance

Norton P (2015) Four paradigms: traffic safety in the twentieth-century United States. *Technology and Culture*. 56(2), SPECIAL ISSUE: (Auto)Mobility, Accidents, and Danger, 319 – 3334. Available at: https://www.jstor.org/stable/24468867?seq=1#metadata_info_tab_contents (Accessed: December 16, 2021).

Nurse JRC, Creese S and de Roure D (2017) Security risk assessment in internet of things systems. *IT Professional* 19(5): 20–26. doi:10.1109/MITP.2017.3680959.

Perrow C (1984) *Normal Accidents: Living with High-Risk Technologies*. Princeton University Press, pp. 1–466.

Pieters W and Coles-Kemp L (2011) Reducing normative conflicts in information security. Proceedings of the 2011 workshop on new security paradigms workshop - NSPW ‘11 [Preprint], (11). doi:10.1145/2073276.

Piètre-Cambacédès L and Chaudet C (2010) The SEMA referential framework: avoiding ambiguities in the terms “security” and “safety”. *International Journal of Critical Infrastructure Protection* 3(2): 55–66.

Pinch TJ and Bijker WE (1984) The social construction of facts and artifacts: or how the sociology of science and the sociology of technology might benefit each other. The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology: Anniversary Edition, 11–44. Available at: https://mitpress.mit.edu/books/social-construction-technological-systems-anniversary-edition (Accessed: December 16, 2021).

Pinch T and Swedberg R (2008). *Living in a Material World: Economic Sociology Meets Science and Technology Studies* (Vol. 1). Cambridge, MA: The MIT Press.

Polleri M (2020) Post-political uncertainties: governing nuclear controversies in post-Fukushima Japan. *Social Studies of Science* 50(4): 567–588.

Reece RP and Stahl BC (2015) The professionalisation of information security: perspectives of UK practitioners. *Computers & Security* 48: 182–195.

Renaud K, Flowerday S, Warkentin M., et al. (2018) Is the responsibilization of the cyber security risk reasonable and judicious? *Computers & Security* 78: 198–211.

Schiølin K (2020) Revolutionary dreams: future essentialism and the sociotechnical imaginary of the fourth industrial revolution in Denmark. *Social Studies of Science* 50(4): 542–566.

Shove E and Trentmann F (2018) *Infrastructures in Practice: The Dynamics of Demand in Networked Societies*. New York: Routledge.

Shreeve B, Hallett J, Edwards M, et al. (2020) ‘So if Mr Blue Head here clicks the link…’ risk thinking in cyber security decision making. *ACM Transactions on Privacy and Security (TOPS)* 24(1): 1–29. doi:10.1145/3419101.

Shukla M, Johnson SD and Jones P (2019) Does the NIS implementation strategy effectively address cyber security risks in the UK? in 2019 International Conference on Cyber Security and Protection of Digital Services, Cyber Security 2019. Institute of Electrical and Electronics Engineers Inc. doi:10.1109/CyberSecPODS.2019.8884963.

Slayton R and Clark-Ginsberg A (2018) Beyond regulatory capture: coproducing expertise for critical infrastructure protection. *Regulation & Governance*, 12(1): 115–130.

Slupska J, Dawson Duckworth SD, Ma L, et al. (2021) Participatory threat modelling: exploring paths to reconfigure cybersecurity. Conference on Human Factors in Computing Systems - Proceedings [Preprint]. doi:10.1145/3411763.3451731.

Spinardi G (2019) Performance-based design, expertise asymmetry, and professionalism: fire safety regulation in the neoliberal era. *Regulation & Governance* 13(4): 520–539.

Stilgoe J (2018) Machine learning, social learning and the governance of self-driving cars. *Social Studies of Science* 48(1): 25–56.

Stilgoe J (2021) How can we know a self-driving car is safe? *Ethics and Information Technology* 2021: 1–13.

Suchman L (2007) *Human-machine Reconfigurations: Plans and Situated Actions*. Cambridge, UK: Cambridge university press.

Tanczer LM, Steenmans I, Elsden M, et al. (2018 Emerging risks in the IoT ecosystem: who's afraid of the big bad smart fridge?. In *Living in the Internet of Things: Cybersecurity of the IoT-2018* (pp. 1 – 19). London: IET

Tatam M, Shanmugam B, Azam S, et al. (2021) A review of threat modelling approaches for APT-style attacks. *Heliyon* 7(1): e05969.

Thekkilakattil A and Dodig-Crnkovic G (2015) Ethics aspects of embedded and cyber-physical systems. Proceedings - International Computer Software and Applications Conference, 2: 39–44. doi:10.1109/COMPSAC.2015.41.

Thomas RJ, Gardiner J, Chothia T, et al. (2020) Catch me if you can: an in-depth study of CVE discovery time and inconsistencies for managing risks in critical infrastructures. CPSIOTSEC 2020 - Proceedings of the 2020 Joint Workshop on CPS and IoT Security and Privacy, 49–60. doi:10.1145/3411498.3419970.

Topping C, Dwyer A, Michalec O, et al. (2021). Beware suppliers bearing gifts!: analysing coverage of supply chain cyber security in critical national infrastructure sectorial and cross-sectorial frameworks. *Computers & Security*, 108, 102324.

Urquhart L and McAuley D (2018) Avoiding the internet of insecure industrial things. *Computer Law and Security Review* 34(3): 450–466.

Wallis T and Johnson C (2020) Implementing the NIS directive, driving cybersecurity improvements for essential services. in 2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 1–10. doi:10.1109/CyberSA49311.2020.9139641.

Wuyts K, Sion L and Joosen W (2020) LINDDUN GO: A lightweight approach to privacy threat modeling. Proceedings - 5th IEEE European Symposium on Security and Privacy Workshops, Euro S and PW 2020, 302–309. doi:10.1109/EUROSPW51379.2020.00047.

Wynne B, Waterton C and Grove-White R (2007) Public perceptions and the nuclear industry in west Cumbria. Available at: http://inis.iaea.org/Search/search.aspx?orig_q=RN:34004547 (Accessed: December 16, 2021).