Radical Systems Thinking and the future role of Computational Modelling in Ergonomics

Citation for published version:
Holman, M, Walker, G, Lansdown, T & Hulme, A 2020, 'Radical Systems Thinking and the future role of Computational Modelling in Ergonomics: An Exploration of Agent-Based Modelling', *Ergonomics*, vol. 63, no. 8, pp. 1057-1074. https://doi.org/10.1080/00140139.2019.1694173

Digital Object Identifier (DOI):
10.1080/00140139.2019.1694173

Published In:
Ergonomics

Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

Radical Systems Thinking and the future role of Computational Modelling in Ergonomics: An Exploration of Agent-Based Modelling

Matt Holman, Guy Walker, Terry Lansdown & Adam Hulme

Radical Systems Thinking and the future role of Computational Modelling in Ergonomics: An Exploration of Agent-Based Modelling

Matt Holman [a], Guy Walker [b], Terry Lansdown [c] and Adam Hulme[d]

[a] *Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh, United Kingdom*

[b] *Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh, United Kingdom*

[c] *School of Social Sciences, Heriot-Watt University, Edinburgh, United Kingdom*

[d] *Centre for Human Factors and Sociotechnical Systems, University of Sunshine Coast, Queensland, Australia*

mdh1@hw.ac.uk[a]

g.h.walker@hw.ac.uk[b]

t.lansdown@hw.ac.uk[c]

ahulme@usc.edu.au[d]

Corresponding Author: Matt Holman [a] – mdh1@hw.ac.uk [a]

Funding

Adam Hulme's contribution was supported by an Australian Research Council (ARC) Discovery Project grant (grant number: DP180100806).

Body Text Word Count (minus in-text references, headings and quotes): 6969

Radical Systems Thinking and the future role of Computational Modelling in Ergonomics

Abstract

We are teetering on the precipice of the imminent fourth industrial revolution. In this new age, systems are set to become more densely intraconnected and interconnected, and massive sociotechnical systems exhibiting unprecedented levels of complexity will increasingly take hold. At the dawning of this new age, the Ergonomics discipline must reflect on its preparedness for tackling problems in these novel systems. This paper engages in this reflection by putting forth a critical commentary on the implication of these changes on the discipline and discusses the utility of our current methods in this new paradigm. A resulting Radical Systems Thinking in Ergonomics Manifesto is put forward - a set of mandates to guide practitioners and researchers in the development of new methods capable of coping with these imminent challenges. From the manifesto are derived a series of capability requirements for future computational modelling approaches in Ergonomics.

Keywords: Ergonomics; Computational Modelling; Systems Thinking; AI; IoT

Practitioner Summary

The goal of this paper is to inspire the Ergonomics community to pursue further applications involving computational modelling approaches such as Agent-Based Modelling. It presents a manifesto for the future of the discipline, and from this the capabilities that future computational modelling approaches need to possess.

Introduction

Hyperbole? Or understatement?

Radical: a term "relating to or affecting the fundamental nature of something; far-reaching; characterized by departure from tradition; innovative or progressive." (Oxford English dictionary; 2008). The term radical is often hyperbole. Very few topics in Ergonomics are truly radical. The implications of the dawning Fourth Industrial Revolution (4IR) are different. Here we have a significant departure from tradition which could fundamentally and perhaps irreversibly affect the nature of Ergonomics as a discipline. A technological paradigm shift is upon us and this paper will argue that Ergonomics, as a discipline, has not been radical enough. The reader is invited on a journey through the 4IR and into the radical systems agenda that will be needed to cope with it.

Article structure

Section 1

The article begins with a brief description of the imminent 4IR and discusses some likely implications of 4IR technology. The concept of 'post-dualist sociotechnical systems' is introduced to describe a new paradigm of work system, expressive of the wider changes imposed by the proliferation of 4IR technology and the subsequent transgression of the wider society-technology dualism. The authors argue that it is this radical new paradigm that the discipline must update its methodological toolkit in preparation for, and that not doing so, could result in Ergonomics becoming an afterthought.

Section 2

A description of the post-dualist landscape is provided throughout section 2 and serves as the conceptual grounding within which a series of broad classes of Ergonomics challenges are framed. Section 2.1 outlines Ergonomics challenges associated with intra-system connection, where links between agents *within* sociotechnical systems are becoming increasingly dense. Section 2.2 contends with increasing inter-system connection at a higher level of abstraction, where links *between* sociotechnical systems are becoming denser and more complex. Finally, section 2.3 outlines the challenges associated with forecasting the behaviour of the 'total system of systems' at the highest level of abstraction (see Figure 1). It is in these subsections that examples of the radical challenges posed by 4IR developments such as artificial intelligence and cyber-physical systems are discussed. Collectively, these examples help to demonstrate that it is only by understanding the evolution of sociotechnical systems into the 4IR that we can effectively call for appropriate methodological advancement. Each subsection wraps up with a critical reflection around the implications of the challenges presented therein on the discipline under the subheadings; *Implications for Ergonomic*s. These subsections also explore the utility of computational modelling approaches (such as ABM) for tackling 4IR challenges and for addressing the wider implications for Ergonomics.

[Figure 1 here]

Section 3

The argument for the need to develop a methodology-based Ergonomics strategy is found towards the latter half of this article. If sociotechnical systems are growing in complexity, then there is a facet of our methodological toolkit that should also dedicate itself to the study of complex, emergent phenomena that result from ever more complex interactions. The authors argue that exploring the use of computational modelling approaches such as ABM to supplement or augment current Ergonomics methods, or indeed to provide the theoretical underpinning of completely new approaches, is a practical first step in realising this. Whilst the focus of this article is on ABM, the general message put forward is agnostic to the specific computational modelling approach(es) adopted.

Finally, to facilitate the exploration into how computational modelling approaches can enhance our methodological toolkit, a new strategic direction; *the Radical Systems Thinking in Ergonomics Manifesto* (RSTEM), is promoted. The paper closes by outlining methodological requirements derived from the RSTEM to encourage the development of methods that will suitably equip the discipline for the radical challenges brought by the forthcoming 4IR.

1. The Fourth Industrial Revolution

The first industrial revolution (1IR, 18th century) was the age of iron, steam, and mechanisation, and was the precursor to the second industrial revolution (2IR, 19th century) which heralded a new age of infrastructure, analogue electronics and cybernetics. Ergonomics itself is a product of this revolution. The third industrial revolution (3IR, 20th century) saw the paradigm shift to digital technologies and computing, and for Ergonomics an increasing interest in, and focus on, systems approaches (Dul et al., 2012; Jenkins, 2009; Naikar, Moylan, & Pearce, 2006; Rasmussen, Pejtersen, & Goodstein, 1994; Thatcher & Yeow, 2016). As the discipline begins to meaningfully grapple with systems challenges, the Fourth Industrial

Revolution (4IR, 21st century) is now upon us.

The 4IR is characterized by the emergence of technologies in a diverse range of fields, including robotics, artificial intelligence (AI), blockchain and cryptocurrency, nanotechnology, quantum computing, biotechnology, digital manufacturing/3D printing and autonomous vehicles and others (Schwab, 2017). 4IR technologies are disrupting industries and companies that are not innovating in line with their contemporaries (Ning & Wang, 2011). The breadth and depth of these changes will deliver radical transformations of ‘entire systems of production, management, and governance’ (Schwab, 2017). A cornerstone of the 4IR paradigm is the interconnectivity of everything and everyone via the ‘Internet of Things’ (IoT); a global platform which hosts a network of interconnected ‘online capable devices’ that gather, store, transmit and receive information.

The IoT is already here, and it is growing rapidly. Far more ‘Things’ are currently connected to the internet than people (Evans, 2011) and they are proliferating within all aspects of industry (the Industrial IoT), commerce and domestic life (Ning & Wang, 2011). Approximately 330 million new devices are connected to the internet every month leading to the number of online capable devices increasing by 31% between 2016 and 2017 (Noto La Diega & Walden, 2016). Some estimates suggest that the IoT will consist of approximately 30 billion objects by 2020. The end result is to create a closely coupled sociotechnical ecosystem on a scale not previously seen in the history of mankind. The IoT’s deep complexity, capability for self-organisation, optimisation, and ‘greater than the sum of its parts’ emergent properties, harbour significant benefits. The IoT might of course, also have profound negative consequences for society ranging from reduced social capital through to threats to national security. However, the focus of this article is to explore the implications of the IoT and 4IR on Ergonomics research and practice moving forwards.

1.1 A Perfect Storm?

Although quality of life has gradually improved throughout previous technological revolutions (Coale, 1989; Szreter & Mooney, 1998), the aggregated latent costs are only now becoming clear through the widespread ecological damage caused by industrialisation. The potential contributions of Ergonomics in meeting the challenge of securing a sustainable future, and the alignment of the discipline's values to this endeavour, have indeed been acknowledged (see Thatcher, 2013; Lange-Morales, Thatcher & García-Acosta, 2014; Thatcher, 2017). The authors wish to strongly emphasise that this notion must be reconciled with those espoused in this paper to ensure the discipline's strategic direction maintains alignment with contemporary global challenges. Whilst our struggle to contain the environmental impacts of previous industrial revolutions continues, the inescapable forces of the free market are propelling us headfirst into the radical 4IR technological paradigm. A perfect storm is on the horizon.

1.2 Ergonomics in the 4IR Paradigm

Conceptually, technology-driven paradigm shifts are nothing new. However, the 4IR is differentiated from previous ones by the sheer speed of technological breakthroughs when compared with previous industrial revolutions (Schwab, 2017). Indeed, this is a feature of Moore's law, which posits that computational power doubles every two years and therefore increases exponentially; a trend which has been observed since Moore first put forward claim in 1965 (Moore, 1965; Lundstrom, 2003) (however there is some indication that this may be slowing down (e.g. Sutardja, 2014)). Moreover, the ubiquity of the adoption of 4IR technologies in industry and commerce, coupled with their disruptive nature, will have tremendous impacts on the way we live, work, and relate to one another.

