Author Accepted Manuscript. This manuscript was accepted for publication by Technological Forecasting and Social Change on November 7, 2021. The paper was published online (with editorial revisions) on November 29, 2021 and is available at Technological Forecasting and Social Change at https://doi.org/10.1016/j.techfore.2021.121347

Policy interactions with research trajectories: The case of *cyber-physical convergence* in manufacturing and industrials

Tausif Bordoloi[a,b], Philip Shapira[a,c] and Paul Mativenga[b]

[a] Manchester Institute of Innovation Research, Alliance Manchester Business School, The University of Manchester, Manchester, M15 6PB, UK

[b] Department of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Manchester, M13 9PL, UK

[c] School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332-0345, USA

Corresponding author email address: tausif.bordoloi@manchester.ac.uk

ORCID ID: Tausif Bordoloi: 0000-0002-2731-8663; Philip Shapira: 0000-0003-2488-5985; Paul Mativenga: 0000-0002-7583-5163

Abstract

From the early 2010s, policymakers and firms in advanced industrial economies began introducing approaches to systemically exploit manufacturing and industrial data (the notion of *cyber-physical convergence*). Three innovation concepts were especially highlighted: Smart Manufacturing, Industrial Internet and Industrie 4.0. In parallel, academics have employed these concepts in numerous ways to promote their work. Despite this broad interest, precise definition and delineation of the *cyber-physical convergence* research domain have received little attention. Also missing is systematic knowledge on the interactions of these concepts with research trajectories. This paper fills these gaps by operationalising a newly constructed definition of *convergence*, and delineating the associated research domain into five data-centric capabilities: Monitoring, Analytics, Modelling-and-Simulation, Transmission and Security. A bibliometric analysis of the domain is then performed for 2010–2019. There are three findings. First, Analytics and Security have assumed strategic positions within the domain, coinciding with a “strategic turn” in policy. Second, backed by concerted policy and funding efforts, growth in Chinese scientific output has outpaced key competitors U.S. and Germany. Finally, the patterns of promoting their works in terms of the three concepts differ significantly among U.S.-, Germany- and China-based authors, which mirrors the different policy discourses prevalent in those countries.

Keywords: Industrie 4.0; Industry 4.0; Smart Manufacturing; Industrial Internet; Digital Manufacturing

1. Introduction

Over the last decade, one trend had stood out in discussions about the future of manufacturing and industrial sectors: the integration of cyber space (as embodied by the Internet and computational capabilities) and physical machinery, resulting in what is called *cyber-physical convergence* (Conti et al., 2012, Ezell, 2018, O'Sullivan et al., 2017). This combination of the digital and the material has been seen by policymakers and firms in advanced industrial economies as central to making tangible gains on innovation, productivity and competitiveness (Aaronson, 2019, Bonvillian, 2012, Pisano and Shih, 2009). Drawn by such prospects, three concepts were coined and championed in the early 2010s – all loosely under the banner of innovation policy: "Smart Manufacturing" as a U.S. Government initiative to revitalise the country's manufacturing sector and strengthen national security (PCAST, 2012); the German public-private coordinated "Industrie 4.0" (also known as Industry 4.0"), which mirrors collective efforts by firms, trade unions, universities and others to strengthen Germany's economy (Kagermann et al., 2013); and the "Industrial Internet", advanced by U.S. private sector firms to boost profitability under challenging economic conditions (Evans and Annunziata, 2012).

A common focus of these concepts is to set enabling conditions, including those that influence the production of scientific knowledge along desired trajectories, for furthering policy goals. Policy interest has been matched by concurrent growth in academic research, where these concepts have been used to frame a list of advanced technologies including augmented reality for simulation (Angrisani et al., 2019), 3D printing for digital fabrication (Ceruti et al., 2019) and blockchain for industrial cybersecurity (Mazzei et al., 2020). Evidence of the increasing scholarly attention being paid to these concepts is also observed in the manner in which they are promoted as a positive and desired attribute of an object of study. For instance, computer scientists have used them in connection to applications of data analytics to improve anomaly detection (Stojanovic et al., 2016), support remote maintenance (Masoni et al., 2017) and measure energy consumption (Qin et al., 2017). Operations researchers have applied them to discuss the advantages of lean production (Mrugalska and Wyrwicka, 2017) and mass customisation (Torn and Vaneker, 2019). Meanwhile, management scholars have associated these concepts with diverse strategic aspects in manufacturing, including firms' dynamic capabilities (Felsberger et al., 2020), business model innovations (Frank et al., 2019) and innovation ecosystems (Benitez et al., 2020).

Yet, while researchers have broadly drawn on Smart Manufacturing, Industrie 4.0 and Industrial Internet policy concepts to underpin their work, there has been little systematic work to define and delineate these particular framings, how they overlap or differ, and how they relate to wider policy trends and interventions. Attempts by the scientometric community to evaluate scientific outputs associated with one or more of these concepts have tended to be *ad hoc* and high-level in scope. For example, one study by Muhuri et al. (2019) queries “industry 4.0” and then performs simple descriptive bibliometrics to analyse literature trends. Trotta and Garengo (2018) expand the search query to include “smart manufacturing”. A recent paper by Meindl et al. (2021) provides a broad overview of research trends associated with four “smart dimensions of Industry 4.0, including manufacturing. These approaches, whilst useful, are not robustly able to measure developments and discreet patterns in *cyber-physical convergence* research output. In part, this reflects a lack of conceptual and boundary consensus regarding the notion of *convergence* itself.

Systematically defining *cyber-physical convergence* is not just an academic exercise. How we define it will influence the approach to evaluate growth and trajectory of research output, which has an important bearing on public and private policy goals (Boswell and Smith, 2017). In this paper, we make three contributions: definitional, methodological and measurement. We put forward a newly constructed definition of *cyber-physical convergence* by drawing on a set of foundational policy and technical documents. Reviewing these documents allows us to examine the technological underpinnings of Smart Manufacturing, Industrie 4.0 and Industrial Internet, and define *cyber-physical convergence* in a manner that is consistent with these underpinnings. Our definition focuses at the level of data generated in manufacturing and industrial processes. With this approach we can separate out signifiers associated with the three concepts. Next, we operationalise this definition using natural language processing methods. In doing so, we set out, for the first time to our knowledge, a boundary around the *cyber-physical convergence* research domain. While this boundary should not be viewed as impermeable, it sufficiently describes and sensitises the domain to the influence of the concepts. This results in the identification and elucidation of five subfields – each associated with a specific data-centric capability: (i) Monitoring, (ii) Analytics, (iii) Modelling-and-Simulation, (iv) Transmission and (v) Security. Our third contribution involves the evaluation of bibliometric performance indicators in the period between 2010 and 2019 to identify developments and patterns in research outputs related to *cyber-physical convergence*.

The paper is structured as follows. Section 2 introduces the notion of *cyber-physical convergence* and defines it according to the three concepts. In Section 3, the definition is operationalised, subfields are identified, and results are analysed. Section 4 presents discussion and concluding remarks.

2. Cyber-physical convergence: Concepts and definitions

2.1. A brief overview of *cyber-physical convergence* and its importance

Cyber-physical convergence is related to a broader concept of technological convergence, which defines a state where boundaries between different technologies or industrial sectors blur (Gauch and Blind, 2015), leading to "*outcomes that in their performance exceed the sum of their parts*" (Hacklin et al., 2009). An early mention of *cyber-physical convergence* in academic research is by Conti et al. (2012) who characterise it as a scenario where physical components pervasively interact with the cyber space via sensing, computing and Internet communication components. The authors argue that as devices (such as smartphones) become more pervasive, they enable the monitoring, collection and analysis of user data to better understand human behaviours. The key feature of this characterisation is represented by the term 'pervasive', which refers to the ability to embed computational capability into physical objects. That is to say, Conti et al's definition directly draws on the idea of 'ubiquitous computing', most often associated with former Xerox computer scientist Mark Weiser. Weiser, in his seminal article "*The Computer for the 21st Century*" (1991) argued that "*specialized elements of hardware and software, connected by wires, radio waves and infrared, will be so ubiquitous that no one will notice their presence*". The modern computer, therefore, would weave into the fabric of everyday human life and become indistinguishable from it. In recent years, Conti et al. (2017) have extended this last point by explicitly including humans ("*Internet of People*") in their *convergence* narrative.

The notion of *cyber-physical convergence* is now increasingly invoked in manufacturing and industrials contexts. O'Sullivan et al. (2017, pp. 336-338) describe three different dimensions of *convergence*: (a) integration of enabling technologies (such as Information and Communication Technology (ICT), nanotechnology, biotechnology and advanced materials) in manufacturing; (b) hybrid manufacturing systems built on a combination of disparate technological foundations, for example, mechatronics (sensing, control, measurement, etc.)

and materials engineering (pressing, extrusion, fabrication, etc.); and (c) application of ICT to enable vertical integration of production lines and units, horizontal integration of value chains, and integration of end-to-end engineering activites across the life-cycle of a product. All three dimensions stress the importance of the Internet in linking physical components ("*Internet of Things*" or *IoT*) distributed throughout value chains, allowing efficient product development, logistics management and services. Jeschke et al. (2017, p. 7) specifically add computer science (data analytics and artificial intelligence) and ergonomic (human-machine interactions) functionalities to the scope of *cyber-physical convergence*. Yet another interpretation is the convergence of Operational Technologies or OT (technologies to control, monitor and automate industrial processes, e.g., Supervisory Control and Data Acquisition) with Information Technologies or IT (data networking, infrastructure and software technologies, e.g., Enterprise Resource Planning), resulting in a plant-to-enterprise integration (Davis et al., 2020).

A broader, macro-level sense of the *convergence* phenomenon is provided by Fort et al. (2018). The authors, in their study on U.S. manufacturing employment, highlight the increase of firms' adoption of computers and electronic networks (that is, the Internet and electronic data exchanges) through the 2000s. In particular, the proportion of plants buying computers increased significantly in 2002. This was accompanied by an analogous growth in the adoption of electronic networks to control or coordinate product shipments. Fort (2017) further indicates that U.S. plants' adoption of electronic networks not just involved the Internet, but also the integration of electronic communication in production processes, resulting in higher labour productivity than non-adopters (Fort et al., 2018). These observations are consistent with a large body of macroeconomic research which has shown that application of ICT in the U.S. nonfarm business sector (includes manufacturing) was a major driver of labour productivity growth between 1995 and 2004 (see, for example, Jorgenson et al., 2008). By comparison, major European economies, in particular, Germany, experienced lower labour productivity gains from ICT in the decade after the mid-1990s (Eicher and Roehn, 2007). More recently, the U.S., European and other advanced economies have all experienced slow productivity growth (Andrews et al., 2016). To reverse such trends, there have been suggestions of greater investments in ICT and its pervasive adoption in manufacturing and industrials (Miller and Atkinson, 2014, Tassey, 2008, Tassey, 2010). These suggestions are emphatically supported in the characterisation of *cyber-physical convergence* according to Smart Manufacturing, Industrie 4.0 and Industrial Internet (Belton et al., 2019, Evans and Annunziata, 2012, Horst and Santiago, 2018).

2.2. Technological underpinnings of the three concepts

We now focus on the technological underpinnings of *cyber-physical convergence* as advanced in the three concepts. To do this, we identified, selected and reviewed a set of policy and technical literature (Appendix 1). We consider this literature to be foundational because: (i) it seeks to frame the concepts through the lens of their original proponents; and (ii) it identifies and describes the key technological steps needed to achieve *convergence* generally, as well as in manufacturing and industrial settings.

Smart Manufacturing materials include reports produced or commissioned by three U.S. Government sources: (a) the President's Council of Advisors on Science and Technology (PCAST) – a science, technology and innovation policy advisory group comprising of experts from industry and academia; (b) the Clean Energy Smart Manufacturing Innovation Institute (CESMII) – a public-private consortium within the U.S. Department of Energy with the mandate to make manufacturing more energy-efficient through the use of cyber-physical capabilities; and (c) the National Institute of Standards and Technology (NIST), which develops technical standards and measurements crucial to a wide range of industries. Resources related to Industrial Internet are those published by General Electric (GE), a U.S. industrial conglomerate, widely credited with coining the concept, and the Industrial Internet Consortium (IIC), an open membership organisation co-founded by GE to catalyse the adoption of Industrial Internet. Industrie 4.0 documents are those by the German National Academy of Science and Engineering (Acatech), whose members have played a fundamental role in championing the Industrie 4.0 narrative within and beyond Germany. After reviewing this compendium of literature, we find that although Smart Manufacturing, Industrie 4.0 and Industrial Internet are guided by different rationales, theorising and practices (the discussion of which is beyond the scope of this paper), they share a common technological base. The next section discusses this in more details.

According to NIST (Lu et al., 2016), the technological building block of Smart Manufacturing (called an Smart Manufacturing Ecosystem) is framed at an organisational level. It covers the entire range of devices, machines and systems within a firm's engineering, production and management functions. Each function generates large amounts of data which is then exploited through a three-step process: acquiring data from devices and machines of diverse supply chain members in different locations; contextualising and analysing the data in cloud-based environments; and visualising and reporting insights to achieve measurable results, for example, optimisation of an additive manufacturing process (CESMII, 2019, Miller, 2018, Schneider, 2018). In a NIST-commissioned study, Gallaher et al. (2016, p. 45) also support a similar data-centric workflow that models, senses, transmits, analyses, communicates and takes action on data. Hence, "*in its simplest form, SM turns data from the manufacturing process into actionable knowledge*."

The technolological unit of Industrie 4.0 (called Cyber-Physical System) is based on a systems level, wherein software is "embedded" into hardware systems (Geisberger and Broy, 2015, pp. 23-26, Hellinger and Seeger, 2011). These systems perform a series of data-driven activities: produce and acquire physical data from the environment; process this data to interpret a situation in relation to pre-defined objectives; and use the knowledge to make decisions in real time (Geisberger and Broy, 2015, p. 64). In comparison, Industrial Internet is characterised at an architectural design level called Industrial Internet System. It consists of three tiers that are interlinked via the Internet: the edge tier collects data from on-field devices and sends it to the platform tier, where the data is consolidated and cleansed. Industrial analytics is then applied to the processed data at the enterprise tier to generate insights and inform business activities (Lin et al., 2015, p. 37).

Where these different concepts overlap, then, is a common framing that data is the fundamental driver of *cyber-physical convergence* (Figure 1). This interpretation of *convergence* follows the etymology of the word which comes from the Latin *convergere* – to come together, to connect machines and boost productivity by exploiting the data they produce. More specifically, in all three concepts, there is an emphasis on a set of distinct, yet overlapping (to a certain extent) capabilities, comprising data flows between different systems, beginning with sensing to collect data all the way through to when it is analysed to generate insights for decision-making. These data-centric capabilities, according to Smart Manufacturing, Industrial Internet and Industrie 4.0, collectively embody and define *convergence* (Table 1).

