Realising the promises of artificial intelligence in manufacturing by enhancing CRISP-DM

Citation for the original published paper (version of record):
Bokrantz, J., Subramaniyan, M., Skoogh, A. (2024). Realising the promises of artificial intelligence in manufacturing by enhancing CRISP-DM. Production Planning and Control, 35(16): 2234-2254.
http://dx.doi.org/10.1080/09537287.2023.2234882

(article starts on next page)

Production Planning & Control
The Management of Operations

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tppc20

Realising the promises of artificial intelligence in manufacturing by enhancing CRISP-DM

Jon Bokrantz, Mukund Subramaniyan & Anders Skoogh

To cite this article: Jon Bokrantz, Mukund Subramaniyan & Anders Skoogh (2023): Realising the promises of artificial intelligence in manufacturing by enhancing CRISP-DM, Production Planning & Control, DOI: 10.1080/09537287.2023.2234882

To link to this article: https://doi.org/10.1080/09537287.2023.2234882

Published online: 18 Jul 2023.

PRODUCTION PLANNING & CONTROL
https://doi.org/10.1080/09537287.2023.2234882

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Realising the promises of artificial intelligence in manufacturing by enhancing CRISP-DM

Jon Bokrantz[a], Mukund Subramaniyan[a,b] and Anders Skoogh[a]

[a]Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden; [b]Insights & Data, Capgemini AB, Gothenburg, Sweden

ABSTRACT

To support manufacturing firms in realising the value of Artificial Intelligence (AI), we embarked on a six-year process of research and practice to enhance the popular and widely used CRISP-DM methodology. We extend CRISP-DM into a continuous, active, and iterative life-cycle of AI solutions by adding the phase of 'Operation and Maintenance' as well as embedding a task-based framework for linking tasks to skills. Our key findings relate to the difficult trade-offs and hidden costs of operating and maintaining AI solutions and managing AI drift, as well as ensuring the presence of domain, data science, and data engineering competence throughout the CRISP-DM phases. Further, we show how data engineering is an essential but often neglected part of the AI workflow, provide novel insights into the trajectory of involvement of the three competences, and illustrate how the enhanced CRISP-DM methodology can be used as a management tool in AI projects.

ARTICLE HISTORY
Received 11 November 2022
Accepted 2 July 2023

KEYWORDS
Artificial intelligence; machine learning; CRISP-DM; manufacturing; production

1. Introduction

Artificial Intelligence (AI) is promised to revolutionise manufacturing. For example, using AI for automated inventory replenishment can yield a close to 40% reduction in penalty costs and up to 60% reduction in misplaced items (Wang, Skeete, and Owusu 2022). Similarly, AI solutions in assembly operations can result in a drastic reduction of energy consumption and excess storage stocks (Manimuthu et al. 2022), and deep learning models can be used to extract hidden insights of machine breakdowns from unstructured text in maintenance logs (Usuga-Cadavid et al. 2022). But, if AI is so powerful, why do so many manufacturing firms struggle so hard in closing the gap between the promises of AI and real productivity gains? One core part of the explanation is that implementing new technology requires complementary intangible investments (Brynjolfsson, Rock, and Syverson 2021). That is, before a manufacturing firm can realise the benefits of AI, it must put and keep in place the essential complements that allow the technology to be effectively leveraged. Such complements are everywhere in modern organisations and can range from new business processes, managerial practices, organisational designs, and even cultural change (Brynjolfsson and Milgrom 2013). Decades of research have validated the importance of complementarities for realising the value of IT (Bresnahan, Brynjolfsson, and Hitt 2002), and recent empirical evidence shows that AI productivity gains are almost entirely limited to manufacturing firms that have the right complements in place (Brynjolfsson, Jin, and McElheran 2021).

To combat this lack of realised benefits from AI implementation, recent literature has put the spotlight on challenges to long-term success such as governance (e.g. structure or roles, practices, and processes) (Schneider et al. 2023), strategy (e.g. access to technology and human capital investment) (Amankwah-Amoah and Lu 2022), management (e.g. defining objectives, setting constraints, choosing training data, providing feedback) and concrete tools that support practitioners in AI scaling (e.g. choosing performance metrics and engaging with stakeholders) (Madaio et al. 2022).

Manufacturing firms that aim to reap the benefits of AI[1] thus face a tough challenge: developing technical AI systems is fast and cheap, but scaling them and realising their long-term value is difficult and expensive (Sculley et al. 2015), in large part due to considerable managerial and organisational challenges (Fosso Wamba et al. 2022). In this article, our central thesis is that value creation from AI in manufacturing can accelerate by means of two interrelated complementarities: (1) operation and maintenance of AI solutions over time, and (2) human skills. First, operating real-world AI solutions entails hidden and difficult trade-offs that must be considered in the long run, many of which result in massive maintenance costs (Sculley et al. 2015). Second, one of the most established, clear-cut, and essential complements to AI is human skills, because even the most sophisticated AI technology has limitations that make humans indispensable for many essential tasks (Brynjolfsson, Rock, and Syverson 2021).

CONTACT Jon Bokrantz jon.bokrantz@chalmers.se

While we focus exclusively on these two, it should be recognised that there may exist a myriad of plausibly important complementarities, but which are beyond the scope of this study.

To support manufacturing firms in managing these complementarities, we believe that one fruitful strategy is to embed hands-on guidelines in existing and widely used engineering methodologies. In particular, there exists a wide range of systematic methodologies that prescribe a sequence of interrelated phases for developing AI solutions to specific practical problems such as CRISP-DM, SEMMA, and KDD (Tripathi et al. 2021). CRISP-DM (Wirth and Hipp 2000) has been extensively used in both industry and academia for successfully developing AI solutions (Schröer, Kruse, and Gómez 2021; Huber et al. 2019) owing to its practical orientation (i.e. explicit focus on business problems, easy to implement, and complete scope), agility (i.e. capable of increasing the velocity of decision-making through quick iterations), and flexibility (i.e. capable of supporting any AI approach whether data-centric, model-centric, or user-centric) (Azevedo and Santos 2008; Dåderman and Rosander 2018). Scholars often explicitly use CRISP-DM to guide their AI solution development (Kharlamov et al. 2020; Jacobsen and Tan 2022) but many also implicitly follow the same or similar phases of CRISP-DM without mentioning the use of a specific methodology.

Yet, while the overall benefits of using CRISP-DM in a manufacturing context are unquestionable, the methodology was initially developed to serve as a generic and systematic approach for conducting data mining projects. Many authors have therefore advocated for continuous development of the methodology to better accommodate technological advancements and more complex problem-solving (Schröer, Kruse, and Gómez 2021). For example, since the original version ends at the deployment stage, scholars have called for research that develops better guidance for practitioners in managing the hidden difficulties and costs associated with operating and maintaining AI solutions over time (Wang, Skeete, and Owusu 2022). Further, owing to the original version's lack of explicit links to the essential human skills that must be present to execute all tasks (Wirth and Hipp 2000), scholars have also called for research that portrays a more holistic perspective on the shared competences that must be involved throughout the entire life-cycle of AI solutions (Huber et al. 2019).

Above all, there is a dire need for more in-depth industry insights about AI usage in real organisations together with clear guidance to managers and practitioners on how to use AI to improve productivity (Fosso Wamba et al. 2022). Of particular value is uncovering the success factors of leading companies and extracting their experiences in developing and implementing real-life AI solutions (Helo and Hao 2022).

Therefore, this article aims to support manufacturing firms in realising the value of AI by enhancing the CRISP-DM methodology. Through a multi-year and multi-method approach that spans six years of immersion in research and practice, we achieve our research aim in two major ways. First, we extend CRISP-DM into a complete life-cycle of AI solutions by adding the phase of 'Operation and Maintenance' of AI solutions in manufacturing environments over time. Second, we embed a task-based framework into CRISP-DM that provides a comprehensive yet intuitive way of linking tasks to skills and ultimately integrating competences held by sets of experts. The implications of the enhanced CRISP-DM are illustrated based on an application of AI-based throughput bottleneck detection in a real-world manufacturing system. We conclude by offering four core recommendations for using CRISP-DM in manufacturing practice.

2. Theoretical background and related work

2.1. CRISP-DM and extensions in manufacturing

In this section, we summarise the CRISP-DM methodology in its original format as well as describe related key extensions of CRISP-DM for manufacturing contexts. For an overview of similar and competing methodologies such as KDD and SEMMA, see Azevedo and Santos (2008) or Dåderman and Rosander (2018). CRISP-DM (Cross Industry Standard Process for Data Mining) was developed as a generic and systematic approach for conducting data mining projects. The original version is extensively described in existing literature (Wirth and Hipp 2000). The goal of introducing CRISP-DM was to serve as a common reference point to discuss data mining, increase the understanding of crucial data mining issues, and fortify data mining as an established engineering practice (Wirth and Hipp 2000). Today, CRISP-DM is widely known as the 'de-facto standard' for applying a process model in data mining projects (Schröer, Kruse, and Gómez 2021). The model represents a systematic process for data mining projects in a hierarchy of four levels: (1) phases, (2) generic tasks, (3) specialised tasks, and (4) process instances. The first level consists of six phases: *business understanding* (understanding and converting project objectives and requirements into a problem definition), *data understanding* (collection and familiarisation with the data), *data preparation* (constructing the final data set), *modelling* (selection, application, and tuning of the model), *evaluation* (assess model performance), and *deployment* (organise, present, and implement the model). The second level consists of generic tasks within each phase, which are intended to be complete and stable and thus apply to all possible data mining situations (e.g. generic task 'determine business objectives' in phase 'business understanding'). The third level consists of specialised tasks that describe how actions in the generic tasks should be carried out in specific contexts (e.g. specialised task 'build clustering model' for generic task 'build model'). The fourth and final level consists of process instances that represent the documentation of actions, decisions, and results during the project. All phases are intended to be iterative and intertwined, and the model was originally envisioned to be applied in a cyclic rather than a waterfall fashion (Wirth and Hipp 2000).

Since the original version of CRISP-DM is intended to be generic, it has been widely adopted across disciplines such as engineering, quality, marketing, and healthcare (Studer et al. 2021). At the same time, the generic nature has driven the development of contextual versions referred to as 'CRISP-

DM extensions'. Such extensions aim to overcome limitations of the original version, tailor it to specific situations, and/or better accommodate recent technological advancements. Specifically, Tripathi et al. (2021) outline two types of extensions: (1) general extensions and (2) industry-specific extensions. The first type focuses on extensions according to general theoretical concepts. For example, Kristoffersen et al. (2019) extend CRISP-DM according to the circular economy concept by introducing an additional phase of data validation and by incorporating analytic profiles. The second type focuses on extensions according to certain industry sectors and/or application areas. For example, Huber et al. (2019) extend CRISP-DM to the manufacturing industry by adding three new phases: technical understanding, technical realisation, and technical implementation.

Concerning the operations and maintenance of AI solutions, Schröer, Kruse, and Gómez (2021) reviewed 24 studies that applied CRISP-DM and observed that the deployment phase was neglected in the vast majority of studies; not a single study fully implemented the solution in an operational environment. They conclude that more research is needed with respect to integrating models and data in operational environments where model performance needs to be continuously monitored and controlled.

Tackling this particular problem, Tripathi et al. (2021) place a special focus on the long-term use of Machine Learning (ML) models in manufacturing and emphasise a wide variety of robustness issues that need to be considered in the deployment phase and beyond. For example, they argue the need for ensuring the models' utility and robustness over time. Due to dynamic changes in the operational environment, the performance of ML models may deteriorate over time, requiring manufacturing companies to monitor changes in terms of e.g. subject, frequency, transition, recurrence, and magnitude. Further, they highlight the need of users (e.g. machine operators) to evaluate the accuracy, interpretability, multiplicity, and transparency of models to ensure that operating personnel continuously understand and trust the results.

Similarly, Studer et al. (2021) explain that CRISP-DM does not cover the application scenario of AI solutions that are used over long periods where the model must be able to adapt to changing environments. Ignoring dynamic changes to the operating environment such as non-stationarity of data, degradation of hardware, and system updates, may cause the AI solution to degrade over time. To cope with this challenge, the authors extend CRISP-DM with a 'Monitoring and Maintenance' phase. This phase includes tasks to monitor input signals, notify when changes have occurred, and re-train the model with new data.

From a human skills perspective, Wirth and Hipp (2000) explain in the original CRISP-DM version that data mining is a creative process that requires several different skills and knowledge, and that success or failure is highly dependent on the teams that carry out the process in practice. This notion has been emphasised in manufacturing-specific extensions. Hiruta et al. (2019) highlight the interplay between data scientists and domain engineers when developing analytics solutions for condition-based maintenance. They argue that both roles are needed but that efficient collaboration is difficult due to differences in backgrounds and challenges in sharing tacit knowledge. Based on CRISP-DM, they develop a three-step procedure (design, evaluation, execution) and an accompanying engineering tool that supports data scientists and domain engineers in jointly developing data analytics solutions.

Similarly, Huber et al. (2019) explain that to implement useful data analytics solutions in manufacturing, cross-disciplinary cooperation of data scientists and domain experts is required. In particular, they highlight the specific difficulties in obtaining and processing data in manufacturing environments that necessitate a deep understanding of production processes as well as machine controls and sensors. To facilitate the cooperation between data scientists and domain experts, they extend CRISP-DM with three new phases: technical understanding, technical realisation, and technical implementation. These phases include domain-specific tasks such as selection of physical parameters, development of data acquisition from machine sensors and interfaces, and actual implementation in operational hardware and software.

