Expert Systems With Applications 216 (2023) 119456

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities

Zohaib Jan[a,*], Farhad Ahamed[a], Wolfgang Mayer[a], Niki Patel[a], Georg Grossmann[a], Markus Stumptner[a], Ana Kuusk[b]

[a] Industrial Artificial Intelligence Research Centre, University of South Australia, Mawson Lakes 5095, Adelaide Australia
[b] BAE Systems, Tonsley, Adelaide, Australia

ARTICLE INFO

Keywords:
Industry 4.0
artificial intelligence
Machine Learning

ABSTRACT

Many industry sectors have been pursuing the adoption of Industry 4.0 (I4.0) ideas and technologies, which promise to realize lean and just-in-time production through digitization and the use of smart machines. This shift is driven by technological advances, including Artificial Intelligence (AI) and machine learning, sensor networks and Internet of Things technologies, cloud computing, additive manufacturing, and the availability of large amounts of data that can be exploited by these technologies. However, the adoption of AI technologies for I4.0 varies considerably among industry sectors. This article complements broader reviews of I4.0 by examining the specific applications of IAI in several industry sectors, highlighting the issues and concerns encountered in and across different industry sectors, and discussing potential solutions that have been introduced along with opportunities and challenges for adoption. In this article, we review the literature to identify common themes and concerns related to the adoption of AI technologies in the context of I4.0 in several industry sectors. AI solutions are discussed in the context of an AI adoption pipeline that spans data collection, processing, model construction, and interpretation of results. Our findings indicate that although different industries share common issues, the adopted solutions are often specific to a particular industry sector, which may be difficult to transfer to other sectors. Moreover, industry sectors may pursue different adoption strategies due to varying experience and maturity of AI practices. These findings may inform managers, practitioners, and decision-makers who are involved in the adaptation of Industry 4.0 transformation in their respective industry sectors.

1. Introduction

Since the advent of the industrial revolution, practitioners have strived to find innovative ways to augment the manufacturing process in favor of production efficiency, reduced cost, and product quality. The industrial revolution has evolved through four main phases, shown in Fig. 1: the first phase focused on automation through steam and water-powered mechanization; the second phase focused on electrification and mass production; the third phase adopted robotics and digital technologies to gain efficiency; the fourth phase has been focusing on cyber-physical systems and artificial intelligence. Whereas the industrial age was predominantly based on a top-down leadership paradigm (Uhl-Bien et al., 2007), the modern information age heavily relies on an ever-increasing amount of data (Baron & Rrustemi) and on making informed decisions utilizing the data in a timely fashion to staying on top of the competition. This consequently has caused another paradigm shift, and now we are entering the "Industry 4.0" (henceforth I4.0) era where technology takes part in most of the processes to automate and augment various industrial processes (Lu, 2017). It has been argued that the rapid evolution of technology and their interconnections will create a fourth industrial revolution, termed Industry 4.0 (Lee & Lim, 2021; Skilton & Hovsepian, 2018).

In the current information age, data has become the most valuable asset for any company seeking to gain a competitive advantage in the industry (Harding et al., 2006; Taranto-Vera et al., 2021). Companies seek to exploit the available data to make data-driven decisions to gain a competitive advantage. In the context of I4.0, the concept of "smart manufacturing" has emerged (Kusiak, 2017), where "smart machines"

* Corresponding author at: University of South Australia, Australia.
E-mail addresses: zohaibjan@gmail.com (Z. Jan), Farhad.Ahamed@westernsydney.edu.au (F. Ahamed), Wolfgang.Mayer@unisa.edu.au (W. Mayer), Niki.Patel@unisa.edu.au (N. Patel), Georg.Grossmann@unisa.edu.au (G. Grossmann), Markus.Stumptner@unisa.edu.au (M. Stumptner), Ana.Kuusk@baesystems.com (A. Kuusk).

https://doi.org/10.1016/j.eswa.2022.119456
Received 18 November 2021; Received in revised form 6 June 2022; Accepted 21 December 2022
Available online 24 December 2022

and "smart processes" learn from data to continuously optimize production processes, largely with little to no human intervention. Industry 4.0 (Rüßmann et al., 2015) is an intersection of various technologies including Big Data and Cloud Computing, the Internet of Things (IoT), and Artificial Intelligence/Machine Learning (Yao et al., 2017), as illustrated in Fig. 2. The combination of these technologies enables capturing and storing data acquired from multiple sources, analyzing it for decision-making, and learning from it (See Fig. 3).

Artificial Intelligence (AI) and machine learning technologies, paired with vast amounts of data gathered through modern digital technologies, have been emerging as one of the cornerstones of the cyber-physical systems underpinning I4.0. Although AI is often used synonymously with machine learning, AI is a wider field that includes aspects of intelligence that include perception, sensing, reasoning, and knowledge representation in addition to the machine learning aspect. Indeed, the term Industrial Artificial Intelligence (IAI) has been associated with applications of AI technologies in the industry (Lee et al., 2018), highlighting the importance of AI as one of the foundations of modern data-driven processes in the industry. The breadth of IAI transcends the machine learning sub-area and includes automated configuration, planning, diagnostics, adaptation, and prognostics required to realize cyber-physical systems. AI-enabled machines and processes have seen rapid growth in the last two decades and have become one of the main contributors to I4.0. Currently, almost every industry sector is pursuing AI-enabled processes, including the manufacturing-, financial-, transport-, healthcare-, and science sectors (Jan & Verma, 2020). AI technologies have profoundly reshaped how some industries operate. For example, the emergence of predictive maintenance (Yan et al., 2017), comprehensive supply chain optimization (Shao et al., 2021), and digital twins have revolutionized the way manufacturing and service industries operate.

Several studies have been published that discuss machine learning in the context of I4.0 (Bertolini et al., 2021; Dalenogare et al., 2018; Lee & Lim, 2021). However, previous studies do not comprehensively cover AI-based methodologies that have been utilized in different industry sectors. There is a clear gap in research about how different industries have utilized AI in order to achieve Industry 4.0, the concerns they encountered, and potential solutions they uncovered. This article fills this gap by highlighting the issues and concerns encountered in different industry sectors, discussing potential solutions that have been introduced, and identifying opportunities that emerged for the adoption of existing solutions and avenues for future research to develop the next generation of solutions. This work complements broader reviews of I4.0 (Meindl et al., 2021) by examining the specific applications of IAI in several industry sectors. This review aims to inform industry practitioners seeking to adopt Industry 4.0 about common concerns and potential IAI solutions that may be adopted to address the concerns. A further distinguishing aspect of this article is that it has been written for a broad audience who may not possess expert knowledge in IAI technologies and their applications. Hence, the discussion of IAI technologies is framed in the context of an IAI pipeline in which different AI technologies can be adopted to address concerns related to data acquisition, processing, modeling, and interpretation of outcomes. This discussion is supported by a technical lexicon that introduces some main AI concepts and technologies that underpin this pipeline.

The research questions adopted in this study are the following:

- RQ1: What are the concerns that current IAI techniques address in I4.0?
- RQ2: Which IAI techniques are used to address which concerns?
- RQ3: What are the challenges to the adoption of current IAI technologies in I4.0?

By answering these research questions, the study makes these main contributions:

- Identify the I4.0 concerns that IAI technology has been addressing in I4.0 initiatives in several industry sectors. The concerns are organized within a general IAI pipeline that spans data acquisition, processing, model construction, and model interpretation activities.
- Associate with each concern the IAI technologies that have been used to form a solution and outline their contributions and limitations.
- Survey the approaches taken to develop and introduce IAI technologies in organizations.
- Distill themes of challenges related to the adoption of IAI technologies that have been identified and describe further research directions.

These contributions may inform managers, practitioners, and decision-makers who are involved in the adaptation of Industry 4.0 transformation in their respective industry sectors to guide them in their selection of IAI technology and approach to the introduction of such technology in their organization. By bringing together different IAI solutions under the umbrella of the IAI pipeline, similarities, and differences among approaches across industry sectors are highlighted, which can serve as a starting point to transfer some of the learnings and potentially also IAI technologies across sectors boundaries.

The remainder of the paper is organized as follows. Section II gives a brief introduction to Artificial Intelligence and Machine Learning and establishes the technical lexicon for the paper. Section III describes the research methodology used to conduct the literature review. Section IV introduces the overarching IAI pipeline and details the findings of the literature study organized by pipeline stage and industry sector. Section V outlines opportunities for future research, and Section VI summarizes the conclusions drawn from this work.

2. Foundations

In this section, an overview of the current state of AI is given and the technical lexicon of the paper is developed. Additionally, an overview of machine learning algorithms is presented.

2.1. *Artificial Intelligence, Machine Learning, and Industrial Artificial Intelligence*

AI can be defined as the intelligence depicted by machines based on mathematical algorithms and statistical analysis from data (Warwick, 2013). The idea of AI was first identified by Alan Turing in his theory

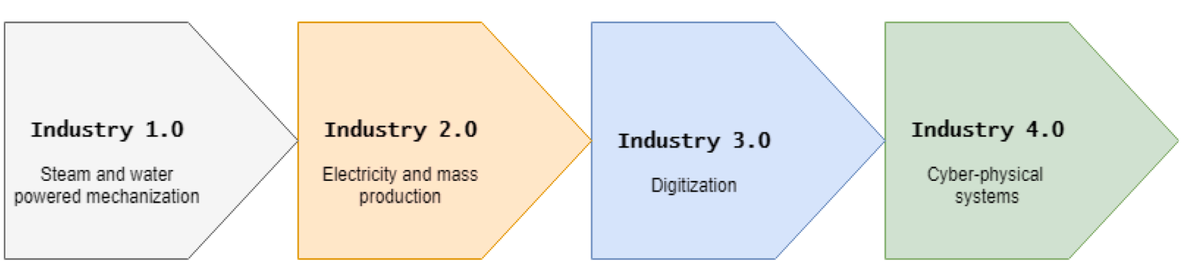

Fig. 1. Evolution of mass industries from industry 1.0 to Industry 4.0

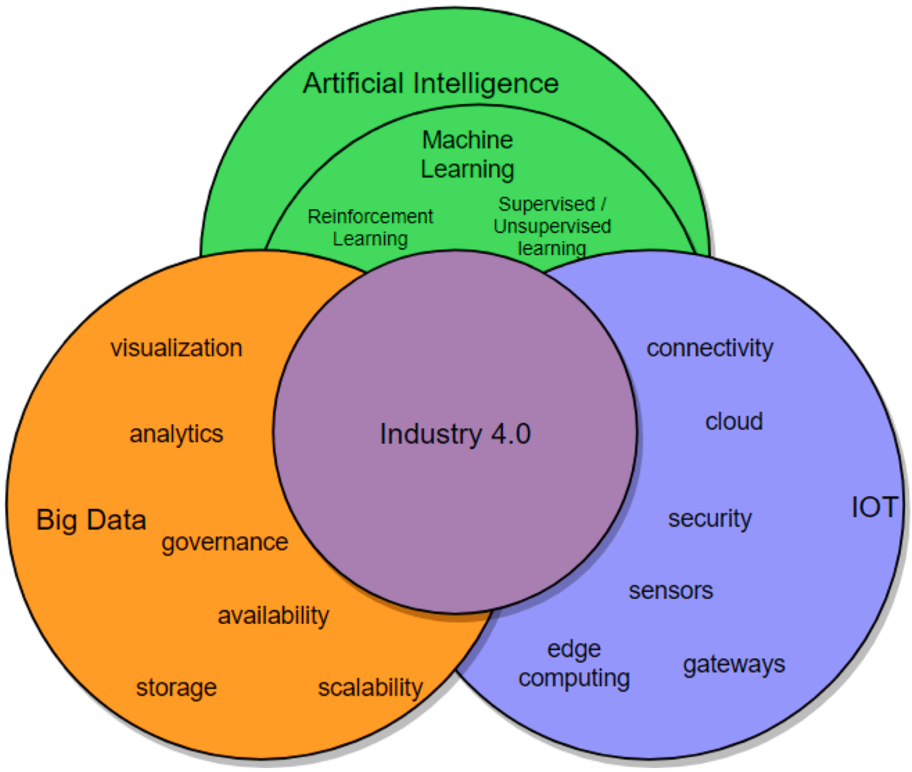

Fig. 2. I4.0 and its enabling technologies

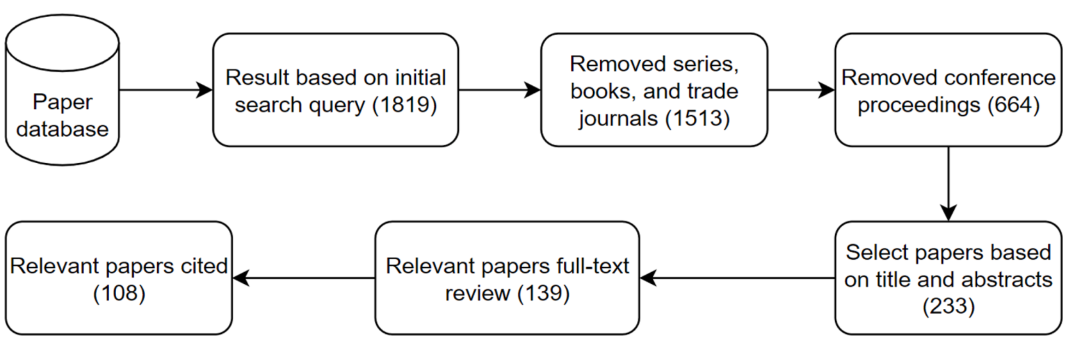

Fig. 3. Research methodology and data extraction

known as the Turing Test in the famous "The Imitation Game" almost 70 years ago (Saygin et al., 2000). Since that time technology has rapidly evolved and matured and we are in an era where machines are achieving better performance in not just mundane tasks but tasks that require intelligent interactions by linking and visualizing data. Machine Learning (ML) is a subfield of AI and is defined as a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict (future) outcomes of interest (Murphy, 2012). Machines are explicitly trained on historical data to learn patterns and therefore predict and trigger data-driven decision outcomes automatically. Often, AI and ML are used synonymously, and an AI-enabled machine is often a machine that utilizes ML algorithms to increase productivity or remove the need for human intervention. This allows industrial applications to utilize the historical data available to integrate into their existing industrial application's ML algorithms to enhance their capabilities. As such in the last decade, owing to the developments in Deep Learning (DL), ML techniques have made significant technological advancements. This is evident in autonomous vehicles (Martínez-Díaz & Soriguera, 2018), to DeepMind which learned to beat humans in video games (Silver et al., 2017). The fourth industrial revolution is more about cyber-physical systems which are formed by seamless integration between the man and the machine. Due to digitization having led to vast amounts of data captured about the real and digital world, ML has been playing a significant role in developing modern "smart" systems. In I4.0, IAI systems make automated decisions based on the contributions that are drawn ML, often combined with knowledge-based AI methods that infuse the ML processes with domain knowledge that is not directly captured in the data. Since IAI systems are heavily reliant on data, having good data as input to the IAI systems is essential. While sensor technology and digitization has provided large amounts of data, additional preparation is often required to improve the quality of the data. AI and ML technology embedded in data processes can assist in the various stages from data collection to data preparation to model training and finally driving business value for an IAI-enabled industry. In the following, some fundamental concepts related to machine learning that are relevant to I4.0 are introduced. The concepts introduced here will form the basis for the subsequent sections in this article.

2.2. Machine Learning (ML) Paradigm

ML algorithms generally fall under one of the three paradigms, which are supervised learning, unsupervised learning, and reinforcement learning (Murphy, 2012).