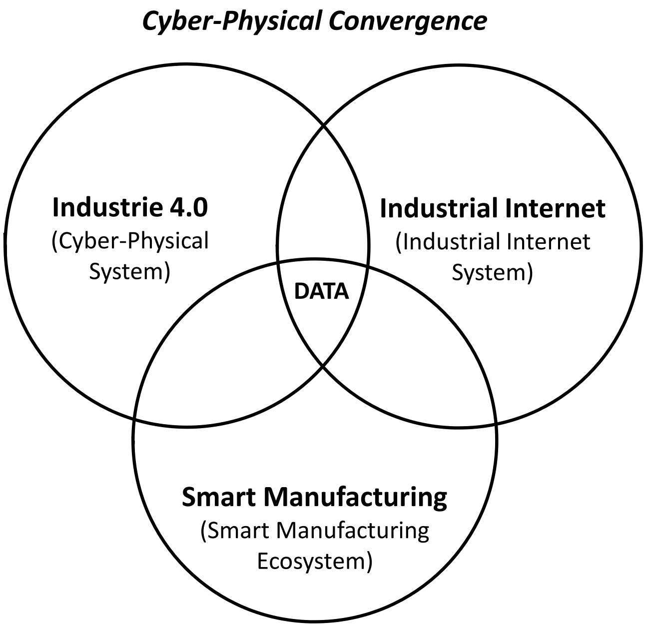

Figure 1: The three concepts overlap at a common data-centric framing of *cyber-physical convergence.*

An important implication of these discussions is that *convergence* does not equate with new technologies. Advanced technologies such as augmented reality, 3D printing or blockchain will not necessarily convey *convergence* unless the data they produce is effectively harnessed and analysed to, say, drive productivity. Conversely, even legacy machinery can be distinguished as convergent if they interact effectively with the cyber space (Kagermann et al., 2013, Schneider, 2018). Given these arguments, we conceive of *cyber-physical convergence* as a "*data-centric workflow that makes use of specific capabilities, starting with data acquisition from physical sources, and then continuing through to the generation of actionable insights*". This definition disentangles random signifiers for the purpose of informing our text-mining and bibliometric research design.

Concept and Authors/Study	Definition (adopted)
Smart Manufacturing Gallaher et al (2016, p. 27) conducted an economic and technological assessment study commissioned by the NIST to examine the role that NIST could play to accelerate Smart Manufacturing development and adoption in the U.S.	Smart Manufacturing comprises six capability areas: (a) modeling and data creation to support design, testing and automation; (b) sensing and monitoring to collect real-time information on processes; (c) transmitting information across multiple platforms and levels of the supply chain; (d) analysing data and trends to support real-time process control and management; (e) communicating information to decision makers to support efficient and/or automated analysis; and (f) determining and implementing required action in a timely and/or closed-loop setting.
Industrial Internet Lin et al (2015, p. 26) of the IIC in a technical paper titled "Industrial Internet Reference Architecture" elaborated the functional dimensions of the Industrial Internet concept as a set of activities that can transform industrial systems on a global scale.	Industrial Internet is enabled by: (a) collecting sensor data from across industrial systems; (b) applying analytics, including models developed through machine learning, to these data, so as to gain insight to a business's operations; and (c) using these insights to help improve decision-making and optimize operations globally through automatic and autonomous orchestration.
Industrie 4.0 Geisberger and Broy (2014, p. 64) describe the technological basis of Industrie 4.0 as part of the "Integrated Research Agenda Cyber-Physical Systems" project initiated by the Acatech and funded by the German Federal Ministry of Education and Research (BMBF).	*Cyber-physical convergence* involves the ability of manufacturing systems to: (a) capture physical data from the environment in parallel via sensors and to merge and process this data – and to do so both locally and globally and in real time; (b) use the information that they have gathered to interpret the situation in terms of predefined goals; (c) detect, interpret, deduce and forecast malfunctions, problems and threats; (d) integrate, regulate, control and interact with components and functions; and (e) carry out globally distributed and networked control and regulation in real time.

Table 1. Definitions of *cyber-physical convergence* according to the three concepts.

3. Text-mining and bibliometrics

In this section, we operationalise our definition of *cyber-physical convergence*. To do this, however, we need to address a more basic question – how exactly do researchers working in the domain use the term ‘data’ in their publications? This question is pertinent because it concerns translating the data-centric capabilities of Table 1 into a controlled lexicon of keywords for use in bibliometric analysis. As emerging scientific domains cannot be accurately described in terms of random signifiers, building such a lexicon helps overcome delineation and measurement issues (Oldham et al., 2012).

3.1. Building a lexicon of relevant keywords

Our lexicon-building strategy combines elements from different scientometric approaches in science and technology. We extract benchmark publications from the scientific publication database Scopus and analyse them with natural language processing tools to produce an initial set of mid- and high-frequency keywords. Capturing high-frequency keywords from publications is an effective means to operationalise emerging technological domains (Hu and Rousseau, 2015, Shapira et al., 2017). Mid-frequency keywords play a complementary role in this process because they reveal not-too-common words in a specific field (Luhn, 1958). The relevance of these keywords with respect to the publication corpus is determined by applying a statistical measure called term frequency-inverse document frequency (TF-IDF), and also by checking their meaningfulness in regards to specific word combinations (phrases). TF-IDF combines both keyword popularity and discrimination to evaluate how important a term is to a publication in a corpus (Chen and Xiao, 2016). Our analysis finds a set of keywords that are not only *cyber-physical convergence*-specific, but also highlights the underlying capabilities that are contributing to it. Next, we test them for their specificity to obtain complementary keywords, which are then added to the set. Thus, following Mogoutov and Kahane (2007), we expand the lexicon using a modular methodology (Figure 2) that better defines the research domain, while fencing out non-relevant publications.

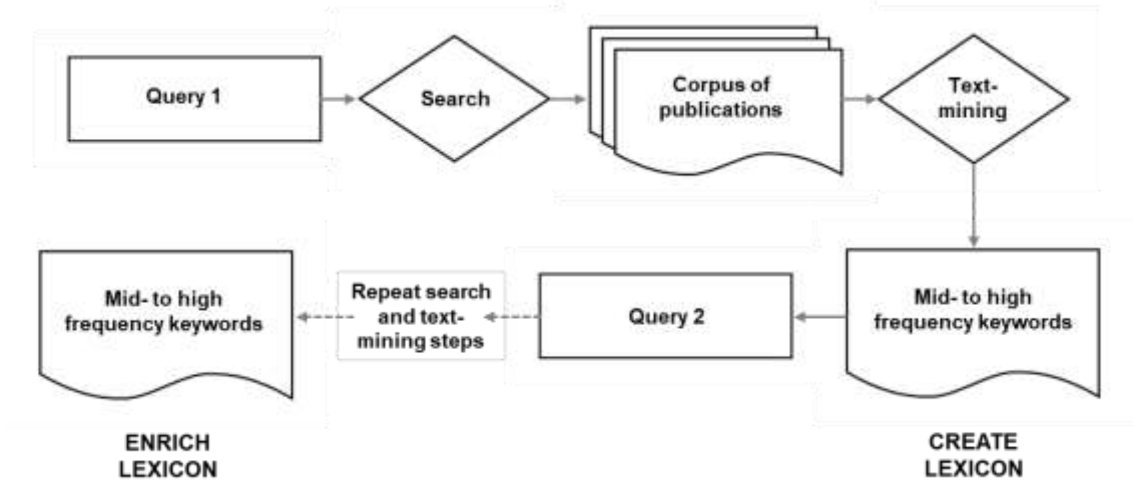

Figure 2: Modular lexicon building methodology. See sections 3.1.1 and 3.1.2 for details on Query 1 and Query 2.

3.1.1. Retrieving benchmark publications

It must be noted that authors in different disciplines have rapidly clutched on the three concepts to promote their research. So, to navigate the search space and retrieve papers that explicitly discuss data-related ideas in the context of Smart Manufacturing, Industrial Internet and Industrie 4.0, we used a first strategy Query 1 (Title = data AND ("industrie 4.0" OR "smart manufacturing" OR "industrial internet")) in Scopus. We selected journal and conference articles in English and German languages from 2010 to 2020 (Scopus provides English translation of a limited number of German publications). This corpus was built in June 2020. We use the term 'data' to consider a wide range of data types, transactions and sources that are associated with this area of activity. For example, data can include sensor measurements about physical elements (e.g. temperature and pressure) and production efficiency (e.g. energy performance). Our search returned 118 publications, consisting of all available publications that demarcated these concepts using 'data'. We then manually checked the titles, abstracts and keywords of the publications, and removed irrelevant ones. The final cleaned dataset contains 96 publications.

3.1.2. Identifying relevant keywords

The titles, abstracts and keywords of the cleaned dataset were analysed using the content analysis tool QDA Miner in combination with its sister text-mining module WORDSTAT. Content analysis has been used as part of broader scientometric strategies to evaluate progress and gaps in emerging technological domains (Rezaeian et al., 2017). QDA Miner provides various statistics such as TF-IDF, term frequency and the number of publications in which terms are found. We extracted individual words and phrases along with their statistical measures. The raw keyword set contained a total of 300 unique keywords. These were manually screened to remove nonsensical and trivial terms, resulting in a parent set of 77 keywords.

These keywords were then subjected to additional processing as follows. First, complementary terms related to the concepts in the parent set – "industry 4.0", "cyber-physical", "iiot", "industrial iot" – were grouped together. To these, we added the terms "factory" and "shop floor" to capture papers that express the same interpretation of *convergence*, even though these papers may not explicitly use the concepts in their title, abstract or keywords. Combining all these terms expands our search string to ("industrie 4.0" OR "industry 4.0" OR "smart manufacturing" OR "cyber manufacturing" OR industrial internet" OR "iiot" OR "industrial iot" OR "cyber physical" OR "cyberphysical" OR "factory" OR "shop floor"). This string (Query 2) represents a reasonable search space from which papers that are judged to embody *convergence* can be extracted.

Second, we selected terms to operationalise *convergence*. We classified the remaining keywords in the parent set according to their frequencies (ranging from 7 to 80) and TF-IDF scores (between 8 and 48). These values were calculated in the above text mining process for all terms in all of the documents in the corpus. We selected terms with frequencies equal to or greater than 30 and TF-IDF weights equal to or greater than 24 because these values were judged to reasonably cover the scale and scope of the *convergence* domain. Also, because the terms are *distinctively frequent*, we viewed them to represent important topics within the corpus. Based on these criteria, eight keywords were selected: "analytics", "transmission", "network", "security", "model", "monitoring", "sensor" and "communication", each representing a specific data-centric capability. The meanings of these keywords were then examined in relation to specific phrases, and also by manually checking titles and abstracts. Association between specific words mean that a distinct aspect of the domain is being discussed. Thus, phrases like "data monitoring", "sensing data", "data acquisition", "data analytics", "data modelling", "data transmission" and "secure communication" (Table 2) help place the keywords in the specified context, and categorise them into five subfields: Monitoring, Analytics, Modelling, Transmission and Security.

Subfield	Keyword	Relevant phrases (Examples)	No. of publications with this keyword	Total frequency of the keyword	TF-IDF score of the keyword
Monitoring	Monitoring	Data monitoring, Tool wear monitoring	12	31	27.9
	Sensor	Sensing data, Data collection (using sensors), Data acquisition (using sensors)	14	33	27.4
Analytics	Analytics	Data analytics, Industrial analytics	24	65	47.8
Modelling	Model	Data modelling, Manufacturing process modelling	20	42	28.4
Transmission	Transmission	Data transmission, Real-time transmission	10	35	34.2
	Network	Wireless network, Software-defined networking	22	51	32.4
	Communication	Data communication, Wireless communication	16	32	24.8
Security	Security	Secure communication, Security in Industrie 4.0	15	38	30.5

Table 2: Subfields with keywords and phrases.

3.2. Enriching the lexicon at subfield levels

These subfields are consistent with the data-centric capabilities indicated in Table 1 (Security is a notable addition). Despite this congruence between lexical and conceptual components, we argue that there is scope for it to be expanded at a subfield level to more comprehensively operationalise *convergence*. To do this, we pursued a modular query enrichment methodology similar to Mogoutov and Kahane (2007), whereby subfield keywords were extracted in another round of querying in Scopus and then tested for inclusion into the lexicon. We briefly define the subfields and discuss the results below.

Monitoring: Monitoring is the starting point in Smart Manufacturing and involves sensing to collect data on physical conditions and industrial processes (Gallaher et al., 2016, Schneider, 2018). Similarly, monitoring is considered as a foundational capability in an Industrial Internet infrastructure because it facilitates the reading of data from sensors (Lin et al., 2015, p. 26). Data collected by sensors is then used to support subsequent applications like analytics and modelling (Gallaher et al., 2016, Wu et al., 2017). However, data collection is not just limited to sensors – it is also gathered from other sources, including traditional automation equipment (e.g. supervisory control and data acquisition systems) as well as more advanced human-machine interfaces (such as augmented reality devices) (see, for example, a review of data collection means and methodologies by Ćwikła (2014)). We tested the keywords "monitor*" and "sens*" by repeating steps 3.1.1. (using query 2) and 3.1.2., and added two relevant keywords to the lexicon: "acquisition" and "collection" (examples of usage of the selected keywords in titles are shown below).

Additional keyword	**Scopus search query and publication example(s)** (TITLE ("monitor*" OR "sens*") AND TITLE ABS (Query 2))
Acquisition	Online **data acquisition** and analysis using multi-sensor network system for **smart manufacturing** (Ohannessian et al., 2019)
Collection	Cloud-enabled Smart **Data Collection** in Shop Floor Environments for **Industry 4.0** (Bosi et al., 2019)

Analytics: According to Industrial Internet, data analytics encompass data science capabilities to "*transform and analyse massive amounts of data from sensors to extract useful information that can deliver specific functions, give operators insightful information and recommendations, and enable real-time business and operational decisions*" (Lin et al., 2015, p. 83). We pursued the approach discussed above, using "analytics" in a Scopus search query. After testing for statistically relevant words, we added the keywords "learning" (as in machine learning and reinforcement learning), "neural network" and "intelligence" (as in artificial intelligence) to our lexicon because they are frequently used in regards to analytic solutions to enhance manufacturing processes. Data analytics in manufacturing includes a range of machine learning techniques (Wuest et al., 2016), which is a branch of artificial intelligence (Lechevalier et al., 2014). One of the most popular machine learning approaches is artificial neural networks (Schwabacher and Goebel, 2007), that can be used to build predictive models using large volumes of manufacturing data (Wu et al., 2017).