Tripathi et al. (2021) complement this perspective and highlight that three key roles need to actively collaborate in CRISP-DM: business experts, data experts, and users. The business experts focus on explaining the business context and goals; the data experts focus on building, evaluating, and scaling the model; the users focus on the model response during daily operations.

2.2. A Task-based framework for AI-skills complementarities

In this section, we describe the theoretical fundamentals of AI-skills complementarities based on economic theory followed by introducing a task-based framework for matching tasks with skills. The link between technological change, labour, and productivity can be understood through task-based economic principles. These principles describe how technology and labour complement each other to realise productivity gains at the level of tasks, and they explain and predict why only some firms are able to reap the benefits from AI (Acemoglu and Restrepo 2018). We explicitly chose an economics lens due to the far-reading implications of AI-skills complementarities and the power of task-based economic principles for describing, explaining, and predicting differences in productivity gains (Acemoglu and Restrepo 2018). Also, a generic task-based framework based on such principles can serve as a comprehensive yet intuitive way of supporting practitioners in linking tasks to skills and subsequently to competences, experts, and jobs that allow manufacturing firms to make full use of AI. Such a framework is presented in Figure 1. Still, it should be recognised that also other theoretical lenses and bodies of literature confirm and similarly demonstrate this need for task-skill matching in AI adoption and use, e.g. dynamic capabilities or technology adoption of information systems.

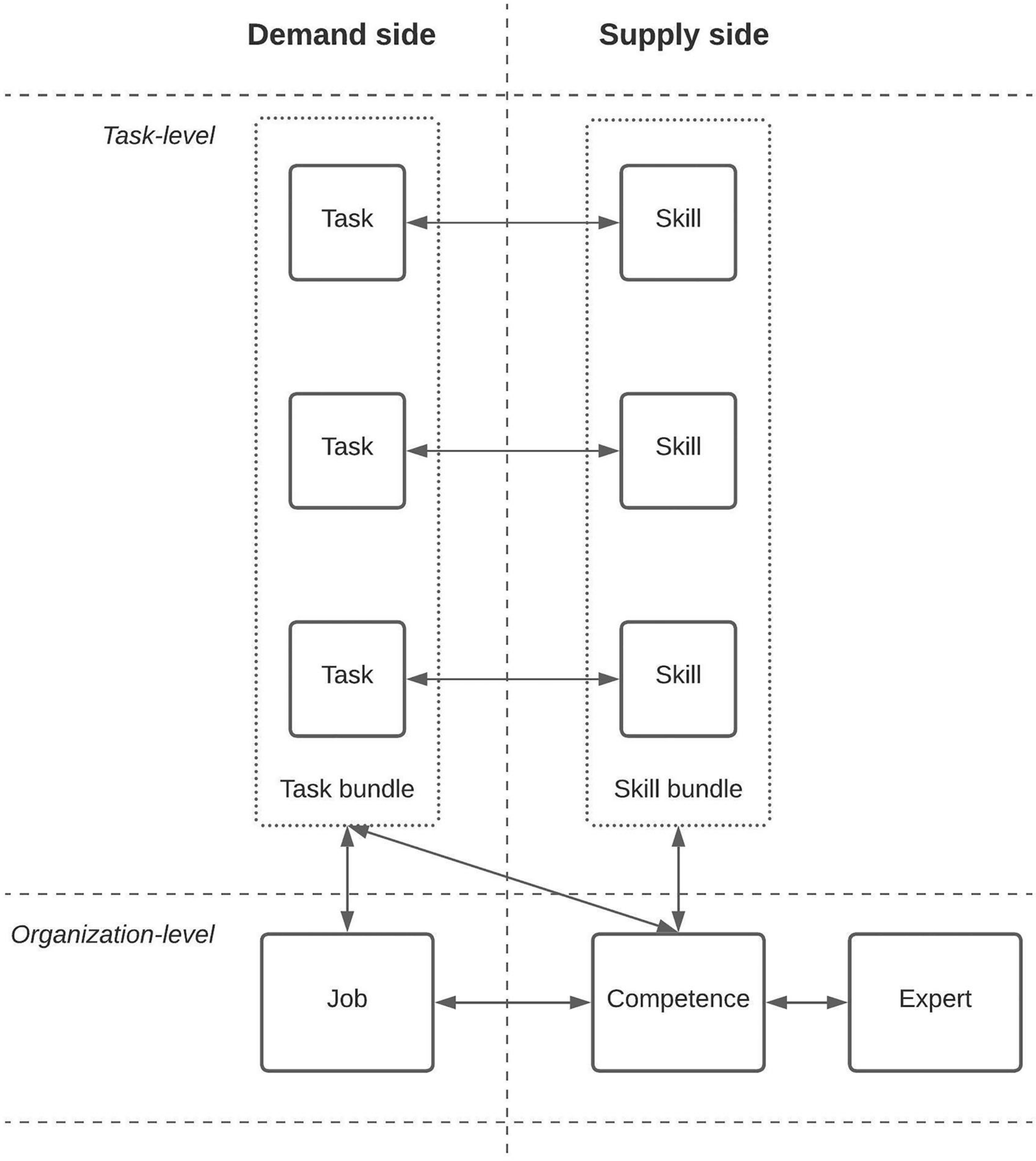

Figure 1. A conceptual task-based framework for matching tasks and skills, Inspired by Rodrigues, Fernández-Macías, and Sostero (2021).

Technological change transforms inputs that are expensive and scarce to become inexpensive and abundant (Benzell and Brynjolfsson 2019). For AI, recent technological advancements have reduced the cost of collecting, storing, and analysing data through more powerful and affordable computing power and scalable ML algorithms (Byrne, Corrado, and Sichel 2018). Thus, foundational AI technology is nowadays widely available to manufacturing firms. This cost reduction leads to AI being used to introduce new tasks and more complex versions of existing tasks (Acemoglu and Restrepo 2018).

Yet, contemporary AI technology has limitations that make humans indispensable, and complementary skills are therefore required for the operation of new tasks. Consequently, workers need to acquire new skills to work on such tasks. But since it takes time to acquire new skills, individuals that know how to leverage AI in manufacturing are not readily available in the labour market (Brynjolfsson, Rock, and Syverson 2021). In essence, as the technology-associated costs of AI fall, the value of complementary skills increases but becomes more scarce. Such an imbalance creates a mismatch between tasks and skills, which acts as a constraint that prevents firms from making full use of AI (Benzell and Brynjolfsson 2019). The dramatic implication is that investments in complementary skills are necessary (critical, essential) to fully realise the value of AI tech. In other words, if the complementary skills are absent, they will act as the weakest link for growth (Brynjolfsson, Rock, and Syverson 2021).

The framework in Figure 1 is grounded in the literature on labour economics more broadly but specifically makes use of the concepts, definitions, and structure in the unified conceptual framework of tasks, skills, and competences in Rodrigues, Fernández-Macías, and Sostero (2021). Figure 1 can be seen as a simplified and contextualised version so as to make it more intuitive and applicable to manufacturing firms. The framework rests on the basic equivalence between tasks and skills and has a quasi-symmetrical structure where tasks and skills cluster into bundles that create jobs and competences. Thus, the framework links concepts at both the demand and supply sides of the labour market.

At the task-level of the framework, we position a set of concepts that substantiates the fundamental relationship between tasks and skills. On the demand side, we find the concepts of *tasks* and *task bundles*. A task is defined as a discrete unit of work activity. A task consists of content together with methods and tools to perform the task (i.e. how and what people do in their work) and thus reflects the smallest unit of labour input. A set of tasks of similar nature create a task bundle, defined as a group of tasks that are clustered together. On the supply side, we have the equivalent concepts of *skills* and *skill bundles*. A skill is defined as the ability to perform a task well, and a skill bundle is defined as a group of skills that are clustered together (Rodrigues, Fernández-Macías, and Sostero 2021).

At the organisational-level of the framework, we position concepts that substantiate how manufacturing firms organise to facilitate a match between tasks and skills. A *job* is defined as a task bundle that is associated with a specific position in an organisation. Tasks are strategically bundled into jobs to maximise the efficiency of workers, which may be influenced by factors such as the firm's business strategy or organisational design. The supply-side equivalent to jobs is *competence,* defined as the ability to do well in a particular task bundle. Here, it is critical to note that while the concept of competence is also related to the concept of a skill bundle, they are distinct. Competence includes more than just a bundle of skills, because 'doing well' also requires knowledge, experience, abilities, and other characteristics of humans. An *expert* is defined as an individual capable of efficiently using the competence required for performing a task bundle. Thus, the concept of an expert links competence to individual employees and adds an efficiency term. That is, an expert can more efficiently leverage the use of competences by also holding a deeper understanding of the technologies, processes, and contexts of a particular domain (e.g. an industry sector or firm-specific business processes) (Rodrigues, Fernández-Macías, and Sostero 2021).

While Figure 1 depicts the framework as static, it is in fact dynamic. Specifically, it allows for understanding AI-skills complementarities by linking it to the implications of technological change on the demand and supply of labour. On the demand side, AI technology can be used to introduce new tasks or more complex versions of existing tasks, which may significantly alter the nature and content of the tasks and task bundles that constitute jobs. On the supply side, this generates a corresponding alteration of skills and skills bundles. This, in turn, may render the competences of existing employees inadequate for doing well in new task bundles, possibly to the point that an individual is no longer an expert for the job. The mismatch between tasks and skills is thus reflected in an imbalance between demand and supply of labour, which results in a bottleneck that prevents firms from realising the benefits of AI tech (Benzell and Brynjolfsson 2019). To re-establish a match between tasks and skills, firms must pursue supply-demand matching strategies by investing in complementary skills (Rodrigues, Fernández-Macías, and Sostero 2021), e.g. supporting existing workers in acquiring new skills through education and training, or hiring new workers that are suitable for the job (Acemoglu and Restrepo 2018). Such complementarities must be re-instated before a manufacturing firm can fully benefit from AI (Brynjolfsson, Jin, and McElheran 2021).

3. Research designs and methods

This study is an analysis of AI in manufacturing over multiple years, grounded on the premise of CRISP-DM as an established engineering methodology yet an incompletely documented phenomenon in practice. Theoretically, this study adopted an orientation towards theory elaboration, which is the process of using pre-existing concepts and models (in this article: CRISP-DM) to collect and organise empirical data for developing new insights and refining existing theory (in this article: AI in manufacturing) (Fisher and Aguinis 2017). Empirically, this study consisted of the execution of three phases ('research', 'transition', and 'practice') spanning a total of six years (2016–2022). The research process emphasised the significance of capturing insights over time to refine and elaborate the understanding of how manufacturing firms can use CRISP-DM to realise value from AI. An overview of the full research process is illustrated in Figure 2. In the forthcoming sections, we provide further details about the methods used throughout the three phases.

3.1. Research phase (2016–2020)

The research phase spanned a total of four years and consisted of collaborative research studies together with three automotive manufacturing firms in Sweden (i.e. co-creating AI solutions using CRISP-DM). The joint goal of all studies was to develop AI algorithms for throughput bottleneck analysis in production systems. The choice of using AI was motivated by the nature of the bottleneck problem, which is characterised by high frequency (i.e. bottlenecks can shift at the time scale of seconds), high impact (i.e. bottlenecks directly impact the throughput of the production line), high complexity (i.e. bottlenecks are determined by interactions between multiple system entities), and high speed in decision-making (i.e. lack of time for manual analysis and requirements on close to real-time decision-support).

The object of study was discrete-part production systems, consisting of a mix of engine machining lines with fully automated CNC machines and fully automated car body welding lines. The layouts included parallel machines with intermediate buffers, serial lines with intermediate buffers, and serial lines with no intermediate buffers. Digital event log data were collected at the machine level. Event log data consists of events (activity information performed by the machine or on the machine) and timestamps (temporal information of the events). A total of six different solutions for bottleneck analysis were developed during the research phase (Solution A-F) (Subramaniyan et al. 2016; Subramaniyan et al. 2019;

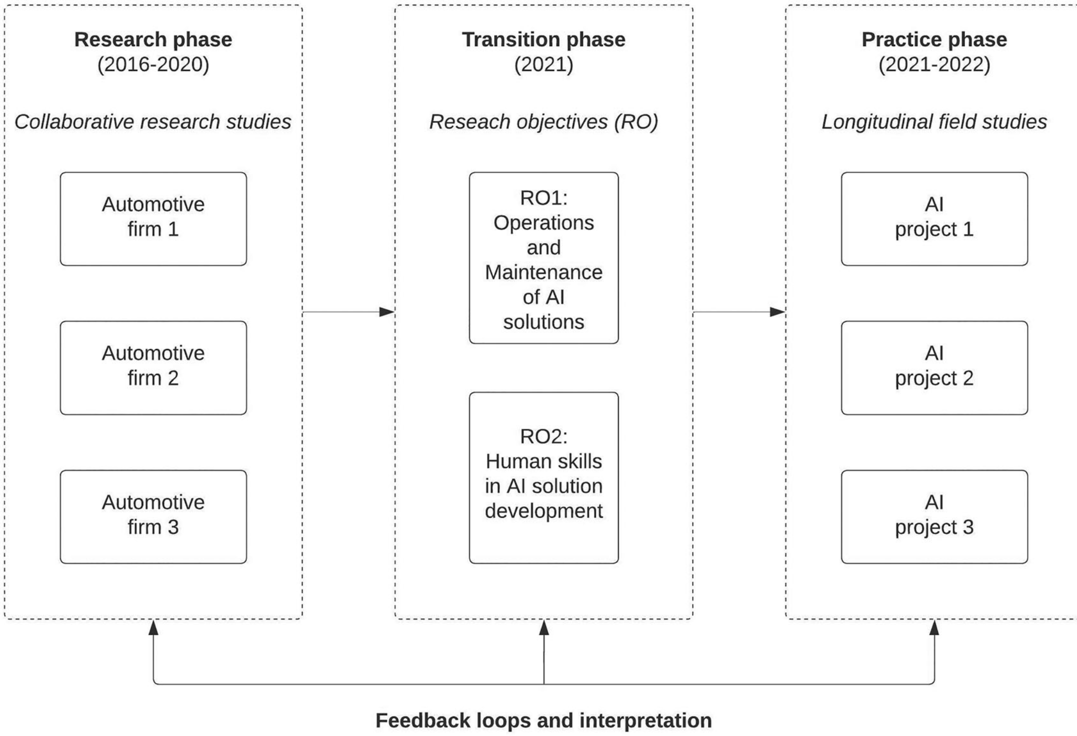

Figure 2. Research process in three phases.