2.3. Supervised Learning

Supervised Learning (SL) is a type of ML in which a model is constructed based on data that include the inputs (features, or independent variables) and known output (target) for a (potentially large) number of examples. SL can be characterized as finding a function f such that $y = f(x)$, where x denotes the input features and y denotes the target output of the model. ML technologies infer the model, f, from data such that the model closely matches the given output y for the given examples x, while generalizing from the specific given examples without sacrificing accuracy. SL algorithms can be classified into two categories known as regression and classification methods. In regression, the output y of the model is a continuous variable, whereas it is a discrete variable in classification.

An example of regression could be house price prediction. For example, given a set of input features such as the number of bedrooms, the number of washrooms, square footage, etc., the output will be house price. A regression SL model will fit a linear function that can best represent the data. This is formulated as follows $y = m(x) + c$. The model will search for the optimum values of the slope m and y-intercept c that can find the minimum of the function $MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}^{i} - y^{i}\right|$, which is the Mean Absolute Error and is also known as the cost function which the model will minimize; $\hat{y}$ is the predicted variable (house price), y is the ground truth (actual house price) and n is the total number of samples in the dataset. Training such a model involves learning the values of m and c so that MAE is minimal over the input data. Generally, every SL model includes training and testing activities.

For complex data, it is not always possible to use linear functions. For example, given the house pricing data, if we keep adding more features such as suburbs, crime rate, weather, etc., it will become difficult to linearly represent the model and a more complex equation will be required. Although more complex models can be created, there is a caveat in that the more complex model tends to capture noise in the training data and fail to generalize to new inputs. Therefore, a balance needs to be maintained between selecting an overly complex model (which may perform well on the training data but fail to generalize to new inputs) and choosing a simpler model (which may not perform as well for the training data but show consistent results for new input). This phenomenon is known as the "bias-variance covariance decomposition" principle (Ren et al., 2016).

Beyond simple linear models, many sophisticated SL models have been developed. These models take advantage of Statistical Learning and use advanced mathematical constructs such as kernel functions. We will not go into further details about the models as it will deviate from the focus of the paper but list a few models that are commonly used for tabular data classification. They are Support Vector Machines (SVM), Decision Trees (DT), k-Nearest Neighbor (KNN), Linear Discriminant Analysis (LDA), Neural Networks (NN), and Naïve Bayes (NB) (Hagenauer & Helbich, 2017). Selecting the right model for the right data is done mainly through trial and error and most of the time there is no right answer (Wolpert & Macready, 1997). Moreover, ensemble models can be constructed by combining several simpler models to form a better-performing model. Ensemble classifiers have seen a lot of development in the last two decades because of their inherent ability to control the bias and variance of classifiers within the ensemble. Some popular ensemble classifiers are Bagging, Random Forest, Ada Boost, and XGBoost (Jan et al., 2020). Time series analysis (TSA) is a form of a model designed to deal with data that changes over time (Wei, 2006). Popular TSA models are Auto-Regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM) and its variants which have been utilized in a lot of works.

The approaches discussed so far are predominantly rooted in statistical learning theory. Neural networks and Deep Learning methods depart from this foundation and instead adopt inspiration from biology, in particular from the way neurons are connected in brain matter. Deep Learning (DL) (Goodfellow et al., 2016) has seen a lot of work in recent years due to the computer's ability to process large quantities of data. Recent algorithmic advances and the availability of large data sets have resulted in DL being adopted increasingly often. Although image processing, computer vision, and natural language processing applications are arguably among the most well-known application domains of DL technology, the technology has become widespread across a wide variety of problems and application domains. Indeed, DL forms a cornerstone of many AI and ML approaches, including supervised, unsupervised, and reinforcement-based machine learning models, where DL techniques serve to detect patterns in data, compute classification and regressions outcomes, and approximate complex decision boundaries. Sophisticated learning algorithms and efficient implementations can often outperform simpler methods. However, DL methods are not always the first choice for a solution, since developing and deploying such technologies requires deep expert knowledge and comprehensive data sets, and it may be difficult to understand and explain how a decision outcome was generated by a model.

2.4. Unsupervised Learning

Unsupervised learning (UL) algorithms are utilized when the correct output, y, is unknown during the training of the model. Grouping and partitioning "similar" data are common examples of such methods. For example, if we have data on customers and we would like to identify how many customer segments there are, we can utilize a clustering algorithm and identify various customer groups (clusters) that have some similarities with each other. A common method for partitioning samples into similar groups is k-means clustering. In this method, samples are grouped based on several cluster centers, and samples that are closest to the center based on a distance formula are allocated to the same cluster. This process is repeated until the minimum of the following equation is found $argmin\left(\sum_{j=1}^{k}\sum_{i=1}^{n}\left|x_i - c_j\right|^2\right)$ where x is a feature vector, k is the number of clusters that are generated (which must set a-priori), n is the total number of samples in the dataset, and c is the centroid (mean) of each cluster. Besides the Euclidean distance, there are other distance measures such as Manhattan distance, Hamming distance, and Minkowski distance to compute the distance of each sample from the cluster center (Kumar et al., 2014). There are advanced clustering methods such as Fuzzy clustering, Density-based clustering, Connectivity-based clustering, and Distribution-based clustering (Xu and Wunsch, 2005).

2.5. Reinforcement Learning

Reinforcement Learning (RL) (Kaelbling et al., 1996) is a type of ML that learns by trial and error from rewards (and punishments) from an "environment". RL algorithms differ from other ML approaches in that they learn by interacting with the environment to incrementally improve their reward outcomes. Instead of mapping data from the input to the output, RL learns a policy, which is a mapping from a system state to the action/output the system will select. RL algorithms can also be utilized on top of SL and UL algorithms to create innovative solutions to problems. For example, one of the most popular RL algorithms is Q-Learning (Watkins & Dayan, 1992), which has been utilized in various fields, including autonomous vehicles where a DL is used in tandem with RL to create self-driving cars (Chishti et al., 2018). Given a scenario where a system can safely interact with an environment, and a "good" interaction is to be determined automatically, RL methods can be applied. One caveat though is mapping the agents to the right environment variables; oftentimes it's not the algorithm that is failing to converge, rather the mapping has not been made correctly.

3. Research Methodology

To answer research questions RQ1 - RQ3, a Systematic Literature Review (SLR) was carried out to review the current literature about Artificial Intelligence in the context of Industry 4.0. This decision was made based on the successful experiences of other researchers while using Scopus (Lu, 2017; Piccarozzi et al., 2018).

3.1. Search String

An SLR of 642 papers collected from the Scopus[1] abstract and citation database and one search engine Google Scholar[2] was conducted to develop an understanding of IAI in the context of I4.0. The following query was used to identify a pool of potentially relevant papers:

KEY ("Industry 4.0")
AND ("artificial intelligence" **OR** "machine learning")
AND PUBYEAR $\geq$ 2000
AND (LIMIT-TO (SRCTYPE,"j"))
AND (LIMIT-TO (LANGUAGE, English))

The query resulted in a total of 642 papers in the different categories given in Table 1. These papers were analyzed in a filtering and data extraction process outlined in subsection C and classified in terms of their application domain and AI/ML area.

3.2. Exclusion and Inclusion criteria

In this study, a similar approach to (Liao et al., 2017) has been adopted to define the exclusion and inclusion criteria to extract relevant information from the extract papers. Publications that discussed Artificial Intelligence in the context of Industry 4.0 and that were published in academic journals were selected for inclusion in the study. Publications in venues other than journals, such as books and book series, trade journals, conference proceedings, editorial materials, and material disseminated via online forums such as blogs and corporate web portals were excluded from consideration. Moreover, publications whose full-text content was inaccessible and those published in languages other than English were also excluded from consideration. Each paper was reviewed by two researchers to determine if it met inclusion and exclusion criteria. Since the term Industry 4.0 was coined in the first decade of the current millennium, papers pre-dating the year 2000 were excluded from consideration. Arguably, the exclusion of publications pre-dating the year 2000 does not impact the findings presented in this article, due to profound technological changes that have transformed AI and Industry 4.0 since then. Indeed, a cursory review of the few older literature mentioning Industry 4.0 and Artificial Intelligence found that these publications were either unrelated to the objectives of this survey or incorrectly dated copies of more recent publications included in the search. The focus of this study was on recent applications of Artificial Intelligence in the industry. Initially, paper titles and abstracts were read to identify their relevancy and a repository of relevant papers was created. Once the repository of papers was established, papers were read in detail. Lastly, papers were organized into four emergent themes identified through the course of this study, which form the basis of further discussion in the relevant section of this article.

Table 1
# of documents collected from various categories

Doc. Type	# of papers
Article	570
Review	76
Editorial	12
Conference Paper	7
Short Survey	3
Note	2
Letter	1

[1] http://www.scopus.com/
[2] http://scholar.google.com/

3.3. Data Extraction

The purpose of data extraction is to identify the relevant papers from the citation database and extract the information needed to answer the research questions. A similar strategy mentioned in (Sony & Naik, 2019) was adopted to extract the relevant information from the paper databases. For this study, information about the types of problem that are being addressed in each paper, the IAI technologies that are used to address the problem, and any challenges and learnings that were identified in each publication were recorded during the data extraction phase. Furthermore, this study was structured based on a methodology outlined in the Preferred Reporting Items for Systematic Literature review and Meta-Analysis (PRISMA) (Moher et al., 2011). Fig. 4 presents the process of study selection adopted in this study.

3.4. Demographics Data

Although, the search query was limited to publications published from the year 2000 onwards, however, the first works in the given area appeared in 2014. Since 2014 there has been a rapid increase in the number of articles published per year as can be seen in Fig. 5. Since this study covers articles until the mid of 2021, therefore, the number of works in 2021 is naturally lower.

3.5. Subject and Journal Data

The query resulted in papers from 24 different subject areas, with the highest number of publications produced in the "engineering" domain, having a total of 384 publications, followed by "computer science" having a total of 317 publications. "Business, management, and accounting" is the third subject area with a total publication count of 116. The details of the number of publications for the other 21 subject areas are given in Table 2.

The search results consisted of publications from around 160 different journals. The journal with the highest number of publications produced is IEEE Access producing a total of 32 publications, with Sensors Switzerland, and IEEE Transactions on Industrial Informatics to follow, publishing a total of 23 and 19 papers, respectively. IEEE access being a multi-disciplinary journal may be one of the reasons for the higher number of publications. A bar chart of 20 journals ranked in order of the number of papers published is given in Fig. 6. Articles that did not mention the methodology or industry type were omitted from the study. However, limitations in this regard should be taken into consideration.

3.6. Preliminary classification

All publications were carefully read and classified in terms of their industry domain, the concern addressed in the publication, and the technological solution presented in the publication. This included the AI/ML approach and any Industry 4.0 technologies, such as Big Data, IoT, Cyber-physical Systems, and Information Systems, that formed part of the solution in each paper. We identified four main categories of concerns and organized them in an IAI data pipeline illustrated in Fig. 7. The pipeline and its stages form the basis for organizing the works surveyed in this review, and it is associated with common technologies and concerns that are prevalent in each step in the pipeline. It comprises

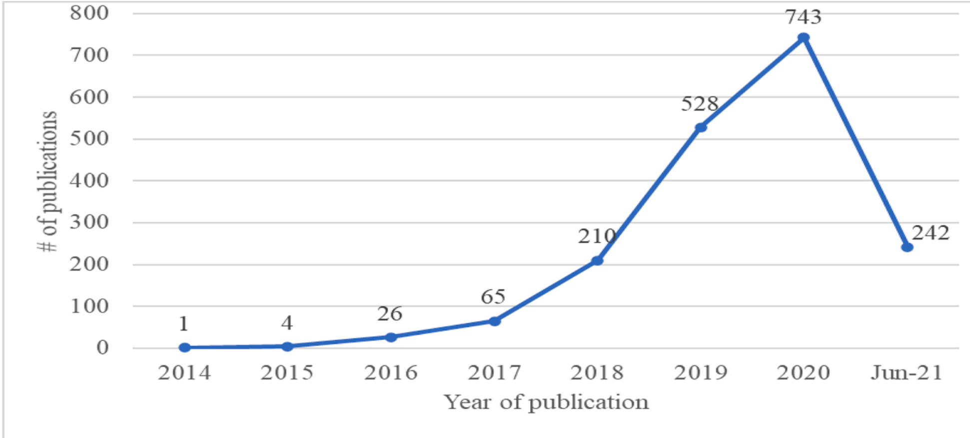

Fig. 4. Publications produced since 2000 in Industry 4.0 and artificial intelligence domain

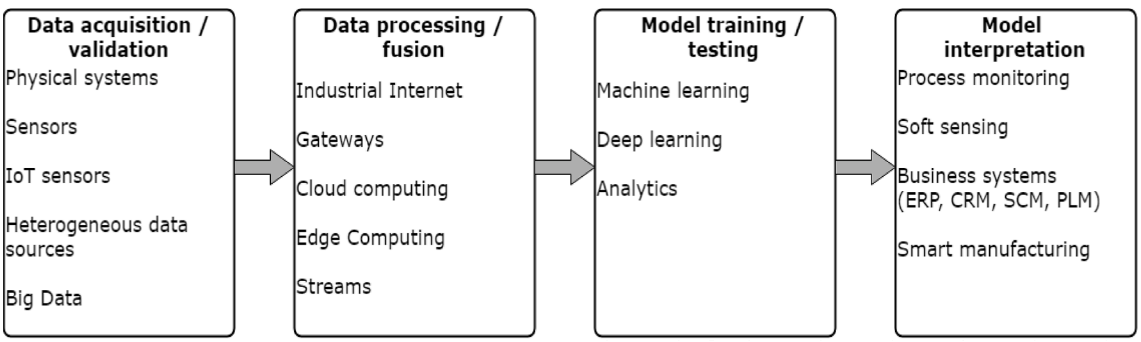

Fig. 5. Industrial AI Data Pipeline

Table 2
Number of papers published in different subject areas since 2000

Subject area	# of pubs.
Engineering	384
Computer Science	317
Business, Management and Accounting	116
Social Sciences	92
Materials Science	84
Physics and Astronomy	63
Decision Sciences	51
Economics, Econometrics and Finance	48
Environmental Science	43
Chemical Engineering	41
Chemistry	41
Energy	41
Biochemistry, Genetics and Molecular Biology	35
Mathematics	34
Arts and Humanities	31
Psychology	16
Medicine	12
Agricultural and Biological Sciences	10
Multidisciplinary	6
Earth and Planetary Sciences	5
Health Professions	4
Pharmacology, Toxicology and Pharmaceutics	3
Immunology and Microbiology	1
Neuroscience	1

the following categories:

1. Data acquisition/validation
a. AI/ML approaches that have been utilized by industrial practitioners to extract data from physical systems, heterogeneous data sources, and big data.
2. Data processing/fusion
b. Data linking, processing, and fusion strategies where AI/ML has assisted in consolidating data from a multitude of data sources.
3. Model training/testing
c. What type of AI/ML models have been utilized and for what purpose and how it has benefitted an industrial application.
4. Model interpretation
d. What strategy has been taken to drive business value by utilizing an AI/ML-enabled system?

In section IV, the concerns and solutions to each concern are discussed, organized by category in the pipeline and industry sector.