Additional Keyword	**Scopus search query and publication example(s)** (TITLE (“analytics”) AND TITLE ABS (Query 2))
Learning	**Machine Learning** approach for Predictive Maintenance in **Industry 4.0** (Paolanti et al., 2018)
Neural network	Cyber-based design for additive manufacturing using **artificial neural networks** for **Industry 4.0** (Elhoone et al., 2020)
Intelligence	Sustainable industrial systems within kernel density analysis of **artificial intelligence** and **industry 4.0** (Soebandrija et al., 2018)

Modelling: Modelling includes operational research techniques that have long been used in manufacturing (Greasley, 2005, Smith et al., 1994) to “*understand the states, conditions and behaviors of the systems under control and those of peer systems by interpreting and correlating data gathered from sensors and peer systems*” (Lin et al., 2015, p. 29). Virtual models built from sensor data provide valuable insights into existing or planned systems and thus enable better decision-making during activities such as shop floor layout optimisation (Eklin et al., 2009), production process control (Iassinovski et al., 2008) and inventory control (Rezg et al., 2005). We replicated the above search approach using “model*” as query and found one statistically relevant keyword “simulation”. Simulation, often used as an interchangeable term with modelling, projects a real-world system in a virtual form (such as interactive 3D charts) to provide interpretative views of the system under different conditions and constraints and hence, support decision-making (Kibira et al., 2015). Discrete event simulation is arguably the most popular operational research technique used in industrial and manufacturing applications (Greasley and Edwards, 2019). While terms such as augmented reality, virtual reality and visualisation did not reach the threshold to be separately listed, these technologies are contained in modelling and simulation (where papers including these keywords are captured using the Scopus search query), particularly in papers concerned with simulating physical and virtual environments.

Additional Keyword	**Scopus search query and publication example(s)** (TITLE (“model*”) AND TITLE ABS (Query 2))
Simulation	An Integrative User-Level Customized Modeling and **Simulation** Environment for **Smart Manufacturing** (Kim et al., 2019)

Transmission: Ubiquitous data transmission serves as the common thread that binds the entire spectrum of *cyber-physical convergence* capabilities, from data sensing to analysis and modelling (Gallaher et al., 2016, Lin et al., 2015). We searched for publications using “transmi*”, “network*” and “communicat*”. The results were text mined as above, leading to the inclusion of “wireless” to our lexicon. Authors working in the domain use “communication” as a broad capability that concerns the transmission of data among distributed hardware (e.g. sensors and actuators) and software resources through the application of wireless communication protocols and associated technologies such as Zigbee, Bluetooth and 6LowPAN (see Meng et al. (2016) for an overview on machine-to-machine communication for industrial applications).

Additional Keyword	**Scopus search query and publication example(s)** (TITLE (“transmi*” OR “network*” OR “communicat*”) AND TITLE ABS (Query 2))
Wireless	**Wireless** Networked Control Systems with Coding-Free Data Transmission for **Industrial IoT** (Liu et al., 2019)

Security: Security encompasses a plethora of cybersecurity-related concepts, technologies and practices to safeguard critical manufacturing assets (ENISA, 2018). Although security is not explicitly indicated in the definitions of *cyber-physical convergence* in Table 1, it has come to form a fundamental pillar in all three concepts. According to Industrial Internet, “*security is the condition of the system being protected from unintended or unauthorized access, change or destruction*” (Schrecker et al., 2016, p. 16). Security is the key system characteristic that most affects the trustworthy treatment of Industrie 4.0 data and protection from cyber attacks (Jänicke et al., 2016). We tested the keyword “secur*” as above and added “attack” to the lexicon (*Note:* Although it has a negative connotation, “attack” was identified as relevant and popular in our text-mining approach because researchers use it in different security-related contexts: “*to block an attack*”, “*encryption-based cyber attack*” or “*securing machines against attacks*”).

Additional Keyword	**Scopus search query and publication example(s)** (TITLE (“secur*”) AND TITLE ABS (Query 2))
Attack	Energy-Based Detection of Defect Injection **Attacks** in **IoT-Enabled Manufacturing** (Monroy et al., 2018)

Our modular approach, thus, yields a total of 16 keywords (Figure 3) that collectively delineate the *cyber-physical convergence* research domain. There were other keywords we considered but did not add to our lexicon because they were judged to be adequately captured by other keywords. For example, the term “predictive” is often used in the contexts of “predictive models” and “predictive analytics”. Our approach also does not take into account technical details of specific capabilities, unless they match our mid-to-high frequency and TF-IDF selection criteria. As a case in point, we include “learning” in analytics but drop related keywords such as “linear regression”. We are aware that our keyword selection leaves out a number of terms that others may consider important. The fact, however, remains, that our approach builds on the notion of *convergence* as characterised in Table 1 and operationalises it with reasonable precision for the next stage of bibliometric analysis.

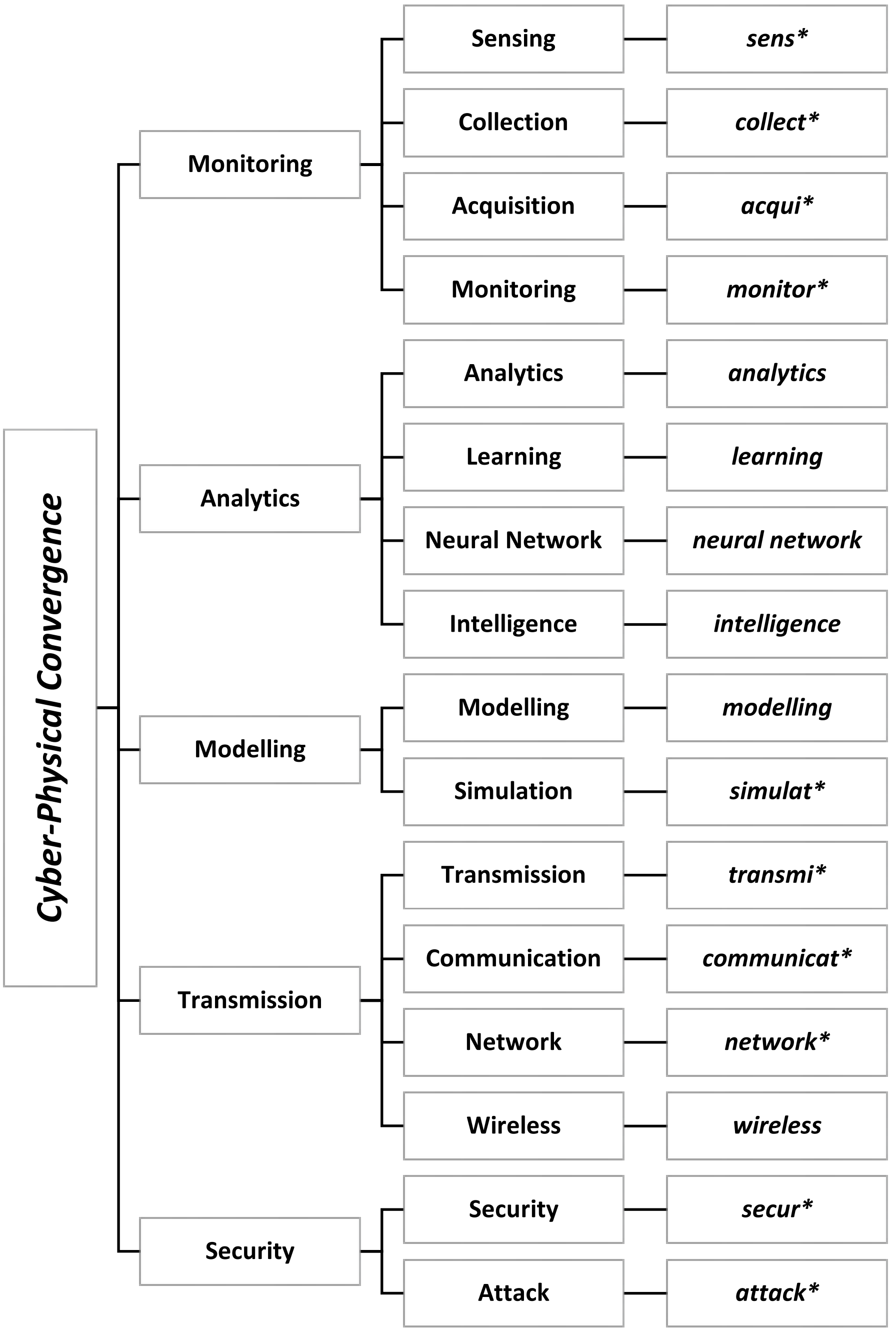

Figure 3: Subfield keywords to operationalise *cyber-physical convergence* and delineate research domain. See section 3.1 for lexicon-building methodology.

3.3. Bibliometric search

The keywords were consolidated into a bibliometric search approach (Table 3) applying appropriate exclusion terms to remove extraneous records. Further discussions on how we added exclusion terms are discussed in Appendix 2.

Subfield	Search Strategies
Monitoring	TITLE (“monitor*” OR “sens*” OR “collect*” OR “acqui*”) AND TITLE-ABS (Query 2) AND NOT TITLE-ABS (“sensitive” OR “sensual” OR “collective*” OR “pollut*” OR “water” OR “agricultur*” OR “oil” OR “sewage” OR “waste” OR “bacteri*” OR “coli” OR “bacillus” OR “cultur*” OR “microb*” OR “literature” OR “survey” OR “review” OR “overview” OR “trend” OR “challenge” OR “opportunit*” OR “cities”)
Analytics	TITLE (“analytics” OR “learning” OR “neural network” OR “intelligence”) AND TITLE-ABS (Query 2) AND NOT TITLE-ABS (“skills” OR “educat*” OR “curricul*” OR “school” OR “training” OR “module” OR “course” OR “graduat*” OR “pollut*” OR “water” OR “agricultur*” OR “oil” OR “sewage” OR “waste” OR “bacteri*” OR “coli” OR “bacillus” OR “cultur*” OR “microb*” OR “literature” OR “survey” OR “review” OR “overview” OR “trend” OR “challenge” OR “opportunit*” OR “cities”)
Modelling-and-Simulation	TITLE (“modelling” OR “simulat*”) AND TITLE-ABS (Query 2) AND NOT TITLE-ABS (“pollut*” OR “water” OR “agricultur*” OR “oil” OR “sewage” OR “waste” OR “literature” OR “survey” OR “review” OR “overview” OR “trend” OR “challenge” OR “opportunit*” OR “cities”)
Transmission	TITLE (“transmi*” OR “communicat*” OR “network*” OR “wireless”) AND TITLE-ABS (Query 2) AND NOT TITLE-ABS (“sens*” OR “neural” OR “bacteri*” OR “coli” OR “bacillus” OR “cultur*” OR “microb*” OR “literature” OR “survey” OR “review” OR “overview” OR “trend” OR “challenge” OR “opportunit*” OR “cities”)
Security	TITLE (“secur*” OR “attack*”) AND TITLE-ABS (Query 2) AND NOT TITLE-ABS (“literature” OR “survey” OR “review” OR “overview” OR “trend” OR “challenge” OR “opportunit*” OR “cities”)

Table 3: Bibliometric search strategies with subfield inclusion and exclusion terms.

While *cyber-physical convergence* can be effectuated across different sectors (e.g., retail and social media) and organisational functions (e.g., digital marketing), the objective of this paper is to investigate the *convergence* research domain associated with manufacturing and industrials. Our scope of manufacturing and industrials includes the following: (1) sectors: automotive (mobility), aerospace and defence, power grids, oil and gas, industrial machinery, and general engineering and production; and (2) functions: manufacturing processes, and factory and shop floor processes). We implemented this scope in our paper by: (a) structuring Query 2 with relevant keywords (e.g., factory, shop floor and cyber manufacturing); (b) applying selective search strategies (see Table 3 and Appendix 2) to eliminate non-relevant sectors such as smart cities, agriculture and sewage treatment; and (c) manually screening the bibliometric search results to weed out non-relevant publications. Our approach (Table 3) offers greater precision than approaches that associate Smart Manufacturing, Industrie 4.0 and Industrial Internet with random signifiers. We consider journal and conference articles that relate to the above subfields in manufacturing and industrials such as power grids and automotive. We applied the above search strategies (in June 2020) to retrieve English and a limited number of German publications (with English translations) in Scopus between the period 2010 and 2019 under the subject areas Computer Science, Engineering, Chemical Engineering, Material Science, Mathematics, Decision Sciences, Energy, Physics and Astronomy, and Business, Management and Accounting. The publications extracted were cleaned using VantagePoint text-mining software and examined manually (mainly titles but at times abstracts to better interpret topics of empirical enquiry). The cleaned dataset included: Monitoring (1360 publications), Analytics (1210), Modelling-and-Simulation (2156), Transmission (1329) and Security (1175) – a total of 7230 publications (*Note:* For a publication judged to be associated with multiple subfields, we assign it equally to all subfields). If we remove duplicate records, our dataset comprises 7141 publications.

3.4. Analysis of results

In this section, we present the results of analyses using the subfield and overall publication datasets derived from the search.

Domain growth and trajectories

Worldwide *cyber-physical convergence* scientific output (N = 7230) encompassing the five subfields grew by almost nine times over the past decade from 217 papers in 2010 to over 1900 papers in 2019 (Figure 4).