The interrelationship of 4IR technologies and their diffuse adoption is arguably thrusting humanity into an era where the dualism of society *and* technology is becoming increasingly one and the same. A paradigm where "society *is* technology, and technology *is* society" is emerging, as the lines between the physical, digital, and biological spheres are increasingly blurred through adoption of the disruptive technologies (Schwab, 2017). The transgression of this dualist society-technology construct is indeed radical at micro, meso and macro levels of analysis. Individual behaviour, organizations and society-at-large will be transformed by 4IR technology - perhaps irreversibly. As a discipline primarily concerned with human well-being, Ergonomics *should* offer value in response to the 4IR, just as it did in previous industrial revolutions. But this objective poses the thorny follow-up question; what *can* it offer in the 4IR paradigm? Further still, where should Ergonomics begin in attempting to address this question?

2. Ergonomics Challenges in a Post-Dualist Paradigm

2.1 Who is the Apex Agent in a Sociotechnical Singularity?

This section delves into the emerging issues and challenges associated with the *intra*-connecting of human and autonomous agents (programs and/or devices which act autonomously and without supervision) within 4IR sociotechnical systems. The argument is that the intra-connecting of human and non-human agents within sociotechnical systems is shifting us into a paradigm where post-dualist sociotechnical systems will dominate. We argue that this paradigm shift will introduce novel challenges which the Ergonomics discipline must respond to.

The embedment of AI into sociotechnical systems signifies a pivotal transformation of the human-technology relationship from Human Computer Interaction (HCI) to Human Machine Cooperation (HMC) (Hoc, 2001). In the HMC paradigm, executive functions are shared between 'socio' and 'technical' elements. These systems will be underpinned by the most advanced and probably most 'intelligent' technology so far invented, whilst the socio construct (e.g. individuals, teams, management structures, etc.) will remain unchanged apart from the changes perpetuated by the technology itself.

Goal directed AI introduces the problem of *instrumental convergence.* Instrumental convergence is the concept that a human AI system programmed with seemingly harmless goals can act in unexpectedly harmful ways. This problem is animated by Bostrom's infamous 'paperclip maximiser' thought experiment, whereby a hypothetical superhuman AI is designed with the sole goal of making as many paperclips as possible. Staying true to this programmed goal, the superhuman AI system identifies the risk that it can be switched off. Thus, the system decides to eradicate humanity so that it cannot be turned off, allowing it to continue making paperclips to fulfil its programmed goal (Bostrom, 2003). This thought experiment represents the criticality of AI systems design and understanding the catastrophic consequences of seemingly small oversights. The problem of instrumental convergence is one consequence of the transformation of non-human agents from subordinate system components to equal and/or superordinate collaborators. As humans cede the control of an increasing number of system functions to AI agents, effective understanding and cooperation becomes vital.

There is risk in placing confidence on the decisions of AI systems when we do not fully understand how the systems are making those decisions. In tandem, AI operating without the capability to understand complex and abstract human constructs such as ethics and values presents risk. In the interconnected post-dualist 4IR paradigm, where HMC systems abound, reciprocal comprehension of human and AI reasoning and behaviour is paramount (Marathe et al., 2018); understanding is a necessary predicator of effective cooperation. But is designing a system architecture for effective HMC enough? What happens in degraded mode operations where the essential functions AI functions suddenly go 'offline', or when the system imperceptibly migrates towards safety boundaries in normal modes of operation when things go awry?

Implications for Ergonomics

Answers to these challenges are elusive and nuanced. On the one hand, post-dualist sociotechnical systems have the potential to bring unlimited virtue in the 4IR paradigm. On the other hand, the system architecture must allow humans to understand and take back control when necessary so that we can avoid the potentially catastrophic risks associated with runaway AI. In addition to the potentially catastrophic safety risks that suboptimal AI integration could impose, there are indeed less severe, yet still undesirable, consequences that could result. For instance, Hancock (2019) outlines examples of undesirable outcomes caused by dysfunctional cybernetics; scenarios where system feedback loops are maladapted to the objective of facilitating user-centred outcomes. An example provided is a teleservice automatically recording calls for quality control, yet no tangible value is experienced by the user during future calls as a result of this function (Hancock, 2019).

The integration of AI into new and existing systems could spawn a range of Ergonomics issues of escalating magnitude, from the irritating to the life-threatening. Can our current set of Ergonomics methods and approaches provide answers in response to a paradigm where humans are transitioning from agents to subjects within emerging 4IR systems? From a theoretical standpoint, this transitioning of status may even undermine the validity of the original sociotechnical system concept, which is predicated on the interaction between people and technology in work places (Trist, 1978). By default, this may expose a foundational precept which underpins modern systems Ergonomics theory (Hendrick, 1991; Clegg, 2000; Wilson, 2000). The conceivability of human-machine cooperation transpiring into machine-human subjugation provokes a refining of the original question. Perhaps '*can* Ergonomics have the answer*?*', is more apposite to the present issue.

Traditional allocation of function analyses (AoF), a widely used method of inquiry, no longer suffices when the technology is autonomous and perhaps even 'superhuman' in its capabilities. An abundance of Ergonomics research has been published regarding the allocation of function between human and conventional computer components within systems (Lee & Moray, 1992; Muir, 1994; Parasuraman & Riley, 1997; Rognin, Salembier, & Zouinar, 2000) and allocation of function analyses are a common task of the HF practitioner (Dearden, Harrison, & Wright, 2000; Lee & Moray, 1992). There is, however, a relative dearth of Ergonomics research in allocation of function between human and autonomous actors within systems (Hancock, 2017). AoF analyses in the AI HMC paradigm requires awareness of the potentially intractable issues caused by instrumental convergence. This is compounded by the requirement to somehow program fuzzy human concepts such as morals and values into the AI components to create a sociotechnical system unified by shared motives (Hancock, Mouloua, & Senders 2009; Dietterich, Thomas & Horvitz, 2015).

A human apex could manage the perils of instrumental convergence by steering the system away from AI decisions that could ultimately threaten human well-being. However, to halt these decision paths in a timely manner, the proposed human apex must somehow comprehend how seemingly inconsequential AI decisions could result in emergent and cascading failures. Moreover, the human apex must be able to understand these consequences at a rapidity at least equal to rate at which the decisions are being made. The likelihood that a conventional AoF analysis would assign control of the off-switch to the non-human agent forebodes serious questions over the relevancy of Ergonomics in the design of such systems.

The integration of AI and perhaps even superintelligence into human systems poses another issue; in a post-dualist sociotechnical system, who or what exactly are we assigning functions to? In post-dualist sociotechnical systems, functions will likely evolve over time as AI and human components continually self-adapt to optimise performance. Traditional Ergonomics AoF analyses which assign functions to largely static components are simply not suited to the inherent dynamism of a post-dualist AI sociotechnical system. The complex and sophisticated challenges nested in these systems warrant the development and deployment of equally sophisticated Ergonomics methods. It seems clear that our traditional approaches, and indeed even our more sophisticated systems methods such as Cognitive Work Analysis (CWA) (Rasmussen et al., 1994; Vicente, 1999), Event Analysis of Systemic Teamwork (EAST) (Walker et al., 2006) and Systems Theoretic Accident Model and Processes (STAMP) (Leveson, 2004) may not go far enough. Or if they do, the practical problems of scaling them to account for massive sociotechnical systems could be an insurmountable barrier in their future uptake. Computational modelling approaches show great promise for this analytical purpose; their facilitation of rapid results generation (Benoit, 2001), scaling potential and capability for modelling a multitude of agent types each with their own goals (Silverman, 2018), constraints and functional dependencies offer the required analytical insight that our current, largely pen and paper, methods do not.

In the Ergonomics literature, there has only been one article that has dedicated its central thesis to the potential benefits of computational modelling for sociotechnical systems design and safety. Hettinger et al. (2015) offer an overview of why ABM is useful to Ergonomics researchers and practitioners and argue the case that such approaches can help to elucidate the immediate impact(s) of making changes to certain system structures and properties. Methodologically, ABMs contain individual micro entities or ‘agents’ that interact based on a number of rules that are assigned *a priori*. Agents are autonomous and goal-directed and can change their internal states to adjust to wider environmental conditions. As a result, the behaviour of the system is governed largely by the collective interaction among self-driven living and/or non-living agents who can exhibit characteristics and features such as memory and intelligence (Hulme et al., 2019).

In drawing on a hypothetical example, the analytical power offered by ABM could allow evaluation of the impact of integrating an AI element before its instantiation within a system. For example, the testing of autonomous vehicle ‘agents’ long before they are introduced into real complex road network systems is an area where computational modelling should be considered. This is especially pertinent at present given the interest around the potential harms of autonomous vehicles, which cannot yet make decisions that match those of a human being in certain situations that require more pragmatism than rule-based programming (Schaefer, Chen, Szalma & Hancock, 2016).

Given the financial and logistic implications associated with ‘real’ experiments, the use of simulation-based methods is highly attractive as useful results can be ascertained without ever having left the lab/research environment. Wilson (2014) argues that departure from a ‘laboratory’ approach is necessary when adopting ‘systems’ Ergonomics research and practice and emphasizes the importance of context for studying and understanding complex interactions within systems. Again, ABM approaches may show promise in this area, as contextual factors can be represented by programmable system parameters. The potential analytical insights afforded by ABM approaches could be crucial given the potentially catastrophic consequences of designing and integrating an AI system without due consideration of context. Whilst the potential of applying computational modelling approaches for AoF has been touched upon by Hettinger et al. (2015), they acknowledge that no current model-based AoF approaches exist (Hettinger et al., 2015). Thus, there is a strong case for the development of an ABM-based AoF approach, capable of predicting emergent phenomena resulting from local human-AI interactions for effective HMC. Analyses of this type could be integral to the appropriate allocation of tasks, controls and affordances between the human and the AI elements, to ensure safe and reliable system performance.

Despite their promise, there is evidence of the failure of adoption of ABM and similar simulation approaches within organisations. A study by Brailsford et al. (2013) aimed to understand the causes of failure of simulation adoption in the NHS. Among the findings were issues related to the need for extensive training, reliance on few trained individuals and dependence on availability of IT equipment. Whilst this study is not directly related to ABM adoption, the results may extrapolate in a broader sense to reveal a potential logistical barrier in their uptake. The combination of these with the oft-time pressured consulting environment within which many Ergonomics practitioners work may represent a significant barrier in the adoption of these methods.