3.7. Paper distribution

The distribution of papers categorized based on the 4 categories is given in Fig. 1. According to the distribution of papers based on the current search criteria, most of the research exists to address the concern of driving business value by utilizing AI through existing systems (Model Interpretation stage in the pipeline). This is well-aligned with the expectation that most industries would be concerned about the return on investment. Acquiring data is the starting point for any industrial practitioner considering migrating to Industry 4.0 enabled systems and due to the rapid evolution of technologies there are more ways to acquire data than ever before. Although this study has provided a holistic

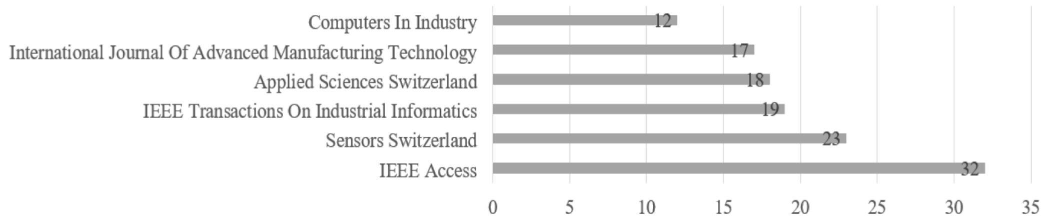

Fig. 6. Top 20 journals with the highest number of papers published in Industry 4.0 and artificial intelligence since 2000 out of 160 total journals

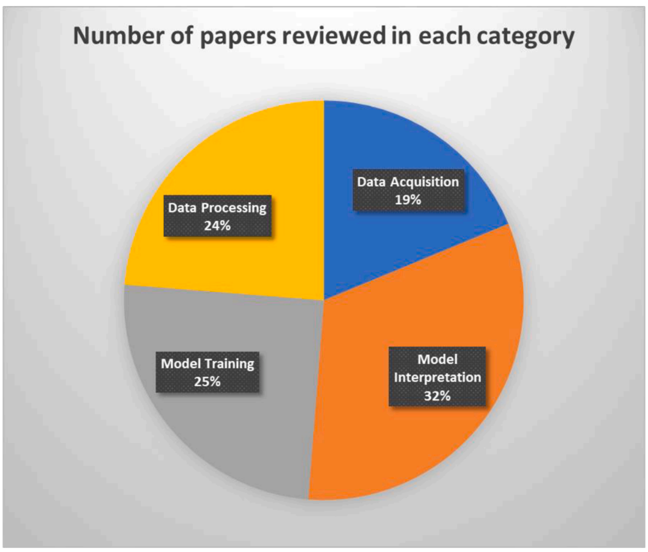

Fig. 7. Distribution of papers

view of the various technologies and approaches needed to enable Industry 4.0, it is believed that a more industry-specific framework must be developed that can be adopted by industrial practitioners to enable Industry 4.0.

4. Limitations

This survey focused on issues, concerns, and solutions that have been reported in articles published in academic journals that are indexed in Scopus and Google Scholar. Issues that practitioners face that may have been reported in other venues, such as magazines, trade journals, white papers published by industry-based organizations, conference presentations, and gray literature may not be reflected in this study. Moreover, the focus of this work was on technological aspects related to Artificial Intelligence with applications in Industry 4.0. The focus on Industry 4.0 and its manifestation in the search criteria, may have resulted in a narrow selection of the literature. However, the breadth of articles included in this survey and the in-depth coverage of artificial intelligence methods for industry 4.0 applications covered in these articles instills confidence that the major areas of concern and related potential solutions could be identified in this article. Other areas of concern, such as managerial, social, and governance-related aspects, were not explicitly included in the study and, therefore, coverage of these concerns may be partial.

5. Findings

Industry 4.0 has not only affected the various industries but has also caused a change in socioeconomic status in society by linking people, machines, and technologies (Rauch et al., 2020). Therefore, this section aims to present the findings gathered from the analysis of the papers and aims to answer research question 3. Different industries are at different stages in the data pipeline categories and, therefore, AI-based approaches that have been utilized by practitioners are discussed in each category are discussed in this section. Moreover, various concerns that have been uncovered through the thematic analysis are also mentioned and summarized in the discussion subsections. Generally, at an abstract level, the concerns are data availability, skill shortage, data security, return on investment, and passive mindset but this study's focus is on the application of AI in an industrial context therefore, not all will be discussed for conciseness purposes.

Based on the four categories Digital Twin has been identified as an umbrella concept in which many of the IAI initiatives in Industry 4.0 fall. A Digital Twin (Lo et al., 2021) is the ability to virtually represent a physical object or a system. Digital Twin is not just a picture or a vivid representation of a system; rather it has the design elements, the build elements, and the operational elements that the physical system exhibits. Digital Twins are adopted to not only replicate data but also to use the data to predict how a new process will perform. Digital Twins are very effective in the manufacturing industry especially in product design as it allows for rapid prototyping, product design changes, and product development at a lower cost. Digital Twin allows practitioners to upgrade their existing processes to AI-enabled processes without changing the physical system as a measure to test the efficacy of the existing system. This helps in "justifying" the efficacy without investing in the capital required for changing the working system. In the next section, a review of current Digital Twin approaches is presented, followed by a summary of Industrial AI concerns. Lastly, the current state of research in the context of 4 categories is discussed and various concerns encountered along with possible solutions are also given.

5.1. *Digital Twin in Industry 4.0*

Digital Twin is at the heart of Industry 4.0 as it replicates physical systems by simulating them in a digital world. This enables industrial practitioners to utilize AI-enabled processes on the simulated system to test their efficacy, without disrupting the running systems. In a research paper (Rojek et al., 2021), it was pointed out that Digital Twin, AI, hybrid simulation, virtual reality, augmented reality, and 3D printing are the driving forces behind Industry 4.0 transformations. It was also stressed that Digital Twins should be adopted by Industry 4.0 practitioners as a starting point to gather data and/or simulate the physical processes. Similarly, in another study (Nakagawa et al., 2021), it was identified that the Digital Twin of factories and their various constituents will accelerate the adoption of Industry 4.0 as it will allow seamless integration of new processes without interruption to existing systems.

In a research paper (Cheng et al., 2020) the importance of Digital Twin in the context of smart manufacturing was discussed. It was argued that Digital Twin is the core enabling technology that can be utilized by the Industrial Internet. The Industrial Internet is the interconnected web of sensors, instruments, and smart devices that are monitored by industrial applications. Therefore, Digital Twin relies heavily on the Industrial Internet to simulate the physical systems more realistically. Another paper (Chen et al., 2021), stressed the benefits of adopting Digital Twin in the wind turbine industry. Digital Twin was adopted to monitor the life cycle of the processes to produce sustainable wind energy. Additionally, it was also pointed out that Digital Twin is a steppingstone towards Industry 5.0 goals, where AI systems will eliminate the need for human interactions, making fully autonomous systems. Lastly, due to the increasing demands for customized products from the customers, the factories need to change delivery models using AI to keep up with the demands. Digital Twins allow industrial practitioners to use virtual production line systems where a customer can provide feedback to improve the product experience and further customize the product in real-time (Wan et al., 2020). A summary of Digital Twin papers is given in Table 3, and it can be noted from the table that different industries share the same goal when it comes to adopting Digital Twin.

6. Industrial AI data pipeline

In this section, based on the four categories identified in Fig. 7, the literature is grouped and discussed. Furthermore, since AI is paving its way in the use of many manufacturing industries to provide sustainable and quality-oriented manufacturing, a discussion is provided on the various industries that have benefitted from adopting "smart" technologies, the solutions that have been uncovered, and the concerns that were encountered are also discussed.

The four stages of the data processing in the data pipeline were presented in Fig. 7. The first stage covers data acquisition and validation from the cyber-physical world, various sensors, and heterogeneous data sources. To develop intelligent models using AI and ML, one of the primary requirements is to collect the relevant data. The data collection and validation stage can be considered the most time and effort-consuming task in the data pipeline process. However, AI/ML can assist to automate the data collection process and can contribute to the big data repository.

In the second stage of the data pipeline, data adaptation and processing are conducted to facilitate machine learning for intelligent systems. The advancement in embedded systems and machine sensing in the industry has resulted in the production of large volumes of machine data. Processing of this data requires a robust infrastructure such as Fog and Edge computing facilities and Industrial Internet. During the data processing stage, AI/ML methods can be utilized for dynamic data linking, data filtering, and data fusion before transferring to the next stage of the I4.0 data pipeline.

In the third stage of the data pipeline, I4.0 takes the benefits of state-of-the-art AI/ML algorithms to develop models and smart systems to solve complex problems. Industry-specific domain experts can assist to develop better AI/ML models in this stage. Various smart pilot systems may emerge in this stage, hence sufficient testing is necessary to select the final smart production system. AI/ML-driven smart model selection and testing are an important part of this stage for I4.0 implementation.

The fourth stage of the data pipeline assists the business to evaluate ROI from the previous stages as well as from the overall I4.0 operations. Process automation and monitoring, manufacturing forecasting, system workload balancing, and performance of connected business systems such as ERP, CRM, SCM, and PLM will be reflected in this stage. Another important aspect of this stage is the understanding of the outcomes from the system to explain which parts of a complex IAI system contributed to an outcome and how they contributed. These questions are important for diagnostic purposes and to build trust in the system.

Table 3
Summary of digital twin papers

Author	Industry focus	Goal
(Lo et al., 2021)	Production Design	To achieve product customization using state-of-the-art technologies
(Rojek et al., 2021)	Machine Modeling	Novel approaches that can be achieved using Digital Twin for customized products in dynamic approaches
(Nakagawa et al., 2021)	Smart Factory	Continuous integration of AI in smart factories in a Digital Twin environment to achieve customized products
(Wan et al., 2020)	Smart Manufacturing	High value manufacturing through predictive simulations in a Digital Twin enabled environment
(Chen et al., 2021)	Wind Turbine	Digital Twin of a wind turbine to achieve sustainable energy production using AI-enabled processes
(Cheng et al., 2020)	Industrial Internet	Digital Twin and Industrial Internet go hand in hand to allow AI-enabled processes for lifecycle management
(Darvishi et al., 2020)	Sensor Tech.	Faulty sensor detection in a Digital Twin enabled environment through the use of Neural Networks
Fuller et al., 2020 (Fuller et al., 2020)	Industry 4.0	Current research issues in Digital Twin in the context of Industry 4.0
(Mao et al., 2019)	Smart Manufacturing	ML-enabled processes are utilized in a Digital Twin environment to identify risk factors and early onset of incidents

6.1. Data acquisition / validation

For any industry to adopt AI/ML in their existing processes the first step is to acquire data. This data can be coming from a multitude of systems, sensors, and other heterogeneous data sources. In this section, existing approaches are discussed that Industry 4.0 practitioners have adopted to gather data from a variety of sources. Acquiring data is a challenging task and a summary of concerns that have been faced by industrial practitioners in acquiring and validating data is as follows:

- A multitude of sensor technologies based on the Internet of Things (IoT) exist, and which one to rely on is a common source of concern.
- Data acquired from sensors is not always reliable and is prone to noise and randomness.
- Acquiring and validating data in a timely manner is a difficult task.

A summary of AI-based solutions that have been utilized in the Industry 4.0 to address the above-mentioned concern is as follows:

- Data is validated by utilizing ML classifiers to classify sensors as faulty.
- Mesh networks based on Long Range Wide Area Network (LORAWAN) and Bluetooth low energy (BLE) are utilized to allow seamless human and machine collaboration.
- Synthetic data is generated to simulate the behavior of physical systems which can later be utilized by ML models for testing purposes.
- Edge computing is a potential solution to gather data with low latencies. Edge computing distributes the computation bringing it closer to the data sources, thereby, improving the response time for real-time systems.

Further elaboration on these concerns and solutions found in literature and new industries that emerged from the conjunction of various technologies are detailed below.

6.2. Industrial big data

Data-driven innovation is the stepping stone of industrial AI and big data's continuous infiltration into industries is the key reason why industries use big data as a tool for smart manufacturing. Therefore, industries rely on big data as means to acquire, analyze, extract, and perform analysis on datasets that are too large to be handled by traditional processes. A paper (Grant, 2021) on various corporate giants such as Capgemini, Deloitte, Management Events, McKinsey, MHI, Microsoft, and PwC outlines various AI-based decision-making systems that are currently in place and it was identified that big data is the core component of most of the systems. It was also concluded that to achieve sustainable smart manufacturing we must utilize state-of-the-art technologies such as IoT, big data, and AI-enabled processes to assimilate an industrial form that is assisted by cognitive decision-making systems. Similarly, in (Andronie et al., 2021) the importance of big data in the context of cyber-physical production systems was discussed. Cyber-physical systems are a core component of Industry 4.0 and without them, it will not be possible for industrial practitioners to adopt Industry 4.0, and cyber-physical systems rely heavily on big data. Lastly, the importance of industrial big data for smart manufacturing was discussed and a reference architecture for the implementation of industrial big data in the marine engineering field was put forward (Zhang et al., 2020).

6.3. Health care sector

Dirty and noisy data can be a big concern, especially in the health-care industry as this may lead to wrong treatments and failed pharmaceutical drugs. Therefore, it is not a surprise that around one-fifth of health executives think that the major drawback for AI to reaching its real potential in healthcare is the lack of clean data (Drew, 2019). To enable the healthcare industry to utilize data effectively an integrated framework that utilized IoTs, data lakes, data analysis, big data, and cloud computing for manufacturing state-of-the-art health monitoring systems was presented (Yu et al., 2021). A significant effort has been made by the researchers to mitigate the data quality issue in health care systems. Moreover, they also proposed a data ingestion procedure designed specifically to manage data security, storage, and management issues. Similarly, in another study (Hosseini et al., 2017) a cloud-computing-based solution was proposed that utilized IoTs, and big data to create a brain-computer interface for seizure predictions in epileptic patients. The data available from IoT sensors is also susceptible to environmental factors. Therefore, sensor validation should be adopted to classify whether the data is noisy or not. Therefore (Darvishi et al., 2020) researchers proposed a three-stage ML-based architecture that utilizes an MLP classifier to detect and isolate faulty sensor patterns.

6.4. Agrarian sector

The agrarian sector also deals with a multitude of data and is one of the key areas where Industry 4.0 is seeing a great deal of application and research due to its potential to promote innovation. This is predominantly due to the rapid evolution of technology, especially IoTs, and their utility in the sector. The use of data effectively and efficiently can improve various practices and operations at individual farms; this has caused the emergence of "Agriculture 4.0". Therefore, in a study (Spanaki et al., 2021) a framework was proposed that states that Data Sharing Agreements (DSA) are a key element when utilizing AI-based technologies for data/process management. They argued that AI-based data management methodologies must be adopted if they are to leverage Industry 4.0 in the agrarian sector. Accordingly, a "smart farm" scenario is the one in which a DSA exists between a farmer and the cloud provider. This essentially allows different actors such as NGOs, research/statistical institutes, and even other farmers to utilize the data to build innovative solutions. Similarly, in (Liu et al., 2020) issues that emerged as the transition began towards Agriculture 4.0 also known as "smart agriculture" were discussed. They also suggested various enabling technologies to gather agriculture data and perform analytics such as IoT, Robotics, AI, big data analytics, and blockchain in the context of Agriculture 4.0. The use of Unmanned Aerial Vehicles (UAV) in tandem with IoT sensors has also been proposed (Lin et al., 2020) for real-time agriculture monitoring systems as an energy-efficient solution. The hierarchical data collection methodology was proposed in the research that allowed the UAVs to move accordingly in a Wireless Ad-hoc Network (WANET) of sensors to save energy consumption.

7. Discussion

A summary of the data acquisition papers, along with the concern(s) that each paper has addressed is given in Table 5. As can be noted from Table 4, the predominant concern in Agriculture 4.0 technologies is data availability, and one key technology to support is IoT. But a common source of concern with IoTs is power efficiency which is mitigated by AI/ML-enabled approaches that efficiently route an IoT network to increase power efficiency. The data acquisition in the smart manufacturing and healthcare industry is leaning toward sensor networks; however, a major concern is validating sensor data. AI/ML-enabled validation approaches have been utilized to mitigate the issues by validating a sensor and its data before utilizing the data.