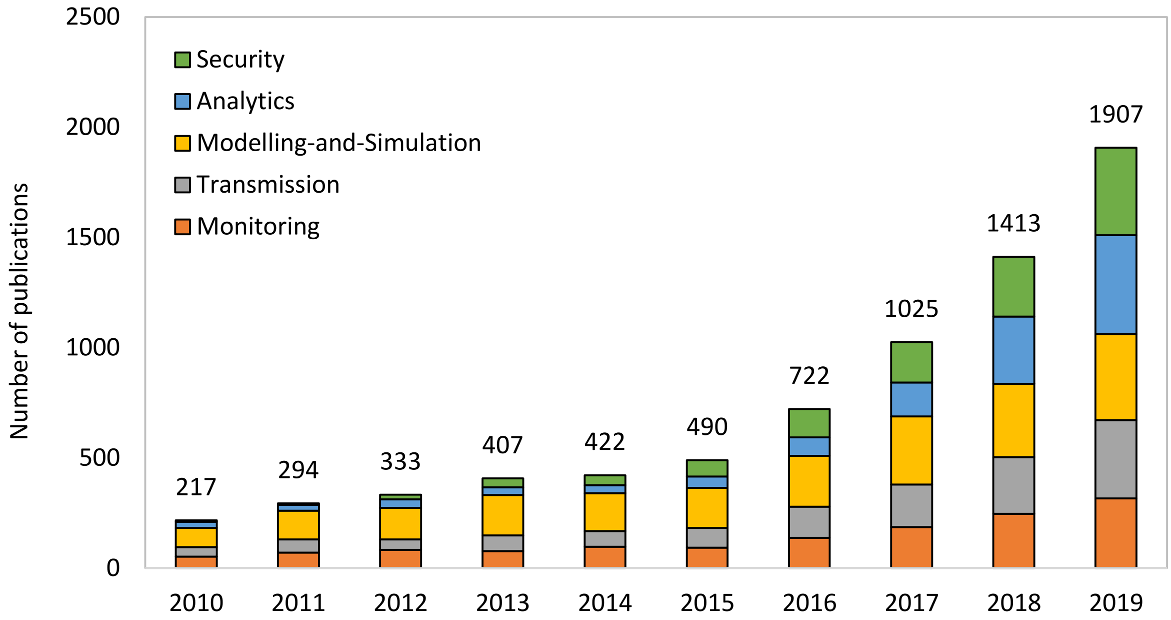

Figure 4: Worldwide *Cyber-physical convergence* publications (*N* = 7230, 2010–2019) encompassing all five subfields. Yearly totals are shown on top of the stacked bars.

The overall publication growth inflection point arrived in 2016, whereby the number of articles in the most recent 4-year period (2016–2019) more than doubled the number of articles in the prior 6-year period (2010–2015). Yet, despite this surge in knowledge production, it is evident that not all subfields have grown at the same pace or contributed equally to the increase. While Modelling-and-Simulation, Monitoring and Transmission collectively formed about 84% of the total number of *cyber-physical convergence* output in 2010, these dropped to 56% in 2019, exhibiting an evolution from telecommunications and operations research to a more computer science orientation, that of Analytics and Security.

To look deeper into relationships between ideas within the subfields and map their temporal shift, we draw on the idea of keyword co-occurrence. The theoretical foundation of co-occurrence analysis is based on “actor network” – the idea that texts produced by scientists serve as a powerful tool to trace the social and epistemic dimensions of their research (Callon et al., 1986). Co-word analyses help produce maps of clustered keywords that shed light on the structure of scientific research (Callon et al., 1991). By comparing cluster changes for different time periods, the evolution of a research field can be assessed objectively (Callon et al., 1983). Cluster changes result from changes in policy, funding and other factors. Thus, co-occurrence mapping serve as a reasonable approach to uncover meaningful insights into the composition and dynamics of the *cyber-physical convergence* domain.

The maps presented in this study are developed using a software tool called VOSviewer, where each keyword is represented as a node and the line (that is, the link) between two nodes represents the number of publications in which they occur together (van Eck and Waltman, 2017). The thicker the line, the stronger will be the link, meaning that they have appeared together in greater number of publications. Furthermore, the distance between two terms indicates the strength of their relation – a shorter distance refers to a larger number of co-occurrences as they reside in similar research fields (Liao et al., 2018). VOSviewer was used to select the top 100 most frequent terms from all keywords (author-supplied and Scopus-indexed) in the dataset. These were manually screened to remove terms that were redundant or irrelevant. The final set (containing around 10 important representative terms for each subfield) was then characterised using VOSviewer into clusters, producing an overlay map to assess the evolution of the research domain.

Figure 5 shows four clusters of co-occurred terms, with Monitoring and Transmission forming a single cluster. This means terms related to these two subfields reside in similar research areas. Keywords co-occurrence shows what the global scientific community considers important at specific periods in time (Waaijer et al., 2010). In the 2015–2016 period, terms related to Monitoring ("sensing", "sensors", "monitoring"), Modelling-and-Simulation ("simulation", "discrete event simulation") and Transmission ("wireless sensor networks", "wireless communications") co-occur with higher frequency than Analytics and Security keywords. Another interesting observation is the emergence of the node "data acquisition" as an extention of "monitoring". This might be an indication of the growing recognition of data collection (using sensors) as a foundational step of *cyber-physical convergence*.

From 2017 onward, co-occurring terms related to Analytics ("machine learning", "artificial intelligence", "deep learning") and Security ("network security", "cyber security", "intrusion detection") started to outweigh terms associated with Monitoring, Modelling-and-Simulation and Transmission. This is because of the steady growth in average number of yearly Analytics and Security publications. Two additional insights can be gained from Figure 5: (a) **External strength:** the more numerous and thicker (that is, stronger) links between terms belonging to the Analytics and Security clusters indicate the growing importance of artificial intelligence and machine learning to address network security and cybersecurity research problems; and (b) **Internal coherence:** the greater density of links that tie the terms within these two individual clusters represents the latters' capacity to maintain themselves and further develop over time (Callon et al., 1991). These features, in turn, highlight the potential of Analytics and Security to drive the scientific and strategic growth of the research domain.

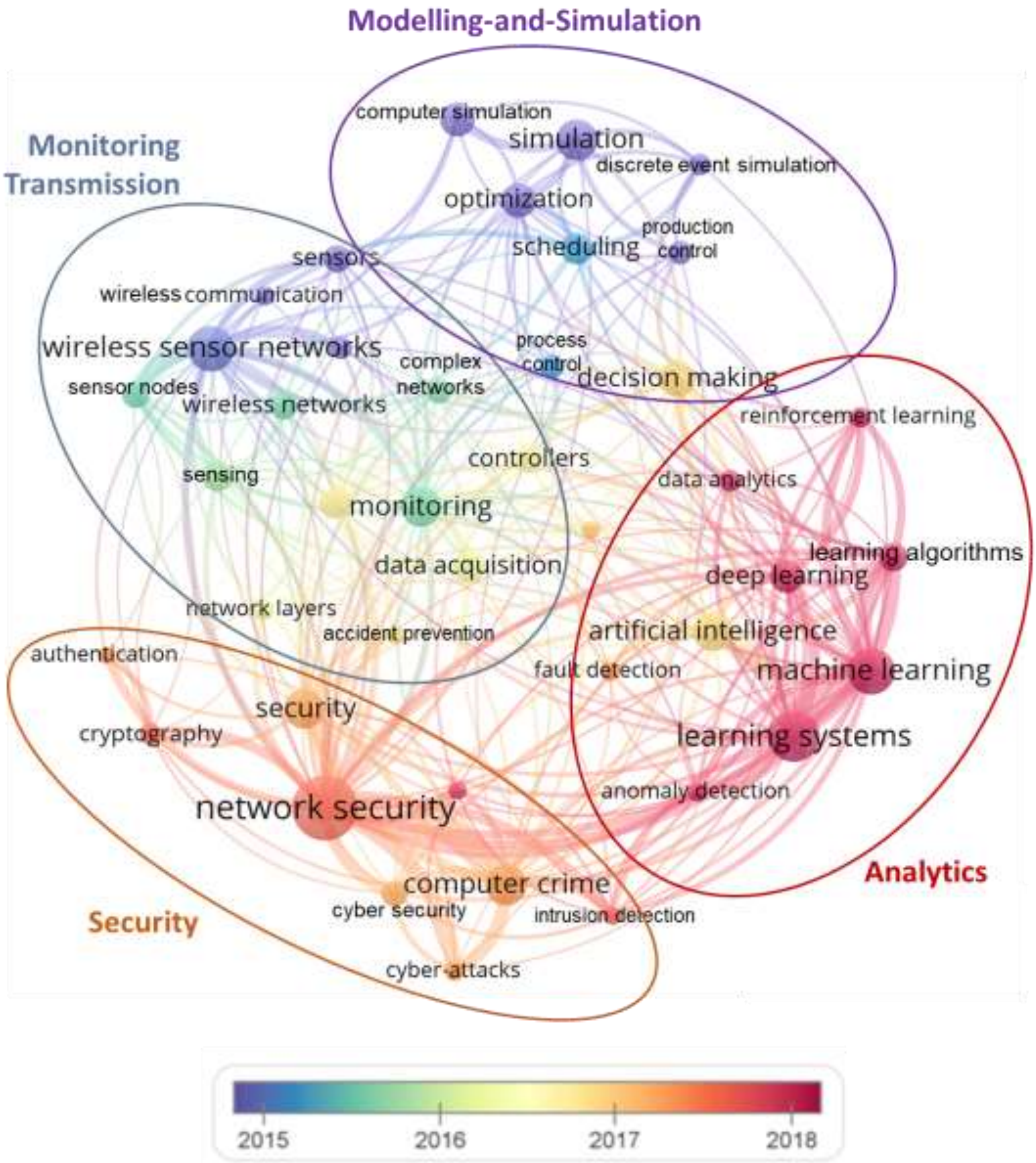

Figure 5: Overlay co-occurrence map of the 50 most frequent terms covering the five subfields. Visualisation with VOSviewer. Node sizes scale with the frequency of terms in the *cyber-physical convergence* dataset. Bigger the node size (and its label), greater is its importance within a research cluster. Node colour shows the relative progression of co-occurring terms over the 2015–2018 period. Between 2015–2016, the average number of publications in which terms related to, for example, Modelling-and-Simulation ("simulation", "discrete event simulation", "scheduling") co-occur outweighs Analytics and Security. This is an indication of research topics that are considered important subjects of empirical investigation at specific points in time. Link widths between terms are proportional to the number of publications in which they co-occur and distance between co-occurred terms show the extent to which they reside in similar research fields.

These evolving dynamics can also be viewed from another perspective – that of the contribution of each subfield to the overall *cyber-physical convergence* research domain. Figure 6 shows papers related to each subfield as a proportion of the total number of *convergence* papers in a year, and how these proportions have changed over the 2010-2019 period. While Monitoring and Transmission have remained largely unchanged, the proportion of Modelling-and-Simulation related publications has been steadily decreasing since 2013. This is in sharp contrast to Analytics, whose proportion has more than doubled during the same period. Indeed, the absolute number of Modelling-and-Simulation publications has recorded an increase of only 17% in 2019 over the previous year, compared to 47% for Analytics (Figure 4). One likely driver of these shifting focus of research could be the application of machine learning to support those activities that had been traditionally carried out by means of simulation, such as process optimisation, shop floor scheduling and production control (Figure 5). This pattern also reflects an emerging literature that incorporates the use of data analytics in discrete event simulation applications (Greasley and Edwards, 2019). Security forms the other significant driver of the domain, forming 21% of the total publications in 2019 (Figure 6). In other words, Security's growth in 2019 is more than 50 times its 2010 output (Figure 4).

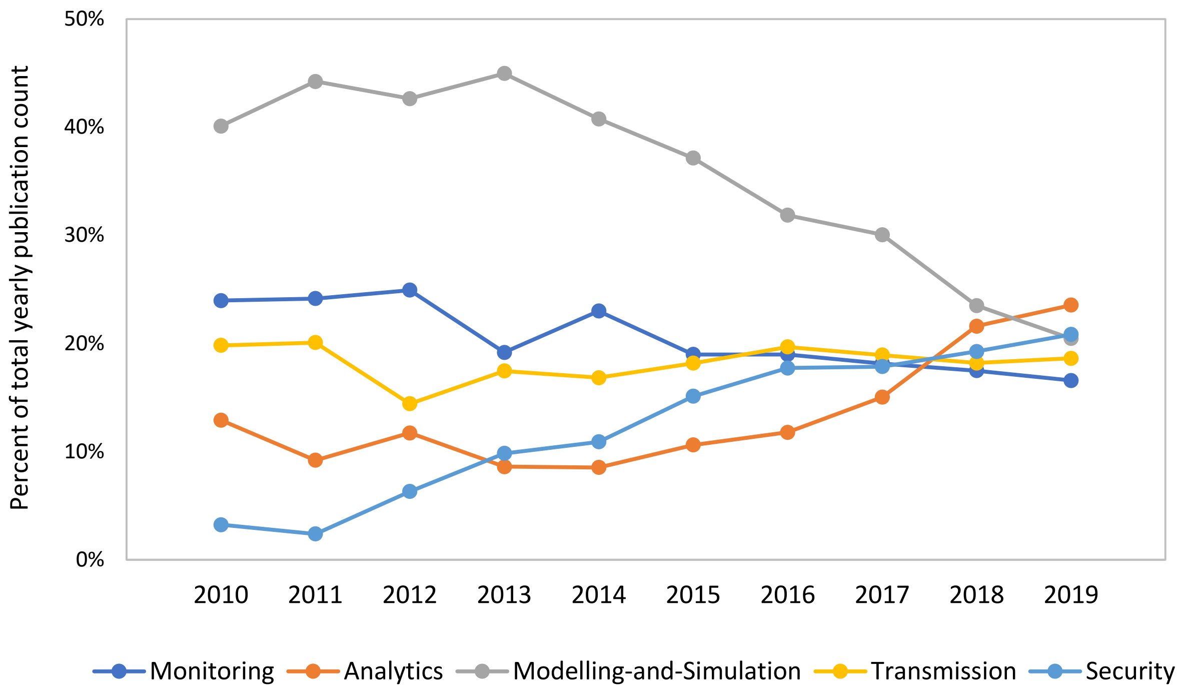

Figure 6: Contribution of each subfield (as a proportion) to the total number of yearly *cyber-physical convergence* publications.

Security forms the other significant driver of the domain, forming 21% of the total publications in 2019, which represents an 18% percentage points increase over 2010. In other words, Security's growth in 2019 is more than 50 times its 2010 output. By comparison, the proportion of Monitoring and Transmission publications with respect to the yearly total has largely remained flat since 2013 onward (Figure 6).

We measured *cyber-physical convergence* publications (as per paper) by year with publications from Scopus for same years, using "manufacturing" and other variations. We find that *cyber-physical convergence* has a much faster growth rate. Publications on *cyber-physical convergence* have become increasingly important as a focus of research, expanding from 1.1% of all publications on manufacturing topics recorded in Scopus in 2010 to 1.7% in 2015, rising to 4.8% in 2019.

National performance indicators

Our *cyber-physical convergence* publication dataset includes publications by authors located in 91 countries, but only 18 countries produce 100 or more papers. Figure 7 shows the yearly publication trend of the top ten countries for the period 2010–2019. This figure is based on at least one author being affiliated with an institute in one of these countries, meaning that an internationally co-authored paper is equally allocated to authors of all affiliated countries. These ten countries account for 66.3% of the 8825 worldwide publications (after allocation). In terms of total number of articles, the U.S. is the leading country, with U.S.-based authors contributing to 17.5% (or 1542 articles) of the worldwide publications. Authors based in China contributed to 14.9% of publications (1312 articles), followed by Germany (10.5% or 923 articles), Italy (4.6% or 409 articles) and U.K. (3.4% or 300 articles).