Table 1. Overview of the research phase.

Research studies	Solution A	Solution B	Solution C	Solution D	Solution E	Solution F
Bottleneck focus	Detect historical long-term bottlenecks	Detect historical long-term bottlenecks	Detect historical short-term bottlenecks	Diagnosis of historical long-term bottlenecks	Prediction of future bottlenecks	Prediction of root causes for future bottlenecks
Firm no.	1, 2, 3	1, 3	1	2	3	3
Model	Descriptive and inferential statistics	Dynamic time wrapping, agglomerative hierarchical clustering	Matrix operations	Descriptive statistics, k-means clustering, visual analytics	Time-series methods, inferential statistics	Time-series, rule-based prescriptive analytics
Output	Set of historical bottleneck machines	Set of historical bottleneck machines	Current bottleneck machines at any time instant	Clusters of different unplanned stops based on their features	Set of future bottleneck machines	Set of root causes and set of prescribed actions for future bottlenecks

Subramaniyan et al. 2020b, 2020a; Subramaniyan, Skoogh, Salomonsson, Bangalore, and Bokrantz 2018; Subramaniyan, Skoogh, Salomonsson, Bangalore, Gopalakrishnan, et al. 2018). Aligned with the 'toolbox' interpretation of AI (see footnote 1), the solutions used a variety of AI techniques for perceiving, processing, learning, and acting from data, ranging from statistics to complex learning. All six solutions were co-created with industry practitioners and explicitly followed the original six-phase version of CRISP-DM. An overview of the research phase is provided in Table 1.

3.2. Transition phase (2021)

The transition phase consisted of approximately six months and started with summarising and documenting the accumulated experiences and learnings from the research phase. During this phase, the research team executed a set of reflection sessions that consisted of reviewing each of the six AI solutions (A-F in Table 1) in three steps: (1) discussing the underlying success factors for the outcomes of the research study, (2) identifying challenges for transferring the specific academic results to industry practice, and (3) abstracting the specific study to generic challenges for making AI solutions successfully operational in manufacturing. A total of four sessions with a duration of 60–120 minutes each were sufficient for analysing the six solutions and formulating two objectives to serve as the focus for the practice phase: (1) operations and maintenance of AI solutions over time and (2) human skills in AI solution development. These objectives focused on using CRISP-DM as a pre-existing model to guide the theoretical elaboration of AI in manufacturing.

3.3. Practice phase (2021–2022)

The practice phase was executed as longitudinal field studies that took inspiration from the clinical research approach

(Karlsson 2010; Schein 1987). This type of inquiry was chosen because of its relevance and ability for deep and rich insights into how a generic phenomenon (i.e. AI) manifests in a certain context (i.e. manufacturing) by studying problems, people, and organisations in real settings (Edmondson and McManus 2007). Further, since academic AI solutions rarely progress to or beyond the deployment stage (Schröer, Kruse, and Gómez 2021), longitudinal and clinical field studies were deemed especially relevant. This approach also allowed for building continuity in the study by maintaining the structure of the core research team throughout the research, transition, and practice phase.

The longitudinal studies were based on sustained participation across three AI projects within a single automotive firm. The firm was relevant because it applies to the theoretical domain of the study (i.e. automotive manufacturing) and it was also directly connected to the firms in the research phase. The automotive sector is also particularly rich in data related to the issue of AI in manufacturing, owing to its advanced manufacturing practices and competitive business environment. To deal with the limitations of drawing inferences from a single firm, our clinical analysis emphasises the generalisation of the issues being studied rather than the generalisation of the specific observations to the population (Karlsson 2010).

In the three projects, one member of the research team adopted a fully professional role with the responsibility of consulting and solving the client firm's needs, thus serving the role of a relative insider with an inside-out perspective (i.e. observing the problem and developing the solution). The remaining research team served the role of relative outsiders with an outside-in perspective (i.e. observing the solution and providing explanations). The continuous dialogue and exchange of information among the research team allowed for concurrently providing information to the client organisation and receiving information for the research aimed at developing academic knowledge (Karlsson 2010).

In contrast to the research studies (Table 1), it was the client organisation that was the problem owner for the field studies, and the firm representatives chose the projects with the greatest value potential. This enables a high degree and involvement in organisational processes (Åhlström and Karlsson 1996) and provided unique opportunities for observing more intricate aspects of the CRISP-DM process. By making observations from within the organisation, the clinical approach overcomes the issues with access that is denied by other approaches (e.g. case studies or surveys) (Sköld and Karlsson 2013). Note that clinical research has subtle but important differences from action research such as the problem owner (company representative vs. researchers) and the data analysis procedures (data feedback and sharing vs. acting and reflecting) (Karlsson 2010).

Specifically, by collaborating and interacting with multiple experts who jointly developed AI solutions using CRISP-DM for extended periods of time, the researcher was able to become close to the organisation and investigate critical issues while they are taking place (Sköld and Karlsson 2007), thereby discovering new insights. The total time period for the three projects was approximately one year, encompassing problem framing, planning, and organisational interactions as well as 12 weeks of formal execution and implementation for each project, respectively. Similar to the research phase, the projects used a variety of AI techniques for perceiving, processing, learning, and acting from data (see footnote 1), ranging from statistics to complex learning. The projects are summarised in Table 2.

While a major advantage of the clinical approach is the ability to focus on the client's needs, the drawback is the restriction on client-researcher confidential data. To extract knowledge from the field studies, the research team met at regular intervals in data feedback sessions (Sköld and Karlsson 2013). The goal of the sessions was to analyse and organise critical incidents in the three projects into topics and categories and make comparisons with existing literature. The individual sessions followed a four-step protocol: (1) the relative insider reported the latest critical incidents and the relative outsiders probed for depth and details in the descriptions, (2) joint discussions about theoretical interpretations, (3) the relative outsiders reported the latest comparisons with literature, and (4) scheduling of the next feedback session. Between each session, the relative insider worked on the projects and the relative outsiders made continuous comparisons with existing theory (see section 3.4). A total of 18 data feedback sessions were conducted during the practice phase, lasting between 60–90 minutes each, with the data being recorded as a research diary.

3.4. Feedback loops and interpretation

Owing to the orientation towards theory elaboration together with the multi-year and multi-method research approach, the findings were synthesised by continuously aggregating the insights and comparing them against existing knowledge (Miles and Huberman 1994). Specifically, the logic and structure of the feedback sessions follow the premises of clinical research where analysis and interpretation are intertwined within the field studies and the feedback process. Thus, scientific results are not developed in a linear fashion from raw data to final interpretation but rather iteratively and non-linearly in parallel processes to

Table 2. Overview of the practice phase.

Field studies	AI project 1	AI project 2	AI project 3
Problem domain	Outbound logistics	Outbound logistics	Predictive maintenance
Focus	Detection of delivery lead time discrepancies	Prediction of carrier route behaviour and delivery lead times	Diagnosis of machine failure root causes
Output	Set of deviations between contracted and actual lead times	Set of predicted carrier lead times with automatic tagging of risky deliveries	Sets of failure modes from upstream/downstream production flow or the equipment itself

the researcher helping the client (Karlsson 2010). That is, the feedback loops from the practice phase (Figure 2) focused on researchers' theoretical elaborations with the existing models, literature, and insights from the research and transition phases as lenses for interpretation (Karlsson 2010).

The goal of this elaboration was to achieve deeper and richer insight into the practical challenges of using CRISP-DM for realising the value of AI in manufacturing and prescribing solutions for overcoming them. Consequently, the findings from the field studies were theoretically validated in relation to existing knowledge (Sköld and Karlsson 2007), and contributions to theory were analysed by comparing emerging insights to the literature (Fisher and Aguinis 2017). Thus, all results presented in the Findings section (section 4), including the AI lifecycle in Figure 3, the tasks and skills in Table 3, as well as the competences, trajectories, and integration in Figures 4–6, and figures, were derived from the data collection, feedback loops, and theoretical interpretations in the research, transition, and practice phase. The output of our research approach is a more elaborate and accurate understanding of AI in manufacturing as well as a refined and extended CRISP-DM methodology that can be used in practice, research, and pedagogy. Thus, the outcomes of the study contribute both to the advancement of scientific knowledge as well as solve the practical concerns of the automotive firms involved.

4. Findings: Enhancing CRISP-DM

The next sections outline our accumulated findings from six years of research and practice on CRISP-DM. We present the results in the form of enhancements to CRISP-DM in line with our two central complementarities: operation and maintenance over time, and human skills. First, we position CRISP-DM as an AI solution life-cycle and describe how manufacturing firms should use it as a day-to-day engineering methodology. We then add and explain the phase of 'Operation and Maintenance' and explain the unique tasks that are needed to operate and maintain AI solutions in manufacturing environments over time. Second, we embed the task-based AI-skills framework into CRISP-DM and describe how it can be used to link tasks to skills and ultimately integrate competences held by sets of experts that allow manufacturing firms to make full use of AI solutions. The tasks and skills are formulated using the Bloom taxonomy (Anderson and Krathwohl 2001). This description is presented in two parts, focusing on the task-level and the organisational-level of the framework, respectively.

4.1. AI solution life-cycle

Figure 3 displays our elaboration of CRISP-DM as encompassing a complete life-cycle of AI solutions – ranging from the initial idea of a practical problem to a fully-functioning solution in everyday operations. Our interpretation depicts CRISP-DM as consisting of seven phases that are logically and practically intertwined: Business understanding, Data understanding, Data preparation, Modelling, Evaluation, Deployment, and Operation and Maintenance. Since the basic structure and content of the six phases remain intact from the original version, we focus on three key messages in our enhanced version of CRISP-DM: (1) using CRISP-DM as a continuous, active, and iterative way of working, (2) the need for an initial conceptual design, and (3) adding 'Operation and Maintenance' as the seventh phase.

Our first key message is that CRISP-DM is not a waterfall methodology that starts with business understanding, and progresses sequentially to data understanding, etcetera. Rather the phases can be executed in a myriad of different sequences or iterative configurations depending on the particular context. This is shown in Figure 3 in the form of a cyclical illustration. Although the execution of specific tasks within each phase may be temporally separated, all phases should be considered constantly and simultaneously. Thus, we see our version of CRISP-DM as a continuous, active, and iterative way of working. Interpreting CRISP-DM in this manner has important implications for how the phases are executed. For example, in one-off AI projects, the business understanding phase typically solely focuses on framing the technical problem to be solved. In a real-world manufacturing organisation, this phase needs to focus on understanding the business problem from multiple lenses such as the nature of the decision-making task (e.g. frequency, time, accuracy), value potentials (e.g. impact in terms of economic, environmental, or social benefits), and human-machine interactions (e.g. ethics, trust, and engagement). In fact, it is critical to even explicitly question whether AI is the right approach to solve the problem in the first place. Furthermore, working continuously with data understanding and preparation implies not only finding specific data sets for a certain model but also systematically developing the infrastructure to enable data to be used for a diversity of AI solutions. Whereas the evaluation phase typically solely focuses on model performance (e.g. precision or recall), manufacturing firms also need to evaluate AI solutions in terms of computational resources (e.g. deep neural nets require much more computational power compared to support vector machines). Similarly, the deployment phase needs to ensure that models effectively scale with the data and that the technical infrastructure facilitates AI solutions to not only continuously learn from new data but also to respond to feedback and adapt.