Although AI/ML-enabled processes have mitigated various cross-industry issues, most of the processes developed are too specific to a problem, therefore the same solutions cannot be replicated in other industries. DL is seeing applications in many industries such as NN for sensor validations by classifying faulty sensor data, CNNs for quality control by capturing product defects through CCTV images, Yolo for virtual cart object detection, etc. DL is a relatively new area that started seeing research and application in recent years. This is predominantly due to the availability of data and availability of computational power. Therefore, a potential future direction is the application of DL to

Table 4
Summary of data acquisition papers

Author	Industry Focus	Concern					Data acquisition		Approach of data extraction
		Data Quality	Data Latency	Data Sharing	Data Avail.	Other	IoT	Big data	
(Spanaki et al., 2021)	Agriculture 4.0			*	*		*	*	DSA framework
(Liu et al., 2020)	Agriculture 4.0	*			*		*		
(Lin et al., 2020)	Agriculture 4.0		*		*	*	*		AI/ML
(Andronie et al., 2021)	Agriculture 4.0				*	*	*		AI/ML
(Rojek et al., 2021)	Machine Modelling				*		*		SIM
(Grant, 2021)	Smart manufacturing					*	*	*	D
(Trinks & Felden, 2018)	Smart manufacturing		*			*	*		E, F
(Garrido-Hidalgo et al., 2018)	Smart manufacturing			*	*	*	*	*	
Yu et al., 2021 (Yu et al., 2021)	Health Care	*	*	*	*	*	*	*	IF, WSN
(Hosseini et al., 2017)	Health Care	*	*			*	*	*	AI/ML
(Yan et al., 2017)	Predictive Maintenance					*	*		AI/ML
(Zonnenshain & Kenett, 2020)	Quality 4.0					*	*	*	
(Zhang et al., 2020)	Smart manufacturing		*	*				*	AI/ML
(Sarfraz et al., 2021)	Society 5.0				*	*	*		DL

DSA: Data sharing agreement, WSN: Wireless Sensor Networks, E: Edge computing, D: Analytics, SIM: Simulated data, DL: Deep Learning, F: Fog computing, IF: Integrated Framework, D: Data analytics

Industry 4.0-specific applications and solutions. DL is used in various domains of Industries, including capturing video and image footage from various locations in manufacturing plants. Data acquired through various sensors (IoTs) or other means in industrial applications are often noisy and voluminous and often a cost is associated with the acquisition process

7.1. Data processing and fusion

Data processing and fusion are important and challenging parts to tackle when adopting in Industry 4.0 enabled industries. AI-assisted technologies can be adopted to improve productivity and sustainability allowing practitioners to reach the Industry 4.0 goal. However, AI-assisted technologies are data-driven, and without the availability of quality data, AI technologies cannot be adopted. Therefore, data storage, processing, analysis, fusion and filtering, and quality enhancing capabilities are required in a smart factory environment to process real-time data from sensor networks and other such sources. Based on the literature, some of the concerns mentioned about data processing and fusion in the context of Industry 4.0 are as follows:

- Security of the data at motion especially in data-sensitive industries such as health and defense industries.
- Data linking and fusion in the unstructured, semi-structured, and structured data environment.
- Filtering noisy data for AI models.
- Data latency and transport issues in capturing real-time data from a variety of sensors.

Various AI-based approaches to mitigate data processing concerns mentioned in the literature are:

- ML classifier to generate Class of Services (CoS) for network traffic.
- AI and novel homomorphic encryption to secure data computation at remote storage (Edge, Cloud).
- An image-based, time-series, audio, and text data fusion domain; DL is being applied using multimodal data to detect the changes in objects and movements from live cameras.
- Applying various ML learning methods for data processing such as representation learning, deep learning, distributed and parallel learning, transfer learning, active learning, and kernel-based learning.
- Network throughput optimization for Fog and Edge computing.
- Enhancing Edge intelligence and reducing data traffic to remote storage.

7.2. Data latency

Edge computing can assist in storing and retrieving data promptly when required, as the processing and storage are brought together. From an Industry 4.0 perspective, edge and fog computing network traffic can be prioritized based on Quality of Service (QoS) and Quality of Experience (QoE), thus, allowing for applications/processes that are of high priority to take precedence. As such in research (Guevara et al., 2020) a classifier was used to determine network traffic QoS values. The QoS parameters represented information about the bandwidth, reliability, security, storage, location, mobility, scalability, delay sensitivity, and loss sensitivity. These features were used as input to train the classifier. Intelligent network firewalls can perform deep packet inspection and can be trained to classify the type of network traffic and the type of application flowing into the network. These smart firewalls can be enhanced to act as dynamic QoS assigners to data traffic. In high-volume data-producing factories, the computation needs to be done near the edge nodes. Therefore, (Pop et al., 2021), a fog architecture for Industry 4.0 using IoT devices was proposed. It was suggested that edge and fog architecture will assist in achieving the Industry 4.0 goal as it will enable the data processing and storage near the endpoints thus reducing the overall latency.

7.3. Data security

A key element for many industries to gather more data is the use of sensor technologies, especially IoT. However, the security of sensor networks has become a key concern for Industry 4.0 practitioners and applications. As the industry is more reliant on smart devices, processes, and environments, there is a growing need for more data. In (El-Hajj et al., 2021) it was proposed that IoT is the key element to bringing an existing system up to Industry 4.0 standards, however, IoT systems are prone to security risks and, therefore, they have proposed an AI and PUF-based (Physical Unclonable Function) scheme to mitigate the issue. Additionally, an AI-based industrial application utilizing blockchain is proposed as an alternative to Public Key Infrastructure (PKI) based authentication (Rathee et al., 2021) for Industry 4.0 applications. Similarly, in (Qu et al., 2020) a federated learning environment with blockchain was suggested to provide more security and privacy that can be also applied to Industry 4.0 networks and platforms. AI/ML strategies have also been proposed to secure networks with smart policies and autonomic network traffic segmentation, such as ML-based classifiers to perform deep packet inspection to automate network security policies, and network segmentation to ensure network reliability (Saghezchi et al., 2018).

7.4. Data transport

Energy efficiency is an issue when it comes to the use of sensor networks. In industrial implementations, wireless sensor networks are required to communicate in low power mode. Since the devices are limited by battery size, energy-efficient AI algorithms, or green algorithms, must be developed to prolong their usage. To cope with energy efficiency, (Pokhrel et al., 2021) a clustering-based routing framework for massive sensor networks in the industrial IoT environment was proposed. The proposed clustering-based algorithm for routing reduced the power consumption of IoT sensors by a significant margin which increased the longevity of the active network. Due to the availability of low-cost processing devices, edge intelligence is also utilized to reduce the data traffic to the network by processing data on edges. AI/ML approaches are assisting edge devices by processing the data locally before sending it to the cloud. Edge intelligence has become an important part of Industry 4.0, which was not a viable option earlier due to the cost of devices and speed of the network. For example (Huang et al., 2019), rapid object recognition in the industrial field for intelligent manufacturing based on edge Intelligence was proposed. Further studies stressing the importance of edge intelligence are mentioned in (Abdulkareem et al., 2019; Deng et al., 2020; Zhou et al., 2019).

8. Discussion

A classification summary of data processing papers is given in Table 5. Many industries relied on edge and fog computing to bring data storage and processing together. This not only enabled the integration of AI/ML- processes but also allowed for real-time processing. Reliance on IoT by industries is increasing daily and an ecosystem of sensors, machines, instruments, and people has given a rise to the industrial internet, where anything and everything is interconnected. Industrial internet, edge computing, and fog computing are the three main technologies when it comes to enabling Industry 4.0. Although various integration frameworks are proposed in research, there is a need for the development of a comprehensive integration framework that can be adopted by cross-industry platforms.

Table 5
Summary of data processing/fusion papers.

Author	Industry Focus	Concern				Data Processing		
		Data Security	Data Quality/Filtering	Data Preservation/Transport	Other	Edge	Fog	Others
(Qu et al., 2020)	Block chain				*			
(Rathee et al., 2021)	Block chain	*				*		
(Zhou et al., 2019)	Edge Intelligence					*		
(Deng et al., 2020)	Edge Intelligence					*		FLF
(Belgiovine et al., 2021)	Edge Intelligence	*	*			*		
(Guevara et al., 2020)	Fog for Industry 4.0	*	*	*			*	
(Abdulkareem et al., 2019)	Fog for Industry 4.0					*	*	
(Azimi et al., 2016)	Fog for Industry 4.0		*				*	
(Li et al., 2018)	Fog for Industry 4.0		*				*	
(El-Hajj et al., 2021)	Industrial Internet	*			*			PUF
(Pokhrel et al., 2021)	Industrial Internet			*				
(Aveleira-Mata et al., 2021)	Industrial Internet	*			*			
Saghezchi et al., 2018)	Industrial Internet	*	*					
(Alsheikh et al., 2014)	Intelligent Manufacturing			*				
(Huang et al., 2019)	Intelligent Manufacturing		*		*			
(Brundage et al., 2018)	Intelligent Manufacturing	*			*			
(Huraj et al., 2021)	Production Line	*				*		
(Wang et al., 2018)	Quality 4.0		*		*			
(Chen et al., 2017)	Smart Factory	*		*		*		
(Shao et al., 2021)	Smart Factory			*		*	*	MSF
(Faheem & Gungor, 2018)	Smart Gird	*	*	*				

MSF: Multi-Stage Implementation Framework, FLF: Federated Learning Framework: PUF: Physical Unclonable Functions

8.1. *Model training and testing*

From the factory management point of view, some of the important functions and capabilities of Industry 4.0 are digitalization of manufacturing and having a single dashboard view of the whole factory units. This allows decision-makers to identify the health of various sections of the production life cycle. The dashboard monitoring, predictive models, smart diagnostics, automated production, supply chain, etc., assist in optimizing the manufacturing process to achieve the Industry 4.0 goal. Moreover, in most of the use cases identified in research various AI models and smart agents work together in a digital environment to achieve the optimal goal of the industry workflow. In this section, various approaches that have been adopted by industrial practitioners to adopt AI/ML-enabled processes in their existing systems are discussed. Based on the literature, some of the concerns mentioned for training AI/ML models for industry 4.0 are as follows:

- Industrial ML models are harder to train because of a lack of data.
- Features are generated from industrial data with minimal domain expertise.
- Real-time learning from streaming data from the industry environment.
- Developing ML models from incomplete and imbalanced data from the sensor network.

AI/ML solutions that have been adopted to address the concerns are as follows:

- Using RL and GA to automate feature engineering and hyper-parameter optimization.
- Using Automated Machine Learning (AutoML) and Neural Architecture Search (NAS) to dynamically produce the best ML model from the dataset.
- Adversarial training of defensive AI and incremental IoT-based OT monitoring systems can assist in training industrial ML models.

8.2. *Industrial feature engineering*

To train and find the right model we start with the right data. Once the data is readily available model parameters can be tuned appropriately and the right model can be trained. Since in industrial applications the data is coming from a multitude of sources, hence the fusion of data can save valuable time before training the model. As such, in (Kuo et al., 2017) dimensionality reduction and feature engineering were applied to collect key information from the dataset to train classifiers for industrial applications. Feature engineering is an important part of model training that is an area of active research in the machine learning community. Selecting important features from heterogeneous data sources is an essential task to accelerate the process of training an appropriate model. In a study (Gondek et al., 2016), feature engineering and data processing was utilized to train a model that can predict future failures for the air pressure system in Scania trucks. The right features were extracted from the data based on histograms and distributions of the groups. A feature selection process was incorporated to rank the features based on the mean precision of the output class and scoring of the feature sets. In (Nath et al., 2021) an AI-enabled feature engineering framework for rotor fault diagnosis was presented. In the research, various AI/ML-based methodologies were utilized to identify important features from the data which consequently enabled them to identify rotor faults. In the automobile industry vibration reduction is a very important part of passenger comfort. Similarly, in a study (Nariman-Zadeh et al., 2010) a vehicle vibration model was proposed to improve the comfort of passengers using optimization algorithms to determine the right input features.

8.3. *Transfer learning and data bias*

Injection molding, also known as 3D printing, is seeing a lot of research due to its ability to incorporate rapid change and allow rapid prototyping. In (Lockner & Hopmann, 2021) an injection molding process was proposed that is enhanced by transfer learning. The research utilized neural networks and genetic algorithms to optimize the molding process to cater to customer demands of low waste and fast throughput. In the injection molding process, AI and genetic algorithms have been utilized to improve the quality, and sustainability of production (Han et al., 2016; Jung et al., 2021). Similarly, (Sanz et al., 2021) a framework for an auto paint shop industry was proposed to manage multiple control systems. The control systems work together to accomplish coordinated painting. This project aims to profit from the cyber-physical systems, computing integration, AI, and IoT inside the Industry 4.0 paradigm in an automotive paint shop process. An assembly assistance platform is an important component of manufacturing, especially in the context of

Industry 4.0. Therefore, in (Gellert et al., 2021) the prediction-based assembly support system was investigated. It was concluded that the context-based predictors partial matching method presents a significant improvement over the existing Markov predictors platform. Lastly, it was pointed out in research (Ahamed & Farid, 2018) that, in mission-critical applications such as personalized healthcare, the bias in data collection and model training is a growing concern. Generally, data is considered biased if there are more samples from one class than another. For example, if we have a dataset of emails and there are more rows about spam emails than non-spam emails, then we will say this data is biased towards spam emails. Consequently, the models that are trained on biased data are also biased. Lastly, a study created a cognitive architecture for an adaptive pipeline in the Industry 4.0 setting to assist the system to select suitable models and algorithms (Fischbach et al., 2020).

8.4. Artificial Neural Networks in various industries

Additive Manufacturing (AM) (Omairi & Ismail, 2021) is a production technology that is seeing a lot of innovation and is anticipated to be vested heavily in various domains including health care, materials engineering, communication, prefabrication construction, aeronautics, power electronics, food, and the space industry. Given the ability to rapidly develop prototypes, AM is at the forefront of Industry 4.0. Fused Filament Fabrication (FFF) is an AM process that uses a thermoplastic material that is fed into a heated printer extruder to form layers. In a study (Oehlmann et al., 2021) the researchers used an Artificial Neural Network (ANN) to predict the pressure within the nozzle of an FFF to optimize the process of printing. A digital twin of an FFF printer was created by equipping it with various sensors and training and testing the performance of the ANN model for prediction on gathering data. They argued that using an ANN to understand the various parameters of the system allowed for a deeper understanding of the process which consequently enabled them to determine the needs to achieve optimum efficiency. A study (Sun et al., 2021) presented an ML-based model to identify defective metal casting in the foundry industry. It was also argued that the implementation of ML-based algorithms in the metal casting industry requires skilled workers who are trained in both data science and materials science and engineering domains. Besides predictive manufacturing, and smart industry AI is paving its way into the chemical industry as well. This encompasses pharmaceuticals and chemical processes across various industries including academia. ML algorithms can learn underlying patterns in data in higher dimensions enabling them to learn many chemical processes that are otherwise hard to identify manually. Because of this, (Kakkar et al., 2021) it was identified that the chemical engineering industry needs the utility of AI/ML knowledge and practitioners are lacking the skill, therefore, they suggested that students pursuing chemical engineering must also learn AI/ML. In another study (Arden et al., 2021) it was discussed the necessity of utilizing Industry 4.0 and AI-enabled processes in pharmaceutical manufacturing. They suggested that the pharma industry must utilize the emerging cutting-edge technologies to evolve to the next generation of manufacturing. Therefore, they aimed to develop an understanding of the current state and the concerns that exist in the pharma industry to adapt to Industry 4.0. It is evident from the literature that absorptive capacity is important for industries for innovation, therefore prior knowledge of AI uses is vital for the industries (Grandinetti, 2016; Van Geenhuizen and Soetanto, 2009). Prior knowledge about the industry domain and AI will shape the industries to adopt other technological advances of Industry 4.0 such as Big data, IoT, Edge computing and it will have a decisive influence to convince the consumers (Paiola et al., 2021). Additionally, it will provide a competitive advantage in the markets for the industries.