Papers with U.S. and German authors have progressively grown since 2011, coinciding with the coining and subsequent institutionalisation of Smart Manufacturing, Industrial Internet and Industrie 4.0 between 2011 and 2014. However, the steepest rise in research output is in China: articles with at least one Chinese author grew by more than 12 times from 29 in 2010 to 370 in 2019, overtaking the U.S. (343 articles) in 2019. Figure 7 shows that Chinese research output hit growth inflection in 2016. This shift resonates with China's Made in China 2025 initiative, a state-led strategy unveiled in 2015, inspired by Germany's Industrie 4.0, to upgrade the country's manufacturing capability by investing in ten priority sectors, including information technology (Belton et al., 2019, Aaronson, 2019).

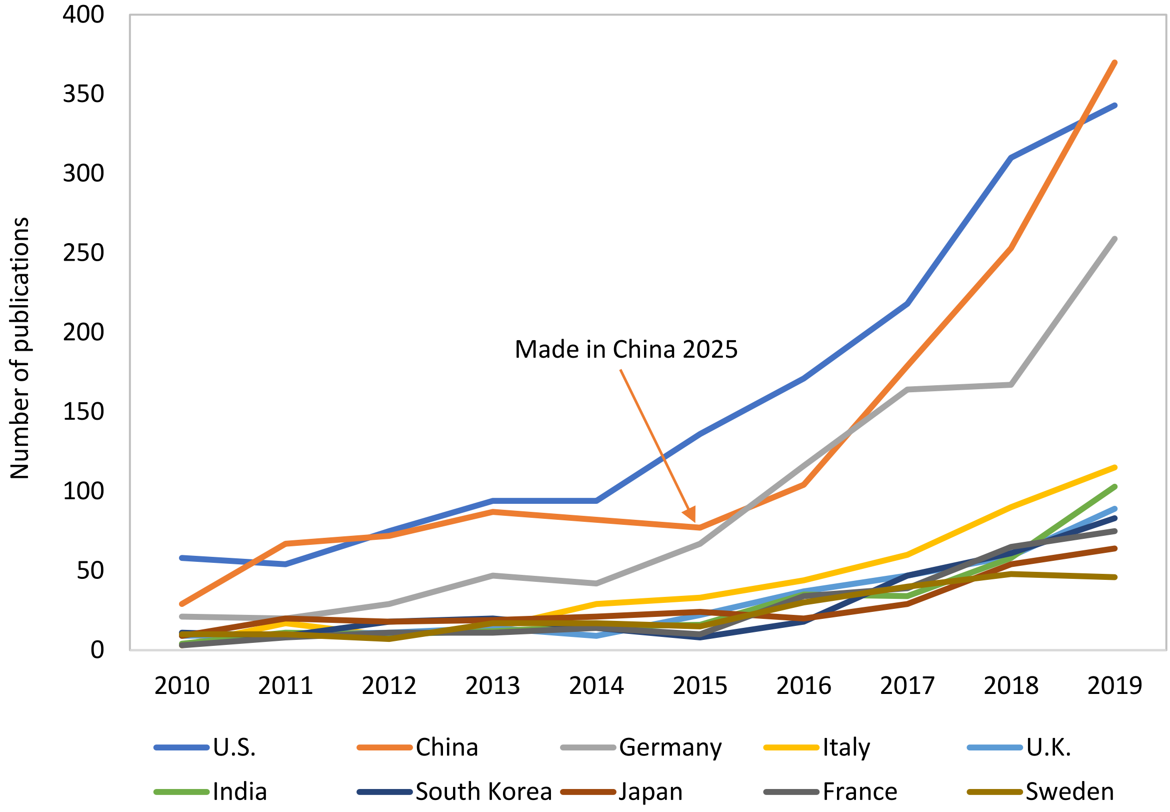

Figure 7: *Cyber-physical convergence* publication trend of the top ten countries by author affiliations. Worldwide publication count exceeds $N = 7230$ due to multiple author affiliations of internationally co-authored papers.

These findings, when compared to growth patterns reported in other scientific fields such as artificial intelligence, reveal additional insights. As Figure 7 shows, Chinese growth inflection point arrived much later in *cyber-physical convergence* (2016) when compared to artificial intelligence (in 2010) (Liu et al., 2021). Chinese *cyber-physical convergence* output overtook the U.S. in mid-2018. By comparison, artificial intelligence publications by authors affiliated with Chinese organisations pulled ahead of U.S.-affiliated authors as early as 2010 (Liu et al., 2021).

While the U.S., China and Germany dominate scientific output in absolute terms (Figure 7), growth in terms of proportion over the 10-year period presents a different picture. According to this measure, Italy demonstrated the largest increase in *cyber-physical convergence* scientific output with a 36.7-fold (or 3667%) rise from three papers in 2010 and to 113 papers in 2019. Italy is followed by India and France, both countries showing about 24-fold increase in that same period. These developments coincide with the introduction of national industrial strategies and programmes in those countries – namely, Industria 4.0 (Italy), Make in India

(India) and Industrie du Futur (France). By comparison, the proportion increase in U.S. (5-fold), Chinese (11-fold) and German (11-fold) publications were relatively modest.

At a subfield level, the bulk of research production between 2010–2019 by the top ten countries is further concentrated at the top three (Figure 8). Thus, publications with at least one U.S.-, China- and Germany-based author account for more than half the total top ten output for all five subfields: Monitoring – 59.4%, Analytics – 65.5%, Modelling-and-Simulation – 63.4%, Transmission – 63.5% and Security – 71.6%.

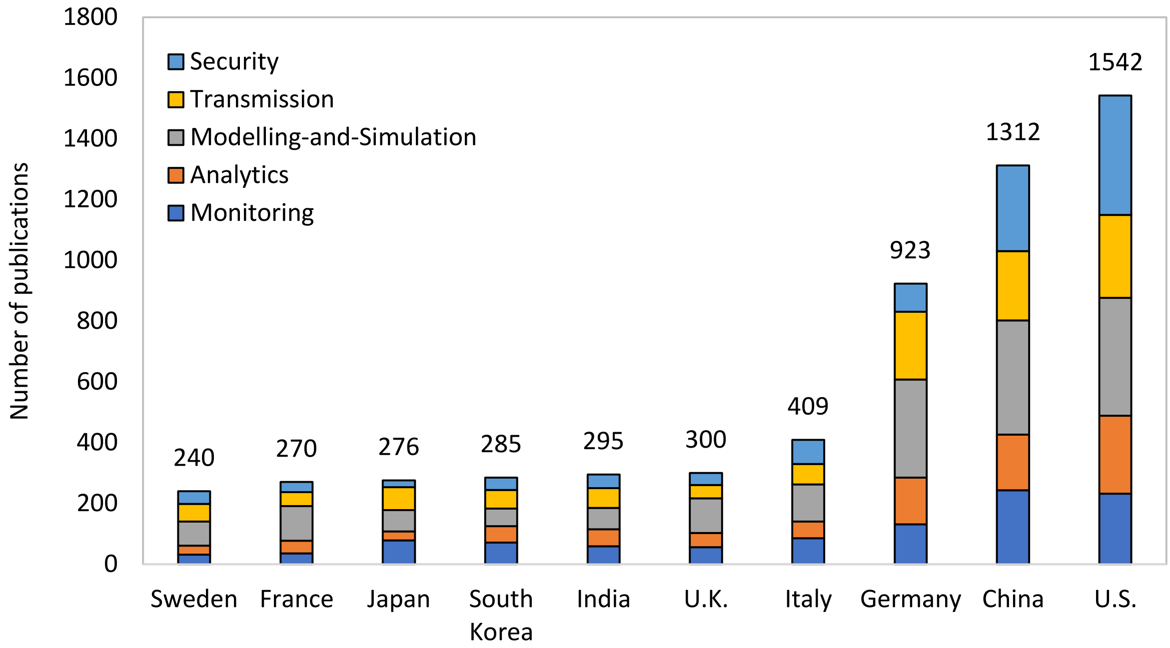

Figure 8: Subfield breakdown of *cyber-physical convergence* publication counts of top ten countries by author affiliations.

Security is the biggest differentiator in U.S. and Chinese cyber-physical convergence research output when compared to articles by Germany-based authors (Figure 9). For China, the contribution of Security-related papers to the total number of yearly publications increased sharply from 3% in 2010 to 32% in 2019.

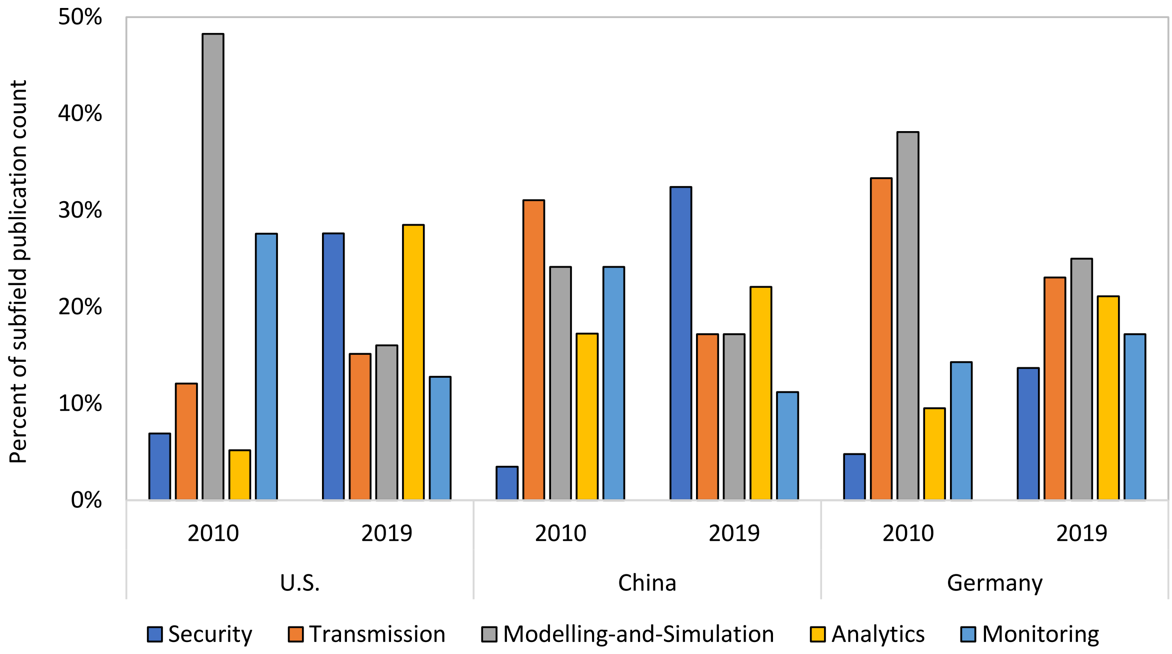

Figure 9: Contribution of each subfield to the total number of *cyber-physical convergence* publications in 2010 and 2019.

As regards to the scientific quality of papers, citation counts are often used to demonstrate research significance (Cole and Cole, 1973, p. 21) and impact (Aksnes et al., 2019). Although not without caveats (e.g., field, national, and reputational influences on citation propensities), the higher the number of citations of a publication, the higher is its perceived quality and influence (Durieux and Gevenois, 2010). We examined highly cited publications (more than 100 citations) for U.S., China and Germany for two time periods: 2010-2014 and 2015-2019 (Table 4). During the first period, authors affiliated with U.S. organisations published 375 papers out of which 21 received more than 100 citations, followed by China (two out of 337 papers) and Germany (one out of 159 papers). Between 2015 and 2019, U.S. publication output was 1167 among which five papers fell within the highly cited category. For Germany, it was 2 out of 764 papers. By comparison, seven papers out of a total of 975 by authors affiliated with Chinese institutes received more than 100 citations, indicating a growing measure of quality and impact of Chinese contributions.

	2010-2014		2015-2019	
	Total no. of publications	Highly-cited publications	Total no. of publications	Highly-cited publications
U.S.	375	21	1167	5
China	337	2	975	7
Germany	159	1	764	2

Table 4: Citation count, *cyber-physical convergence* publications (highly citated publications have more than 100 citations).

An important way in which research trajectory changes – certain topics are more preferred than others – is through the agency of funding (Franzoni et al., 2011). Because resources are limited, new areas compete with and may grow at the expense of established research fields. External funding is particularly important to scale up emerging scientific fields, be it in terms of expanding collaborations, producing academic papers or attracting attention. The influence of financial incentives on research is perhaps most overt when research is directly commissioned by sources with specific policy interests (Boswell and Smith, 2017). For our *cyber-physical convergence* publication dataset, we analyse funding acknowledgement data of articles following Wang and Shapira (2011). Of the 7858 articles in our dataset, around 40% or 3101 papers acknowledge one or more funders. The growth inflection point occurred in 2016, whereby the percentage of such papers (as a fraction of the 3101 dataset) increased from 8% in 2010–2015 to 80% in 2016–2019. The highest growth, both in absolute and proportionate terms, occurred in Security where the number of articles grew from one in 2010 to 260 in 2019, followed by Analytics (increased from three articles in 2010 to 240 in 2019).

The types of sponsor acknowledged include research councils, regional and federal government departments and agencies, universities, scholarship and fellowship programmes, foundations and firms. There were more than 6500 names (and their variants) of sponsors. Different variants (abbreviations, acronyms, spelling errors) of sponsors were cleaned (manually and using VantagePoint) and then combined. After two rounds of cleaning, we zoomed in on the top ten sponsors in terms of articles with funding acknowledgement (Figure10). All sponsors in the top ten group funded more papers in the 2016–2019 period than in the prior 2010–2015 period. These ten sponsors accounted for 53.9% of the funded articles (or 1466 articles) published between 2010 and 2019, within which a group of five funders is acknowledged in about 83% of the articles. The National Natural Science Foundation of China (NSFC) leads with 427 articles, followed by the European Union or EU (308 articles), the National Science Foundation or NSF (288 articles) and the two German sponsors – the Federal Ministry of Education and Research or BMBF (101 articles) and the German Research Foundation or DFG (81 articles). Overall, the US has three funders among the top ten global sponsors (including U.S. Department of Energy and Defense Advanced Research Projects Agency), Germany has two, China, South Korea, Taiwan, the U.K. and the EU have one each.