Our second key message is the need to consider the entire AI solution life-cycle and all of the seven phases even in the early phases of design. Like all design endeavours, the degrees of freedom are highest at the start but decrease with time, whereas the cost of making changes has the opposite trajectory. The final characteristics and performance of an operational AI solution are path-dependent on the design choices that are made at an early stage. Thus, the best opportunity to manage trade-offs in the solution design

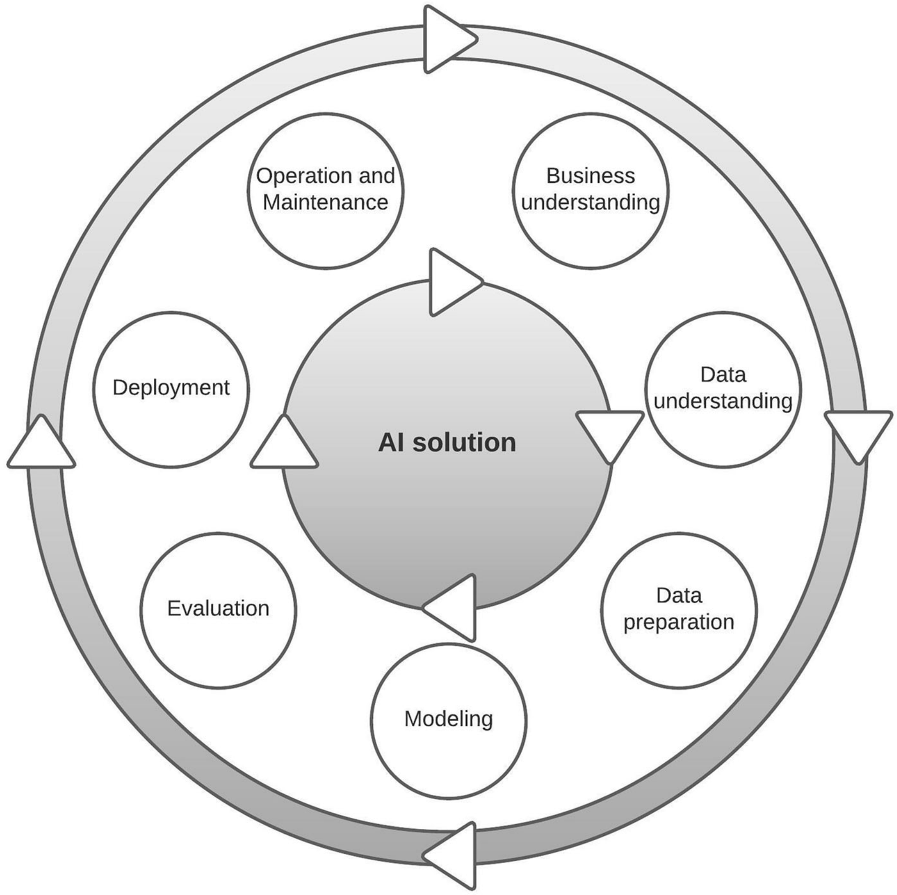

Figure 3. AI solution life-cycle.

is as early as possible. We therefore recommend manufacturing firms to start by creating an *initial conceptual design* of the entire AI solution. That is, start by logically reflecting on all of the seven phases in Figure 3 and ideally make a simple yet holistic drawing of the complete AI solution. This drawing should list and sketch all major design choices throughout the seven phases, even if it is only possible to do so roughly and imperfectly. Then, and only then, we recommend proceeding by formally executing the core tasks of the methodology (Table 3). A specific reason why we recommend this initial conceptual design is the fact that the hidden costs of operation and maintenance may completely alter the structural design of the AI solutions. For example, minor design choices in feature selection and modelling might come with a massive increase in operational system complexity. Even tiny changes such as adding one or two small data dependencies may render the final AI solution unfeasible, impractical, or overly costly. Thus, if these hidden costs are neglected during the early design and development of AI solutions, value creation will be severely constrained.

Our third key message is the addition of Operation and Maintenance as the seventh phase in CRISP-DM. After an AI solution has been successfully deployed, it is not automatically time to move on to the next challenge. In contrast, the AI life-cycle continues long after deployment. To ensure stable performance over time, AI solutions need to be carefully operated and maintained. That is, the AI solution needs to be tuned, updated, and sometimes even re-designed to meet performance targets in the long run. Failing to perform this phase is likely to result in that AI performance deteriorates over time, which may lead to inaccurate insights and false conclusions that negatively impact the production process. The underlying reason is that AI models are probabilistic and work under a set of given assumptions. These assumptions stem from the data distributions or the production processes that are encountered during the training of the model (in the Modelling phase). In other words, when AI solutions are put into continuous operations in manufacturing environments, the future data will not look like the past data used to build the AI model. This leads to the inevitable risk of *AI drift*.

AI drift consists of two forms: data drift and concept drift. Data drift occurs when training data used in the modelling phase no longer adequately represent the states of reality, which can occur for two broad reasons: (1) stochastic changes in production system behaviour (e.g. machine degradation or dynamic changes in machine interactions) that alter the stationarity and distribution of the data, or (2) data corruption (e.g. erroneous sensor output or errors incurred through data pipeline transfer) that introduces faulty data points. Concept drift occurs when the underlying dynamics of the production process are changed, e.g. introducing new products, installing new machines, or altering the production flow. Over time, AI drift may lead to that AI performance deteriorates, sometimes

Table 3. CRISP-DM task-level framework.

Phase	Generic tasks	Generic skills
Business understanding	Deciding the objective function of the solution	Assessing factory-level objectives and detecting value potentials in manufacturing
	Selecting the target manufacturing problem	Analysing value potentials and articulating connections to objectives
	Measuring the baseline performance of the manufacturing process	Arguing for suitable performance indicators and calculating baselines
	Deciding the level of desired process improvement for success	Linking management goals to factory-level goals
	Determining whether AI is the right approach for the problem	Explaining and presenting decision-making structures for manufacturing problems
	Formalising the AI solution in terms of its intended features and use	Assessing current organisational practices for manufacturing problems; formulating and reviewing requirement specifications for AI solutions
Data understanding	Deciding the data processing approach based on the target manufacturing problem (e.g. batch or stream)	Assessing data processing types for manufacturing problems
	Procuring the data by establishing data pipelines from different sources	Determining suitable data types and data sources for manufacturing problems
	Storing the data in a relevant place (e.g. database or data lake)	Executing data migration and storage procedures
	Establishing data pipeline from data storage to AI workbench	Designing and creating data pipelines
Data preparation	Selecting the features of the data based on the target manufacturing problem	Comparing methods to solve manufacturing problems and converting their characteristics into model features
	Preparing and labelling the data according to the target manufacturing problem	Interpreting meta-data and categorising labels
	Evaluating data quality and formalising improvement actions	Identifying, adapting, and reviewing data quality measures for manufacturing problems
	Deciding the sampling strategy of the data for AI training and testing	Evaluating strategies for model training and testing
	Formalising conceptual designs of different AI solutions to be trained and tested	Designing viable conceptual AI solutions for manufacturing problems
	Establishing performance metrics for the AI solution based on the target manufacturing problem	Comparing and evaluating the appropriateness of AI evaluation metrics for manufacturing problems
Modelling	Training different AI solutions on the training data sample	Assessing viable AI solutions and adapting model parameters to manufacturing problems
	Testing different AI solutions on the testing data sample	Reviewing knowledge from training data and implementing models to unseen testing data
Evaluation	Conducting internal evaluation of the different AI solutions based on the AI performance metrics	Interpreting AI performance and distinguishing adaptations to model parameters
	Conducting external evaluation of the different AI solutions relative to the baseline performance of the manufacturing process	Designing and implementing experimental procedures for manufacturing problems; calculating and interpreting AI evaluation metrics relative to current organisational practices
	Evaluating trade-offs based on internal and external evaluation and choosing the final AI solution	Linking knowledge from AI solution evaluation to the objective function; interpreting AI evaluation results and comparing the performance of viable AI solutions
Deployment	Establishing data pipelines from AI workbench to dashboards for visualisation of AI insights	Designing, creating, and optimising data pipelines
	Validating the AI solution in its intended environment using offline testing	Assessing the usability of AI solutions from multiple stakeholders
	Tuning the AI solution parameters (if necessary)	Assessing user feedback and determining actions for AI solution optimisation
	Embedding AI solution in live operating environment at scale	Designing, building, and implementing scalable User Experience designs for AI solutions
	Optimising the AI solution for insights consumption from end-users	Assessing user feedback on AI solutions from multiple stakeholders
Operation and Maintenance	Continuously consuming AI insights and taking relevant actions to improve the target manufacturing problem	Interpreting AI insights and directing actions to manufacturing problems
	Continuously collecting user feedback and taking relevant actions on the AI solution (e.g. human-in-the-loop)	Evaluating actions taken towards manufacturing problems and directing AI solution improvements
	Establishing metrics to monitor AI performance in the operating environment	Proposing continuous AI metrics for manufacturing problems
	Deciding on the frequency of monitoring of data and concept drift	Creating and implementing management practices for data and concept drift
	Continuously taking actions for mitigating data and concept drift and ensuring AI solution performance over time	Distinguishing data and concept drift and articulating changes to AI solution design

even to the point that the AI solution no longer holds adequate predictive power for interpreting unfamiliar data.

AI drift can be addressed by regularly monitoring AI performance, e.g. deterioration of predictive accuracy. Various diagnosis methods can be used to diagnose drift such as sequential analysis method (e.g. systematic analysis of the data flow through pipelines), time distribution-based methods (e.g. randomly testing the distribution of the data at different time steps using statistical methods), and automated checks to detect the corruption of the data at source (e.g.

automatically detect if the sensor data is missing or out of range). If drift is detected, multiple options can be pursued to restore the performance of the model. For example, in instances of minor drift, one-shot retraining on new data (e.g. using the latest time window of the data) or incremental retraining on new data (e.g. collecting the data from the recent time window after drift has been detected and appending it to the training data set) may be sufficient to ensure that the model parameters adequately reflect the new data distributions. However, instances of major drift are often complicated by a lack of sufficient data that reflects the updated production system behaviour, requiring the AI solution to be partly re-designed (e.g. creating ensemble models for increased robustness) or even fully re-designed (i.e. essentially bringing the AI solution life cycle back to square one).

Thus, carefully operating and maintaining AI solutions over time is necessary to prevent AI drift and consistently ensure a desirable level of AI performance. This phase thereby encompasses the use of a variety of maintenance policies to ensure the functioning of the AI solution and its supporting infrastructure (Khazraei and Deuse 2011). For example, policies corresponding to time-based and/or condition-based maintenance could be used to maintain the pipelines that supply AI solutions with data (e.g. machine to cloud interfaces), and design-out maintenance could be used to eliminate recurrent AI drift by re-designing the commissioned, in-service AI solution.

4.2. Task-level framework

In this section, we embed the task-based AI-skills framework (Figure 1) into our enhanced CRISP-DM version (Figure 3) to create a holistic and practical model that supports manufacturing firms in making full use of AI. We present the task-level framework in a schematic and tabular format to substantiate the relationship between tasks and skills (Table 3). On the left side of the table, we illustrate the demand-side in the form of the CRISP-DM phases. Each of the seven CRISP-DM phases represents a task bundle. On the right side of the table, we illustrate the supply-side in the form of the corresponding skill bundle for each phase, respectively. This framework thereby provides a comprehensive yet intuitive way of linking tasks to skills.

In Table 3, we list a set of generic tasks and skills that we consider indispensable for developing AI solutions in manufacturing. That is, while the tasks and skills are described in a generic manner, they are indeed specific to AI solution development for manufacturing scenarios. Note that Table 3 is not intended to be an exhaustive list but rather aims to cover the most important and stable tasks and skills that apply to most manufacturing contexts. Thus, individual manufacturing firms may complement their use of the model by also listing specialised tasks and skills together with documentation of process instances that apply specifically to their firm's context. The tasks and skills in Table 3 were derived from the data collection, feedback loops, and theoretical interpretations in the research, transition, and practice phase (see sections 3.1 to 3.4).

4.3. Organisational-level framework

In this section, we describe the organisational-level of the framework by highlighting how the tasks and skills in Table 2 emerge into competences, experts, and jobs. The findings in this section (including Figures 4–6) were derived from using the task-based framework in Figure 1 as a theoretical lens for interpreting the research, transition, and practice phase in the feedback loops (see sections 3.1 to 3.4). From the tasks and skills across the seven CRISP-DM phases in Table 3, three distinct competences emerge: *domain, data science*, and *data engineering*.

Domain competence is specific to the tasks of the problem domain and accumulated through prior learning from working within that domain. For example, it may consist of knowledge of production flows, previous experience with industry best practices, and an in-depth understanding of manufacturing process technologies. This competence is critical for identifying and formulating manufacturing problems, finding relevant data, providing input to AI solution development from a user perspective, and taking actions based on algorithmic insights during operations.

Data science competence is specific to the tasks of extracting knowledge from data and accumulated through prior learning from using AI technologies. For example, it may consist of deep knowledge of ML algorithms and data wrangling techniques, programming in AI software, and an understanding of evaluation strategies and metrics for AI performance. This competence is critical for designing, developing, and validating AI solutions that serve the needs of users.

Data engineering competence is specific to the tasks of building infrastructure that facilitates the collection, curation, consumption, and control of data. For example, it may consist of knowledge of IT architectures, an understanding of database systems, and experience in establishing efficient data pipelines. This competence is critical for developing the technical infrastructure that enables access to valuable data sources as well as facilitates final AI solutions to be effectively used in operational settings.