8.5. AI in smart grid

A Smart Grid is an electrical grid that utilizes a variety of technologies to increase the sustainability and efficiency of systems. In (Kotsiopoulos et al., 2021) various AI/ML algorithms including their applicability to existing industrial processes was discussed. Different AI/ML algorithms and their associated utilities are Energy consumption forecasting – Bayesian Networks; Power management systems – K-Nearest Neighbours; Control procedures – Multiple Linear Regression; Simulating platforms for electric grids – Gradient Boosted; Short term load prediction – Locally weighted training; Load estimation – K-means; Theft detection, price, and load forecasting – Autoencoders; Cyber-attacks – Recurrent neural networks; Intrusion detection systems – Multilayer perceptron.

The United Nations' 17 Sustainable Development Goals (Petillion et al., 2019) state that the energy industry needs to adopt digital transformation and Industry 4.0 to allow for growth towards more flexible power systems. Digital transformation not only allows for sustainable development but also provides future opportunities for the industry. To achieve digital transformations, a study (Giraldo et al., 2021) proposed a framework that can be utilized by energy industries to achieve digital transformation. It was suggested that it is pivotal for energy industries to move forward towards Industry 4.0 if they want to achieve sustainable growth. A 14-step process was provided to adopt the digital transformation. Digital transformation is heavily reliant on modern cutting-edge technologies such as cloud computing, IoT, AI, etc. An example of a successful digital transformation of a power plant is described by (Ashraf et al., 2021) where power curves were generated with a 95% accuracy using ANNs. They proposed that simulating data allowed them to utilize state-of-the-art AI algorithms to achieve the Industry 4.0 goal.

9. Discussion

A classification summary of model training and testing papers is presented in Table 6. Several studies applied supervised ML in various industries, especially for AM. This is predominantly because AM has data available and has adopted data-enabled processes to optimize production. Most of the studies mentioned in Table 7 faced concerns when integrating AI/ML models and some of the concerns reported were model bias, lack of data, and data linking. Based on the synthesis, it can be noted that supervised learning is the most common in the industry, followed by deep learning, with reinforcement learning being used the least. Since reinforcement learning is very situational, it is not adopted as much as supervised learning.

Allowing seamless integration without interruptions to running processes in the production system will be a concern using supervised ML. Reinforcement models are difficult to train for industrial applications and the cost associated with them is high. Also, the cost associated with a false positive is even higher, for example, an AI-trained autonomous car. In the foundry industry, developing AI-driven models is difficult due to the nature of the data that are unbalanced, heterogeneous, and limited in sample size. Furthermore, in the foundry industry, the data is semi-supervised, therefore supervised ML is not preferable. It will be also challenging to apply unsupervised ML due to the limited sample size.

9.1. Model interpretation

Model interpretation in this study is referred to as the ability of a business owner to derive value from AI/ML-enabled processes. When ML models are adopted by industrial practitioners, it is a challenging task to derive business value from them. For example, in some industries, such as the Airline industry, false positives can be very cost intensive. Some of the concerns that have been drawn are as follows:

Table 6
Summary of model training and testing papers.

Author	Industry Focus	Modeling Concerns					AI/ML approach
		Lack of data	Biased model	Data Linking	AI/ML Integration	Other	S/U/D/R/T/O
(Oehlmann et al., 2021)	Additive Manufacturing	*			*	*	U, D
(Sun et al., 2021)	Additive Manufacturing	*	*		*		S
(Omairi & Ismail, 2021)	Additive Manufacturing				*	*	U/D/O
(Huang et al., 2018)	Additive Manufacturing				*		D/R
(Radel et al., 2019)	Additive Manufacturing				*	*	O
(Gellert et al., 2021)	Assembly Line					*	S
(Sanz et al., 2021)	Automotive			*	*		S/U
(Pinheiro et al., 2018)	Block Chain					*	O
(Kakkar et al., 2021)	Chemical Engineering				*		O
(Pop et al., 2021)	Edge Intelligence					*	O
Giraldo et al., 2021)	Energy				*	*	O
(Ahamed & Farid, 2018)	Health Care		*				S
(Nariman-Zadeh et al., 2010)	Industrial Optimization				*		S
(Lockner & Hopmann, 2021)	Additive Manufacturing	*					U/D/R
(Han et al., 2016)	Operation Management		*				S
(Arden et al., 2021)	Pharmaceutical				*		O
(Fischbach et al., 2020)	Production System				*		S
(Gondek et al., 2016)	Production System	*					S
(Nath et al., 2021)	Smart Manufacturing			*	*		S
(Kuo et al., 2017)	Smart Manufacturing				*		D

S: Supervised Learning, U: Unsupervised Learning, D: Deep Learning, R: Reinforcement Learning, T: Transfer Learning, O: Other.

Table 7
Summary of model interpretation papers

Author	Industry Focus	Model interpretation					Business value driven by
		PM	SS	BS	SM	Other	
Alsheibani et al., 2020 [113]	AI strategy					*	
Lipton, 2018 [112]	AI strategy					*	
Kartanaitė et al., 2021 [117]	Financial industry						Supervised ML
Forrester, 1999 (Forrester, 1999)	Industrial Enterprise	*	*	*	*		
Liebrecht et al., 2021 [134]	Industrial Enterprise					*	
Kudelina et al., 2021 [122]	Industrial Enterprise	*	*	*	*	*	DL, GA
Mehta et al., 2021 [131]	Oil and Gas			*		*	Stakeholder royalty transactions framework
Ashraf et al., 2021 [132]	Power Plant						DL
Tanuska et al., 2021 [119]	Predictive Maintenance		*		*		DL
Marino et al., 2020 [121]	Predictive Maintenance		*		*		Supervised ML
Benbarrad et al., 2021 [123]	Predictive Maintenance				*	*	DL
Züfle et al., 2022 [124]	Predictive Maintenance				*	*	RF
Wang et al., 2021 [130]	Predictive Maintenance	*	*				Supervised ML
Mohan et al., 2021 [118]	Predictive Maintenance		*			*	DL
Yadav et al., 2021 [129]	Quality 4.0					*	Quality control framework
Kotsiopoulos et al., 2021 [106]	Smart grid	*				*	
Petillion et al., 2019 [107]	Smart Manufacturing	*	*		*		DL
(Bécue et al., 2021)	Smart manufacturing					*	Adversarial AI
(Gulivindala et al., 2021)	Smart manufacturing						GAN
Singer and Cohen, 2021 [126]	Smart manufacturing						Supervised ML
Jung et al., 2021 (Jung et al., 2021)	Smart manufacturing						DL
Fukuyama, 2018 [127]	Society 5.0					*	
Mhlanga, 2021 [115]	Socio-economics					*	
Muslikhin et al., 2021 [128]	Supermarket 4.0	*					DL
Xu, 2020 [125]	Systems science in AI					*	
Matheri et al., 2021 [116]	Wastewater	*		*			DL

PM: Process monitoring, SS: Soft sensing, BS: Business systems, SM: Smart manufacturing, GAN: Genetic Adversarial Networks, DL: Deep Learning

- How to properly interpret the ML model.
- How to drive business value from the model.

Among the various solutions that have been identified, the predominant ones are:

- A "six key steps'' framework was identified, using which businesses can drive value by using AI (Alsheibani et al., 2020). The six key steps are summarized as follows:
  o AI compatibility with existing systems.
  o AI skill set available.
  o Effective use and availability of data.
  o The business case aligned with AI use.
  o Gains/goals to be achieved by using AI are identified.
  o Business owners are supporting AI owners to make the move.
- For organizations to utilize or rely on AI, existing Information Systems (IS) must be integrated with industrial systems to allow them to extract value from them efficiently (Scheer, 2012).

In this section, we discuss the current research standings, and the various approaches industrial practitioners have adopted to integrate AI/ML-based systems in their current processes, as well as innovative solutions that have been reached.

9.2. Continuous integration of AI in existing systems

Continuous Integration (CI) is a development practice that is often adapted in modern applications to integrate changes from multiple sources into a project that is already in production. It is mainly a DevOps best practice that allows collaboration between different developers. It has been identified (Rojek et al., 2021) that AI/CI is a necessity for industries to enable and utilize AI in their existing processes to transition into an Industry 4.0. AI/CI also allows the practitioners to overcome the barriers faced by industrial practitioners by incorporating AI into existing systems without disrupting production and enhancing business value. Besides the impact of AI and Industry 4.0 on the technological advancement of companies, according to (Mhlanga, 2021) together they both can play a major role in the socio-economic development of a country as well. Education is a major variable that contributes to the overall development of a country, and AI can play a major role in it by bringing adaptive learning together with Industry 4.0, bringing education to remote regions. This is an integration of Education 4.0 and Industry 4.0, where AI-enabled education services will be made available to rural areas with minimal need for human interaction. Practitioners in other industries, such as Wastewater Treatment Plant (WWTP) (Matheri et al., 2021), have derived business value by utilizing ANNs to predict trace metals. It was suggested that due to the chemical composition and complexity of various processes that undergo in a WWTP, an AI-enabled process will be able to leverage the data and improve the efficiency of the existing system. The financial industry has also benefited from the integration and adaptation of AI-enabled systems. For example, (Kartanaitė et al., 2021) it was aimed to define and summarize the main financial/economic forecasting methods for production companies that can enable them to adopt Industry 4.0. The model was using features based on company size, production planning, supply and demands, and industry type. The average accuracy of 96% was achieved using AI for financial forecasting. Four financial forecasting groups that were classified in the research are production, customer and demand, industry, and media information.

9.3. Productive Maintenance

Although in industry 3.0 a Total Productive Maintenance (TPM) checklist was implemented adhering to which allowed downtime reduction by 53% and breakdown occurrences by 54%. The key performance metrics that were identified were Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR), and Overall Equipment Effectiveness (OEE). However, it is nearly impossible to achieve zero downtime in production. In Industry 4.0, utilizing state-of-the-art data-driven approaches has allowed practitioners to predict the likelihood of downtime using historical data. Also known as Predictive Maintenance, where alternatives or backups are arranged beforehand so as to achieve nearly zero downtime in production. ML-based models have seen a lot of integration in predictive maintenance, especially because of the availability of data from a variety of sources. The same data can be utilized to predict failures, errors, and other variables that can be utilized in maintenance. In (Mohan et al., 2021) an ML-based approach for predictive maintenance was proposed. A data-driven regression system was proposed that transforms industry 3.0 to Industry 4.0 with minimalistic changes to the existing system. According to the analysis, the main contributing factor in the major breakdown was oil contamination in the machinery setup. Therefore, regular measurement of oil contaminations using IoT sensors allowed them to achieve zero downtime.

A study (Tanuska et al., 2021) presented a system to prevent accidental carrier bearings damage on an assembly conveyor. The system utilized LoRaWAN (Adelantado et al., 2017) to gather data from sound analysis and temperature monitoring. Data was used to train a neural network to predict anomalies in the system. Similarly, a novel approach to fault detection in distributed industrial 4.0 systems based on various ML models was presented (Marino et al., 2020). In the research, it was suggested that the use of lightweight ML classifiers as it allows to distribute computing to edge nodes in a sensor network. This enabled the proposed approach to be implemented in a resource-constrained environment. In (Kudelina et al., 2021) the current state of ML-based fault diagnostic techniques in electrical fields for predictive maintenance was reviewed. They argued that reliance on artificial intelligence-based condition monitoring is increasing as computational abilities are increasing. They also suggested that, unlike conventional diagnosis, which was made by onboard devices, AI-enabled processes can run on the cloud and are not restrained. Consequently, this will improve the reliability of industrial systems and processes. In another study (Benbarrad et al., 2021) a review of various computer vision algorithms that can be utilized in process prediction and quality control. They also proposed a framework that utilized IoT and edge computing to enable process prediction using transfer learning. ML-based algorithms have also been employed to detect the current production and machine degradation in power generators. This was achieved by utilizing and capturing multiple streams of data generated by various sensors. The sensor data contained information about machine velocity, position, vibration, and temperature of various machine parts, which was then utilized by the ML algorithm for analysis (Züfle et al., 2022).

9.4. Closed-loop systems

Control theory, cybernetics, and other enabling engineering systems have been the key components of Industry 4.0. The overlap between the various systems enabled the development of systems that utilized state-of-the-art technologies to achieve smart manufacturing. This has led to the creation of closed-loop systems, also known as cybernetic systems. Generally, a closed-loop system is a system in which the output of the system is utilized as an input or “feedback” as the system progresses or evolves to be more efficient (Xu, 2020) based on the feedback. Recent advancements in IoT sensors and data technologies have caused a lean in Industry 4.0 towards closed-loop systems. Closed-loop systems are effective in defect prevention especially in manufacturing systems to achieve minimal downtime (Radel et al., 2019). (Singer & Cohen, 2021) proposed a framework that utilized data from sensors to train ML models for anomaly detection in a closed-loop system. They suggested 3 components in their framework: firstly, the automatic process control component that includes a variety of control factor values for various processes; secondly, the process component that measures and estimates various process states; lastly, the self-adaptive process switching component that uses various ML models to identify the best process state that should be taken. This allows the system to achieve self-adaptive control as the ML model covers various patterns and dependencies present within the data. The system also evolves as more data is fed and achieves better performance.

9.5. AI and society

Society 5.0 (Fukuyama, 2018) is a concept that was first drafted in the 5th Science and Technology Basic Plan by the Council for Science, Technology, and Innovation in January 2016. After which it was identified as one of the growth strategies for Japan. The goal of Society 5.0 is to create a human-centric society in which both the financial development and cultural challenges are conquered, and individuals can partake in a life that is completely dynamic and agreeable. Society 5.0 is also considered Industry 5.0 where the technology is seemingly integrated with humans in every facet of life. In (Muslikhin et al., 2021) an automated product picking by robots to fulfill online orders using the state-of-the-art real-time object detection algorithm Yolo (Huang et al., 2018) was proposed. It was also suggested that automated order fulfilling in a store will help in achieving Society 5.0 goals. Similarly, a concept was introduced by amazon known as Supermarket 4.0 called amazon go[3] which is a new kind of supermarket that utilizes deep learning, computer vision, and sensor technology to allow grab-and-go shopping. It was their initiative to contribute to Society 5.0 as an Industry 4.0 revolution. Customers entering an amazon go store use a QR code using the amazon go app. Any item a customer picks is automatically added to their virtual cart and removed if they put it back. In the end, all they do is walk out and the store will automatically charge them based on their virtual cart. It is believed that Society 5.0 will be the interdependence of man and machine using cognitive computing and human intelligence. This will allow for mass customization and personalization with seamless and continuous communication at the heart of it. COVID -19 is also considered a driving force that pushed industries to adopt Society 5.0 (Sarfraz et al., 2021) as a step to adopt Industry 4.0. It was mainly because of working online that mass communication strategies were adopted.

9.6. Quality control

Quality 4.0 (Zonnenshain & Kenett, 2020) is the utilization of Industry 4.0 innovations, for example, IoT, ML, AI, Big-Data, and Robotics to guarantee quality improvements. While Quality 4.0 spotlights innovation, measures, cooperation, culture, etc., traditional quality control focused exclusively on products. Lean Six Sigma (LSS) has been utilized extensively in the past quality control measures by businesses to improve performance by increasing collaboration. Therefore, Critical Success Factors (CSF) for LSS in the context of Quality 4.0 were discussed in the research (Yadav et al., 2021). It was argued that CSFs for LSS Quality 4.0 are not clearly defined and there was a gap in research. Therefore, they contributed CSF for LSS using Quality 4.0 in various organizations.