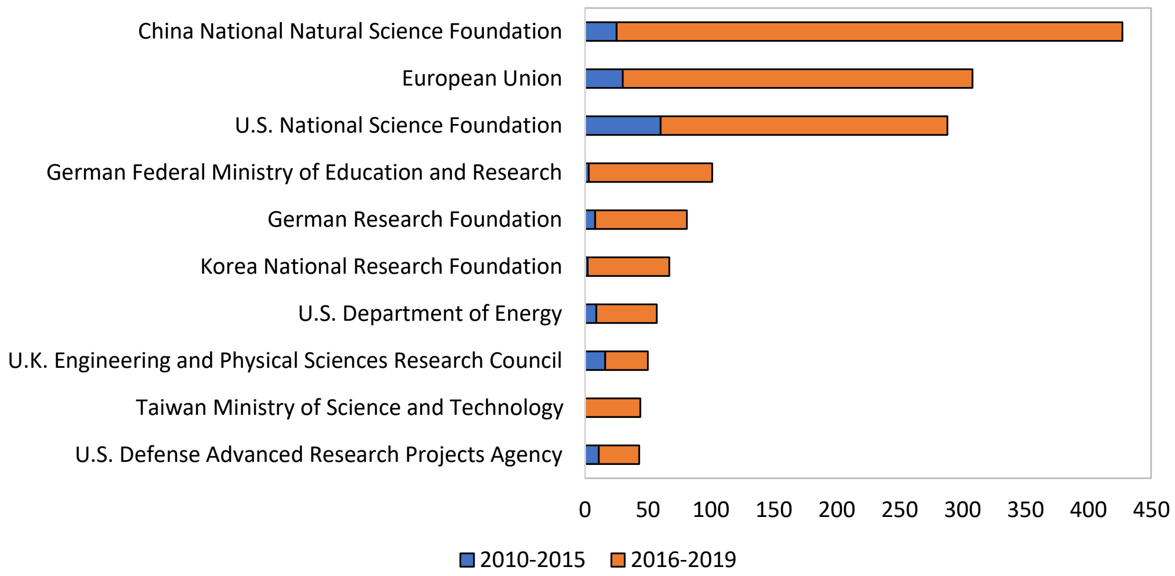

Figure 10: Top ten research sponsors of *cyber-physical convergence* publications (with one or more funding acknowledgements, 2010–2019).

The number of papers with acknowledged NSFC sponsorship has increased by more than 15 times since the inflection point in 2016 and overtaken both the EU and the NSF by 2017. In addition to NSFC, which is managed by the Ministry of Science and Technology of China, two other notable Chinese funders in the top 20 research sponsors are the National Basic Research Program of China (973 Program) and the China Postdoctoral Science Foundation. Backed by such support, China's output in *cyber-physical convergence* research has increased by over 300% from 77 articles in 2015 to 370 in 2019.

Comparing the recent with the previous year, the biggest growth in the number of funded papers occurred in NSFC-sponsored Security articles – 71 articles in 2019. This mirrors China's broader trend to regulate digital activities and counter threats in manufacturing through the coming into effect of its new Cybersecurity Law in June 2017 (Belton et al., 2019). The growth in Security articles also coincides with a steady rise in total funding for "*information* *and network security*" projects under the following NSFC programmes: General Program, Young Scientists Fund and Fund for Less Developed Regions (Table 5).

NSFC Funding Programmes	2015	2016	2017	2018
General Program Projects	12.9	12.8	13.9	14.1
Young Scientists Fund	4.7	4.6	5.8	5.7
Fund for Less Developed Regions	1.2	1.1	1.1	1.1
Total	18.9	18.5	20.8	20.9

Table 5: NSFC funding for "*information and network security*" projects[1] (In millions of US Dollars).

In line with these developments, but more specifically for manufacturing, the NSFC–Guangdong Province joint funding programme was initiated in 2015 to support basic research in "*security characteristics of discrete manufacturing networks based on industrial big data*" (NSFC, 2019). A second strategic area for NSFC sponsorship is Analytics. The number of Analytics articles acknowledging NSFC funding increased from one in 2015 to 44 in 2019. Comparable to this growth are: (a) the increase in NSFC's support (under General Program Projects) for artificial intelligence (AI) applications from USD 12.9 million in 2015 to USD 13.9 million in 2017; and (b) the creation of the NSFC–Zhejiang joint fund programme for the period 2015 – 2019 to support the integration of AI into manufacturing operations (NSFC, 2019).

The steady increase in Analytics and Security articles acknowledging NSF funding reflects the United States' broader R&D priorities. In October 2016, the U.S. National Artificial Intelligence Research and Development Strategic Plan was advanced to invest in basic AI R&D that may impact a wide range of industries, including manufacuring. Since then, the NSF has funded/co-funded research in artificial intelligence, notably the Real-Time Machine Learning (RTML) programme with the Defense Advanced Research Projects Agency (NITRD, 2019). The aim of this programme is to explore "*high-performance, energy-efficient hardware for real-time machine-learning*" and spans 1017 active awards of approximately USD 400 million (as of August 3, 2020)[2]. Similarly, the 2011 report, "*Trustworthy Cyberspace: Strategic Plan for the Federal Cybersecurity Research and Development Program*" laid the foundation to make cybersecurity a cross-cutting research agenda across multiple sectors such as

[1] Funding data collected from *NSFC Guide to Program* reports for the years 2015–2018 (latest available data). Currency converted from Chinese Yuan (CNY) to US Dollar (USD) at 1 CHY = 0.14276 USD as on 30 July 2020

[2] Total dollar amounts and number of awards calculated from the NSF RTML Active Awards database: https://www.nsf.gov/awards/award_visualization.jsp?org=NSF&pims_id=505640&ProgEleCode=7564,7798&from=fund

manufacturing. To support this plan, the NSF through its flagship Secure and Trustworthy Cyberspace (SaTC) programme (Amla et al., 2012), has invested around USD 430 million in 950 active awards (as of August 3, 2020)[3]. A more recent 2018 report called "*Strategy for American Leadership in Advanced Manufacturing*" builds on earlier versions of PCAST recommendations and identifies cybersecurity as a priority area. In accordance with this report, the NSF has launched its Future Manufacturing programme that aims to fund basic research in cybersecurity and analytics in 2020 by awarding grants of USD 500,000 to 2 million per year for up to five years (NSF, 2020).

Policy labelling in research output

So far, we have seen that since the coining of the three concepts, *cyber-physical convergence* research output has grown, albeit at varying pace, magnitude and direction, in their countries of origin as well as other countries, some of which have introduced their own versions of the concepts. At least part of this growth has been driven by concerted government-led funding efforts. While policy influence on knowledge is most overt when research is funded, there are other, subtler ways in which such influence may be felt (Boswell and Smith, 2017). A rich body of literature in science policy, sociology of science and science and technology studies (see for a critical review Boswell and Smith, 2017) argues that scientific knowledge production picks up signals from different public and private sources to secure legitimacy and support. Authors label new or existing work with the hope or expectation of their research being found serendipitously by funding bodies (Mogoutov and Kahane, 2007). Labels have also been used as "umbrella terms" (terms of unclear scope that may be useful in gaining support in emerging and strategic fields, for example, nanotechnology) to signal changes in research content caused by direct or indirect external intervention (Rip and Voß, 2013). In this section, we take a closer look into policy labelling in *cyber-physical convergence* research, and what this reflects about policy discourses and approaches in different countries. In doing so, we contribute to "policy shapes knowledge" debates (Boswell and Smith, 2017), which have important implications for the development of a research domain (Merz and Sormani, 2016).

We focus on keywords that are both exact and synonymous matches to Smart Manufacturing, Industrie 4.0 and Industrial Internet. Hence, Industrial Internet includes the variants "industrial

[3] Total dollar amounts and number of awards calculated from the NSF SaTC Active Awards database: https://www.nsf.gov/awards/award_visualization.jsp?org=NSF&pims_id=504709&ProgEleCode=8060&from=fund

internet”, “industrial iot”, “iiot” and “internet of things”, Smart Manufacturing includes “smart manufacturing”, “advanced manufacturing” and “cyber manufacturing” and Industrie 4.0 includes “industry 4.0” and “industrie 4.0”. We argue that specific keywords related to Smart Manufacturing, Industrie 4.0 and Industrial Internet act as satisfying proxies of policy influence on research growth. These terms are extracted from titles, abstracts and author keywords of the *cyber-physical convergence* dataset. Publications with multiple terms are assigned equally to all indicated terms. This process yielded a total of 1653 publications spanning the top ten countries by author affiliations for the period 2010–2019. Table 6 presents the breakdown of labelled publications.

Countries	Smart Manufacturing	Industrie 4.0	Industrial Internet
United States	77	61	156
China	49	75	166
Germany	25	259	90
Italy	10	110	50
United Kingdom	17	48	48
India	11	27	78
South Korea	20	25	55
Japan	8	15	38
France	6	36	38
Sweden	3	23	29

Table 6: Number of labelled publications, top ten countries by author affiliation.

As might be expected, U.S.- and Germany-based authors lead in labelling their research in terms of their respective national industrial policies. However, only about 5% (or 77 papers) of all publications by U.S.-based authors are labelled as “smart manufacturing” compared to about 28% or 256 “industry 4.0” labelled articles by German authors. These findings are insightful from the perspective of manufacturing innovation approaches in these two countries. While Industrie 4.0 is characterised by a state-coordinated strategy, the U.S. approach is market-driven with less stricter government mandates (Belton et al., 2019). Consequently, German policy emphasis has been on developing a consistent industrial discourse through concerted efforts of different stakeholders (Horst and Santiago, 2018). Over time, as our findings suggest, this discourse has become reinforced in German scientific output through greater reference to Industrie 4.0. For the U.S., the impetus for growth in knowledge production has come through a constellation of discourses. This originally included Smart Manufacturing as a data-centric approach to achieve energy efficiency in production. Smart Manufacturing itself was carved out of a broader, whole-of-government initiative called Advanced Manufacturing to boost U.S. national security and manufacturing competitiveness (PCAST, 2014). In its most recent embodiment, this constellation includes Future Manufacturing: "*manufacturing that either does not exist today or exists only at such small scales that it is not viable*" (NSF, 2020). Such an evolving discourse could mean usage of Smart Manufacturing has not achieved particular prominence among U.S.-affiliated authors. This is because authors continuously experiment with different labelling strategies to seek attention for their scientific endeavours (Rip and Voß, 2013).

"Industrial internet", on the other hand, shows wider acceptance and usage by all of the top ten countries, particularly by China-based researchers (N = 166 or about 13% of all Chinese publications). This is likely due to the morphing of "industrial internet" into a more general terminology that indicates a technological enabler of Industrie 4.0 (Jeschke et al., 2017). For China, developing indigenous industrial internet platforms is a central priority of its top-down, state-led Made in China 2025 programme (Arcesati et al., 2020). The Chinese Ministry of Industry and Information Technology established (and governs) the public-private Alliance of Industrial Internet in 2016 (along the lines of the U.S. Industrial Internet Consortium) to facilitate collaborative research in this space. In packaging their research using "industrial internet" labels, China-based authors also seem to compete for interactions with NSFC, which has actively endorsed funding opportunities to create a cross-sectoral industrial internet platform (NSFC, 2019).

4. Discussion and conclusions

Cyber-physical convergence – the pervasive integration of data into manufacturing and industrial processes – has assumed increasing importance among policymakers for its perceived ability to drive productivity growth, strengthen national security and address other ideological and business concerns (Evans and Annunziata, 2012, Kagermann et al., 2013, PCAST, 2011). This has spurred government and private interventions in many countries under the banner of industrial and/or innovation policy concepts. Policy interest, in parallel, has been matched as well as influenced by growth in academic research. Yet, as a domain of study, *cyber-physical convergence* has itself received little attention, particularly in terms of delineation, operationalisation and measurement.

The analysis in this paper has presented a newly constructed definition of *cyber-physical convergence* that is closely aligned with three pioneering concepts: the U.S. state-managed Smart Manufacturing, the U.S. private sector-led Industrial Internet, and the German public-private coordinated Industrie 4.0. These concepts are pioneering because they recognised and institutionalised the emerging importance of *convergence,* and set the scene for other countries to follow suit. We defined *convergence* as a "*data-centric workflow that makes use of specific capabilities, starting with data acquisition from physical sources, and then continuing through to the generation of actionable insights*". A core set of policy and technical documents related to the three concepts (Appendix 1) was the main source, based on which we conceived our definition. This definition appreciates the fact that data is now an increasingly valuable asset for manufacturing firms and will dictate interactions between humans, systems and firms (Belton et al., 2019).

We then built a lexicon of relevant keywords and progressively enriched it using natural language processing approaches. Because our characterisation of *convergence* starts at the data level, it could be more precisely – and originally – operationalised. This resulted in the original demarcation of the *convergence* research domain into five subfields: (i) Monitoring, (ii) Analytics, (iii) Modelling-and-Simulation, (iv) Transmission and (v) Security. Finally, we conducted a targeted bibliometric search, and measured growth and trajectories of the domain in the period 2010–2019. In tracking these dynamics, we zoomed in on the internal structures of the subfields (in terms of their degree of temporal development), national outputs (supported by funding structures) and policy labelling.

In summary, these analyses lead to three perspectives on the interactions of policy with the growth and trajectories of academic research. Our first perspective examines how the overall growth of the *cyber-physical convergence* research domain varies by subfields. We find the relative contributions of the subfields to the overall output converging over time, an indication of the influence of external stimuli to trigger a greater preference of certain topics. A key message of this perspective is that, beginning the mid-2010s, Analytics and Security have assumed central positions within the *convergence* research domain. Particularly the latter, even though it was not explicitly included in the definitions of *convergence* advanced by the three concepts (Table 1), has grown by leaps and bounds in terms of scientific output. These dynamics seem to coincide with a "strategic turn" in science policy, as policymakers now consider artificial intelligence and cybersecurity central to national industrial growth (Arcesati et al., 2020, PCAST, 2020).

Our second perspective analyses how exactly national-level interventions have contributed to this evolution of the domain. Funding represents perhaps the most direct mechanism through which policy shapes the production of knowledge (Boswell and Smith, 2017). Within this governance system, research funding agencies act as key mediators (Gläser and Laudel, 2016). While the shift towards Analytics has been largely driven by U.S., China and Germany affiliated authors, it is the former two that have emerged as foremost in Security. These efforts, as we have seen, are backed by concerted U.S. and Chinese funding programmes. We find that Smart Manufacturing, Industrial Internet and Industrie 4.0 have become a reference point for Chinese policy makers and funding agencies, leading to a "funding race" in which China has significantly scaled up and pulled ahead in funding Analytics and Security research.