2.2 Cyber-physical system of systems

The inter-connecting of ostensibly disparate systems is being facilitated by the spread of IoT- based Cyber-Physical systems. Cyber Physical Systems (CPSs) are highly sophisticated systems characterised by the synergy of computational and physical processes, where various classes of components including hardware, software, and other computational elements seamlessly integrate to form a connected whole (Colombo et al., 2017; Jazdi, 2014; Lee, 2015; Leitao et al., 2016). Instantiations of CPSs abound throughout a range of sectors, with examples including smart energy grids, autonomous automobile systems, networked medical monitoring equipment, process control systems, and robotics systems (Loukas, 2015).

The rapid and widespread adoption of IoT devices is ushering in an era where cloud-based CPSs will likely dominate industrial, commercial and domestic spheres (Colombo et al., 2017). Cloud-based industrial CPSs have already materialised in the transport, energy and manufacturing sectors (Colombo et al., 2014; Wahlster, 2012; Wu, Rosen, Wang, & Schaefer, 2015). The breadth and depth of this IoT 'system of systems' will enable a multiplicity of routes for local perturbations to cascade into massive systemic effects. The 2016 Dyn cyberattack offered an alarming insight into the vulnerability of the increasingly vast and interconnected IoT network. Approximately 100,000 IoT devices (including printers, baby monitors, security cameras, and electronic gates) were hacked using Mirai, a type of malware that turns networked devices into 'bots' whose behaviour can be coordinated into a 'botnet' to carry out Distributed Denial of Service (DDoS) attacks to disrupt the normal flow of internet traffic (Hilton, 2016; Scott Sr & Summit, 2016). The Dyn DDoS attacks affected major service platforms such as Facebook, Amazon and Twitter, including the web services of major networks such as BBC, Fox News and The Guardian. Whilst the Dyn attack provided an alarming insight into the vulnerability of networked IoT systems, its impacts were not directly couched in 'physical space' and were thus largely benign (James & Spaniel, 2016). Arguably, the symbolic impact of how easily IoT devices were compromised was far greater, as it brought to attention the viability of a novel and disturbing class of threat; Cyber-Physical Attacks (CPA)s.

Cyber-physical attacks are cyber-attacks which manifest impacts in physical space. CPAs are perpetrated through the compromising of cyber-components embedded within networked IoT devices and systems. Once compromised, these devices can be controlled and operated to directly affect physical space. Components embedded in critical infrastructure systems such as smart grids, gas pipelines and transport networks are obvious targets, as are components embedded in simpler but equally critical systems such as medical devices, vehicles and industrial machinery (Loukas, 2015).

Components in 'high impact' critical infrastructure are not the only target. Even seemingly benign household devices can be compromised to initiate a CPA - in the 4IR paradigm, the thermostat is not just a simple temperature regulating device; it may also be the tool of choice for the arsonist. Perhaps the most alarming prospect is the potential for simple household IoT devices to be hacked and coordinated to instigate a CPA at a massive scale; as indicated by the 2016 Dyn attack, simple devices could be hacked and coordinated to initiate far reaching CPAs. Smart cities, intelligent transport networks and other networked mega-systems are particularly vulnerable in this radical new paradigm.

Implications for Ergonomics

Transgression of the cyber-physical dualism means that the potential impacts of CPAs, and the routes for perpetrating them, are virtually limitless. The expanding cyber-physical network of IoT technology and devices provide a complex substructure where emergent systemic risks can propagate rapidly and imperceptibly. Whilst the concept of latent threats is well grounded in Ergonomics and has formed the basis of many of our methods (e.g. TRACEr (Shorrock and Kirwan; 2002), HERA (Kirwan; 1998b), HET (Stanton et al.; 2009)) the scale and complexity of interconnected CPSs are simply too great for these reductionist approaches to provide meaningful insights (Waterson et al., 2015; Salmon, Walker, Goode, & Stanton, 2017; Walker, Stanton, & Jenkins, 2017).

What can Ergonomics offer up in response to the scale and complexity of interconnected CPSs? The assiduous practitioner would likely turn to scalable formative methods such as CWA, Functional Resonance Analysis Method (FRAM) (Hollnagel, 2017) and EAST which focus on system constraints, functional dependencies, team performance or the exploration of novel behaviours (Kant, 2017; Naikar, 2017). However, the sheer scale of networked 4IR systems (and indeed the emerging global system of systems) would make the use of these approaches insufficient for understanding the time-dependent nature of the complex interactions that define overall system behaviour. Even if unlimited time and labour were to be granted for analysis, it is not apparent that these approaches could offer exhaustive and meaningful insights throughout CPSs of such an unprecedented scale and complexity. For example, the inherent dynamism of an interconnected CPS would mean that system properties change at a rate far greater than what could be reasonably modelled through these analyses. Moreover, the multiplicity of dynamic interactions throughout multiple levels of abstract system hierarchy and the various timespans across which these occur is may be immeasurable.

Whilst the aforementioned 'systems' approaches were conceived to analyse 'systems' and are, therefore, by their definition, anti-reductionist, they are still predicated on the ontological dualism of systems being comprised of constraints *and* local behaviours. The delineation between these higher-order constraints and local behaviours, (i.e. meso and micro ontologies) become less and less meaningful as cyber and physical space converge into a unified construct. Integration of 4IR technology propounds a tightness of coupling between phenomena across these traditionally distinct ontological strata, where local behaviours can impact higher-order constraints, and vice versa at an almost imperceptible rate. In this paradigm, their delineation and treatment as separate constructs is almost meaningless; they are virtually indivisible (see Figure 2).

[Figure 2 here]

The suspicion is that in the 4IR post-dualist paradigm where massive CPSs abound, even our most sophisticated Ergonomics 'systems' methods will be limited both practically and ontologically. There are currently no methods in the Ergonomics toolkit that can simulate complex systems behaviours and dynamically understand how systems evolve/change over time.

As already mentioned, computational methods such as ABM have demonstrated their capability to provide forecasts and insights into the behaviour and dynamics of large non-linear human systems (Bonabeau, 2002; Berry, Kiel & Elliot, 2002). ABM offers the analyst the capability to study complex systems by adjusting system rules and parameters, allowing them to derive operational hypotheses of the modelled system (or parts of it) prior to its implementation. Augmenting or supplementing traditional Ergonomics methods with agent-based simulations could therefore provide the analytical enhancement needed for Ergonomics practitioners to address emergent risks within systems Ergonomics problems in the 4IR age. The ability to predict and/or understand certain system patterns and behaviours will become ever more challenging as technological complexity proliferates. Methods such as ABM that have the capability to model the dynamic interactions between people, artefacts, technologies, services, and policies are required, especially given that small changes in, and between complex systems can often result in larger, and sometimes unintended consequences. Moreover, ABMs can be theoretically scaled to match the size of any Ergonomics problem. These types of insights, applied to large scale systems, are critical for designing resilient systems in the 4IR paradigm especially in account of the vulnerability of CPSs as exposed by the 2016 Dyn cyber-attacks. As a type of in silico experiment that can test a range of hypothetical ‘what if’ scenarios, there is a pressing need to explore ABM and other computational modelling approaches in the Ergonomics space.

Encouragingly, research has already been carried out in this area. For instance, Patriarca et al. (2017) introduced a computational extension of the FRAM. This study involved supplementing FRAM with a Monte Carlo analysis to assess potential system and performance variability in Air Traffic Management. Furthermore, combinations of FRAM with other quantitative approaches have been attempted such as Fault Tree Analysis (Toroody, 2016) and Bayesian Networks (Toroody, 2017). Whilst these applications show promise, the FRAM is predicated on functions and interactions being known prior to the analysis. This may limit the utility of computational and quantitative extensions of FRAM for analysing 4IR systems which, due to their high complexity, are becoming increasingly opaque.

2.3 Situated AI and the Distributed Singularity

The technological singularity is the hypothesis that the unleashing of AI systems with superhuman intelligence will set off a cascade of exponential technological growth, resulting in irreversible changes to human civilization (Muehlhauser & Salamon, 2012). The singularity hypothesis posits that a superhuman AI system would self-amplify its power and functional capabilities. This positive feedback loop (i.e. more computing power means more rapid self-optimisation and so on) would scale super linearly and could be barely perceptible to humanity. It is argued that a technological singularity, could bring forth bottomless virtue, or it could be an extinction scenario (Vinge, 1993; Kurzweil, 2014; Bostrom, 2003; Goertzel, 2012; Hancock, 2017).

Notions of this scenario are quelled by criticisms of the hypothesis. For instance, the theory of embodied cognition (Clark, 1999; Steels & Brooks, 2018) offers hope. The theory of embodied cognition posits that true intelligence cannot exist separately from a body and rather, can only be instantiated via some sort of physical interface (i.e. through a ‘body’). Without this interface, the ‘intelligence’ has limited functional utility in the ‘real world’ as it cannot gather and maintain sufficiently accurate models of the world and would thus pose minimal threat. The theory of embodied cognition underpins Brooks’ theory of Situated AI (SAI) (Brooks, 1990). SAI postulates that true intelligence can only emerge organically through interaction with the ‘real world’. Brooks argues that drawing discrete representations of the ‘real world’, from the ‘real world’, and assembling these representations into a unified construct generates and sustains an emergent, real-time dynamic representation of the ‘real world’. Brooks argues that this emergent construct allows the system to operate effectively in complex environments.

More pragmatically, the emergence of a real-world representation would be dependent on the continuous sensing and monitoring of the external environment through the modular array of information gathering sensors. Information drawn from these distributed sensors converge to create an emergent embodied cognition (see Figure 3). Proponents of the SAI approach who oppose notions of the technological singularity hypothesis argue that humanity is perhaps decades away from creating a physical interface sophisticated enough to achieve true intelligence and thus, the technological singularity scenario is extremely unlike to occur (Brooks, 1990; Heylighen, 2012; Steels & Brooks, 2018).