Product Lifecycle Management (PLM) is an industry that deals with a multitude of stakeholders from engineering to management throughout. Since the industry is moving rapidly towards smart manufacturing and quality products PLM is also leveraging AI in its processes. Therefore, a roadmap for the implementation of AI in the context of PLM along with various challenges, opportunities, and potential research is presented in (Wang et al., 2021). Injection Molding (IM) is a manufacturing industry in which parts are produced by injecting molten materials into a mold. Quality control is a major concern in the IM manufacturing industry and in (Jung et al., 2021) application of various ML-based algorithms to predict the quality of the molding process is presented. According to the study, autoencoders outperformed other classifiers in terms of predictive accuracy.

9.6.1. Other industries

Smart manufacturing is another area that is heavily researched and impacted by the integration of Industry 4.0. Due to the increase in demand for customized products from the customers, factories need to change delivery models, by adapting AI-enabled processes to provide high-value manufacturing (Wan et al., 2020). One key area is the Oil and Gas sector (OaG) industry where they must manage a multitude of stakeholders ranging from landowners to operators, to governmental bodies. Making it difficult to manage the supply chain effectively and efficiently. Therefore, to derive business value by adapting Industry 4.0. In (Mehta et al., 2021) a blockchain-based transactions scheme was proposed that securely executes transactions among various stakeholders involved in OaG industries. Similarly in the power industry, AI has been utilized (Ashraf, Uddin, Farooq, Riaz, Ahmad, Kamal, Anwar, El-Sherbeeny, Khan, & Hafeez, 2021) to generate power curves by training an ANN on pre-processed operational data of a power plant. It was proposed that data-driven construction of power curves can be treated as Industry 4.0 data analytics and not only promotes energy-efficient solutions but can do so at lower costs. Genetic Algorithms (Gulivindala et al.) have also been adopted in manufacturing industries typically in optimal disassemble sequence plan (DSP) for real-time product performance evaluation. This allows for the utility of traditional AI algorithms to utilize the data from IoT sensors to enhance production quality. Lastly, in (Liebrecht et al., 2021) a methodology to support the decision-making of industrial companies for Industry 4.0. application in a production environment is presented. Big Data, IoT, VR, and AR are the most important technologies to enable Industry 4.0. They have also provided a recommendation for Industry 4.0 implementation roadmap with various readiness indicators and preconditions to reach the goal of Industry 4.0.

9.6.2. Discussion

A summary of the papers that discussed the issues of driving business value by upgrading to Industry 4.0 by utilizing AI/ML-enabled processes is given in Table 7. A common issue when employing ML models is model selection, and it was found to be a common issue in most of the works presented in this study. Since this is a well-known issue in the ML community, it has been omitted from the table. There are many existing automated model selection tools/libraries available that automate the process of selecting the right model for the right data, an example of one is AutoML. A common source of concern for industries when adopting Industry 4.0 is the return on investment. It was found that there is an opportunity for industries to create business value by relying on predictive maintenance, additive manufacturing, or generally smart manufacturing, and AI has a role to enable that. But certain obstacles such as return on investment become an obstacle for small to mid-tier industries, and there is a potential for future research there.

Table 7 summarizes how various industries have driven value from utilizing AI/ML-enabled processes integrated with their existing systems. AM or 3D printing industry where AI-enabled processes typically ANNs have been utilized to monitor nozzle pressure for quality control. The power industry has driven business value by utilizing AI-enabled processes to simulate power generation curves which helped them in optimizing their existing processes to improve efficiency. Due to the flexibility of use, there are almost infinite applications of DL in the industry. Many industries are relying on DL-based models integrated within their existing systems to improve sustainability and increase efficiency. Some studies have proposed frameworks for industrial AI integration; however, these frameworks are very domain-specific, and a potential direction would be a generalized industrial framework. Various sensors are used to monitor product manufacturing and the wastewater industry and to perform predictive maintenance. These industries utilize DL. However, sensor data can be unreliable due to environmental contamination during the lifecycle of the instruments. Therefore, data validation methodologies need to be adopted to classify sensors according to various readings.

9.6.3. Ethics in Artificial Intelligence

The identified literature from the survey did not include work that discusses ethics. However, there is an increased concern about ethics in AI and there are standardization bodies that start to address those concerns and incorporate them into standards (Winfield et al, 2021), (Standards Australia, 2020). Jobin et al. (2019) identified the following main principles of ethics in AI through a literature survey: Transparency, Justice/Fairness, Non-Maleficence, Responsibility, Privacy, Beneficence, Freedom & Autonomy, Trust, Sustainability, Dignity, and Solidarity. We discuss here briefly the top five principles which were mentioned as being significant by more than half of the literature investigated by (Jobin et al., 2019):

[3] Amazon Go is a state-of-the-art store that features grab and go shopping with minimum human intervention needed. (https://www.amazon.com/b?ie=UTF8&node=16008589011)

- **Transparency**: is about explaining, understanding, and interpreting AI algorithms. This will be relevant for Industry 4.0 if AI algorithms are used for decision-making. It is important to understand why algorithms come to certain decisions and this needs to be communicated to end-users as part of transparency.
- **Justice and Fairness**: deals with fairness, consistency, inclusion, equality, equity, and related topics. It could be relevant for Industry 4.0 if AI is for example applied in Human Resources of manufacturing and recruiting new workforce.
- **Non-maleficence**: refers to security, safety, harm, protection, and similar topics. This could be relevant in different ways for Industry 4.0. For example, from a manufacturing point of view, does the manufacturing process create products that are safe and do not cause harm. It can also be relevant for the manufacturing process itself, for example, is the application of robots on a shop floor safe and does not cause harm to employees.
- **Responsibility**: This principle is relevant when dealing with accountability, liability, and acting with integrity. If AI algorithms are applied in an Industry 4.0 context, then it must be specified who is responsible for them and for the decisions made based on the AI algorithms.
- **Privacy:** The last main principle deals with privacy and personal or private information. Industry 4.0 covers mainly the corporate space and privacy might be less relevant for the manufacturing process itself which does not involve personal or private information. However, with the increased connection of products through the Internet of Things (IoT) and the collection of customer data, privacy is relevant and should be considered in the design of products and services.

There are standardization bodies such as IEEE, ISO/IEC, and Standards Australia, which are creating new standards for addressing ethical concerns. The IEEE Standards Association (SA) is working on various recommendations including *ethically aligned Design*[4], and *ethics in the development of AI and Autonomous Systems*[5]. The ISO/IEC is covering ethics in the JTC 1/SC 42 on Artificial Intelligence[6] and developed a transparency taxonomy, an overview of ethical and societal concerns, and investigates risk management, controllability, treatment of unwanted bias, and trustworthiness in artificial intelligence. Standards Australia published a roadmap in consultations with a broad cross-section of stakeholders at the beginning of 2020[7] (Standards Australia, 2020) and highlight the importance of ethics in alignment with IEEE and ISO/IEC.

9.6.4. Summary of Industrial Artificial Intelligence Concerns

In this section, a summary of the concerns that have been drawn from previous subsections which cause an industry to resist the change towards adopting AI-enabled processes to transition towards Industry 4.0 (Bécue et al., 2021). These concerns also answer research question 3. At an abstract level, the IAI concerns can be divided into 5 generic categories which are data availability, skill shortage, data security, return on investment, and passive mindset (Stanisławski & Szymonik, 2021; Yao, 2017). These concerns are summarized in the subsections to follow. Since this study is more about applications of AI in an industrial context some of the concerns raised here are supported by our literature, whereas, others mentioned such as passive mindset, and skill-shortage are not discussed in much of the works mentioned earlier and are not the focus of this study.

[4] https://standards.ieee.org/industry-connections/ec/ead1e-infographic/
[5] https://standards.ieee.org/news/2016/ethically_aligned_design/
[6] https://www.iso.org/committee/6794475/x/catalogue/
[7] https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx

9.6.5. Data Availability

One of the biggest concerns when it comes to adapting AI-enabled processes is the availability of data. The industrial system relies heavily on data from multiple sources such as IoT sensors etc (Darvishi et al., 2020). The concern here is the data is often noisy and voluminous and therefore, a lot of time needs to be vested to prepare the data before performing any form of analysis on it. Additionally, it is oftentimes very costly to generate training data where the preciseness of data is important. Moreover, to enable faster and more accurate judgment based on heterogeneous data environments (image, video, database, text, audio, and others) additional processes must be relied on which is also a concern for high-value manufacturing and production environments. Therefore, to ensure data reliability practitioners suggested the use of ML-based algorithms (classifiers) that can classify the input data as reliable or not and filter unreliable data preventing it to be used in inference (Fuller et al., 2020). AI can assist to address the concern of data shortage in some cases using two possible means (a) when some sample data is available, synthetic data can be generated based on SMOTE, GAN, and other related algorithms and develop an AI model from that data, (b) Generate a whole dataset based on some pre-conditions, domain, and ranges by applying statistical methods and create AI model for these virtual scenarios and outcomes.

9.6.6. Skill Shortage

To upgrade an industrial process to an AI-enabled industrial process a lot of technical manpower is required. Not only people who are domain experts are needed but additional staff such as data scientists, AI engineers, data engineers, etc., are needed to understand the system and its components. Also, depending on the industry, especially data-sensitive sectors such as defense sectors it is not always possible to find the right people according to compliance requirements (Mao et al., 2019). It was identified that to cope with skill shortage one solution is to have the existing people get the AI/ML skills specifically so that they have the necessary knowledge to innovate and update existing systems (Kakkar et al., 2021).

9.6.7. Data Security

As with the evolution of technology malicious users are also acquiring the skills and tools that are obtained by cybersecurity professionals. Especially with the advent of IoT sensitive data is monitored by sensors such as health care systems which transmit patient sensitive data and can be stolen for wrongful purposes. Another source of concern is adversarial attacks where an AI-based model is fooled by giving wrongful inputs to the system. Although a lot of research is underway to prevent malicious use and improve security, however, security still is a major source of resistance for industries that rely heavily on sensitive data (Brundage et al., 2018). Physical unclonable function (PUF) (Yoon et al., 2020) that physically provides a digital fingerprint of the object. PUF allows for mutual authentication between devices in a large-scale IoT environment. Therefore, a scheme of arbiter PUFs to establish the identities of IoT devices in an Industry 4.0 setting is proposed as a solution in the literature (El-Hajj et al., 2021). Blockchain-based IoT device security research is also underway and requires further development (Rathee et al., 2021). IIoT is used for production lines prone to Distributed Denial of Services (DDoS). Huraj et al. proposed ML-based DDoS mitigation for IoT-based production lines (Huraj et al., 2021).

9.6.8. Return on Investment

Another cause of concern is the cost associated when it comes to adopting AI-based processes. The two obvious costs are the cost of change, and the cost of talent, however, as Yao (Yao, 2017) suggested the cost associated with false positives. Yao notes that "if an analytical system on a plane determines an engine is faulty, specialist technicians and engineers must be dispatched to remove and repair the faulty part. Simultaneously, a loaner engine must be provided so the airline can keep up flight operations. The entire deal can easily surpass $200,000." This as opposed to faulty prediction on an Amazon online store giving wrong recommendations to a customer is unsurmountable. As with operational cost, the solution that was identified to minimize cost is to use simulated data and draw inference from the simulated data before an actual system can be adopted.

9.6.9. Passive Mindset

According to a study (Stanisławski & Szymonik, 2021) conducted in Poland regarding industries' willingness to adopt Industry 4.0 and intelligent systems that utilize AI. It was found that many companies believe that IAI will not give them any market advantage. The focus of the study was on six key areas IoT, Blockchain, SCADA (Supervisory Control and Data Acquisition), and SMAC (Social, Mobile, Analytics, Cloud), and their impact on the market. It was concluded that only large enterprises are willing to implement IAI and that is predominantly because they have the capital to invest in the change. Most micro-business is not interested because they do not see tangible benefits in near future, also, they do not have the resources to implement the change.

10. Opportunities for future research

Industry 4.0 is an ecosystem of a variety of interfaces of information with individuals, processes, administrations, frameworks, and IoT-empowered modern resources across the digital and the physical world. The objective is to make the most out of data. For prospective practitioners, I4.0 portrays a future condition of the industry described by exhaustive digitized processes. Although a significant amount of research has been conducted in the field, there still exists potential opportunities for future research. One critical requirement as identified in this research is the availability of a generalized framework that allows the application IAI approaches across different industries. Since most of the applications are specific to a certain industry or domain, it is difficult to reuse them in other contexts. Thus, a generalized framework is desired that enables industrial practitioners to utilize or customize existing approaches to address their specific needs. IAI adaptation requires heavy reliance on technology adaptation which means that a stable IT infrastructure and matching skillset are present. Therefore, research into the required skill set and required technologies needed for IAI adoption will assist prospective practitioners to know about the right people and the right technology to acquire before they decide to opt for IAI technology adoption. Since introducing IAI technology in an organization requires significant IAI expertise, research into automating technology deployment and model maintenance is an opportunity to increase the efficiency of the adoption process and lower the barriers to entry. This is especially true for industries that are reluctant to adopt IAI, as a common misconception is that IAI requires significant investment which may not deliver sufficient returns relative to other potential initiatives. Additionally, a significant amount of research is present that details the ongoing processes that exist for the creation of information highway, however, there is a need for further investigations on ways to improve the supply chain in continuous economies.

Furthermore, despite the applicability and utility of AI in various industries, there is still a need to increase the trust of society on AI. As mentioned by (Black et al., 2022), AI systems are getting used to coping with systems that involve a lot of variability and stochasticity where there exists a vast range of opportunities. For example, consider the number of variables that exist when making autonomous vehicles. Therefore, there is a need for testing AI systems at a level where confidence has been obtained that the system will respond adequately in all possible situations. However, due to the nature of many AI technologies, they are inherently difficult to test, since many of the models are essentially "black boxes" that convey little insight as to why certain outcomes are generated. The stochastic nature of training and impossibility of conductive exhaustive testing for systems with many high-dimensional inputs further limits the confidence that can be gained through classical testing techniques. Modern software testing practices, such as metamorphic testing can reduce the effort required to assure basic properties are upheld. However, further development of frameworks for explainability and trust of AI systems is direly needed especially in IAI systems. Kumar et al 2019, proposed an ontological framework for Industry 4.0 that addresses the concerns that are faced within the industry. They mentioned the use of robotics has precipitated, among others, and that there is a dire need to develop an interoperable communication model to interconnect various components in a "smart factory" effectively. Therefore, the study proposed a modern-day's perspective of the state of ontologies for Industry 4.0 and opinions both present ontological frameworks and ontological standardization efforts in that discipline. Lastly, Winfield et al. 2021, discussed transparency in AI and how transparency affects different stakeholders in the system. Transparency isn't always a singular asset of structures that could meet the desires of all stakeholders. For example, a consumer does not require the same degree of understanding of a robot compared to an engineer who maintains it. Equally, a user may also require reasons for behavioral aspects that would be apparent and obvious to engineers. Thus, to tackle this issue, P7001 defines five distinct corporations of stakeholders. These stakeholders are categorized into two groups: non-professionals, such as customers, consumers, and other general users, and professionals, such as experts including engineers, business users, and other professionals. These stakeholders are detailed in the P7001 standard and further research is needed to apply it to the industry 4.0 context. Based on the findings, potential opportunities for future research exist in generating a generalized framework for industrial integrations that can allow the utility of AI/ML techniques in existing systems across industries.

11. Conclusion

In this study, a review of the current state of Artificial Intelligence in the context of Industry 4.0 is conducted. The paper presented a pipeline that outlines the various stages of transition from industry to Industry 4.0 and Artificial Intelligence and the predominant concerns and potential solutions related to each step in the pipeline. The findings arising from this review inform researchers by highlighting trends of AI technology development and adoption in several industry sectors, and by identifying remaining areas of concern for the development and adoption of AI technologies for Industry 4.0 applications. Moreover, this work summarized several AI-powered solutions for common concerns in several industries. End-users and industry professionals can draw upon these findings to learn which AI technologies are typically used to address which issues and identify opportunities for adopting technologies and solutions that may already be in use in other industry sectors. In this context, our findings may inform the selection of appropriate AI strategies and technologies.