On the left side of Figure 4, the three types of competences are illustrated as distinct entities that encapsulate sets of matching task and skill bundles. On the right side of Figure 3, the overlapping area between the three competences represents the shared entity that encompasses the core competences that are needed for developing AI solutions in manufacturing. We considered these three competences as non-substitutable for developing AI solutions in manufacturing. That is, each competence is necessary but not sufficient for AI success. This implies that the absence of any of the competences acts as a bottleneck that guarantees failure, but that the presence of all these competences does not guarantee success because also other competences play a role. Thus, individual

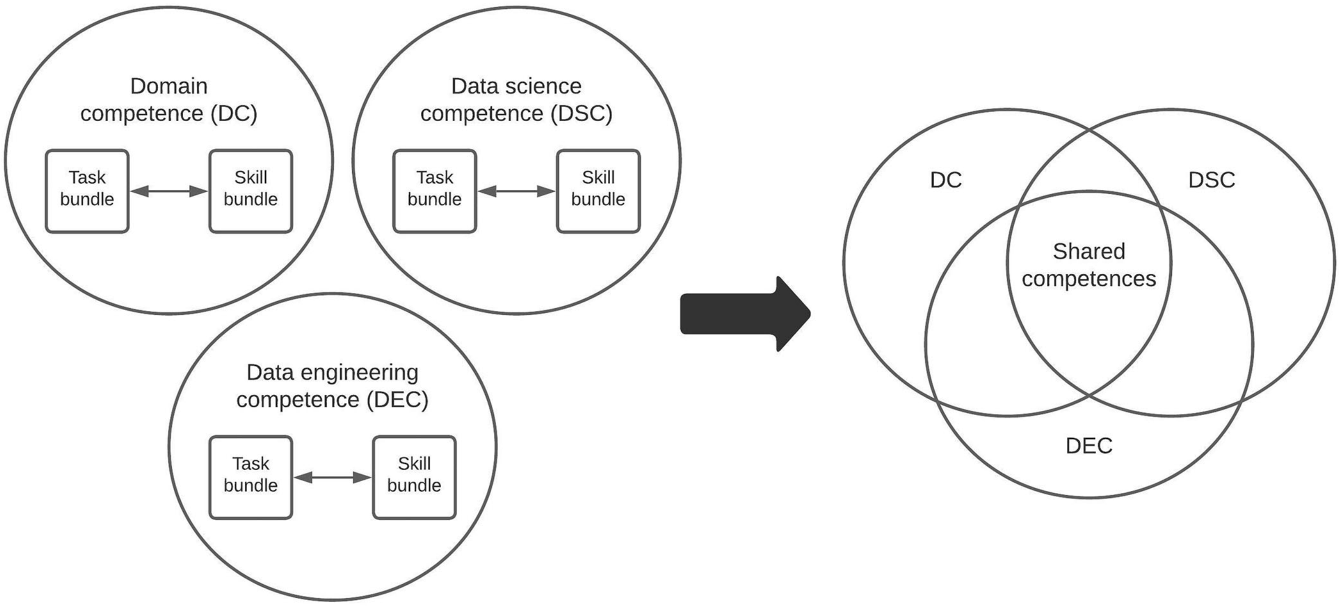

Figure 4. Domain, data science, and data engineering competence.

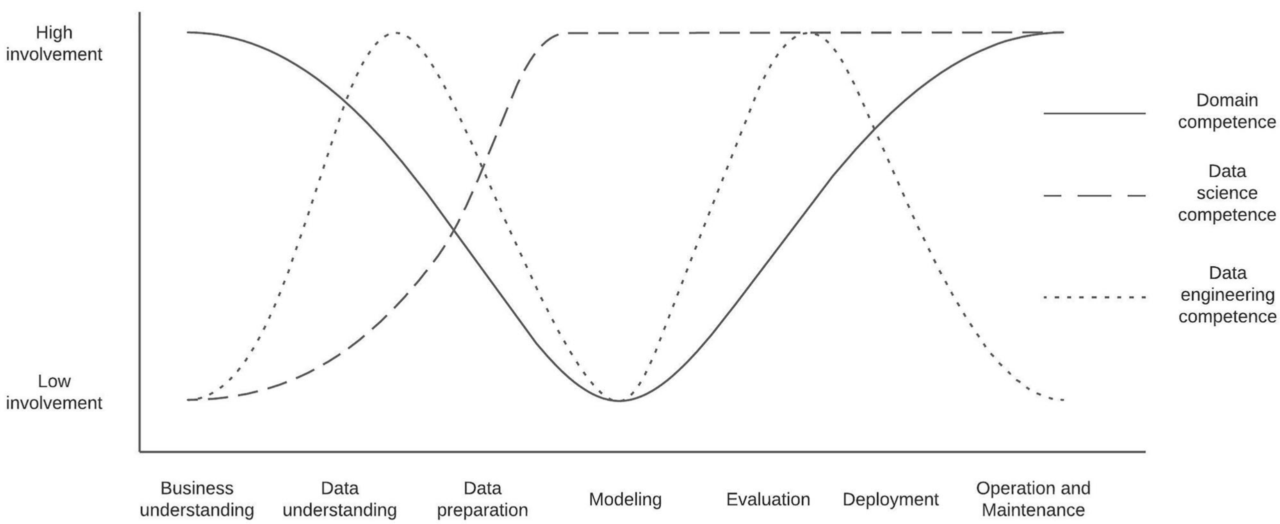

Figure 5. Trajectories of involvement from the domain, data science, and data engineering competence throughout the seven CRISP-DM phases.

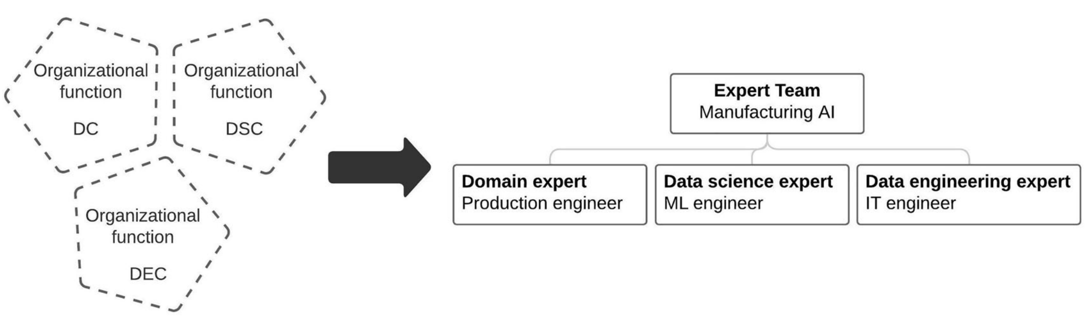

Figure 6. Cross-functional integration of experts.

manufacturing firms may also include other competences that are beneficial to developing AI solutions in their context.

All three competences must be involved throughout the CRISP-DM process. However, the degree of involvement among the competences varies across the different phases, resulting in that the competences interact differently with each other over time. In Figure 5, we schematically illustrate the trajectories of involvement for the domain-, data science-, and data engineering competence. The involvement of domain competence follows a u-shaped curve; starting at a high level, reducing in the middle, and returning to high levels.

This reflects how domain competence is central for identifying and framing the problem as well as interpreting and taking action on AI insights during operations. The involvement of data science competence follows a logistic growth curve; starting at a low level, increasing rapidly, and maintaining high throughout the process. This reflects how data science competence is central to the core design and development of the AI solution and making sure it consistently works as intended during operations. The involvement of data engineering follows a bimodal curve that peaks twice during the process. This reflects how data engineering comes into place when creating the data pipelines for the AI solution and facilitating its deployment and implementation in operational settings.

The three competences are embodied by experts, which are individuals that hold jobs tied to a manufacturing firm's organisation. In most cases, each unique competence is held by different individuals, but in special cases, all competences can also be held by a single individual. The key for manufacturing firms is therefore to ensure the presence of domain experts, data science experts, and data engineering experts. These individuals are typically tied to specific positions within distinct organisational functions (left side of Figure 6). For example, domain experts can typically be production or maintenance engineers working in the operations function, data science experts can be ML engineers working in the data analytics function, and data engineering experts can be information systems engineers working in the IT function. Consequently, manufacturing firms are advised to establish cross-functional integration of these experts. With integration, we mean that distinct and interdependent experts constitute a unified whole by coordinating, communicating, and collaborating throughout the CRISP-DM process. While we show this as an internal-functional relationship on the right side of Figure 5, it could also be external-functional or a mix of both. Cross-functional integration can be achieved through a variety of integrative mechanisms, such as formalisation and cross-unit structures. In Figure 6, we exemplify the use of multi-disciplinary teams as the integrative mechanism.

The integration of experts into cross-functional teams also allows for addressing conflicting or contradictory objectives and interpretations. For example, data scientists might strive to maximise model precision whereas domain experts might satisfy with false positives in favour of model simplicity, and data engineers might prefer flexibility in data structure or type to facilitate data flows through pipelines whereas data scientists often seek to reduce risks by working with pre-defined data formats. Close communication, collaboration, and coordination between experts allow for resolving conflicts through mutual agreement, and introducing additional liaison roles into the team can further strengthen expert integration.

5. Illustration: CRISP-DM for data-driven bottleneck detection

To showcase the implications of our enhanced CRISP-DM methodology, we now provide an illustration of our empirical findings. The intention is to provide a broadly accessible illustration that helps manufacturing firms in realising the value of AI in their own organisations. The illustration consists of a post-hoc application of the enhanced CRISP-DM methodology to one of the studies in the research phase (Figure 2 and Table 1): the generic hierarchical clustering approach for detecting bottlenecks in manufacturing (Subramaniyan et al. 2020b).

This approach was chosen not to highlight specific prior work in the research phase, but for the following main reasons. Due to the nature of the client-researcher confidentiality of the field studies (see section 3.3), using a published article as the focal case for illustration allows us to provide a more detailed and elaborate description that supports manufacturing practitioners. Further, in-depth information about the tasks, skills, and other nuances required to develop the complete AI solution remains hidden in the original article. But as members of the research team that conducted the study and published the original article, we are able to synthesise the original research with the empirical findings from the field studies to elaborate on these hidden aspects of the CRISP-DM process. Note that we do not intend to provide a fully exhaustive demonstration of every detail, rather the goal is to illustrate how the enhanced CRISP-DM methodology can be used in manufacturing practice.

We present the illustration in four steps. In section 5.1, we describe the empirical setting of the study and outline the initial conceptual design of the AI solution. In section 5.2, we describe in tabular format the tasks and skills needed to execute all seven phases of the CRISP-DM methodology. In section 5.3, we show how the team of domain, data science, and data engineering experts was formed and organised throughout the project. In section 5.4, we explain how the competences are related to the technical details of the AI solution architecture.

5.1. Empirical setting and initial conceptual design

The empirical setting was an automotive manufacturing firm in Sweden and a production system consisting of 13 Computer Numerical Controlled (CNC) machines in a serial line flow (see section 3.1). Event log data was collected from a Manufacturing Execution System (MES). The manufacturing firm was interested in detecting bottleneck machines in the production line. Consequently, the study aimed to develop an AI solution for detecting historical long-term bottleneck machines in the production system. While there are various methods for identifying bottlenecks, many of them have limited applicability in practice due to difficulties in managing data from time-varying dynamic systems and underlying statistical assumptions (e.g. statistical distribution, autocorrelation, cross-correlation, and stationarity). The lead researcher realised that unsupervised ML techniques such as hierarchical clustering of time series could overcome these limitations. Hence, the project started with the creation of an initial conceptual design of a six-step AI solution by reflecting on the complete AI life-cycle (see Figure 7). Based on this initial design, the team was ready to proceed by executing the formal phases of the CRISP-DM methodology. For additional details regarding data collection, selection of bottleneck detection method, data pre-processing, algorithm

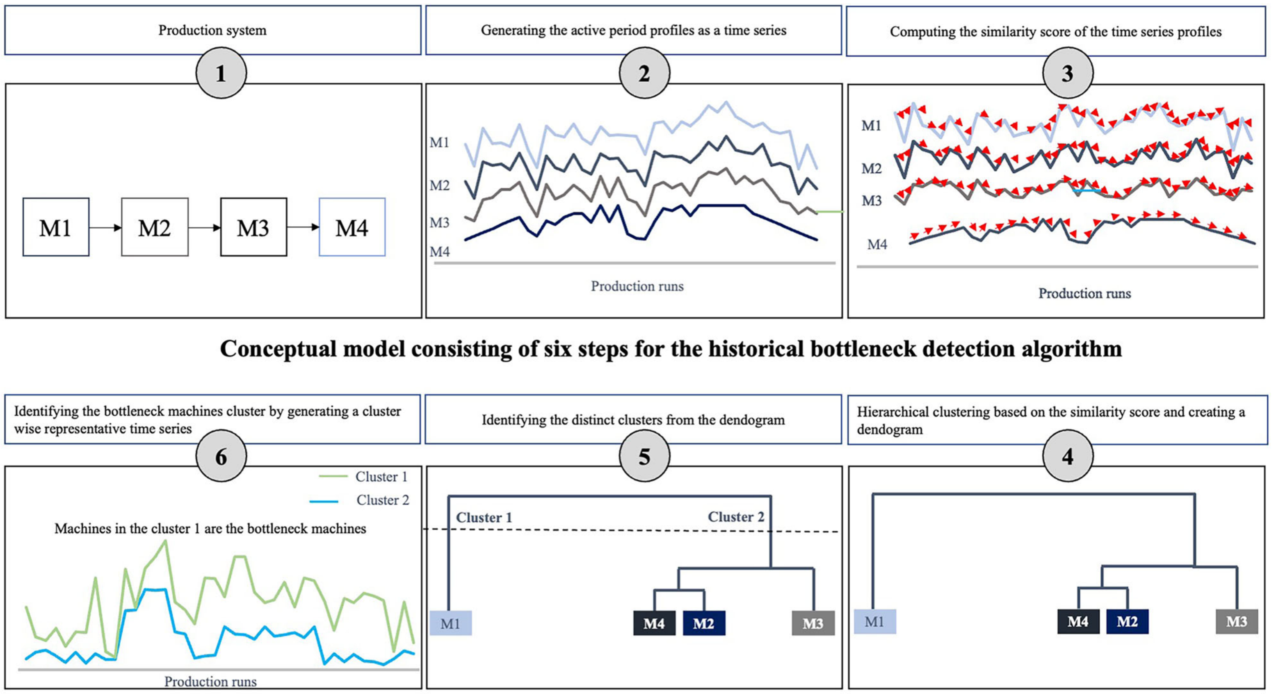

Figure 7. Initial conceptual design of the AI solution for bottleneck detection.

architecture, and cluster interpretation, see pages 146–154 in Subramaniyan et al. (2020b).