[Figure 3 here]

Whilst the theory of SAI provides some assurance against the likelihood of a technological singularity in the near future, it is difficult to ignore the resemblance of the distributed network of IoT capable devices to a 'global SAI' system of sorts. A global SAI, enabled by the breadth and depth of the pervasion of IoT devices which gather and transmit data, represents a networked system of decomposed intelligent agents whose collective activity creates a form of emergent intelligence. This idea (which, in fact, predates 4IR technology) has been referred to as the Global Brain (Bernstein, Klein, & Malone, 2012; Mayer- Kress & Barczys, 1995). In this concept, the requisite 'body' already exists as the totality of human and technological 'sensors' working in tandem to continuously gather and transmit information about the real world, whilst the internet is the functional architecture which links these 'sensors' into an interconnected network. The Global Brain hypothesis neatly circumvents the premise that true intelligence requires equally powerful artificial sensorimotor systems; they are already here.

The accelerating uptake of IoT devices, and subsequent spread of the Global Brain could induce a 'distributed singularity', which adapts by propagating the technologies, ideas, links, institutions, and agents exhibiting the most utility for its self-growth and preservation. In this hypothesis, this network would adapt by coordinating the activities of its agents to increase its own intelligence and accelerate the expansion of the network (Heylighen et al., 2012). Arguably, this hypothesis is already an actuality. The interconnected network of IoT devices is continuously adapting by tracking consumer behaviour and delivering targeted advertising to encourage the purchasing of more IoT devices, accelerating both the spread and density of the network. Furthermore, market forces are compelling companies in virtually every industry to innovate or fail. Is the Global Brain inevitable? This emergent meta-system comprising the totality of post-dualist sociotechnical systems, simultaneously holds an accurate and current representation of the world and, perhaps most ominously, of itself.

Implications for Ergonomics

The possibility that an emergent massive and singular meta-system will serve as the context in which Ergonomics problems will reside could pose an intractable challenge to the discipline. To provide value in this context, Ergonomics practitioners would require methodologies capable of analysis which could be scaled to massive systems spanning vast geographical regions. Challenges associated with the exploration of Ergonomics problems which could arise as a result of the seemingly intelligent behaviour of an embodied meta-system could be accounted for through the application of computational modelling approaches. For instance, ABM could be used to create simulations of rule-based agents to provide a powerful model of the emergent behaviour exhibited by this meta-system. Discussions on the practicalities of creating a model of such scale are beyond the scope of this paper, however the use of ABM could provide a distilled view of a complex global system from which Ergonomics insights not otherwise realised through traditional analyses could be gleaned. Whilst there are fundamental constraints associated with scaling computational models to analyse global systems, agent-based models pitched at a global scale have indeed been attempted. For example, Cederman (1997) used ABM to model the behaviour of nation states. Cederman assigned drives and motivations to the nation states (agents), and explored their patterns of interaction in an attempt to elicit higher-order emergent phenomena. ABMs comprising billions of agents have also been generated, with Parker and

Epstein's (2011) development of the 'Global-Scale Agent Model' to simulate the outbreak of a disease epidemic and Callegari et al's (2013) development of a global model applied across evolutionary time spans to analyse human population dispersion (Callegari et al's (2013).

These examples indicate that sheer scale should not, at least theoretically, preclude the use of computational modelling approaches such as ABM to derive Ergonomics insights from massive sociotechnical systems should they emerge. In the meantime, it is these types of higher-order insight that could provide Ergonomics practitioners with the analytical power to tackle pressing yet elusive macro-Ergonomics problems such as organisational drift, work as done vs work as imagined and emergent risks (e.g. Dekker, 2016).

For instance, Read et al. (in press) propose an approach which combines ABM with CWA to identify and manage emergent risks associated with the integration of autonomous vehicles on to legacy road networks. Approaches which exploit the strengths of computational modelling and Ergonomics systems methods could yield a variety of insights not possible using traditional Ergonomics methods in isolation, such as the prediction of the impact of emergent phenomena on higher-order systems constraints. Agent-based simulations could be used to understand emergent phenomena as a result of local agent behaviour and interaction. The analyst could then use these insights to explore the wider impacts of emergence on the configuration of the systems constraints using CWA. The resultant changes to system constraints could then be reprogrammed as new agent rules to model changes to local behaviours, yielding new emergent properties. This approach could, theoretically, capture the recursive interplay between local behaviour, emergent phenomena and system constraints, potentially providing a variety of insights analogous to problematic Ergonomics concepts such as organizational drift and the migration of systems towards safety boundaries (see figure 4).

[Figure 4 here]

A key benefit of this approach, and one that makes it apt to take on the challenges described throughout this paper, is the potential to model system dynamism - a feature of complexity that our current systems Ergonomics methods are not well suited to. Along with aiding the discipline in tackling current intractable problems, these types of insight are paramount to the discipline's contribution towards the design and integration of safe systems in the 4IR paradigm. The manifold risks associated with the challenges described, highlight the pressing need to predict and understand the impact of emergent phenomena. The approach put forward by Read et al. (in press) is an excellent example of how Ergonomics methods can be augmented with computational modelling to provide greater insight into current and future complex sociotechnical systems.

Further research is required to explore how current and future Ergonomics methods can be underpinned by computational modelling approaches such as ABM to simulate and analyse the features of the emerging meta-system of systems and the densely connected sociotechnical systems nested within it. In the absence of such capability, Ergonomics may find itself in less and less demand, as the utility diminishes in the face of ever-increasing complexity of 4IR sociotechnical systems (Salmon et al., 2017). This leaves us questioning why a client seeking value-adding insights into a ‘radical’ 4IR problem would turn to Ergonomics for the answer. The lack of an immediate and credible justification for engaging Ergonomics practitioners to tackle ‘systems problems’ in the 4IR technological paradigm is ominous. The unfortunate reality for Ergonomics practitioners is that traditional Ergonomics methods may be simply inadequate, both practically and ontologically, to tackle problems residing in post-dualist sociotechnical systems throughout the 4IR. Whilst ABM and indeed other computational modelling methods should not be treated as a panacea, exploration of their use to supplement or augment our current systems methods, or underpin new methods represent a possible antidote to our current sticking points and to the novel challenges associated with deeply complex 4IR systems.

3. State of the science

What value can Ergonomics methodology offer in an era dominated by post-dualist 4IR sociotechnical systems? The answer that can be gleaned from the sections above is, “not much”. A growing body of research is already conveying doubts about the fitness for purpose of Ergonomics methods for prevailing 3IR systems, let alone what will undoubtedly come in the 4IR (Cornelissen, Salmon, McClure, & Stanton, 2013; Dekker, 2013; Salmon et al., 2017; Stanton, 2014).

If this body of research paints a worrying picture now, it paints an even more serious one for the future. The ontological incompatibility of our 2IR/3IR methodology with the post-dualist 4IR paradigm potentially threatens the validity, and indeed the existence, of the discipline as a whole. Reasons to believe that insights offered by traditional Ergonomics methods would be in demand, or indeed useful, in the 4IR paradigm are sparse.

Whilst some recent Ergonomics research acknowledges the implications of the 4IR and a broad, if undefined, call to arms has been instituted (Dekker, Hancock, & Wilkin, 2013; Hancock, 2017; Salmon et al., 2017; Thatcher & Yeow, 2016; Walker, 2015; Walker, Salmon, Bedinger, & Stanton, 2017), there has been little concurrent progress in the development of new methodology with the sophistication required to cope with complexity inherent in 4IR technological systems.

3.1 Strategic directions

Over the years, several attempts to define future strategic directions in Ergonomics have been made in response to global challenges (albeit, some of these have been implicit). Moray's 1995 paper *Ergonomics and the global problems of the twenty-first century* (Moray, 1995), contended with emerging global issues contemporaneous with its time, including; water issues; food issues; energy issues; urbanisation; violence and terrorism; pollution and waste; and health and medicine. In response, Moray (1995) put forward the following 'programme' for Ergonomics;

- "Design for people's needs rather than their wants. Do not seek to support maximized economic growth.
- Perform conceptual task analyses on the problems that are present and approaching, and develop a new understanding of the social, cultural, and individual well-springs of behaviour, and co-operate with other social science disciplines.
- Use the knowledge of cognitive and motivational psychology in co-operation with engineers and economists to design appropriate 'affordances'.

- Design the negative feedback control systems needed at all levels and time scales of society including 'sensors', and 'effectors' to control collective behaviour."

(Moray, 1995)

A recent review by Thatcher et al. (Thatcher, Waterson, Todd, & Moray, 2017) concluded that of Moray's seven global challenges, energy and health and medicine issues are the only challenges to have been overtly addressed by Ergonomics over the past twenty-five years and remark that the remaining issues 'have become more interrelated and complex, requiring different approaches from the Ergonomics discipline' (Thatcher et al., 2017).

For the most recent and still prevailing Ergonomics strategy, we turn to Dul et. al. (2012). Dul et. al (2012) outline a strategy for the discipline with the following objective: "*To strengthen the demand for and the application of high-quality HFE*" (Dul et. al, 2012; page 389). To achieve this aim, Dul et. al (2012) propose a series of strategic actions. These are shown in Table 1 below.

[Table 1 here]

Dul et. al (2012) acknowledge the importance of aligning strategy with developments in the 'external world'. They list the following six developments as the key challenges which Ergonomics must respond to;

- "Global change of work systems;
- Cultural diversity
- Ageing
- Information and communication technology (ICT)
- Enhanced competitiveness and the need for innovation
- Sustainability and corporate social responsibility."

(Dul et al., 2012)

Whilst their strategic actions (see Table 1) are valid and appropriate for the developments in the 'external world' contemporary of their time, they were arguably conceived at the dawning of a new 'wave of innovation'. Figure 5 illustrates the concept of technological innovations arriving in 'waves' of increasing frequency and magnitude (and accounts for the seemingly exponential evolution of the 4IR). Pondering on the 'external developments' which Dul et. al (2012) predicate their strategy on, it can be argued that the strategy is couched somewhere in the 5th wave of innovation, which is dominated by digital technology characteristic of the third industrial revolution. As the discipline continues to navigate the transition from the 2IR to the 3IR, the 4IR is already arriving. The increasing frequency and magnitude of innovations has outpaced even our most recent overt Ergonomics strategy.