Through a systematic literature review (SLR) it was observed that research in IAI is seeing a positive trend as more industries are leaning towards AI-enabled processes to achieve sustainable production, improved quality control, zero downtime, customized products, and decrease production/system costs. Various concerns that exist in the four categories, namely data collection, preparation, model construction, and model interpretation, that have been faced by industrial practitioners and solutions adopted were discussed. With regards to RQ1, it was found that the specific concerns addressed by different industries, such as automating data collection and cleaning, quality control, etc share broad similarities, but the nature of the solutions is often too specific to the industry or even individual corporations to be shared widely across industry sectors.

With regards to RQ2, it was found that many of the technological solutions included machine learning methods that rested on technologies including Decision Trees and Deep Learning. However, the specific data preparation and model interpretation steps surrounding the model construction algorithms varied considerably among the examined solutions. This trend towards the integration of AI/ML algorithms in a variety of precise industries speaks for the capability of AI/ML methodologies and the possibilities to achieve innovation. It is also important to note that the enabling technologies have reached a level of maturity, and some of the obstacles present are the availability, acquisition ability, and quality of data. However, as mentioned before, these problems can be solved by adopting state-of-the-art ML-based algorithms that can identify and assist in data acquisition, validation, and processing. The real task, however, is to identify the right algorithm for the right problem. Hence, it is hoped that this review, in which categorization of the various industries is established, the methodologies that they have been adopted, and the solutions that have been found could assist prospective practitioners. One aspect that was noted from the survey was that reliance on various ML open datasets allowed practitioners, especially the power industry to simulate their processes without investing in the cost. Therefore, further creation and sharing of datasets in other industries will allow practitioners to test the efficacy of adopting IAI "digitally" and can assist in making the right decisions.

With regards to RQ3, challenges impeding on the adoption of IAI solutions can be found in the technical and the organizational domain. In the technical domain, the aforementioned challenges related to data availability and quality are common, whereas concerns about the investment and skills required to successfully adopt IAI solutions are commonly cited. One solution that has been found to be widely recognized is the Digital Twin. Digital Twin along with AI not only enables industries to replicate their processes and systems but also provides the data that can be further exploited to test innovative data-enabled solutions. However, as per the SLR, only 8% of the papers in this survey utilized DT and AI as means to achieve IAI which indicates there is still a significant need for research by the industries to adopt DT and AI as the steppingstone toward IAI. One of the key reasons for the industries to be reluctant to move towards Industry 4.0 as identified in the research is the return on investment. Major corporations are sufficiently well-resourced to spend the capital to adopt advanced IAI, whereas there is a perception among small and medium businesses that IAI is too costly to implement or will not deliver a significant-enough benefit to them. Among other key concerns such as domain knowledge, technical expertise, lack of data, false positives, and security are preventing industries from adopting IAI. Similarly, based on the distribution of data based on the IAI pipeline proposed in the study, 14% of papers addressed data acquisition, 18% addressed data processing, 19% addressed model training and testing and 24% addressed model interpretation concerns. The model interpretation which is basically how an industry can drive business value by utilizing IAI is the most researched topic in this work. Given that the end objective of any system is to drive business value, therefore, it is understandable. The main concern is to integrate AI-enabled processes into the existing systems and balance out domain knowledge vs business value vs technical expertise. This is where continuous integration of AI was proposed earlier which has a lot of potential for future research in an industrial context.

Several frameworks have been identified in this study that have been proposed by industrial practitioners for AI integration. However, there is a need for a generalized framework that can be adopted by industrial practitioners from different sectors seeking to adopt AI in their industrial processes. Industry 4.0 is a huge domain and encompasses a plethora of industries and although the study has narrowed down the scope to AI/ML, however, the research is still a vast topic to cover. This study was mainly exploratory and aimed to present an introduction to the topic for non-specialists who wish to obtain a broad overview of the technologies and research aspects that pertain to Industrial Artificial Intelligence in the context of Industry 4.0.

CRediT authorship contribution statement

Zohaib Jan: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing. **Farhad Ahamed:** Investigation, Writing – original draft, Writing – review & editing. **Wolfgang Mayer:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Niki Patel:** Investigation, Writing – original draft. **Georg Grossmann:** Writing – review & editing. **Markus Stumptner:** Funding acquisition. **Ana Kuusk:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This study was co-funded by the Department of Industry, Science, Energy and Resources (Innovative Manufacturing CRC Ltd), and ASC Shipbuilding Pty. Ltd. (IMCRC/ASB/190720) in partnership with the University of South Australia.

References

Abdulkareem, K. H., Mohammed, M. A., Gunasekaran, S. S., Al-Mhiqani, M. N., Mutlag, A. A., Mostafa, S. A., Ali, N. S., & Ibrahim, D. A. (2019). A review of fog computing and machine learning: Concepts, applications, concerns, and open issues. *Ieee Access, 7*, 153123–153140.

Adelantado, F., Vilajosana, X., Tuset-Peiro, P., Martinez, B., Melia-Segui, J., & Watteyne, T. (2017). Understanding the limits of LoRaWAN. *IEEE Communications magazine, 55*(9), 34–40.

Ahamed, F., & Farid, F. (2018). Applying internet of things and machine-learning for personalized healthcare: Issues and challenges. 2018 International Conference on Machine Learning and Data Engineering (iCMLDE).

Alsheibani, S. A., Cheung, Y., Messom, C., & Alhosni, M. (2020). *Winning ai strategy: Six-steps to create value from artificial intelligence.* AMCIS.

Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H.-P. (2014). Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *IEEE Communications Surveys & Tutorials, 16*(4), 1996–2018.

Andronie, M., Lăzăroiu, G., Iatagan, M., Hurloiu, I., & Dijmărescu, I. (2021). Sustainable cyber-physical production systems in big data-driven smart urban economy: A systematic literature review. *Sustainability, 13*(2), 751.

Arden, N. S., Fisher, A. C., Tyner, K., Lawrence, X. Y., Lee, S. L., & Kopcha, M. (2021). Industry 4.0 for Pharmaceutical Manufacturing: Preparing for the Smart Factories of the Future. *International Journal of Pharmaceutics, 120554*.

Ashraf, W. M., Uddin, G. M., Farooq, M., Riaz, F., Ahmad, H. A., Kamal, A. H., Anwar, S., El-Sherbeeny, A. M., Khan, M. H., & Hafeez, N. (2021). Construction of operational data-driven power curve of a generator by industry 4.0 data analytics. *Energies, 14*(5), 1227.

Ashraf, W. M., Uddin, G. M., Farooq, M., Riaz, F., Ahmad, H. A., Kamal, A. H., Anwar, S., El-Sherbeeny, A. M., Khan, M. H., Hafeez, N., Ali, A., Samee, A., Naeem, M. A., Jamil, A., Hassan, H. A., Muneeb, M., Chaudhary, I. A., Sosnowski, M., & Krzywanski, J. (2021). Construction of operational data-driven power curve of a generator by industry 4.0 data analytics. *Energies, 14*(5), 1227.

Aveleira-Mata, J., Muñoz-Castañeda, Á. L., García-Ordás, M. T., Benavides-Cuellar, C., Benítez-Andrades, J. A., & Alaiz-Moretón, H. (2021). Prototipo de IDS para detección de intrusiones con modelos de machine learning en sistemas IoT de la Industria 4.0. *DYNA-Ingeniería e Industria, 96*(3).

Azimi, I., Anzanpour, A., Rahmani, A. M., Liljeberg, P., & Salakoski, T. (2016). Medical warning system based on Internet of Things using fog computing. 2016 international workshop on big data and information security (IWBIS).

Baron, M., & Rrustemi, A. Artificial Intelligence Tools Versus Practice in.

Bécue, A., Praça, I., & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artificial Intelligence Review, 54*(5), 3849–3886.

Belgiovine, M., Sankhe, K., Bocanegra, C., Roy, D., & Chowdhury, K. R. (2021). Deep learning at the edge for channel estimation in beyond-5G massive MIMO. *IEEE Wireless Communications, 28*(2), 19–25.

Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks, 10*(1), 7.

Bertolini, M., Mezzogori, D., Neroni, M., & Zammori, F. (2021). Machine Learning for industrial applications: A comprehensive literature review. *Expert Systems with Applications, 114820*.

Black, R., Davenport, J., Olszewska, J., Rößler, J., Smith, A. L., & Wright, J. (2022). Artificial Intelligence and Software Testing: Building Systems You Can Trust, BCS Press, ISBN 9781780175768.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., & Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.*

Chen, B., Wan, J., Shu, L., Li, P., Mukherjee, M., & Yin, B. (2017). Smart factory of industry 4.0: Key technologies, application case, and challenges. *Ieee Access, 6*, 6505–6519.

Chen, X., Eder, M. A., Shihavuddin, A., & Zheng, D. (2021). A human-cyber-physical system toward intelligent wind turbine operation and maintenance. *Sustainability, 13* (2), 561.

Cheng, J., Zhang, H., Tao, F., & Juang, C.-F. (2020). DT-II: Digital twin enhanced Industrial Internet reference framework towards smart manufacturing. *Robotics and Computer-Integrated Manufacturing, 62*, Article 101881.

Chishti, S. O., Riaz, S., BilalZaib, M., & Nauman, M. (2018). Self-driving cars using CNN and Q-learning. 2018 IEEE 21st International Multi-Topic Conference (INMIC).

Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of Industry 4.0 technologies for industrial performance. *International Journal of Production Economics, 204*, 383–394.

Darvishi, H., Ciuonzo, D., Eide, E. R., & Rossi, P. S. (2020). Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture. *IEEE Sensors Journal, 21*(4), 4827–4838.

Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., & Zomaya, A. Y. (2020). Edge intelligence: The confluence of edge computing and artificial intelligence. *IEEE Internet of Things Journal, 7*(8), 7457–7469.

Drew, L. (2019). *How healthcare data will make or break healthcare AI*. Accenture. Retrieved 11/08/2021 from https://www.accenture.com/us-en/blogs/insight-driven-health/how-healthcare-data-will-make-or-break-healthcare-ai.

El-Hajj, M., Fadlallah, A., Chamoun, M., & Serhrouchni, A. (2021). A taxonomy of PUF Schemes with a novel Arbiter-based PUF resisting machine learning attacks. *Computer Networks, 108133*.

Faheem, M., & Gungor, V. C. (2018). Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0. *Applied Soft Computing, 68*, 910–922.

Fischbach, A., Strohschein, J., Bunte, A., Stork, J., Faeskorn-Woyke, H., Moriz, N., & Bartz-Beielstein, T. (2020). CAAI—a cognitive architecture to introduce artificial intelligence in cyber-physical production systems. *The International Journal of Advanced Manufacturing Technology, 111*(1), 609–626.

Forrester, P. (1999). *Business Process Engineering: Reference Models for Industrial Enterprises*. Integrated Manufacturing Systems.

Fukuyama, M. (2018). Society 5.0: Aiming for a new human-centered society. *Japan Spotlight, 1*, 47–50.

Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *Ieee Access, 8*, 108952–108971.

Garrido-Hidalgo, C., Hortelano, D., Roda-Sanchez, L., Olivares, T., Ruiz, M. C., & Lopez, V. (2018). IoT heterogeneous mesh network deployment for human-in-the-loop challenges towards a social and sustainable Industry 4.0. *Ieee Access, 6*, 28417–28437.

Gellert, A., Precup, S.-A., Pirvu, B.-C., Fiore, U., Zamfirescu, C.-B., & Palmieri, F. (2021). An empirical evaluation of prediction by partial matching in assembly assistance systems. *Applied Sciences, 11*(7), 3278.

Giraldo, S., la Rotta, D., Nieto-Londoño, C., Vásquez, R. E., & Escudero-Atehortúa, A. (2021). Digital transformation of energy companies: A colombian case study. *Energies, 14*(9), 2523.

Gondek, C., Hafner, D., & Sampson, O. R. (2016). Prediction of failures in the air pressure system of scania trucks using a random forest and feature engineering. International Symposium on Intelligent Data Analysis.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Grant, E. (2021). Big data-driven innovation, deep learning-assisted smart process planning, and product decision-making information systems in sustainable industry 4.0. *Economics, Management, and Financial Markets, 16*(1), 9–19.

Guevara, J. C., Torres, R. D. S., & da Fonseca, N. L. (2020). On the classification of fog computing applications: A machine learning perspective. *Journal of Network and Computer Applications, 159*, Article 102596.

Gulivindala, A. K., Bahubalendruni, M. R., Chandrasekar, R., Ahmed, E., Abidi, M. H., & Al-Ahmari, A. Automated Disassembly Sequence Prediction for Industry 4.0 Using Enhanced Genetic Algorithm.

Gulivindala, A. K., Bahubalendruni, M. R., Chandrasekar, R., Ahmed, E., Abidi, M. H., & Al-Ahmari, A. (2021). Automated disassembly sequence prediction for industry 4.0 using enhanced genetic algorithm. *CMC-Computers Materials & Continua, 69*(2), 2531–2548.

Hagenauer, J., & Helbich, M. (2017). A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications, 78*, 273–282.

Han, B., Liu, C., & Zhang, W. (2016). A method to measure the resilience of algorithm for operation management. *IFAC-PapersOnLine, 49*(12), 1442–1447.

Harding, J. A., Shahbaz, M., & Kusiak, A. (2006). Data mining in manufacturing: a review.

Hosseini, M.-P., Pompili, D., Elisevich, K., & Soltanian-Zadeh, H. (2017). Optimized deep learning for EEG big data and seizure prediction BCI via internet of things. *IEEE Transactions on Big Data, 3*(4), 392–404.

Huang, R., Gu, J., Sun, X., Hou, Y., & Uddin, S. (2019). A rapid recognition method for electronic components based on the improved YOLO-V3 network. *Electronics, 8*(8), 825.

Huang, R., Pedoeem, J., & Chen, C. (2018). YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. 2018 IEEE International Conference on Big Data (Big Data).

Huraj, L., Horak, T., Strelec, P., & Tanuska, P. (2021). Mitigation against DDoS attacks on an IoT-based production line using machine learning. *Applied Sciences, 11*(4), 1847.

Jan, M. Z., Munoz, J. C., & Ali, M. A. (2020). A novel method for creating an optimized ensemble classifier by introducing cluster size reduction and diversity. *IEEE Transactions on Knowledge and Data Engineering*.

Jan, Z., & Verma, B. (2020). Multicluster class-balanced ensemble. *IEEE Transactions on Neural Networks and Learning Systems.*

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(2019), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jung, H., Jeon, J., Choi, D., & Park, J.-Y. (2021). Application of machine learning techniques in injection molding quality prediction: Implications on sustainable manufacturing industry. *Sustainability, 13*(8), 4120.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4*, 237–285.

Kakkar, S., Kwapinski, W., Howard, C. A., & Kumar, K. V. (2021). Deep neural networks in chemical engineering classrooms to accurately model adsorption equilibrium data. *Education for Chemical Engineers*.

Kartanaitė, I., Kovalov, B. L., Kubatko, O. V., & Kruŝinskas, R. (2021). Financial modeling trends for production companies in the context of Industry 4.0.

Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., & Tzovaras, D. (2021). Machine learning and deep learning in smart manufacturing: The smart grid paradigm. *Computer Science Review, 40*, Article 100341.

Kudelina, K., Vaimann, T., Asad, B., Rassõlkin, A., Kallaste, A., & Demidova, G. (2021). Trends and challenges in intelligent condition monitoring of electrical machines using machine learning. *Applied Sciences, 11*(6), 2761.