5.2. Tasks and skills for historical bottleneck detection

In Table 4, we outline the tasks and skills needed throughout the seven CRISP-DM phases for developing an AI solution capable of detecting historical throughput bottlenecks using event log data. For the illustrative demonstration, the task and skills in Table 4 were derived from the original article and the authors' own post-hoc reflections, and they were assessed using the empirical findings of the present study (section 4) as a theoretical lens.

5.3. Competences and expert team for bottleneck detection

The project team consisted of nine core individuals that jointly possessed the shared domain, data science, and data engineering competence: six researchers from two universities specialising in industrial engineering, computer science, and data management, and three engineers from the manufacturing firm specialising in production and maintenance engineering as well as information systems. The integrative mechanism used was the formation of a dedicated team for the duration of the project. The structure and members of the team are illustrated in Figure 8.

5.4. Relationships between AI solution architecture and competences

The expert team (Figure 8) encompassed the shared domain, data science, and data engineering competence that enabled a match between the task bundles and skill bundles throughout the CRISP-DM process (Table 4). To further illustrate the crucial role of all three competences, we now exemplify relationships between the technical AI solution architecture and the competences. The AI solution architecture for bottleneck detection proposed in Subramaniyan et al. (2020b) consists of first collecting the event log data files of the different machines followed by generating the active period time series profile for each machine. The algorithm then uses Dynamic Time Wrapping (DTW) to compare the similarity between the time series. DTW was chosen due to its superior abilities in dealing with multi-dimensional and multi-variate time series with high efficiency and robustness. DTW also solves the computational challenge stemming from the dynamic nature of comparing each element of a time series with the corresponding element of another time series. Still, DTW needs sufficient computational power to compute the pair-wise similarity score between different time series. Thereafter, the algorithm applies a complete-linkage agglomerative hierarchical clustering technique to generate clusters of machines with similar dynamic profiles that enable the detection of bottlenecks.

All three competences were necessary for realising the technical structure of the bottleneck detection solution. For example, domain competence was crucial for assessing the practical relevance of the solution in light of factory physics, especially to ensure that the algorithm correctly captures the shifting nature of bottlenecks where the bottleneck moves between the machines in the system along with time. This understanding of shiftiness in the real-world system was essential to the choice of DTW as the most suitable similarity measure. However, since DTW is computationally challenging, data engineering competence was crucial in designing the recommended number of computing notes (e.g. in Microsoft Azure) and balancing this against computing costs.

Table 4. Tasks and skills for developing the AI solution for bottleneck detection.

Phase	Tasks	Skills
Business understanding	The objective function is to increase the throughput of the production line	Assessing objectives in the focal factory and detecting value potentials in throughput improvements
	The target problem is to identify historical bottlenecks in the production line	Analysing factory dynamics and articulating connections to bottleneck analysis
	The baseline performance is the average historical throughput of the production line (number of products produced)	Arguing for throughput as a suitable performance indicator for bottleneck analysis and calculating the baseline
	The desired process improvement is a 10% increase in throughput, matching the tact time of the product demand	Linking management goals to measurable throughput improvements
	Bottleneck analysis is a high frequency problem that influences daily operations; a high impact problem that determines the throughput of the production line; and a time-consuming decision-making process	Explaining and presenting the decision-making structure of bottleneck analysis in the focal factory; arguing for the need for an AI solution
	Current practice of manual analysis using value stream mapping leads to a lack of granular insights into system dynamics leads to vulnerable assumptions and ambiguous decisions	Critically assessing current bottleneck management practices in the focal factory
	The AI solution should be capable of identifying the bottleneck in any historical time frame; the AI solution should avoid vulnerable assumptions and provide granular insights into system dynamics	Formulating requirement specifications for AI solution for bottleneck analysis; reviewing requirement specifications relative to management goals and improvement goals
Data understanding	A batch approach for data processing is suitable for historical bottleneck analysis	Assessing data processing types for historical bottleneck analysis
	Procuring two years of historical event log data from the MES system	Determining the suitable data type and data source for bottleneck analysis
	Storing event log data in Microsoft Azure data lake	Executing data migration and storage procedures from MES system to Microsoft Azure
	Establishing a data pipeline between the Azure data lake and an AI workbench in Snowflake and Python	Designing and creating data pipelines in Azure, Python, and Snowflake; programming in Python and Snowflake
Data preparation	Feature #1 - Date and time-stamp of machine events; Feature #2 - Machine states (encoded as Andon lights); Feature #3 - Alarm information for the corresponding machine states	Comparing bottleneck analysis methods and converting their characteristics into model features
	Creating labels of machine states by interpreting the Andon lights and corresponding alarm information	Interpreting meta-data and categorising labels according to the selected bottleneck analysis method
	Labelling the procured data set with machine states	Relating data points to label definitions
	Within each machine, assessing missingness and errors in the recording of date and time-stamps	Identifying and adapting data quality measures to bottleneck analysis
	Across each machine, comparing the time stamps and assessing the flow of production	Reviewing data sets in relation to data quality measures
	Choosing a 50/50 split for training and testing data (one year of data each)	Evaluating strategies for bottleneck analysis model training and testing
	Formalising three conceptual designs of AI solutions to train and test; AI solution #1 - Clustering (hierarchical) and classification (rule-based); AI solution #2 - Regression (linear) and classification (hypothesis testing); AI solution #3 - Rule-based classification (maximum active period)	Designing viable conceptual AI solutions for bottleneck analysis
	Choosing confusion matrix for evaluating classification models; focusing on precision and recall	Comparing relevant AI evaluation metrics and evaluating the appropriateness for bottleneck analysis
Modelling	Modelling the time and machine states information from the procured and labelled data set	Distinguishing temporal information of test data and machine states for bottleneck analysis
	Applying the different AI solutions to the training data (e.g. hierarchical clustering and rule-based classification)	Executing Python-based AI libraries suitable for the temporal information and set of viable AI solutions for bottleneck analysis
	Tuning the model parameters based on training performance (e.g. distance metric, cluster generation technique, linkage method)	Assessing viable AI solutions and adapting model parameters to the specific bottleneck problem
	Applying the different AI solutions to the testing data	Reviewing knowledge gained from training data and implementing bottleneck analysis models to unseen testing data
	Tuning model parameters based on testing performance	Interpreting AI performance on testing data and distinguishing suitable adaptations to bottleneck analysis model parameters
Evaluation	Evaluating the precision and recall of all AI solutions; placing particular emphasis on recall (decision-makers can analyse false positives and decide on suitable strategies)	Calculating and interpreting AI evaluation metrics relative to bottleneck management practices in the focal factory
	Implementing the AI solutions and observing if the objective function is met (i.e. incremental increase of throughput from use)	Interpreting AI model results relative to bottleneck management practices and improvement goals in the focal factory; designing and implementing experimental procedures for bottleneck improvement procedures
	Comparing the AI performance results from internal and external evaluation	Linking knowledge from AI solution evaluation to the throughput of the production line

(*continued*)

Table 4. Continued.

Phase	Tasks	Skills
	Choosing Solution #1 (clustering and classification) as the final AI solution	Interpreting AI evaluation results and comparing the performance of viable AI solutions
Deployment	Establishing a data pipeline from the AI workbench (Snowflake and Python) to Power BI	Designing, creating, and optimising data pipelines for latency performance in Snowflake, Python, and Power BI; devising suitable visualisation techniques
	Conducting offline testing of the AI solution in its intended environment	Assessing the usability of AI solutions from multiple stakeholders
	Conducting user interaction tests (e.g. evaluate latency); collecting and evaluating user feedback (e.g. changing dendrograms to text messages that state which machine was the bottleneck); updating model parameters based on user feedback (e.g. assessing latency relative to the choice of cluster generation technique)	Assessing user feedback and determining actions for AI solution optimisation
	Implementing the AI solution in the live environment to all intended users in their desired interfaces (e.g. laptops, mobile phones, tablets)	Building and implementing User Experience design for the AI solution; designing and creating scalable infrastructure for the AI solution
	Collecting and evaluating user feedback in the live environment (e.g. evaluate speech recognition to deliver voice message of the current bottleneck)	Assessing the usability of the AI solution from multiple stakeholders
Operation and Maintenance	Continuously consuming AI insights and taking relevant actions on the identified bottleneck machine(s) (e.g. dynamic buffers in front of bottleneck machines)	Interpreting AI insights and directing actions to mitigate or eliminate bottlenecks
	Continuously collecting user feedback and taking relevant actions on the AI solution (e.g. improving recall by reducing missed actual bottlenecks)	Evaluating actions taken towards the bottleneck and devising AI solution improvements
	Regularly evaluating data drift by checking elbow plots	Reviewing bottleneck behaviour and attributing factory dynamics to data drift
	Regularly evaluating concept drift by assessing formal records (e.g. introducing new products or changing the flow)	Reviewing bottleneck behaviour and attributing factory dynamics to concept drift
	Establishing routines and alert system for evaluation of data and concept drift every six months	Creating and implementing management practices for data and concept drift
	If data drift is observed, go back to the modelling phase and re-evaluate the AI solution	Distinguishing data drift and articulating changes to AI model parameters
	If concept drift is observed, go back to business understanding and re-evaluate the need to design a new AI solution	Distinguishing concept drift and articulating changes to AI solution design and/or target bottleneck problem

This competence was also necessary for designing and establishing seamless data pipelines (e.g. from Snowflake to Azure) that fed the AI solution with high-quality data at the right time.

Further, a key interplay between domain and data science competence occurred in the clustering step. Here, data science competence was the key to designing the algorithm so that it starts with each machine as a separate cluster followed by linking the different units to form distinct clusters. The domain competence was key to selecting the number of clusters and visually interpreting the representative time series for each cluster and detecting the cluster containing the bottleneck machine. Furthermore, in cases where additional computational power cannot be provided, trade-offs between the choice of similarity measure, computational power, and accuracy of bottleneck detection needed to be managed. Such trade-offs required a combination of all three competences to modify the algorithm semantics and assign them with desired accuracy and necessary computational power.

In essence, these relationships further illustrate how domain, data science, and data engineering competence are individually necessary for the creation of impactful AI solutions in manufacturing. Thus, when all competences are present, it is possible to solve real-world manufacturing problems with AI solutions.

6. Discussion

Motivated by the slow pace of diffusion and lack of substantial success in realising AI productivity gains within manufacturing, we report a six-year analysis of AI in manufacturing that yielded a subsequent enhancement of the CRISP-DM methodology. With a focus on two key complementarities – operation and maintenance of AI solutions over time and human skills – we support manufacturing firms to become better equipped to reap the benefits of AI in their organisations. The enhanced CRISP-DM methodology proposed in this article has a range of academic and practical implications.