[Figure 5 here]

Moreover, whilst Dul et al. (2012) appreciate the need for Ergonomics to confront system complexity and acknowledge the recent progress in doing so, they do not explicitly address the contrasting stagnation of methodological development – a theme also brought to attention by Thatcher et al. (2018). A related criticism is their failure to discuss how Ergonomics methods should align with sociotechnical values to maximise their contribution towards optimal system design – a proposition put forward by Read et al. (2015). These omissions are especially pertinent as we enter an era where methodological innovation is needed now more than ever. To illustrate this point, Figure 6 plots the number of Ergonomics methods developed per decade, from the 1950s to the 2010s (Stanton et al., 2017)[1]

[Figure 6 here]

The spike of development in the 1980s and 1990s reflects where the discipline's core analytical capability lies; in the 2IR/3IR paradigm. As discussed at length throughout section 2, our 2IR/3IR methods could be in less and less demand in a post-dualist 4IR paradigm and thus, in order for the strategy put forward by Dul et al. (2012); "*to strengthen the demand for and the application of high-quality HFE"* to be actualised in the 4IR, the development of appropriate methodology is imperative.

3.2 The Radical Systems Thinking in Ergonomics Manifesto

Inspired by both Dul et al's strategy (2012) and the broad call to arms put forward by recent Ergonomics research (Dekker et al., 2013; Davis et al., 2014; Hancock, 2017; Salmon et al., 2017; Thatcher & Yeow, 2016; Walker, 2016; Walker et al., 2017) this paper concludes by putting forth the '*Radical Systems Thinking in Ergonomics*

[1] Only methods with discrete authors and dates are included in this table. The resultant data represents 79 out of the 104 methods described in (N. A. Stanton et al., 2017).

Manifesto (RSTEM) in the attempt to reattune the strategic direction of the discipline with the grand challenges of the 4IR era.

The RSTEM operates under the well-grounded presupposition that current Ergonomics methodologies do not tend to cope well with ‘systems Ergonomics’ problems nested in highly complex sociotechnical systems (Woods & Dekker, 2000; Walker et al., 2017; Salmon et al., 2017). In keeping with the moral pursuit of a harmonious and sustainable future for all (Dekker et al., 2013; Thatcher & Yeow, 2016), the RSTEM places an injunction on the discipline to develop the capability of coping with the unprecedented complexity exhibited by the systems and challenges described throughout section 2.

The core tenets of the RSTEM are as follows;

- To face up to the emerging radical systems of the 4IR by adopting a commensurate ‘radical systems thinking’ approach to Ergonomics practice
- To explore how computational modelling methods (e.g., ABM) could supplement and/or augment our existing suite of systems methods in preparation for the paradigm shifting challenges presented by the 4IR
- To inspire the development of new methods, underpinned by computational modelling approaches, capable of coping with broad classes of problems of a potentially ‘massive’ scale likely to arise in the envisioned future.

Drawing directly from the core tenets of the RSTEM listed above, and in account of broad challenges presented throughout this paper, an initial set of capability requirements for new methods is set out below;

- Methods should be capable of modelling interaction between humans and autonomous agents, including the effects of this interaction on the whole system, for safe and effective allocation of function in post-dualistic sociotechnical systems
- Methods should be capable of modelling emergence in complex interconnected 4IR systems to forecast and address latent risks
- Methods should be capable of performing multilevel analysis to account for the scale of massive sociotechnical systems.

It is envisioned that Ergonomics methods developed in line with these requirements will enable the discipline to stay consistent with its core purpose of optimising human well-being and overall system performance throughout the 4IR.

4. Conclusion

The grand challenges outlined in this paper are deliberately broad. Statement of specific issues has been intentionally avoided. Myriad Ergonomics problems will emerge from scales of human-technology systems from 1-to-1 interaction, to the embedding of 4IR technologies into organizations and sociotechnical systems, all the way up to the new technology driven metaphysics of society-at-large. In response to the increasing complexity of systems at these broad levels of abstraction, Ergonomics should not equip itself for specific problems, but rather, for broad classes of contemporary problems and those likely to arise in the envisioned future. This proactive approach gives confidence that additions to our new methodological toolkit could be applied to a range of specific problems as and when they arise.

The essence of the RSTEM, and the derived methodological requirements, is that as problems evolve, so should our methodological approaches. The authors strongly urge Ergonomics researchers and practitioners alike to adopt and promote the core tenets of the RSTEM. Without the rapid development of appropriate methodology, Ergonomics could be in danger of becoming an afterthought in the 4IR. The complexity instantiated by 4IR technologies seems to be outpacing the realisation of Dul et. al's strategic direction. As the status of Ergonomics as a value-adding discipline comes into question, the reformist animus driving Dul et. al's aim 'to strengthen the demand for and the application of high-quality HFE' becomes ever dimmer. Irrespective of the 4IR and the growing levels of sociotechnical complexity, the time has come to broaden our methodological horizon and embrace new challenges. The imperative brought forth by the RSTEM seeks to ignite this, with the long-term aim of sealing Ergonomics as an indispensable practice in the 4IR paradigm. New types of Ergonomics methods are needed now more than ever and the de facto RSTEM for of the future needs to be radical. It's not hyperbole, it is fact. It may actually be an understatement.

Bibliography

Bernstein, Abraham, Mark Klein, and Thomas W. Malone. 2012. "Programming the global brain." *Communications of the ACM*55 (5): 41-43.

Berry, Brian JL, L. Douglas Kiel, and Euel Elliott. 2002 "Adaptive agents, intelligence, and emergent human organization: Capturing complexity through agent-based modeling." *Proceedings of the National Academy of Sciences* 99 (3): 7187-7188.

Benoit, Kenneth. 2001. "Simulation methodologies for political scientists." *The political methodologist* 10 (1): 12-16.

Brailsford, Sally C., Timothy B. Bolt, Giovanni Bucci, T. M. Chaussalet, N. A. Connell, Paul R. Harper, Jonathan H. Klein, Martin Pitt, and M. Taylor. 2013. "Overcoming the barriers: a qualitative study of simulation adoption in the NHS." *Journal of the Operational Research Society* 64 (2): 157-168.

Bonabeau, Eric. 2002. "Agent-based modeling: Methods and techniques for simulating human systems." *Proceedings of the national academy of sciences* 99 (3): 7280-7287.

Bostrom, Nick. 2003. "Are we living in a computer simulation?." *The Philosophical Quarterly* 53 (211): 243-255.

Brooks, Rodney A. 1990. "Elephants don't play chess." *Robotics and autonomous systems* 6 (1-2): 3-15.

Callegari, Simone, John David Weissmann, Natalie Tkachenko, Wesley P. Petersen, George Lake, Marcia Ponce De Leon, and Christoph PE Zollikofer. 2013. "An agent-based model of human dispersals at a global scale." *Advances in Complex Systems* 16 (4-5): 1350023.

Cederman, Lars-Erik. 1997. *Emergent actors in world politics: how states and nations develop and dissolve*. Vol. 2. Princeton University Press.

Clark, Andy. 1999. "An embodied cognitive science?." *Trends in cognitive sciences* 3 (9): 345-351.

Clegg, Chris W. 2000. "Sociotechnical principles for system design." *Applied ergonomics* 31 (5): 463-477.

Coale, Ansley J. 1989. "Demographic transition." In *Social economics*, pp. 16-23. Palgrave Macmillan, London.

Colombo, Armando W., Thomas Bangemann, Statmatis Karnouskos, Jerker Delsing, Petr Stluka, Robert Harrison, Francois Jammes, and Jose L. Lastra. 2014. "Industrial cloud-based cyber-physical systems." *The IMC-AESOP Approach* 22.

Colombo, Armando W., Stamatis Karnouskos, Okyay Kaynak, Yang Shi, and Shen Yin. 2017. "Industrial cyberphysical systems: A backbone of the fourth industrial revolution." *IEEE Industrial Electronics Magazine* 11 (1): 6-16.

Cornelissen, Miranda, Paul M. Salmon, Roderick McClure, and Neville A. Stanton. 2013. "Using cognitive work analysis and the strategies analysis diagram to understand variability in road user behaviour at intersections." *Ergonomics* 56 (5): 764-780.

Davis, Matthew C., Rose Challenger, Dharshana NW Jayewardene, and Chris W. Clegg. 2014. "Advancing socio-technical systems thinking: A call for bravery." *Applied ergonomics* 45 (2): 171-180.

Dearden, Andy, Michael Harrison, and Peter Wright. 2000. "Allocation of function: scenarios, context and the economics of effort." *International Journal of Human-Computer Studies* 52 (2): 289-318.

Dekker, Sidney WA. 2013 "On the epistemology and ethics of communicating a Cartesian consciousness." *Safety science* 56: 96-99.

Dekker, Sidney WA, Peter A. Hancock, and Peter Wilkin. 2013. "Ergonomics and sustainability: towards an embrace of complexity and emergence." *Ergonomics* 56 (3): 357-364.

Dekker, Sidney. 2016. *Drift into failure: From hunting broken components to understanding complex systems*. CRC Press.

Dietterich, Thomas G., and Eric Horvitz. 2015. "Rise of concerns about AI: Reflections and directions." *Communications of the ACM* 58 (10): 38-40.

Dul, Jan, Ralph Bruder, Peter Buckle, Pascale Carayon, Pierre Falzon, William S. Marras, John R. Wilson, and Bas van der Doelen. 2012. "A strategy for human factors/ergonomics: developing the discipline and profession." *Ergonomics* 55 (4): 377-395.

Evans, Dave. 2011. "The internet of things: How the next evolution of the internet is changing everything." *CISCO white paper* (2): 1-11.

Goertzel, Ben. 2012. "Should humanity build a global AI nanny to delay the singularity until it's better understood?." *Journal of consciousness studies* 19 (1-2): 96-111.

Hancock, P. A. 2019. "The humane use of human beings?." *Applied ergonomics* 79: 91-97.