Our final perspective sheds light on the extent to which the three concepts are embedded – as labels – in *cyber-physical convergence* research. Labelling research using policy terminology is a subtle, but important, way to validate discourses, gain recognition and mobilise resources (Rip and Voß, 2013). In comparing the labelling patterns of Smart Manufacturing, Industrial Internet and Industrie 4.0, we find that Industrie 4.0, because of its unified and consistent discourse, has become more persistent in German scientific content when compared to Smart Manufacturing in U.S. knowledge output. By comparison, the "industrial internet" label has achieved wider acceptance, particularly in China, which coincides with China's centrally coordinated R&D programmes to develop an indigenous industrial internet. Indeed, these observations support the assessment that different types of policy approaches and discourses have different macro-level manifestations in research (Gläser and Laudel, 2016).

Shifting attention to the limitations of our paper, we acknowledge the potential imperfections associated with defining and then operationalising the notion of *cyber-physical convergence*. This paper focuses on *cyber-physical convergence* in manufacturing and industrials sectors, so the findings may not be relevant to other sectors such as services and retail. Our attempt in this paper has been to shed light on the potential influence of policy on research by intrinsically building up the research domain itself from original policy guidelines. Even then, we recognise that establishing direct causal links is problematic because of the many complex variables between the two (Gläser and Laudel, 2016, Leydesdorff, 1989), and also due to the multiple units of analysis (scientists, documents and cognitive content) of the scientific enterprise itself. As such, deeper insights of the mechanisms through which policy influences scientific output would help in the better interpretation of results. While our definition of *convergence* directly draws on conceptions advocated by the original proponents of Smart Manufacturing, Industrial Internet and Industrie 4.0, there may be differences regarding what other experts and stakeholders consider important as to the definition. The choice of keywords to delineate the research domain may also be disputed. Although we have carefully selected and enriched the keywords using a modular text-mining approach, we may have dropped certain terms that others will consider. Then there are the usual limitations associated with Scopus in terms of the scope of indexed publications and completeness of funding acknowledgements.

Notwithstanding these limitations, our paper has moved the policy–research nexus forward by systematically defining and chalking out – for the first time to our knowledge – a boundary around the *cyber-physical convergence* research domain. The use of a controlled vocabulary helps navigate through "*uninformatively circular*" and random usage of concepts such as Industrial Internet and Industrie 4.0 (Boyes et al., 2018) that have rapidly swamped the domain. Moreover, our analysis opens a new viewpoint for understanding the influence of policy on the growth and directionality of academic research. In so doing, the paper offers insights for academia, policy and practice, particularly at a time when innovation in manufacturing and industrials has assumed new prominence.

We envisage several opportunities for future research. First, it will be useful to continue refining our subfield operationalisation and bibliometric strategies by testing and adding relevant keywords. Second, including terms related to policy concepts that have emerged after Industrie 4.0, Smart Manufacturing and Industrial Internet will expand the scope of the current research design. Finally, in this paper, we found differences between the labelling patterns of U.S., China- and Germany-based authors. Yet our knowledge of the causal impacts of policy on the direction or quality of scientific output is limited. Therefore, more detailed analyses of our results at a micro-level (scientists and experts as the units of analysis), applying qualitative or sociological approaches such as case studies (Leydesdorff, 1989, Rip, 1997) could provide a better understanding of causality in the direction or quality of research. Also, increasing access to funding data may provide further insights into the influence of policy on *cyber-physical convergence* research.

Appendix 1

Concepts	Policy and technical reports
Smart Manufacturing	*Report to the President on Ensuring American Leadership in Advanced Manufacturing* (PCAST, 2011)
	Report to the President: Accelerating U.S. Advanced Manufacturing (PCAST, 2014)
	Economic Analysis of Technology Infrastructure Needs for Advanced Manufacturing - Smart Manufacturing (Gallaher et al., 2016)
	An Introduction to Smart Manufacturing: Lighting the Path to the Value of your Data (CESMII, 2019)
Industrial Internet	*Industrial Internet: Pushing the Boundaries of Minds and Machines* (Evans and Annunziata, 2012)
	Industrial Internet Reference Architecture (Lin et al., 2015)
Industrie 4.0	*Cyber-Physical Systems. Driving Force for Innovation in Mobility, Health, Energy and Production* (Hellinger and Seeger, 2011)
	Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry (Kagermann et al., 2013)
	Living in a Networked World: Integrated Research Agenda Cyber-Physical Systems (agendaCPS) (Geisberger and Broy, 2015)

Appendix 2

Exclusion terms were introduced to remove extraneous articles associated with the subfield keywords. For all five subfields, we excluded the terms “survey”, “review”, “overview”, “trend”, “challenge”, “opportunity” and “landscape” to eliminate articles that summarise current states of research or forecast future research areas. This is because the focus of this paper is to systematically define the technical and technological boundaries of the research domain. The terms “bacteria”, “microbe”, “coli” and “culture” were added to avoid records associated with transmission and monitoring of, for example, microbes. Similarly, “cities”, “agriculture”, “water”, “waste”, “garbage” and “sewage” were included to eliminate as they captured articles outside our scope of manufacturing and industrials (namely, urban and regional planning, crop protection and farming and waste management).

Subfield keywords	Specific exclusion terms	Comments
(“monitor*” OR “sens*” OR “collect*” OR “acqui*”)	“sensitive”, “sensual”	Eliminates extraneous records associated with the truncation “sens*”
(“analytics” OR “learning” OR “neural network” OR “intelligence”)	“skills”, “educat*”, “curricul*”, “school”, “training”, “module”, “course”, “graduat*”	Excludes extraneous articles associated with the keyword “learning” and “intelligence”
(“transmi*” OR “communicat*” OR “network*” OR “wireless”)	“sens*”, “neural”	Excludes articles associated with the truncation “sens*” and the keyword “neural” because they are captured under Analytics

References

AARONSON, S. A. 2019. The Trade Impact of Smart Factories. *Smart Factories: Issues of Information Governance.* Manufacturing Policy Initiative, Indiana University.

AKSNES, D. W., LANGFELDT, L. & WOUTERS, P. 2019. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. *Sage Open,* 9**,** 2158244019829575.

AMLA, N., ATLURI, V., EPSTEIN, J., GREENSPAN, S., MUHLBERGER, P., PIOTROWSKI, V., POLLINGTON, A., THOMPSON, K., TIAN, Z. & WEBER, S. 2012. Toward a secure and trustworthy cyberspace. National Science Foundation.

ANDREWS, D., CRISCUOLO, C. & GAL, P. 2016. The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy. *OECD Productivity Working Papers.* Paris: OECD.

ANGRISANI, L., ARPAIA, P., ESPOSITO, A. & MOCCALDI, N. 2019. A Wearable Brain-Computer Interface Instrument for Augmented Reality-based Inspection in Industry 4.0. *IEEE Transactions on Instrumentation and Measurement*.

ARCESATI, R., HOLZMANN, A., MAO, Y., NYAMDORJ, M., SHI-KUPFER, K., VON CARNAP, K. & WESSLING, C. 2020. China's Digital Platform Economy: Assessing Developments Towards Industry 4.0. Mercator Institute for China Studies.

BELTON, K. B., AUDRETSCH, D. B., GRAHAM, J. D. & RUPP, J. A. 2019. Who Will Set the Rules for Smart Factories? *Issues in Science and Technology,* 35**,** 70-76.

BENITEZ, G. B., AYALA, N. F. & FRANK, A. G. 2020. Industry 4.0 Innovation Ecosystems: An Evolutionary Perspective on Value Cocreation. *International Journal of Production Economics***,** 107735.

BONVILLIAN, W. B. 2012. Reinventing American Manufacturing: The Role of Innovation. *MIT Press*.

BOSI, F., CORRADI, A., FOSCHINI, L., MONTI, S., PATERA, L., POLI, L. & SOLIMANDO, M. Cloud-enabled Smart Data Collection in Shop Floor Environments for Industry 4.0. 2019 15th IEEE International Workshop on Factory Communication Systems, 2019. IEEE, 1-8.

BOSWELL, C. & SMITH, K. 2017. Rethinking policy 'impact': four models of research-policy relations. *Palgrave Communications,* 3**,** 1-10.

BOYES, H., HALLAQ, B., CUNNINGHAM, J. & WATSON, T. 2018. The Industrial Internet of Things (IIoT): An Analysis Framework. *Computers in Industry,* 101**,** 1-12.

CALLON, M., COURTIAL, J.-P., TURNER, W. A. & BAUIN, S. 1983. From Translations to Problematic Networks: An Introduction to Co-Word Analysis. *Social Science Information,* 22**,** 191-235.

CALLON, M., COURTIAL, J. P. & LAVILLE, F. 1991. Co-Word Analysis as a Tool for Describing the Network of Interactions Between Basic and Technological Research: The Case of Polymer Chemsitry. *Scientometrics,* 22**,** 155-205.

CALLON, M., RIP, A. & LAW, J. 1986. *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, Springer.

CERUTI, A., MARZOCCA, P., LIVERANI, A. & BIL, C. 2019. Maintenance in Aeronautics in an Industry 4.0 Context: The Role of Augmented Reality and Additive Manufacturing. *Journal of Computational Design and Engineering,* 6**,** 516-526.

CESMII. 2019. *An Introduction to Smart Manufacturing: Lighting the Path to the Value of your Data* [Online]. Clean Energy Smart Manufacturing Innovation Institute.

Available: https://www.osti.gov/servlets/purl/1506430 [Accessed 8 December 2019].
CHEN, G. & XIAO, L. 2016. Selecting Publication Keywords for Domain Analysis in Bibliometrics: A Comparison of Three Methods. *Journal of Informetrics,* 10**,** 212-223.
COLE, J. R. & COLE, S. 1973. *Social Stratification in Science,* Chicago, Illinois, The University of Chicago Press.
CONTI, M., DAS, S. K., BISDIKIAN, C., KUMAR, M., NI, L. M., PASSARELLA, A., ROUSSOS, G., TRÖSTER, G., TSUDIK, G. & ZAMBONELLI, F. 2012. Looking Ahead in Pervasive Computing: Challenges and Opportunities in the Era of Cyber–Physical Convergence. *Pervasive and Mobile Computing,* 8**,** 2-21.
CONTI, M., PASSARELLA, A. & DAS, S. K. 2017. The Internet of People (IoP): A New Wave in Pervasive Mobile Computing. *Pervasive and Mobile Computing,* 41**,** 1-27.
ĆWIKŁA, G. Methods of Manufacturing Data Acquisition for Production Management - A Review. Advanced Materials Research, 2014. Trans Tech Publications, 618-623.
DAVIS, J., MALKANI, H., DYCK, J., KORAMBATH, P. & WISE, J. 2020. Cyberinfrastructure for the democratization of smart manufacturing. *In:* SOROUSH, M., BALEDA, M. & EDGAR, T. (eds.) *Smart Manufacturing: Concepts and Methods.* Amsterdam: Elsevier.
DURIEUX, V. & GEVENOIS, P. A. 2010. Bibliometric indicators: quality measurements of scientific publication. *Radiology,* 255**,** 342-351.
EICHER, T. S. & ROEHN, O. 2007. Sources of the German Productivity Demise: Tracing the Effects of Industry ‐ Level Information and Communication Technology Investment. *German Economic Review,* 8**,** 211-236.
EKLIN, M., ARZI, Y. & SHTUB, A. 2009. Model for Cost Estimation in a Finite-Capacity Stochastic Environment Based on Shop Floor Optimization Combined with Simulation. *European Journal of Operational Research,* 194**,** 294-306.
ELHOONE, H., ZHANG, T., ANWAR, M. & DESAI, S. 2020. Cyber-Based Design for Additive Manufacturing Using Artificial Neural Networks for Industry 4.0. *International Journal of Production Research,* 58**,** 2841-2861.
ENISA 2018. Good Practices for Security of Internet of Things in the context of Smart Manufacturing. European Union Agency for Cybersecurity.
EVANS, P. & ANNUNZIATA, M. 2012. Industrial Internet: Pushing the Boundaries of Minds and Machines. GE.
EZELL, S. J. 2018. Why Manufacturing Digitalization Matters and How Countries Are Supporting It. Information Technology and Innovation Foundation.
FELSBERGER, A., QAISER, F. H., CHOUDHARY, A. & REINER, G. 2020. The impact of Industry 4.0 on the reconciliation of dynamic capabilities: evidence from the European manufacturing industries. *Production Planning & Control***,** 1-24.
FORT, T. C. 2017. Technology and Production Fragmentation: Domestic Versus Foreign Sourcing. *The Review of Economic Studies,* 84**,** 650-687.
FORT, T. C., PIERCE, J. R. & SCHOTT, P. K. 2018. New Perspectives on the Decline of US Manufacturing Employment. *Journal of Economic Perspectives,* 32**,** 47-72.
FRANK, A. G., MENDES, G. H., AYALA, N. F. & GHEZZI, A. 2019. Servitization and Industry 4.0 convergence in the digital transformation of product firms: A business model innovation perspective. *Technological Forecasting and Social Change,* 141**,** 341-351.
FRANZONI, C., SCELLATO, G. & STEPHAN, P. 2011. Changing Incentives to Publish. *Science,* 333**,** 702-703.

GALLAHER, M., OLIVER, Z., RIETH, K. & O'CONNOR, A. 2016. Economic Analysis of Technology Infrastructure Needs for Advanced Manufacturing - Smart Manufacturing. National Institute of Standards and Technology.

GAUCH, S. & BLIND, K. 2015. Technological Convergence and the Absorptive Capacity of Standardisation. *Technological Forecasting and Social Change,* 91**,** 236-249.

GEISBERGER, E. & BROY, M. 2015. Living in a Networked World: Integrated Research Agenda Cyber-Physical Systems (agendaCPS). Acatech.

GLÄSER, J. & LAUDEL, G. 2016. Governing Science: How Science Policy Shapes Research Content. *European Journal of Sociology,* 57**,** 117-168.

GREASLEY, A. 2005. Using System Dynamics in a Discrete-Event Simulation Study of a Manufacturing Plant. *International Journal of Operations & Production Management,* 25**,** 534-548.

GREASLEY, A. & EDWARDS, J. S. 2019. Enhancing Discrete Event Simulation with Big Data Analytics: A Review. *Journal of the Operational Research Society***,** 1-21.

HACKLIN, F., MARXT, C. & FAHRNI, F. 2009. Coevolutionary cycles of convergence: An extrapolation from the ICT industry. *Technological Forecasting and Social Change,* 76**,** 723-736.