6.1. Academic implications

We enhance the CRISP-DM methodology in three major ways, each having clear implications for industrial engineering as well as operations and supply chain management. First, our elaboration of CRISP-DM as a complete life-cycle of AI solutions (Figure 3) strongly emphasises the need for a holistic view of the design, development, and implementation of AI solutions in manufacturing. Whereas most applications of CRISP-DM in manufacturing are one-off projects executed sequentially (Schröer, Kruse, and Gómez 2021), we highlight the full and cyclic nature of the AI solution life-cycle. This life-cycle

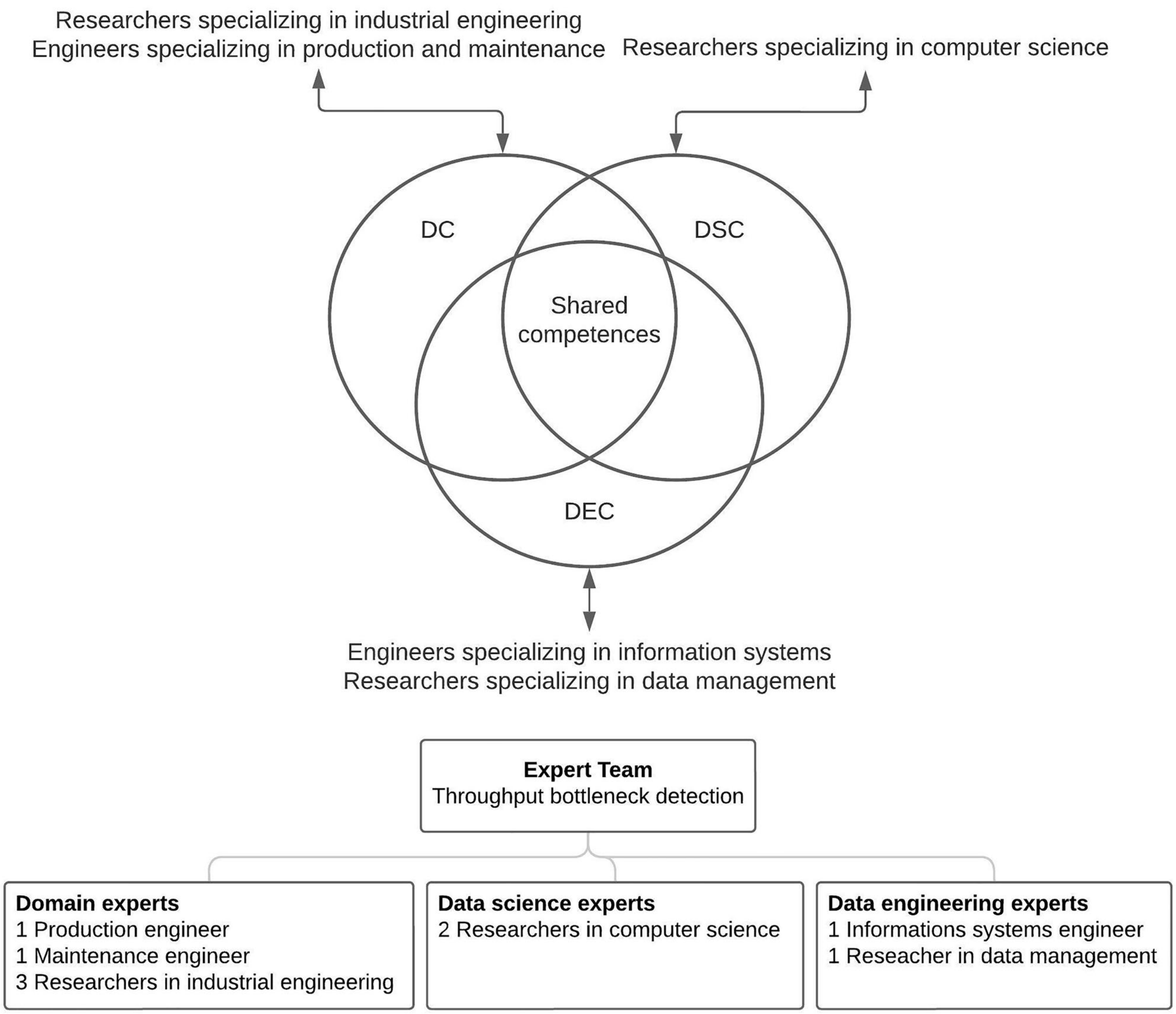

Figure 8. Competences and expert team for developing the AI solution for bottleneck analysis.

encompasses much more than the technical development of prediction models. It puts an equally strong emphasis on identifying and framing manufacturing problems as AI problems, drafting conceptual designs, fostering the usability of AI solutions to shop floor practitioners, and enabling efficient human-AI interactions during daily operations.

Second, by adding 'Operations and Maintenance' as the seventh phase of CRISP-DM, we put the spotlight on the difficult trade-offs and hidden costs of operating and maintaining AI solutions over time. By incorporating such challenges into an explicit part of CRISP-DM, it is possible to overcome the caveat that most AI solutions in manufacturing get stuck in the deployment phase (Schröer, Kruse, and Gómez 2021; Studer et al. 2021). Specifically, researchers need to carefully consider techniques and practices for managing data and concept drift, otherwise, AI performance risks deteriorating over time and incurring hidden maintenance costs (Sculley et al. 2015; Tripathi et al. 2021; Dreyfus et al. 2022).

Third, by embedding a novel task-based framework within CRISP-DM (Figure 1), we provide clarity in the meaning and relationships between tasks, skills, competences, experts, and jobs. By mapping tasks and skills (Table 3), we show that AI solution development in manufacturing encompasses a lot more than sophisticated data wrangling. While existing research has emphasised the need for combined domain and data science competence (Li et al. 2021; Hiruta et al. 2019; Huber et al. 2019), we firmly position and explain the critical role of data engineering competence (Figure 4). Further, we provide a novel visualisation of the trajectories of all three competences throughout the CRISP-DM phases (Figure 5). Thus, any endeavour to develop AI solutions in manufacturing needs to ensure the presence of all three competences and the corresponding integration of experts to enable successful value creation from AI (Figure 6).

6.2. Practical implications

Our enhanced CRISP-DM version has two major implications for manufacturing practice. First, we hope to put the last nail in the coffin of viewing AI solution development in general and CRISP-DM in particular as a waterfall process executed in dedicated, one-off projects. We strongly encourage manufacturing firms to embrace the full life-cycle of AI solutions and use the enhanced CRISP-DM as a continuous, active, and iterative way of working. Putting and keeping organisational practices in place that fully encompass the AI solution life-cycle is key to the successful realisation of AI productivity gains in manufacturing. Moreover, extending CRISP-DM into operations and maintenance will help managers budget for AI investments and more accurately evaluate Return on Investment (ROI). At present, ROI is typically evaluated shortly after the AI solution is deployed. However, due to the risk of AI drift, the performance of the AI solution may quickly fluctuate or change, affecting the possibility to realise projected productivity gains. Thus, managers need to allocate additional budgets for the maintenance of AI solutions to ensure their long-term value potential.

In effect, including Operations & Maintenance as a formal part of CRISP-DM stimulates a more holistic and long-term approach to developing and scaling AI solutions by incentivizing manufacturing firms to revise their *mindset* (e.g. considering AI maintenance already from the start to support technology scaling and avoiding unexpected and expensive mistakes), *systems* (e.g. tech tools to maintain and ensure the smooth working of the algorithms), and *process* (e.g. defined ways of working for detecting and rectifying data and concept drift). Thus, manufacturing firms can use CRISP-DM to scale AI solutions by establishing repeatable procedures for ensuring robustness (e.g. handling variations of incoming data), maintainability and flexibility (e.g. making offline changes and efficiently deploying them live), and budgeting (e.g. anticipating and allocating costs to AI maintenance).

Second, embedding the task-based framework within CRISP-DM facilitates managers to design and assemble manufacturing analytics teams that encompass all necessary competences as well as staffing projects and organisational functions with key experts. The framework can also be used as a strategic tool to assess the current state of existing competences (in-house and/or externally sourced) and continuously monitor the match between tasks and skills in line with technological change. For example, manufacturing firms can use it to identify imbalances between the demand (tasks, jobs) and supply (skills, competences, experts) of labour that are specific to manufacturing AI in their own context. If mismatches between tasks and skills are detected, managers can pursue supply-demand matching strategies through human capital investments (e.g. education, training, or hiring). The description of tasks and skills in Tables 3 and 4 may also be useful for formulating desired job descriptions and distributing workplace ads. Here, it is important to distinguish between the terminology for the generic competences (i.e. domain, data science, data engineering) and the terminology for the professional roles embodying the competences (e.g. AI-oriented roles with various task-skill profiles are referred to using multiple and potentially overlapping terms such as 'ML Engineer', 'AI Engineer', 'DL Engineer', 'Software Engineers', or 'Data Analyst') (Meesters, Heck, and Serebrenik 2022).

6.3. Future research

Although our enhanced CRISP-DM methodology takes a leap forward in realising the value of AI in manufacturing, there are still improvement potentials and avenues for future research. To stimulate the progression of the field, we particularly suggest the following five directions.

First, much more emphasis needs to be put on research that specifically targets the 'Operations and Maintenance' phase of the CRISP-DM methodology (Figure 3). Specifically, there is a dire need to develop methods, tools, and techniques for managing AI drift in manufacturing settings. At present, there are no clear guidelines or standardised procedures for monitoring and acting against data and concept drift (Table 3). For example, should data scientists manually monitor AI performance over time, or could models be embedded with the capability of automatically detecting drift?

Second, further research is needed on the behavioural aspects of AI in manufacturing. While current research has made substantial progress in the technical aspects of solving practical problems with industrial data and ML algorithms, there is plenty of work to do to figure out best practices for supporting end users in consuming and acting on AI insights. One direction is to continue research on superimposed model functions such as explainability (Rudin 2019; Barredo Arrieta et al. 2020), yet such technical aspects should ideally also be linked to human behaviour. For example, one fruitful avenue would be to study how model transparency and reliability influence how manufacturing practitioners develop trust in AI technology (Glikson and Woolley 2020).

Third, while our mapping of tasks and skills (Table 3) aims to cover the most important and stable characteristics of AI solution development in manufacturing, it is not intended to be exhaustive. Researchers can therefore delve even deeper by also identifying specialised tasks and facilitating efficient documentation of process instances that widely apply in manufacturing contexts. As the embedded task-based framework (Figure 1) bridges the economic and educational literature on both the supply- and demand-side of the labour market, our enhanced CRISP-DM methodology may also be used as inspiration for developing novel curricula, courses, and teaching and learning activities in both higher education and professional education aimed at diffusing and accelerating AI in manufacturing.

Fourth, as this research took its departure from independent and stand-alone AI solutions (e.g. one specific solution for long-term bottleneck detection), future research should be devoted to expanding both the CRISP-DM methodology and its inherent tasks and skills for end-to-end AI solutions, e.g. when multiple models are combined or when one model provides input into to another model (i.e. interaction). While such solutions are still exceptionally rare in manufacturing, they are likely to be much more prevalent in the future. CRISP-DM could be a viable methodology for communicating one model output to another model as well as facilitating the thinking about required data engineering procedures to make such communication possible. Further, it can be anticipated that such solutions would broadly require the same core competences executing the same tasks, but that new and other potentially necessary competences might emerge.

Fifth, as Generative AI (GAI), e.g. ChatGPT by Open AI, and associated transformer models have recently seen significant advancement and gained extreme popularity and hype, both in society at large and in manufacturing, it is imperative that future research investigates how CRISP-DM can be used and/or adapted to accommodate GAI development as well as help to excel its diffusion and impact in manufacturing.

7. Conclusions

To stimulate productivity gains from AI, we embarked on a multi-year and multi-method process that comprised a total of six years of research and practice. Our analysis provided novel insights into the management of AI complementarities in manufacturing and laid the foundation for enhancing the popular and widely used CRISP-DM methodology. Specifically, we extend CRISP-DM into a full life-cycle of AI solutions by adding the phase of 'Operation and Maintenance'. We also embed a task-based framework into CRISP-DM that provide guidance for linking tasks to skills and integrating competences held by experts. Using the enhanced CRISP-DM methodology will support manufacturing firms in becoming better equipped to reap the benefits of AI. We also open the avenues for new research on the challenges inherent to operating and maintaining AI solutions over time, managing AI drift, and ensuring the presence of necessary competences to enable successful value creation from AI.

We conclude by providing the following four core recommendations. First, we advise manufacturing firms to use the enhanced version of CRISP-DM as a continuous, active, and iterative way of working. Embrace the full life-cycle of AI solutions and recognise how all seven phases are intertwined. Second, recognise the critical role of operating and maintaining AI solutions over time. Realise that decisions made in early phases of AI design may result in hidden maintenance costs and acknowledge that deterioration of AI performance due to AI drift needs to be carefully managed. Third, for any endeavour to develop AI solutions, carefully map tasks with skills and recognise that the presence of domain, data science, and data engineering competence is necessary for success. Make sure to staff AI development endeavours with a suitable set of experts that are cross-functionally integrated. Fourth, use the CRISP-DM methodology as a guiding framework and strategic management tool to monitor the match between tasks and skills. When imbalances are detected, pursue supply-demand matching strategies through human capital investments.

Note

1. In this article, we focus on the practical usefulness of AI in manufacturing, i.e., that AI solutions can be adapted contextually and effectively used by practitioners to achieve desired outcomes by being both useful (capable of creating benefits) and usable (target outcomes that can be influenced). Further, in light of the lack of a consensus definition of AI, we uphold a 'toolbox interpretation' of AI and refer to the concept in a broad sense; encompassing a variety of methods, tools, and techniques for perceiving, processing, learning, and acting from data. Thus, Machine Learning (ML) and Deep Learning (DL) are subsets of AI. For discussions on conceptualizing AI in manufacturing, see e.g., Subramaniyan et al. (2021) or Helo and Hao (2022).

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This work has been performed within the Sustainable Production Initiative and the Production Area of Advance at Chalmers. The support is greatly appreciated.

Disclosure statement

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.

Notes on contributors

Jon Bokrantz is a researcher at the Department of Industrial and Materials Science at Chalmers University of Technology. His research focuses on production and operations management with a special emphasis on industrial maintenance. His research interests include the interplay of technology, people, and organisation, especially in the context of advancing and diffusing digital technologies to maximise operational performance.