Hancock, Peter A., Mustapha Mouloua, and John W. Senders. 2009. "On the philosophical foundations of the distracted driver and driving distraction." *Driver distraction: Theory, effects, and mitigation*: 11-30.

Hancock, Peter. A. 2017. "Imposing limits on autonomous systems." *Ergonomics* 60 (2): 284-291.

Hendrick, Hal W. 1991. "Ergonomics in organizational design and management." *Ergonomics* 34 (6): 743-756.

Hettinger, Lawrence J., Alex Kirlik, Yang Miah Goh and Petert Buckle. 2015. "Modelling and simulation of complex sociotechnical systems: Envisioning and analysing work environments". *Ergonomics*, 58 (4):600-614.

Heylighen, Francis. 2012 "Brain in a vat cannot break out." *Journal of Consciousness Studies* 19 (1-2): 126-142.

Heylighen, Francis, Evo Busseniers, Viktoras Veitas, Clément Vidal, and David R. Weinbaum. 2012. "Foundations for a Mathematical Model of the Global Brain: architecture, components, and specifications." (4)

Hilton, Scott. 2016. "Dyn analysis summary of friday october 21 attack." *URL: https://dyn. com/blog/dyn-analysis-summary-of-friday-october-21-attack/*.

Hoc, Jean-Michel. 2001. "Towards a cognitive approach to human–machine cooperation in dynamic situations." *International journal of human-computer studies* 54 (4): 509-540.

Hollnagel, Erik. 2017. *FRAM: the functional resonance analysis method: modelling complex socio-technical systems*. CRC Press.

Hulme, Adam, Jason Thompson, Rasmus Oestergaard Nielsen, Gemma JM Read, and Paul M. Salmon. 2019. "Towards a complex systems approach in sports injury research: simulating running-related injury development with agent-based modelling." *British journal of sports medicine* 53 (9): 560-569.

Jazdi, Nasser. 2014. "Cyber physical systems in the context of Industry 4.0." In *2014 IEEE international conference on automation, quality and testing, robotics*, pp. 1-4. IEEE.

Jenkins, Daniel P. 2009. *Cognitive work analysis: coping with complexity*. Ashgate Publishing, Ltd.

Kant, Vivek. 2017. "Extending Cognitive Work Analysis for embodiment: ecological psychology, activity theory and Worker Competency Analysis." *Theoretical Issues in Ergonomics Science* 18 (6): 548-572.

Kirwan, Barry. 1998. "Human error identification techniques for risk assessment of high risk systems—Part 2: towards a framework approach." *Applied ergonomics* 29 (5): 299-318.

Kurzweil, Ray. 2014. "The singularity is near." In *Ethics and emerging technologies*, pp. 393-406. Palgrave Macmillan, London.

Lee, Edward. 2015. "The past, present and future of cyber-physical systems: A focus on models." *Sensors* 15 (3): 4837-4869.

Lee, John, and Neville Moray. 1992. "Trust, control strategies and allocation of function in human-machine systems." *Ergonomics* 35 (10): 1243-1270.

Leitao, Paulo, Stamatis Karnouskos, Luis Ribeiro, Jay Lee, Thomas Strasser, and Armando W. Colombo. 2016. "Smart agents in industrial cyber–physical systems." *Proceedings of the IEEE* 104 (5): 1086-1101.

Leveson, Nancy. 2004. "A new accident model for engineering safer systems." *Safety science* 42 (4): 237-270.

Loukas, George. 2015. *Cyber-physical attacks: A growing invisible threat*. Butterworth-Heinemann.

Lundstrom, Mark. 2003. "Moore's law forever?." *Science* 299 (5604): 210-211.

Marathe, Amar R., Jason S. Metcalfe, Brent J. Lance, Jamie R. Lukos, David Jangraw, Kuan-Ting Lai, Jonathan Touryan et al. 2018. "The privileged sensing framework: A principled approach to improved human-autonomy integration." *Theoretical Issues in Ergonomics Science* 19 (3): 283-320.

Mayer- Kress, Gottfried, and Cathleen Barczys. 1995. "The global brain as an emergent structure from the Worldwide Computing Network, and its implications for modeling." *The information society* 11 (1): 1-27.

Moore, Gordon E. 1965. "Cramming more components onto integrated circuits.": 114-117.

Moray, Neville. 1995. "Ergonomics and the global problems of the twenty-first century." *Ergonomics* 38 (8): 1691-1707.

Muehlhauser, Luke, and Anna Salamon. 2012. "Intelligence explosion: Evidence and import." In *Singularity Hypotheses*, pp. 15-42. Springer, Berlin, Heidelberg.

Muir, Bonnie M. "Trust in automation: Part I. 1994. Theoretical issues in the study of trust and human intervention in automated systems." *Ergonomics* 37 (11): 1905-1922.

Naikar, Neelam. 2017. "Cognitive work analysis: An influential legacy extending beyond human factors and engineering." *Applied ergonomics* 59: 528-540.

Naikar, Neelam, Anna Moylan, and Brett Pearce. 2006. "Analysing activity in complex systems with cognitive work analysis: concepts, guidelines and case study for control task analysis." *Theoretical Issues in Ergonomics Science* 7 (4): 371-394.

Natural Capitalism Solutions. 2005. "The Natural Edge Project." *Prospering in a Carbon Constrained World: Profitable Opportunities for Greenhouse Gas Emissions Reduction*.

Ning, Huansheng, and Ziou Wang. 2011. "Future internet of things architecture: like mankind neural system or social organization framework?." *IEEE Communications Letters* 15 (4): 461-463.

Noto La Diega, Guido, and Ian Walden. 2016. "Contracting for the 'Internet of Things': Looking into the Nest." *Queen Mary School of Law Legal Studies Research Paper* 219.

Parasuraman, Raja, and Victor Riley. 1997. "Humans and automation: Use, misuse, disuse, abuse." *Human factors* 39 (2): 230-253.

Parker, Jon, and Joshua M. Epstein. 2011. "A distributed platform for global-scale agent-based models of disease transmission." *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 22 (1): 2.

Patriarca, Riccardo, Giulio Di Gravio, Francesco Costantino, and Massimo Tronci. 2017. "The Functional Resonance Analysis Method for a systemic risk based environmental auditing in a sinter plant: A semi-quantitative approach." *Environmental Impact Assessment Review* 63: 72-86.

Rasmussen, Jens, Annelise Mark Pejtersen, and Len P. Goodstein. 1994. "*Cognitive systems engineering*.": 813-816.

Read, Gemma J.M., Paul M. Salmon, Michael G. Lenné, and Neville A. Stanton. 2015. "Designing sociotechnical systems with cognitive work analysis: putting theory back into practice." *Ergonomics* 58 (5): 822-851.

Read, Gemma. J.M., Thompson, Jason. & Salmon, Paul. M. (In press). "Using cognitive work analysis to inform agent-based modelling of automated driving". In Verhagen, H., Borit, M., Bravo, G. & Williams, N. (Eds). 2018 "Advances in

Social Simulation: Looking in the Mirror". *Proceedings of the European Social Simulation Association Conference*. European Switzerland: Springer.

Rognin, Laurence, Pascal Salembier, and Moustapha Zouinar. 2000. "Cooperation, reliability of socio-technical systems and allocation of function." *International Journal of Human-Computer Studies* 52 (2): 357-379.

Salmon, Paul M., Guy H. Walker, Gemma J. M. Read, Natassia Goode, and Neville A. Stanton. 2017. "Fitting methods to paradigms: are ergonomics methods fit for systems thinking?." *Ergonomics* 60 (2): 194-205.

Schaefer, Kristin E., Jessie YC Chen, James L. Szalma, and Peter A. Hancock. 2016. "A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems." *Human factors* 58 (3): 377-400.

Schwab, Klaus. 2017. *The fourth industrial revolution*. Currency.

Scott, James, and Drew Spaniel. 2016. "Rise of the machines: The dyn attack was just a practice run." *Institute for Critical Infrastructure Technology (ICIT)*: 60.

Shorrock, Steven T., and Barry Kirwan. 2002. "Development and application of a human error identification tool for air traffic control." *Applied ergonomics* 33 (4): 319-336.

Silverman, Eric. 2018. *Methodological Investigations in Agent-Based Modelling: With Applications for the Social Sciences*: Vol. 13. Springer.

Stanton, Neville A., Paul Salmon, Don Harris, Andrew Marshall, Jason Demagalski, Mark S. Young, Thomas Waldmann, and Sidney Dekker. 2009. "Predicting pilot error: testing a new methodology and a multi-methods and analysts approach." *Applied ergonomics* 40 (3): 464-471.

Stanton, Neville A., Paul M. Salmon, Laura A. Rafferty, Guy H. Walker, Chris Baber, and Daniel P. Jenkins. 2017. *Human factors methods: a practical guide for engineering and design*. CRC Press.

Stanton, Neville A. 2014. "Representing distributed cognition in complex systems: how a submarine returns to periscope depth." *Ergonomics* 57 (3): 403-418.

Steels, Luc, and Rodney Brooks. 2018. *The artificial life route to artificial intelligence: Building embodied, situated agents*. Routledge.

Sutardja, Sehat. 2014. "Slowing of Moore's law signals the beginning of smart everything." In *2014 44th European Solid State Device Research Conference (ESSDERC)* 7-8.

Szreter, Simon, and Graham Mooney. 1998. "Urbanization, mortality, and the standard of living debate: new estimates of the expectation of life at birth in nineteenth-century British cities." *The Economic History Review* 51 (1): 84-112.

Thatcher, Andrew, Patrick Waterson, Andrew Todd, and Neville Moray. 2018. "State of science: ergonomics and global issues." *Ergonomics* 61 (2): 197-213.

Thatcher, Andrew, and Paul HP Yeow. 2016. "A sustainable system of systems approach: a new HFE paradigm." *Ergonomics* 59 (2): 167-178.

Toroody, Ahmad Bahoo, Mohammad Mahdy Abaei, and Reza Gholamnia. 2016 "Conceptual compression discussion on a multi-linear (FTA) and systematic (FRAM) method in an offshore operation's accident modeling." *International journal of occupational safety and ergonomics* 22 (4): 532-540.