Kumar, V., Chhabra, J. K., & Kumar, D. (2014). Performance evaluation of distance metrics in the clustering algorithms. *INFOCOMP Journal of Computer Science, 13*(1), 38–52.

Kumar, V. R. S., Khamis, A., Fiorini, S., Carbonera, J. L., Alarcos, A. O., Habib, M., … Olszewska, J. I. (2019). Ontologies for industry 4.0. *The Knowledge Engineering Review, 34*.

Kuo, C.-J., Ting, K.-C., Chen, Y.-C., Yang, D.-L., & Chen, H.-M. (2017). Automatic machine status prediction in the era of Industry 4.0: Case study of machines in a spring factory. *Journal of Systems Architecture, 81*, 44–53.

Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature News, 544*(7648), 23.

Lee, C., & Lim, C. (2021). From technological development to social advance: A review of Industry 4.0 through machine learning. *Technological Forecasting and Social Change, 167*, Article 120653.

Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. *Manufacturing letters, 18*, 20–23.

Li, L., Ota, K., & Dong, M. (2018). Deep learning for smart industry: Efficient manufacture inspection system with fog computing. *IEEE Transactions on Industrial Informatics, 14*(10), 4665–4673.

Liao, Y., Deschamps, F., Loures, E. D. F. R., & Ramos, L. F. P. (2017). Past, present and future of Industry 4.0-a systematic literature review and research agenda proposal. *International journal of production research, 55*(12), 3609–3629.

Liebrecht, C., Kandler, M., Lang, M., Schaumann, S., Stricker, N., Wuest, T., & Lanza, G. (2021). Decision support for the implementation of Industry 4.0 methods: Toolbox, Assessment and Implementation Sequences for Industry 4.0. *Journal of Manufacturing Systems, 58*, 412–430.

Lin, C., Han, G., Qi, X., Du, J., Xu, T., & Martinez-Garcia, M. (2020). Energy-optimal data collection for UAV-aided industrial WSN-based agricultural monitoring system: A clustering compressed sampling approach. *IEEE Transactions on Industrial Informatics.*

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue, 16*(3), 31–57.

Liu, Y., Ma, X., Shu, L., Hancke, G. P., & Abu-Mahfouz, A. M. (2020). From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges. *IEEE Transactions on Industrial Informatics.*

Lo, C., Chen, C., & Zhong, R. Y. (2021). A review of digital twin in product design and development. *Advanced Engineering Informatics, 48*, Article 101297.

Lockner, Y., & Hopmann, C. (2021). Induced network-based transfer learning in injection molding for process modelling and optimization with artificial neural networks. *The International Journal of Advanced Manufacturing Technology, 112*(11), 3501–3513.

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of industrial information integration, 6*, 1–10.

Van Geenhuizen, M., & Soetanto, D. P. (2009). (2009) "Academic spin-offs at different ages: A case study in search of key obstacles to growth". *Technovation, 29*(10), 671–681.

Mao, S., Wang, B., Tang, Y., & Qian, F. (2019). Opportunities and challenges of artificial intelligence for green manufacturing in the process industry. *Engineering, 5*(6), 995–1002.

Marino, R., Wisultschew, C., Otero, A., Lanza-Gutierrez, J. M., Portilla, J., & de la Torre, E. (2020). A machine-learning-based distributed system for fault diagnosis with scalable detection quality in industrial IoT. *IEEE Internet of Things Journal, 8*(6), 4339–4352.

Martínez-Díaz, M., & Soriguera, F. (2018). Autonomous vehicles: Theoretical and practical challenges. *Transportation Research Procedia, 33*, 275–282.

Matheri, A. N., Ntuli, F., Ngila, J. C., Seodigeng, T., & Zvinowanda, C. (2021). Performance prediction of trace metals and cod in wastewater treatment using artificial neural network. *Computers & Chemical Engineering, 149*, Article 107308.

Mehta, D., Tanwar, S., Bodkhe, U., Shukla, A., & Kumar, N. (2021). Blockchain-based royalty contract transactions scheme for Industry 4.0 supply-chain management. *Information Processing & Management, 58*(4), Article 102586.
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*, Article 120784.
Mhlanga, D. (2021). Artificial Intelligence in the Industry 4.0, and Its Impact on Poverty, Innovation, Infrastructure Development, and the Sustainable Development Goals. *Lessons from Emerging Economies? Sustainability, 13*(11), 5788.
Mohan, T. R., Roselyn, J. P., Uthra, R. A., Devaraj, D., & Umachandran, K. (2021). Intelligent machine learning based total productive maintenance approach for achieving zero downtime in industrial machinery. *Computers & Industrial Engineering, 107267*.
Moher, D., Altman, D. G., Liberati, A., & Tetzlaff, J. (2011). PRISMA statement. *Epidemiology, 22*(1), 128. author reply 128.
Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT press.
Muslikhin, M., Horng, J. R., Yang, S. Y., Wang, M. S., & Awaluddin, B. A. (2021). An artificial intelligence of things-based picking algorithm for online shop in the society 5.0's Context. *Sensors (Basel), 21*(8), 2813.
Nakagawa, E. Y., Antonino, P. O., Schnicke, F., Kuhn, T., & Liggesmeyer, P. (2021). Continuous Systems and Software Engineering for Industry 4.0: A disruptive view. *Information and Software Technology, 135*, Article 106562.
Nariman-Zadeh, N., Salehpour, M., Jamali, A., & Haghgoo, E. (2010). Pareto optimization of a five-degree of freedom vehicle vibration model using a multi-objective uniform-diversity genetic algorithm (MUGA). *Engineering Applications of Artificial Intelligence, 23*(4), 543–551.
Nath, A. G., Udmale, S. S., & Singh, S. K. (2021). Role of artificial intelligence in rotor fault diagnosis: A comprehensive review. *Artificial Intelligence Review, 54*(4), 2609–2668.
Oehlmann, P., Osswald, P., Blanco, J. C., Friedrich, M., Rietzel, D., & Witt, G. (2021). Modeling fused filament fabrication using artificial neural networks. *Production Engineering, 15*(3), 467–478.
Omairi, A., & Ismail, Z. H. (2021). Towards machine learning for error compensation in additive manufacturing. *Applied Sciences, 11*(5), 2375.
Paiola, M., Schiavone, F., Khvatova, T., Grandinetti, R. (2021). "Prior knowledge, industry 4.0 and digital servitization. An inductive framework", Technological Forecasting and Social Change,Volume 171,2021,120963.
Petillion, R. J., Freeman, T. K., & McNeil, W. S. (2019). United nations sustainable development goals as a thematic framework for an introductory chemistry curriculum. *Journal of Chemical Education, 96*(12), 2845–2851.
Piccarozzi, M., Aquilani, B., & Gatti, C. (2018). Industry 4.0 in management studies: A systematic literature review. *Sustainability, 10*(10), 3821.
Pinheiro, P., Macedo, M., Barbosa, R., Santos, R., & Novais, P. (2018). Multi-agent systems approach to industry 4.0: Enabling collaboration considering a blockchain for knowledge representation. International Conference on Practical Applications of Agents and Multi-Agent Systems.
Pokhrel, S. R., Verma, S., Garg, S., Sharma, A. K., & Choi, J. (2021). An efficient clustering framework for massive sensor networking in industrial IoT. *IEEE Transactions on Industrial Informatics*.
Pop, P., Zarrin, B., Barzegaran, M., Schulte, S., Punnekkat, S., Ruh, J., & Steiner, W. (2021). The FORA fog computing platform for industrial IoT. *Information Systems, 98*, Article 101727.
Qu, Y., Pokhrel, S. R., Garg, S., Gao, L., & Xiang, Y. (2020). A blockchained federated learning framework for cognitive computing in industry 4.0 networks. *IEEE Transactions on Industrial Informatics, 17*(4), 2964–2973.
Grandinetti, R., (2016). "Absorptive capacity and knowledge management in small and medium enterprises", Knowl. Manage. Rese. Practice, 14 (2) (2016), pp. 159-168.
Radel, S., Diourte, A., Soulié, F., Company, O., & Bordreuil, C. (2019). Skeleton arc additive manufacturing with closed loop control. *Additive Manufacturing, 26*, 106–116.
Rathee, G., Balasaraswathi, M., Chandran, K. P., Gupta, S. D., & Boopathi, C. S. (2021). A secure IoT sensors communication in industry 4.0 using blockchain technology. *Journal of Ambient Intelligence and Humanized Computing, 12*(1), 533–545.
Rauch, E., Linder, C., & Dallasega, P. (2020). Anthropocentric perspective of production before and within Industry 4.0. *Computers & Industrial Engineering, 139*, Article 105644.
Ren, Y., Zhang, L., & Suganthan, P. N. (2016). Ensemble classification and regression-recent developments, applications and future directions [Review Article]. *IEEE Computational Intelligence Magazine, 11*(1), 41–53.
Rojek, I., Macko, M., Mikołajewski, D., Sága, M., & Burczyński, T. (2021). Modern methods in the field of machine modelling and simulation as a research and practical issue related to industry 4.0. In *Bulletin of the Polish Academy of Sciences: Technical Sciences* (pp. e136717–e).
Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, *9*(1), 54-89.
Saghezchi, F. B., Mantas, G., Ribeiro, J., Esfahani, A., Alizadeh, H., Bastos, J., & Rodriguez, J. (2018). Machine learning to automate network segregation for enhanced security in industry 4.0. *International Conference on Broadband Communications, Networks and Systems*.
Sanz, E., Blesa, J., & Puig, V. (2021). BiDrac industry 4.0 framework: Application to an automotive paint shop process. *Control Engineering Practice, 109*, Article 104757.
Sarfraz, Z., Sarfraz, A., Iftikar, H. M., & Akhund, R. (2021). Is covid-19 pushing us to the fifth industrial revolution (Society 5.0)? *Pakistan journal of medical sciences*, *37*(2), 591.
Saygin, A. P., Cicekli, I., & Akman, V. (2000). Turing test: 50 years later. *Minds and Machines, 10*(4), 463–518.
Scheer, A.-W. (2012). *Business process engineering: Reference models for industrial enterprises*. Springer Science & Business Media.
Shao, X. F., Liu, W., Li, Y., Chaudhry, H. R., & Yue, X. G. (2021). Multistage implementation framework for smart supply chain management under industry 4.0. *Technol Forecast Soc Change*, *162*, 120354.
Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., & Graepel, T. (2017). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*.
Singer, G., & Cohen, Y. (2021). A framework for smart control using machine-learning modeling for processes with closed-loop control in Industry 4.0. *Engineering Applications of Artificial Intelligence, 102*, Article 104236.
Skilton, M., & Hovsepian, F. (2018). *The 4th industrial revolution*. Springer.
Sony, M., & Naik, S. (2019). Key ingredients for evaluating Industry 4.0 readiness for organizations: A literature review. *Benchmarking: An International Journal*.
Spanaki, K., Karafili, E., & Despoudi, S. (2021). AI applications of data sharing in agriculture 4.0: A framework for role-based data access control. *International Journal of Information Management, 59*, Article 102350.
Stanisławski, R., & Szymonik, A. (2021). Impact of selected intelligent systems in logistics on the creation of a sustainable market position of manufacturing companies in poland in the context of Industry 4.0. *Sustainability, 13*(7), 3996.
Standards Australia (2020). "An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard" URL: https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx. Accessed on 5 June 2022.
Sun, N., Kopper, A., Karkare, R., Paffenroth, R. C., & Apelian, D. (2021). Machine learning pathway for harnessing knowledge and data in material processing. *International Journal of Metalcasting, 15*(2), 398–410.
Tanuska, P., Spendla, L., Kebisek, M., Duris, R., & Stremy, M. (2021). Smart anomaly detection and prediction for assembly process maintenance in compliance with industry 4.0. *Sensors (Basel), 21*(7), 2376.
Taranto-Vera, G., Galindo-Villardón, P., Merchán-Sánchez-Jara, J., Salazar-Pozo, J., Moreno-Salazar, A., & Salazar-Villalva, V. (2021). Algorithms and software for data mining and machine learning: A critical comparative view from a systematic review of the literature. *The Journal of Supercomputing*, 1–33.
Trinks, S., & Felden, C. (2018). Edge computing architecture to support real-time analytic applications: A state-of-the-art within the application area of smart factory and industry 4.0. 2018 IEEE International Conference on Big Data (Big Data).
Uhl-Bien, M., Marion, R., & McKelvey, B. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The leadership quarterly, 18*(4), 298–318.
Wan, J., Li, X., Dai, H.-N., Kusiak, A., Martínez-García, M., & Li, D. (2020). Artificial-intelligence-driven customized manufacturing factory: Key technologies, applications, and challenges. *Proceedings of the IEEE, 109*(4), 377–398.
Wang, K., Zhang, G., Leng, Y., & Leung, H. (2018). Synthetic aperture radar image generation with deep generative models. *IEEE Geoscience and Remote Sensing Letters, 16*(6), 912–916.
Wang, L., Liu, Z., Liu, A., & Tao, F. (2021). Artificial intelligence in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, 1–26.
Warwick, K. (2013). *Artificial intelligence: The basics*. Routledge.
Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine Learning, 8*(3–4), 279–292.
Wei, W. W. (2006). Time series analysis. In *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*.
Winfield, A. F., Booth, S., Dennis, L. A., Egawa, T., Hastie, H., Jacobs, N., … Watson, E. (2021). IEEE P7001: A proposed standard on transparency. *Frontiers in Robotics and A, I*, 225.
Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82.
Xu, L. D. (2020). The contribution of systems science to Industry 4.0. *Systems Research and Behavioral Science, 37*(4), 618–631.
Xu, R., Wunsch, D., & 2nd.. (2005). Survey of clustering algorithms. *IEEE Trans Neural Netw, 16*(3), 645–678.
Yadav, N., Shankar, R., & Singh, S. P. (2021). Critical success factors for lean six sigma in quality 4.0. *International Journal of Quality and Service Sciences*.
Yan, J., Meng, Y., Lu, L., & Li, L. (2017). Industrial big data in an industry 4.0 environment: Challenges, schemes, and applications for predictive maintenance. *IEEE Access, 5*, 23484–23491.
Yao, M. (2017). *4 Unique Challenges Of Industrial Artificial Intelligence*. Forbes. Retrieved 12/07/2021 from https://www.forbes.com/sites/mariyayao/2017/04/14/unique-challenges-of-industrial-artificial-intelligence-general-electric/?sh=23c4c7c21305.
Yao, X., Zhou, J., Zhang, J., & Boër, C. R. (2017). From intelligent manufacturing to smart manufacturing for industry 4.0 driven by next-generation artificial intelligence and further on. *2017 5th international conference on enterprise systems (ES)*.
Yoon, S., Kim, B., Kang, Y., & Choi, D. (2020). PUF-based Authentication Scheme for IoT Devices. *2020 International Conference on Information and Communication Technology Convergence (ICTC)*.
Yu, W., Liu, Y., Dillon, T., Rahayu, W., & Mostafa, F. (2021). An Integrated framework for health state monitoring in a smart factory employing IoT and big data techniques. *IEEE Internet of Things Journal*.
Zhang, X., Ming, X., & Yin, D. (2020). Application of industrial big data for smart manufacturing in product-service system based on system engineering using fuzzy DEMATEL. *Journal of Cleaner Production, 265*, Article 121863.

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE, 107*(8), 1738–1762.

Zonnenshain, A., & Kenett, R. S. (2020). Quality 4.0—the challenging future of quality engineering. *Quality Engineering, 32*(4), 614–626.

Züfle, M., Moog, F., Lesch, V., Krupitzer, C., & Kounev, S. (2022). A machine learning-based workflow for automatic detection of anomalies in machine tools. *ISA transactions, 125*, 445–458.