Mukund Subramaniyan is working as a consultant at Capgemini, advising companies in the intersection of operations and Artificial Intelligence (AI). Specifically, helping manufacturing companies transform operations by deploying AI solutions and ensuring these solutions go from pilot to full-scale adoption and scaling. He is also a visiting researcher at the Department of Industrial and Materials Science at Chalmers University of Technology. His research interests include the application of AI within operations and supply chain management.

Anders Skoogh is a Professor at the Department of Industrial and Material Science at Chalmers University of Technology. He is a research group leader for Production Service & Maintenance Systems. Anders is also the director of Chalmers' Masters' program in Production Engineering and a board member of the think-tank Sustainability Circle. Before starting his research career, he accumulated industrial experience from being a logistics developer at Volvo Cars.

ORCID

Jon Bokrantz http://orcid.org/0000-0003-4930-7786

Data availability statement

Due to the nature of the study (i.e. privacy restrictions on production system event log data and client-researcher confidential data in clinical field studies), participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

References

Acemoglu, Daron, and Pascual Restrepo. 2018. Artificial intelligence, automation and work. National Bureau of Economic Research.

Amankwah-Amoah, Joseph, and Yingfa Lu. 2022. "Harnessing AI for Business Development: A Review of Drivers and Challenges in Africa." *Production Planning & Control*. https://doi.org/10.1080/09537287.2022.2069049

Anderson, Lorin W., and David R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*: New York: Longman.

Barredo Arrieta, Alejandro, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, et al. 2020. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI." *Information Fusion* 58: 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Azevedo, Ana, and Manuel Filipe Santos. 2008. *KDD, SEMMA and CRISP-DM: A Parallel Overview*. Amsterdam: IADS-DM.

Benzell, Seth G., and Erik Brynjolfsson. 2019. *Digital Abundance and Scarce Genius: Implications for Wages, Interest Rates, and Growth*. National Bureau of Economic Research.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *The Quarterly Journal of Economics* 117 (1): 339–376. https://doi.org/10.1162/003355302753399526

Brynjolfsson, Erik, Wang Jin, and Kristina McElheran. 2021. "The Power of Prediction: predictive Analytics, Workplace Complements, and Business Performance." *Business Economics* 56 (4): 217–239. https://doi.org/10.1057/s11369-021-00224-5

Brynjolfsson, Erik, and Paul Milgrom. 2013. "Complementarity in Organizations." In *The Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 11–55. New Jersey: Princeton University Press.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies." *American Economic Journal: Macroeconomics* 13 (1): 333–372. https://doi.org/10.1257/mac.20180386

Byrne, David, Carol Corrado, and Daniel E. Sichel. 2018. *The Rise of Cloud Computing: Minding Your P's, Q's and K's.* National Bureau of Economic Research.

Dreyfus, Paul-Arthur, Antoine Pélissier, Foivos Psarommatis, and Dimitris Kiritsis. 2022. "Data-Based Model Maintenance in the Era of Industry 4.0: A Methodology." *Journal of Manufacturing Systems* 63: 304–316. https://doi.org/10.1016/j.jmsy.2022.03.015

Dåderman, Antonia, and Sara Rosander. 2018. *Evaluating Frameworks for Implementing Machine Learning in Signal Processing: A Comparative Study of CRISP-DM, SEMMA and KDD*.

Edmondson, Amy C., and Stacy E. McManus. 2007. "Methodological Fit in Management Field Research." *Academy of Management Review* 32 (4): 1246–1264. https://doi.org/10.5465/amr.2007.26586086

Fisher, Greg, and Herman Aguinis. 2017. "Using Theory Elaboration to Make Theoretical Advancements." *Organizational Research Methods* 20 (3): 438–464. https://doi.org/10.1177/1094428116689707

Fosso Wamba, Samuel, Maciel M. Queiroz, Cameron Guthrie, and Ashley Braganza. 2022. "Industry Experiences of Artificial Intelligence (AI): Benefits and Challenges in Operations and Supply Chain Management." *Production Planning & Control* 33 (16): 1493–1497. https://doi.org/10.1080/09537287.2021.1882695

Glikson, Ella, and Anita Williams Woolley. 2020. "Human Trust in Artificial Intelligence: Review of Empirical Research." *Academy of Management Annals* 14 (2): 627–660. https://doi.org/10.5465/annals.2018.0057

Helo, Petri, and Yuqiuge Hao. 2022. "Artificial Intelligence in Operations Management and Supply Chain Management: An Exploratory Case Study." *Production Planning & Control* 33 (16): 1573–1590. https://doi.org/10.1080/09537287.2021.1882690

Hiruta, Tomoaki, Takayuki Uchida, Shinya Yuda, and Yasushi Umeda. 2019. "A Design Method of Data Analytics Process for Condition Based Maintenance." *CIRP Annals* 68 (1): 145–148. https://doi.org/10.1016/j.cirp.2019.04.049

Huber, Steffen, Hajo Wiemer, Dorothea Schneider, and Steffen Ihlenfeldt. 2019. "DMME: Data Mining Methodology for Engineering Applications–a Holistic Extension to the CRISP-DM Model." *Procedia CIRP* 79: 403–408. https://doi.org/10.1016/j.procir.2019.02.106

Jacobsen, Hendrik, and Kim Hua Tan. 2022. "Improving Food Safety through Data Pattern Discovery in a Sensor-Based Monitoring System." *Production Planning & Control* 33 (16): 1548–1558. https://doi.org/10.1080/09537287.2021.1882691

Karlsson, Christer. 2010. *Researching Operations Management*. New York: Routledge.

Kharlamov, Alexander A., Luís Miguel, D. F. Ferreira, and Janet Godsell. 2020. "Developing a Framework to Support Strategic Supply Chain Segmentation Decisions: A Case Study." *Production Planning & Control* 31 (16): 1349–1362. https://doi.org/10.1080/09537287.2019.1707896

Khazraei, Khashayar, and Jochen Deuse. 2011. "A Strategic Standpoint on Maintenance Taxonomy." *Journal of Facilities Management* 9 (2): 96–113. https://doi.org/10.1108/14725961111128452

Kristoffersen, Eivind, Oluseun Omotola Aremu, Fenna Blomsma, Patrick Mikalef, and Jingyue Li. 2019. *Exploring the Relationship between Data Science and Circular Economy: An Enhanced CRISP-DM Process Model. Paper Presented at the Conference on e-Business, e-Services and e-Society*.

Li, Guoyan, Chenxi Yuan, Sagar Kamarthi, Mohsen Moghaddam, and Xiaoning Jin. 2021. "Data Science Skills and Domain Knowledge Requirements in the Manufacturing Industry: A Gap Analysis." *Journal of Manufacturing Systems* 60: 692–706. https://doi.org/10.1016/j.jmsy.2021.07.007

Madaio, Michael, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 2022. "Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support." *Proceedings of the ACM on Human-Computer Interaction* 6 (CSCW1): 1–26. https://doi.org/10.1145/3512899

Manimuthu, Arunmozhi, V. G. Venkatesh, V. Raja Sreedharan, and Venkatesh Mani. 2022. "Modelling and Analysis of Artificial Intelligence for Commercial Vehicle Assembly Process in VUCA World: A Case Study." *International Journal of Production Research* 60 (14): 4529–4547. https://doi.org/10.1080/00207543.2021.1910361

Meesters, Marcel, Petra Heck, and Alexander Serebrenik. 2022. "What is an AI Engineer? An Empirical Analysis of Job Ads in The Netherlands. Paper Presented at the." Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI.

Miles, Matthew B., and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. London: Sage.

Rodrigues, Margarida, Enrique Fernández-Macías, and Matteo Sostero. 2021. "A Unified Conceptual Framework of Tasks, Skills and Competences." JRC Working Papers Series on Labour, Education and Technology.

Rudin, Cynthia. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models instead." *Nature Machine Intelligence* 1 (5): 206–215. https://doi.org/10.1038/s42256-019-0048-x

Schein, Edgar H. 1987. *The Clinical Perspective in Fieldwork*: Sage Publications, Inc.

Schneider, Johannes, Rene Abraham, Christian Meske, and Jan Vom Brocke. 2023. "Artificial Intelligence Governance for Businesses." *Information Systems Management* 40 (3): 229–249. https://doi.org/10.1080/10580530.2022.2085825

Schröer, Christoph, Felix Kruse, and Jorge Marx Gómez. 2021. "A Systematic Literature Review on Applying CRISP-DM Process Model." *Procedia Computer Science* 181: 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Sculley, David, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. 2015. "Hidden Technical Debt in Machine Learning Systems." *Advances in Neural Information Processing Systems* 28: 1–9.

Sköld, Martin, and Christer Karlsson. 2007. "Multibranded Platform Development: A Corporate Strategy with Multimanagerial Challenges." *Journal of Product Innovation Management* 24 (6): 554–566. https://doi.org/10.1111/j.1540-5885.2007.00271.x

Sköld, Martin, and Christer Karlsson. 2013. "Stratifying the Development of Product Platforms: requirements for Resources, Organization, and Management Styles." *Journal of Product Innovation Management* 30: 62–76. https://doi.org/10.1111/jpim.12064

Studer, Stefan, Thanh Binh Bui, Christian Drescher, Alexander Hanuschkin, Ludwig Winkler, Steven Peters, and Klaus-Robert Müller. 2021. "Towards CRISP-ML (Q): a Machine Learning Process Model with Quality Assurance Methodology." *Machine Learning and Knowledge Extraction* 3 (2): 392–413. https://doi.org/10.3390/make3020020

Subramaniyan, Mukund, Anders Skoogh, Jon Bokrantz, Muhammad Azam Sheikh, Matthias Thürer, and Qing Chang. 2021. "Artificial Intelligence for Throughput Bottleneck Analysis–State-of-the-Art and Future Directions." *Journal of Manufacturing Systems* 60: 734–751. https://doi.org/10.1016/j.jmsy.2021.07.021

Subramaniyan, Mukund, Anders Skoogh, Maheshwaran Gopalakrishnan, Hans Salomonsson, Atieh Hanna, and Dan Lämkull. 2016. "An Algorithm for Data-Driven Shifting Bottleneck Detection." *Cogent Engineering* 3 (1): 1239516. https://doi.org/10.1080/23311916.2016.1239516

Subramaniyan, Mukund, Anders Skoogh, Azam Sheikh Muhammad, Jon Bokrantz, and Ebru Turanoğlu Bekar. 2019. "A Prognostic Algorithm to Prescribe Improvement Measures on Throughput Bottlenecks." *Journal of Manufacturing Systems* 53: 271–281. https://doi.org/10.1016/j.jmsy.2019.07.004

Subramaniyan, Mukund, Anders Skoogh, Azam Sheikh Muhammad, Jon Bokrantz, Björn Johansson, and Christoph Roser. 2020a. "A Data-Driven Approach to Diagnosing Throughput Bottlenecks from a Maintenance Perspective." *Computers & Industrial Engineering* 150: 106851. https://doi.org/10.1016/j.cie.2020.106851

Subramaniyan, Mukund, Anders Skoogh, Azam Sheikh Muhammad, Jon Bokrantz, Björn Johansson, and Christoph Roser. 2020b. "A Generic Hierarchical Clustering Approach for Detecting Bottlenecks in Manufacturing." *Journal of Manufacturing Systems* 55: 143–158. https://doi.org/10.1016/j.jmsy.2020.02.011

Subramaniyan, Mukund, Anders Skoogh, Hans Salomonsson, Pramod Bangalore, and Jon Bokrantz. 2018. "A Data-Driven Algorithm to Predict Throughput Bottlenecks in a Production System Based on Active Periods of the Machines." *Computers & Industrial Engineering* 125: 533–544. https://doi.org/10.1016/j.cie.2018.04.024

Subramaniyan, Mukund, Anders Skoogh, Hans Salomonsson, Pramod Bangalore, Maheshwaran Gopalakrishnan, and Azam Sheikh Muhammad. 2018. "Data-Driven Algorithm for Throughput Bottleneck Analysis of Production Systems." *Production & Manufacturing Research* 6 (1): 225–246. https://doi.org/10.1080/21693277.2018.1496491

Tripathi, Shailesh, David Muhr, Manuel Brunner, Herbert Jodlbauer, Matthias Dehmer, and Frank Emmert-Streib. 2021. "Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing." *Frontiers in Artificial Intelligence* 4: 22. https://doi.org/10.3389/frai.2021.576892

Usuga-Cadavid, Juan Pablo, Samir Lamouri, Bernard Grabot, and Arnaud Fortin. 2022. "Using Deep Learning to Value Free-Form Text Data for Predictive Maintenance." *International Journal of Production Research* 60 (14): 4548–4575. https://doi.org/10.1080/00207543.2021.1951868

Wang, Yingli, Jean-Paul Skeete, and Gilbert Owusu. 2022. "Understanding the Implications of Artificial Intelligence on Field Service Operations: A Case Study of BT." *Production Planning & Control* 33 (16): 1591–1607. https://doi.org/10.1080/09537287.2021.1882694

Wirth, Rüdiger, and Jochen Hipp. 2000. "CRISP-DM: Towards a Standard Process Model for Data Mining." Paper Presented at the Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, Crowne Plaza Midland Hotel, Manchester, UK, April 11–13.

Åhlström, Pär, and Christer Karlsson. 1996. "Change Processes towards Lean Production." *International Journal of Operations & Production Management* 16 (11): 42–56. https://doi.org/10.1108/01443579610131447