Toroody, Ahmad Bahoo, Farshad Bahoo Toroody, and Filippo De Carlo. 2017. "Development of a risk based methodology to consider influence of human failure in industrial plants operation." *Summer School "Francesco Turco*.

Trist, Eric L. 1978. "On socio-technical systems." *Sociotechnical systems: A sourcebook*: 43-57.

Vicente, Kim J. 1999. *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. CRC Press.

Vinge, Vernor. 1993. "The coming technological singularity: How to survive in the post-human era." *Science Fiction Criticism: An Anthology of Essential Writings*: 352-363.

Wahlster, Wolfgang. 2012. "From industry 1.0 to industry 4.0: Towards the 4th industrial revolution." In *Forum Business meets Research*.

Walker, Guy. "Fortune favours the bold." 2016. *Theoretical Issues in Ergonomics Science* 17 (4): 452-458.

Walker, Guy H., Huw Gibson, Neville A. Stanton, Chris Baber, Paul Salmon, and Damian Green. 2006. "Event analysis of systemic teamwork (EAST): a novel integration of ergonomics methods to analyse C4i activity." *Ergonomics* 49, (12-13): 1345-1369.

Walker, Guy H., Paul M. Salmon, Melissa Bedinger, and Neville A. Stanton. 2017. "Quantum ergonomics: shifting the paradigm of the systems agenda." *Ergonomics* 60 (2): 157-166.

Walker, Guy H., Neville A. Stanton, and Daniel P. Jenkins. 2017. *Command and control: the sociotechnical perspective*. CRC Press.

Waterson, Patrick, Michelle M. Robertson, Nancy J. Cooke, Laura Militello, Emilie Roth, and Neville A. Stanton. 2015. "Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety." *Ergonomics* 58 (4): 565-599.

Wilson, John R. 2000. "Fundamentals of ergonomics in theory and practice." *Applied ergonomics* 31 (6): 557-567.

Wilson, John R. 2014. "Fundamentals of systems ergonomics/human factors." *Applied ergonomics* 45 (1): 5-13.

Wu, Dazhong, David W. Rosen, Lihui Wang, and Dirk Schaefer. 2015. "Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation." *Computer-Aided Design* 59: 1-14.

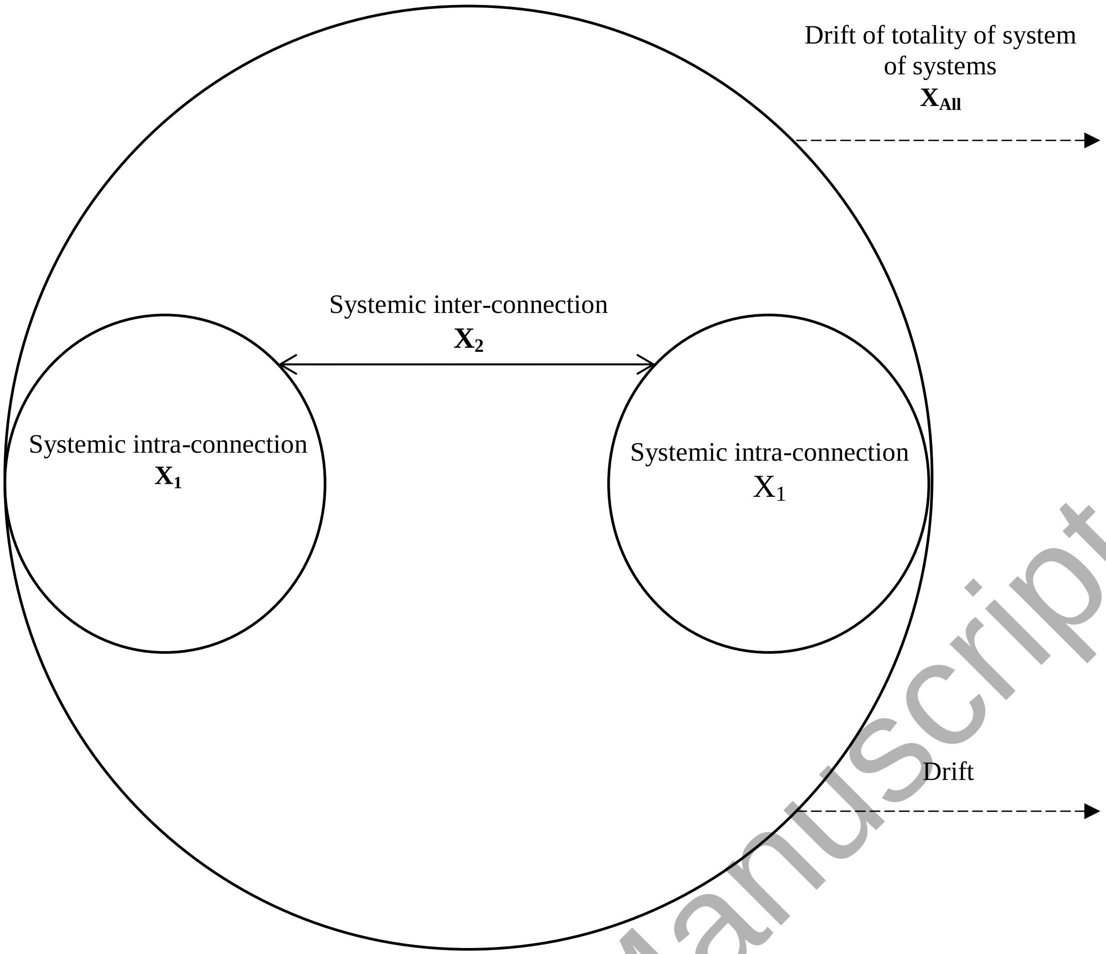

***Figure 1** - Post-Dualist Construct: X1 - The impacts of the integration of 4IR technology applications within sociotechnical systems is making interactions at the intra-system level more complex. Underpinned by a dense network of 4IR technology, perturbations can amplify and spread rapidly and imperceptibly. X2 – Integration of 4IR technology is thrusting us into a paradigm where the coupling of once disparate systems is rapidly becoming an infinitely complex, and perhaps global, system of systems. Perturbations in one system could feasibly propagate through the whole system of systems in this paradigm. For example, the 2016 Dyn Cyber attacks demonstrated how the convergence between 'cyber' and the 'physical' systems, and their dense interconnection, could expose us to catastrophe. X3 - At the highest level of ontological abstraction, the totality of all systems (and all socio and technical elements within this higher-order construct) will drift in ways that we cannot forecast, or perhaps even fathom.*

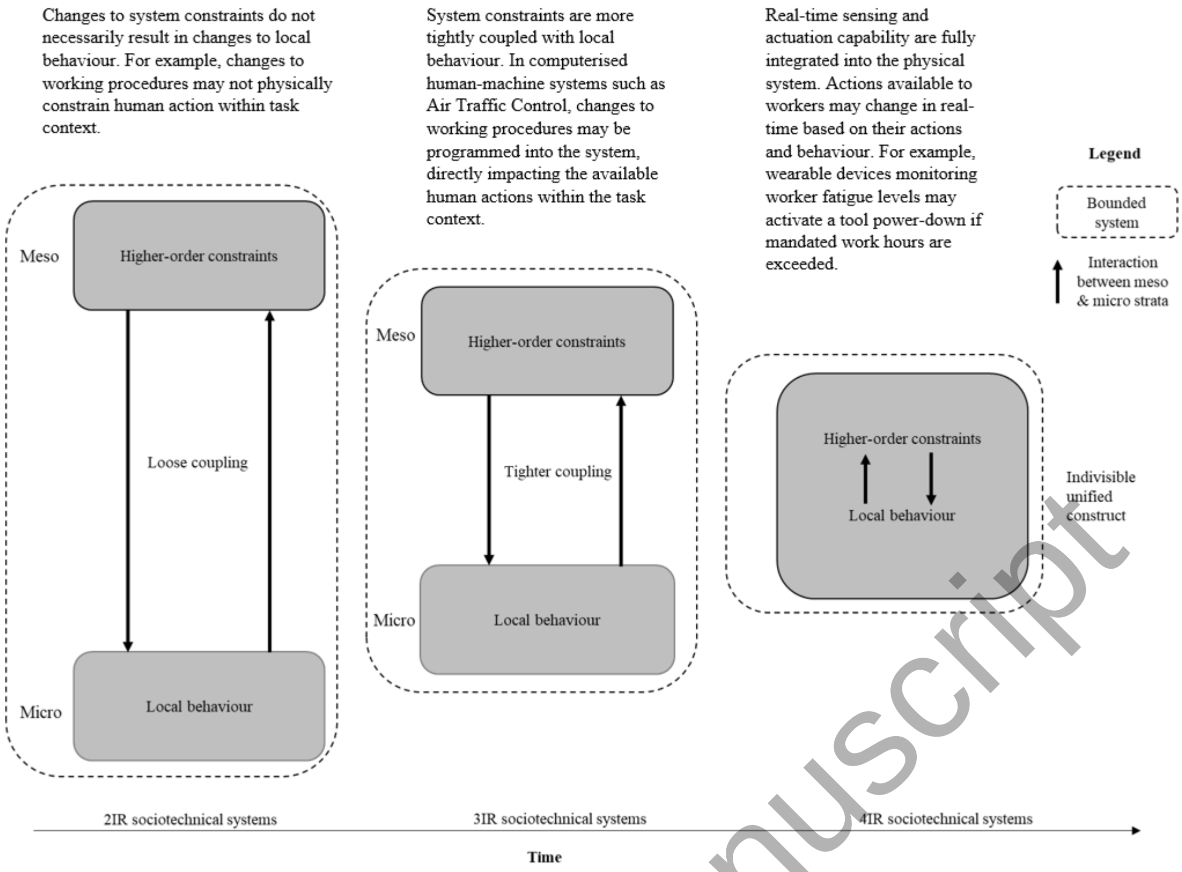

Figure 2 -Meso and micro constructs become unified in the 4IR technological paradigm