HELLINGER, A. & SEEGER, H. 2011. Cyber-Physical Systems. Driving Force for Innovation in Mobility, Health, Energy and Production. Acatech.

HORST, J. & SANTIAGO, F. 2018. What Can Policymakers Learn From Germany's Industrie 4.0 Development Strategy? : United Nations Industrial Development Organization.

HU, X. & ROUSSEAU, R. 2015. From a Word to a World: The Current Situation in the Interdisciplinary Field of Synthetic Biology. *PeerJ,* 3.

IASSINOVSKI, S., ARTIBA, A. & FAGNART, C. 2008. A Generic Production Rules-Based System for On-Line Simulation, Decision Making and Discrete Process Control. *International Journal of Production Economics,* 112**,** 62-76.

JÄNICKE, L., JOCHEM, M., KLASEN, W., KOSCH, B., KRAMMEL, M., LINKE, L., MEHRFELD, J., SANDNER, M., TEUSCHER, A., WALLOSCHKE, T. & ZIMMERMANN, S. 2016. IT Security in Industrie 4.0: First Steps Towards Secure Production. Plattform Industrie 4.0.

JESCHKE, S., BRECHER, C., MEISEN, T., ÖZDEMIR, D. & ESCHERT, T. 2017. Industrial Internet of Things and Cyber Manufacturing Systems. *Industrial Internet of Things.* Springer.

JORGENSON, D. W., HO, M. S. & STIROH, K. J. 2008. A Retrospective Look at the U.S. Productivity Growth Resurgence. *Journal of Economic perspectives,* 22**,** 3-24.

KAGERMANN, H., WAHLSTER, W. & HELBIG, J. 2013. Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry. Acatech and Forschungsunion.

KIBIRA, D., HATIM, Q., KUMARA, S. & SHAO, G. Integrating Data Analytics and Simulation Methods to Support Manufacturing Decision Making. 2015 Winter Simulation Conference, 2015. IEEE, 2100-2111.

KIM, B. S., JIN, Y. & NAM, S. 2019. An Integrative User-Level Customized Modeling and Simulation Environment for Smart Manufacturing. *IEEE Access,* 7**,** 186637-186645.

LECHEVALIER, D., NARAYANAN, A. & RACHURI, S. Towards a Domain-Specific Framework for Predictive Analytics in Manufacturing. 2014 IEEE International Conference on Big Data, 2014. IEEE, 987-995.

LEYDESDORFF, L. 1989. The relations between qualitative theory and scientometric methods in science and technology studies: introduction to the topical issue. *Scientometrics,* 15**,** 333-347.

LIAO, H., TANG, M., LUO, L., LI, C., CHICLANA, F. & ZENG, X.-J. 2018. A Bibliometric Analysis and Visualization of Medical Big Data Research. *Sustainability,* 10**,** 166.

LIN, S.-W., MILLER, B., DURAND, J., JOSHI, R., DIDIER, P., CHIGANI, A., TORENBEEK, R., DUGGAL, D., MARTIN, R. & BLEAKLEY, G. 2015. Industrial Internet Reference Architecture. Industrial Internet Consortium.

LIU, N., SHAPIRA, P. & YUE, X. 2021. Tracking developments in artificial intelligence research: constructing and applying a new search strategy. *Scientometrics,* 126**,** 3153-3192.

LIU, W., POPOVSKI, P., LI, Y. & VUCETIC, B. 2019. Wireless Networked Control Systems with Coding-Free Data Transmission for Industrial IoT. *IEEE Internet of Things Journal*.

LU, Y., MORRIS, K. & FRECHETTE, S. 2016. Current Standards Landscape for Smart Manufacturing Systems. National Institute of Standards and Technology.

LUHN, H. P. 1958. The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development,* 2**,** 159-165.

MASONI, R., FERRISE, F., BORDEGONI, M., GATTULLO, M., UVA, A. E., FIORENTINO, M., CARRABBA, E. & DI DONATO, M. 2017. Supporting Remote Maintenance in Industry 4.0 Through Augmented Reality. *Procedia Manufacturing,* 11**,** 1296-1302.

MAZZEI, D., BALDI, G., FANTONI, G., MONTELISCIANI, G., PITASI, A., RICCI, L. & RIZZELLO, L. 2020. A Blockchain Tokenizer for Industrial IOT trustless applications. *Future Generation Computer Systems,* 105**,** 432-445.

MEINDL, B., AYALA, N. F., MENDONÇA, J. & FRANK, A. G. 2021. The four smarts of Industry 4.0: Evolution of ten years of research and future perspectives. *Technological Forecasting and Social Change,* 168**,** 120784.

MENG, Z., WU, Z., MUVIANTO, C. & GRAY, J. 2016. A Data-Oriented M2M Messaging Mechanism for Industrial IoT Applications. *IEEE Internet of Things Journal,* 4**,** 236-246.

MERZ, M. & SORMANI, P. 2016. *The Local Configuration of New Research Fields: On Regional and National Diversity*, Springer International Publishing.

MILLER, B. & ATKINSON, R. D. 2014. Raising European Productivity Growth Through ICT. Information Technology and Innovation Foundation.

MILLER, S. 2018. *Smart Manufacturing and the Clean Energy Smart Manufacturing Innovation Institute* [Online]. Clean Energy Smart Manufacturing Innovation Institute. Available: https://cdn.osisoft.com/osi/presentations/2018-uc-san-francisco/UC18NA-D2TR04-Texas-A-M-University-SMiller-Clean-Energy-Smart-Manufacturing-Innovation-Institute.pdf [Accessed 17 November 2020].

MOGOUTOV, A. & KAHANE, B. 2007. Data Search Strategy for Science and Technology Emergence: A Scalable and Evolutionary Query for Nanotechnology Tracking. *Research Policy,* 36**,** 893-903.

MONROY, S. A. S., LI, M. & LI, P. Energy-Based Detection of Defect Injection Attacks in IoT-Enabled Manufacturing. 2018 IEEE Global Communications Conference, 2018. IEEE.

MRUGALSKA, B. & WYRWICKA, M. K. 2017. Towards Lean Production in Industry 4.0. *Procedia Engineering,* 182**,** 466-473.

MUHURI, P. K., SHUKLA, A. K. & ABRAHAM, A. 2019. Industry 4.0: A Bibliometric Analysis and Detailed Overview. *Engineering Applications of Artificial Intelligence,* 78**,** 218-235.

NITRD 2019. The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update. Networking and Information Technology Research and Development.

NSF. 2020. *Future Manufacturing - Program Solicitation* [Online]. National Science Foundation. Available: https://www.nsf.gov/pubs/2020/nsf20552/nsf20552.htm [Accessed 10 August 2020].

NSFC 2019. National Natural Science Fund Guide to Programs 2019. National Natural Science Foundation of China.

O'SULLIVAN, E., LÓPEZ-GÓMEZ, C. & HEAD, P. L. U. 2017. An international review of emerging manufacturing R&D priorities and policies for the next production revolution. *The Next Production Revolution.* Organisation for Economic Co-operation and Development.

OHANNESSIAN, H., WARSCHEWSKE, F. & WOITAG, M. Online Data Acquisition and Analysis Using Multi-Sensor Network System for Smart Manufacturing. Smart SysTech 2019; European Conference on Smart Objects, Systems and Technologies, 2019. VDE, 1-7.

OLDHAM, P., HALL, S. & BURTON, G. 2012. Synthetic Biology: Mapping the Scientific Landscape. *PLoS One,* 7.

PAOLANTI, M., ROMEO, L., FELICETTI, A., MANCINI, A., FRONTONI, E. & LONCARSKI, J. Machine Learning Approach for Predictive Maintenance in Industry 4.0. 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications, 2018. IEEE, 1-6.

PCAST 2011. Report to the President on Ensuring American Leadership in Advanced Manufacturing. President's Council of Advisors on Science and Technology.

PCAST 2012. Report to the President on Capturing Domestic Competitive Advantage in Advanced Manufacturing. President's Council of Advisors on Science and Technology.

PCAST 2014. Report to the President: Accelerating U.S. Advanced Manufacturing. President's Council of Advisors on Science and Technology.

PCAST 2020. Recommendations for Strengthening American Leadership in Industries of the Future. President's Council of Advisors on Science and Technology.

PISANO, G. P. & SHIH, W. C. 2009. Restoring American Competitiveness. *Harvard Business Review,* 87**,** 114-125.

QIN, J., LIU, Y. & GROSVENOR, R. Data Analytics for Energy Consumption of Digital Manufacturing Systems Using Internet of Things Method. 2017 13th IEEE Conference on Automation Science and Engineering (CASE), 2017. IEEE, 482-487.

REZAEIAN, M., MONTAZERI, H. & LOONEN, R. 2017. Science foresight using life-cycle analysis, text mining and clustering: A case study on natural ventilation. *Technological Forecasting and Social Change,* 118**,** 270-280.

REZG, N., CHELBI, A. & XIE, X. 2005. Modeling and Optimizing a Joint Inventory Control and Preventive Maintenance Strategy for a Randomly Failing Production Unit: Analytical and Simulation Approaches. *International Journal of Computer Integrated Manufacturing,* 18**,** 225-235.

RIP, A. 1997. Qualitative conditions of scientometrics: The new challenges. *Scientometrics,* 38**,** 7-26.

RIP, A. & VOß, J.-P. 2013. Umbrella terms as mediators in the governance of emerging science and technology. *Science, Technology & Innovation Studies,* 9**,** 39-59.

SCHNEIDER, D. 2018. *The Process of Smart Manufacturing* [Online]. Clean Energy Smart Manufacturing Innovation Institute. Available: http://mepol.org/wp-content/uploads/2018/06/Mfg-Summit-18-Dean-Schneider-presentation.pdf [Accessed 8 January 2020].

SCHRECKER, S., SOROUSH, H., MOLINA, J., LEBLANC, J., HIRSCH, F., BUCHHEIT, M., GINTER, A., MARTIN, R., BANAVARA, H. & ESWARAHALLY, S. 2016. Industrial Internet of Things Volume G4: Security Framework. Industrial Internet Consortium.

SCHWABACHER, M. & GOEBEL, K. A Survey of Artificial Intelligence for Prognostics. AAAI Fall Symposium: Artificial Intelligence for Prognostics, 2007. 108-115.

SHAPIRA, P., KWON, S. & YOUTIE, J. 2017. Tracking the Emergence of Synthetic Biology. *Scientometrics,* 112**,** 1439-1469.

SMITH, J. S., WYSK, R. A., STURROCK, D. T., RAMASWAMY, S. E., SMITH, G. D. & JOSHI, S. B. Discrete Event Simulation for Shop Floor Control. Proceedings of Winter Simulation Conference, 1994. IEEE, 962-969.

SOEBANDRIJA, K. E. N., HUTABARAT, D. P. & KHAIR, F. Sustainable Industrial Systems Within Kernel Density Analysis of Artificial Intelligence and Industry 4.0. IOP Conference Series: Earth and Environmental Science, 2018. IOP Publishing, 012040.

STOJANOVIC, L., DINIC, M., STOJANOVIC, N. & STOJADINOVIC, A. Big-Data-Driven Anomaly Detection in Industry (4.0): An Approach and a Case Study. 2016 IEEE International Conference on Big Data, 2016. IEEE, 1647-1652.

TASSEY, G. 2008. Globalization of Technology-Based Growth: The Policy Imperative. *The Journal of Technology Transfer,* 33**,** 560-578.

TASSEY, G. 2010. Rationales and mechanisms for revitalizing US manufacturing R&D strategies. *The Journal of Technology Transfer,* 35**,** 283-333.

TORN, I. & VANEKER, T. H. 2019. Mass Personalization with Industry 4.0 by SMEs: A Concept for Collaborative Networks. *Procedia Manufacturing,* 28**,** 135-141.

TROTTA, D. & GARENGO, P. Industry 4.0 Key Research Topics: A Bibliometric Review. 2018 7th international conference on industrial technology and management (ICITM), 2018. IEEE, 113-117.

VAN ECK, N. J. & WALTMAN, L. 2017. *VOSviewer Manual version 1.6.6* [Online]. Universiteit Leiden. Available: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.6.pdf [Accessed 15 June 2020].

WAAIJER, C. J., VAN BOCHOVE, C. A. & VAN ECK, N. J. 2010. Journal Editorials Give Indication of Driving Science Issues. *Nature,* 463**,** 157-157.

WANG, J. & SHAPIRA, P. 2011. Funding acknowledgement analysis: an enhanced tool to investigate research sponsorship impacts: the case of nanotechnology. *Scientometrics,* 87**,** 563-586.

WEISER, M. 1991. The Computer for the 21st Century. *Scientific American.* Scientific American.

WU, D., JENNINGS, C., TERPENNY, J., GAO, R. X. & KUMARA, S. 2017. A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. *Journal of Manufacturing Science and Engineering,* 139.

WUEST, T., WEIMER, D., IRGENS, C. & THOBEN, K.-D. 2016. Machine Learning in Manufacturing: Advantages, Challenges, and Applications. *Production & Manufacturing Research,* 4**,** 23-45.

[a]**Tausif Bordoloi** is a final-year PhD researcher at Manchester Institute of Innovation Research, Alliance Manchester Business School, The University of Manchester, and the Department of Mechanical, Aerospace and Civil Engineering, The University of Manchester.

His doctoral research centres on two areas: the nexus of policy and research in manufacturing, and the adoption of digital manufacturing technologies. He has an MBA from the University of Exeter Business School, where he was a Chevening Scholar – the UK Government's flagship scholarship programme for international students. He previously held strategy, innovation and finance roles in biotechnology and automotive industries.

Philip Shapira is Professor of Innovation Management and Policy, Manchester Institute of Innovation Research, Alliance Manchester Business School, The University of Manchester, and a Professor of Public Policy at Georgia Institute of Technology. His interests encompass science and technology, innovation management, manufacturing strategies, emerging technologies, and responsible innovation. He is a Co-Investigator and Lead for Responsible Research and Innovation with the Manchester Synthetic Biology Research Centre. Recent outputs have examined institutions for technology diffusion and the next production revolution. He is a Fellow of the American Association for the Advancement of Science.

Paul Mativenga holds Chair in Multi-scale & Sustainable Manufacturing at The University of Manchester. His research interests include innovative manufacturing technologies and resource efficient or "green" manufacturing. Research outputs have focused on understanding how to make manufacturing industries technologically and economically competitive, while operating within the boundaries of environmentally friendly manufacture. He is a Member of the International Academy of Production Engineering, College International pour la Recherche en Productique.