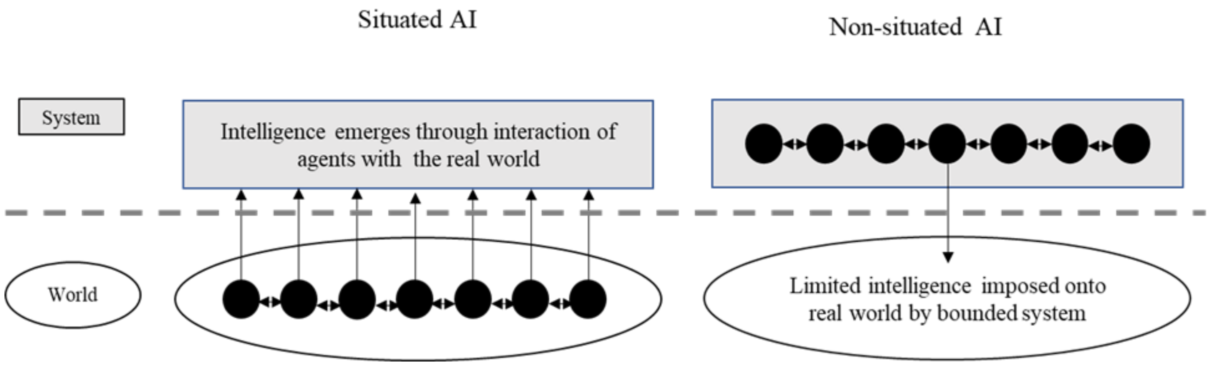

Figure 3 - Situated AI vs Non-situated AI. True intelligence is an emergent property which manifests as a result of local interaction of discrete and simple agents distributed across the 'real world'. Proponents of the SAI theory argue that SAI systems 'draw' intelligence from the real-world. By contrast, Non-situated AI, which underpins the singularity hypothesis, imposes intelligence on the 'real world'. Opposers of the non-SAI approach argue that these systems do not establish or maintain continuous and 'live' model of the real-world and thus has limited operational effectiveness in the real-world.

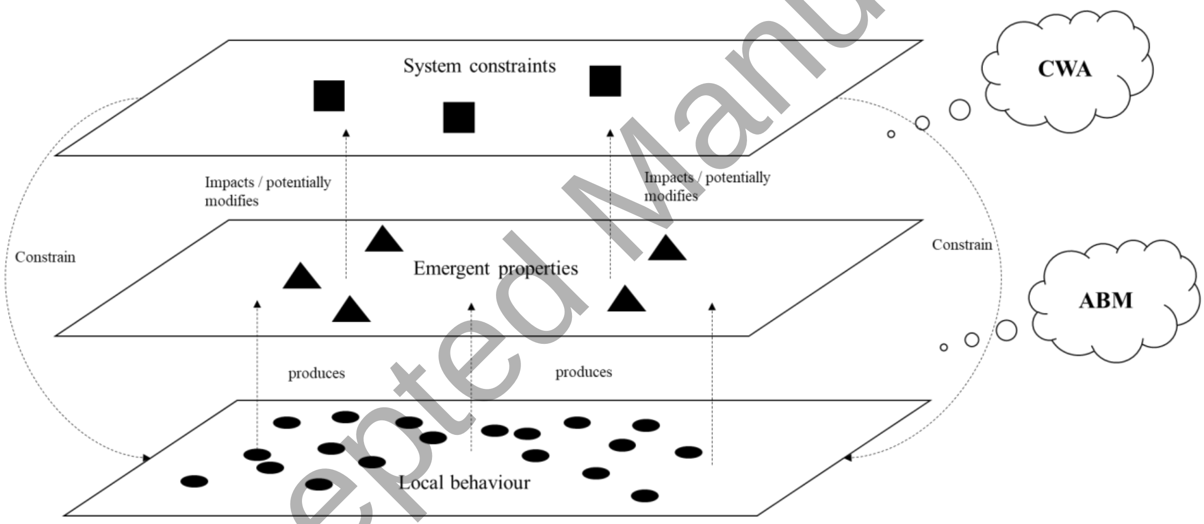

Figure 4 Combined ABM and CWA approach to model emergence and its impacts on higher-order system constraints.

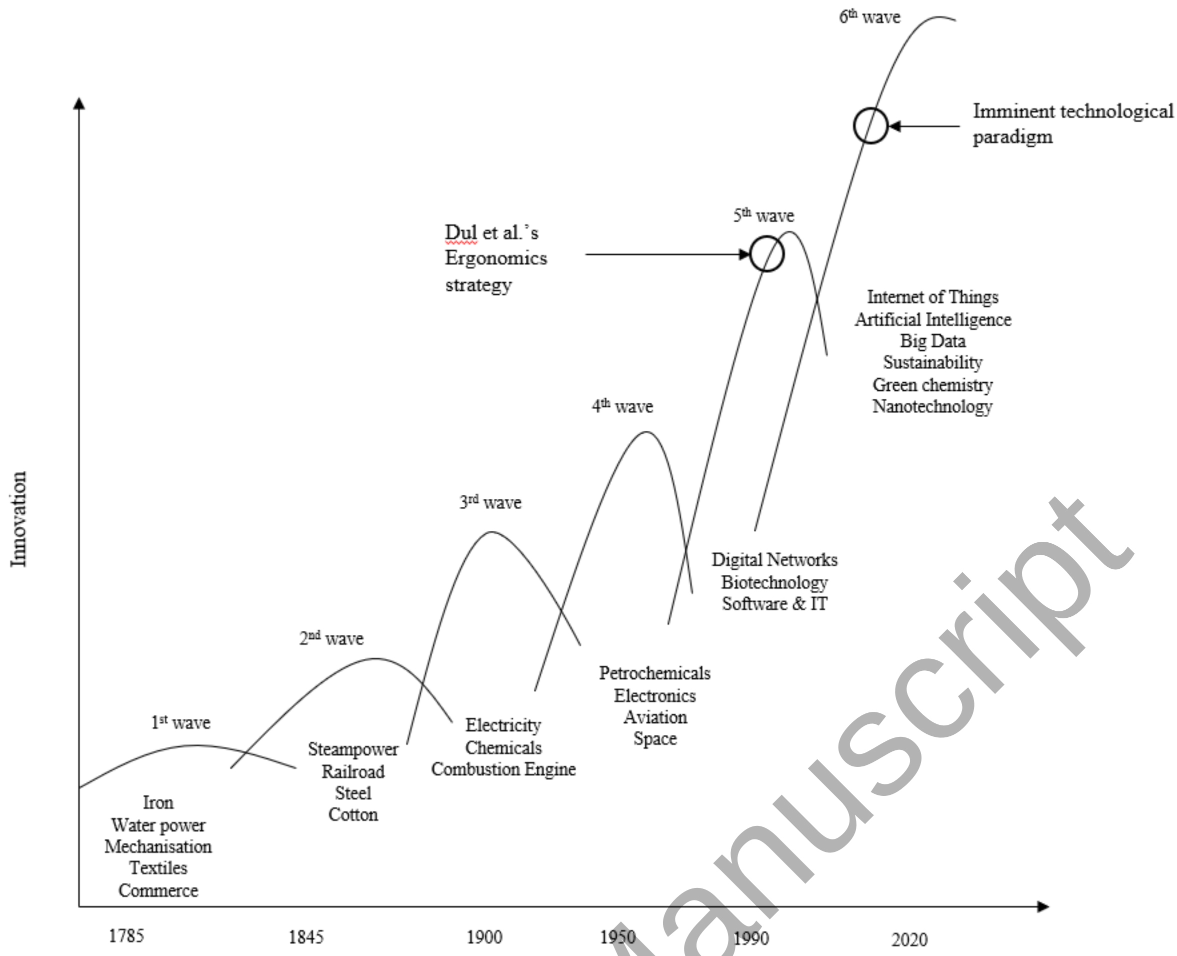

Figure 5 - Waves of Innovation (adapted from Natural Capitalism Solutions, 2005).

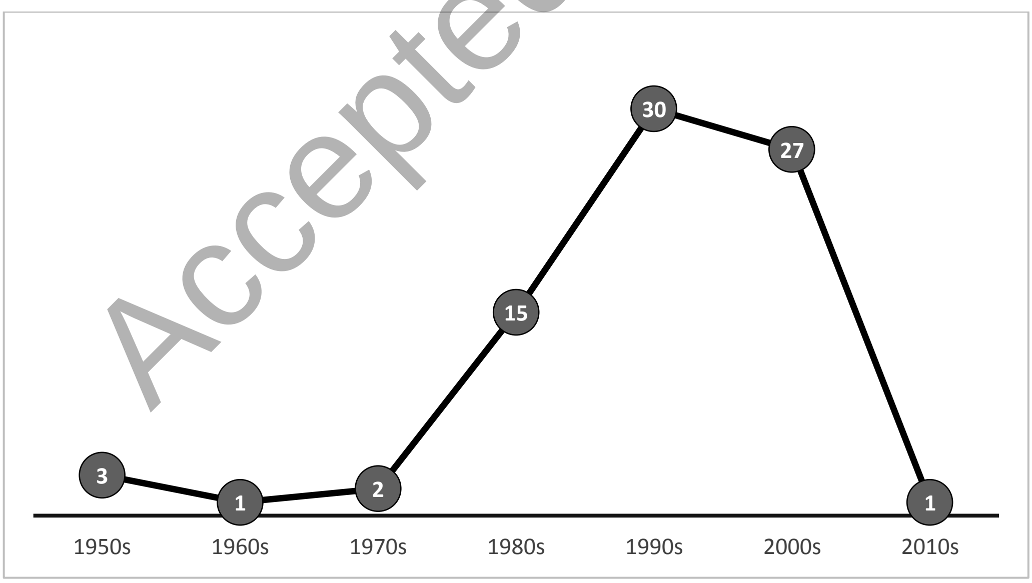

Figure 6 - Methods Developed per Decade

Strengthen the demand for HFE	Strengthen application of high-quality HFE
Communicating with dominant stakeholders	Promoting the education of high quality HFE
Building strategic partnerships	Ensuring high quality standards of HFE applications
Educating (future) stakeholders	Promoting HFE research excellence at universities and other organisations

Table 1 - Dul et al's strategic actions