A Meta-Review on Artificial Intelligence in Product Creation

Ruslan Bernijazov[1], Alexander Dicks[2], Roman Dumitrescu[1,3], Marc Foullois[1], Jonas Hanselle[3*], Eyke Hüllermeier[4], Gökce Karakaya[1], PatrickKödding[3], Volker Lohweg[2], Manuel Malatyali[3], Friedhelm Meyer auf der Heide[3], Melina Panzner[3], Christian Soltenborn[3]

[1]Fraunhofer Institute for Mechatronic Systems Design IEM

[2]inIT - Institute Industrial IT, Technische Hochschule Ostwestfalen-Lippe

[3]Paderborn University, Heinz Nixdorf Institute

[4]Institute of Informatics, University of Munich (LMU)

jonas.hanselle@uni-paderborn.de[3]

Abstract

Product Creation (PC) refers to the process of planning and developing a product as well as related services from the initial idea until manufacturing and distribution. Throughout this process, there are numerous tasks that depend on human expertise and are typically undertaken by experienced practitioners. As the field of Artificial Intelligence (AI) continues to evolve and finds its way into the manufacturing sector, there exist many possibilities for an application of AI in order to assist in solving aforementioned tasks. In this work, we provide a comprehensive overview of the current state of the art of the use of AI in PC.

In detail, we analyze 40 existing surveys on AI in PC and 94 case studies in order to find out which areas of PC are primarily addressed by current research in this field, how mature the discussed AI methods are, and to which extent data-centric approaches are utilized in current research.

1 Introduction

Compared to low-wage countries, the focus of highly industrialized countries such as Germany is not on production, but on the development of new products. This poses its very own challenges, partly because product complexity has increased significantly in many domains. One way to deal with this is digitization of resources and processes in many companies, typically generating large volumes of data. Thus, it seems a natural choice to apply increasingly powerful AI methods to these data with the aim of improving productivity in PC.

However, for small and medium-sized enterprises, this is a major challenge in many cases, as they generally do not have the necessary expertise and resources to exploit the potential of AI. The government funded project *AI Marketplace* [AI Marketplace, 2021] therefore has the goal of developing a platform that gives such companies intelligent access to AI experts and their solutions. However, in order to develop such a platform, it is necessary to have an overview of the current state of the use of AI solutions in the context of PC.

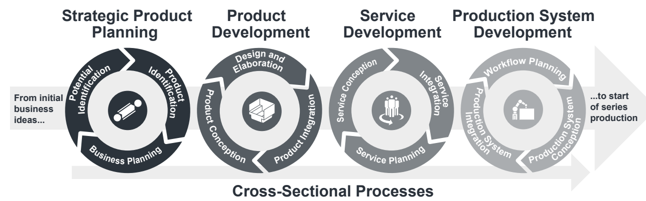

Figure 1: Illustration of the 4-cycle model of PC. Adapted from [Gausemeier *et al.*, 2018].

To obtain this overview, we therefore systematically reviewed relevant literature on AI in PC, guided mainly by two questions: First, we wanted to find out which use cases are addressed by AI in PC; for this, we classified the papers according to a *4-cycle reference model* of PC. Second, we wanted to learn which AI methods are applied in PC; we approached this by investigating whether the AI methods were data-centric and by quantifying the maturity of the AI methods. This paper is the result of this literature review and, to the best of our knowledge, the first comprehensive overview of AI in PC.

2 Background: Product Creation

This section provides a brief overview of the main subaspects with the corresponding literature references. For more details, we refer to Gausemeier *et al.* [2018, 2014], the referenced literature, and the corresponding references therein.

The area of PC bundles all the activities required to create an interdisciplinary technical product or any associated service. It extends from the generation of initial business ideas to the start of series production. The 4-cycle model of PC is a reference model to divide the activities of PC (see Figure 1). The reference model divides the activities into the cycles of Strategic Product Planning, Product Development, Service Development, and Production System Development.

The *first cycle (Strategic Product Planning)* ranges from the procedure of finding prospective success potentials until start of development. It includes the tasks of Potential Iden-

[*]Contact Author

tification, Product Identification, and Business Planning. The *second cycle (Product Development)* comprises the Interdisciplinary Product Conception, the design, and the elaboration in the respective areas of expertise as well as the integration of the results into an overall solution. The goal of the *third cycle (Service Development)* is the conversion of a service idea into a market service [Bullinger and Scheer, 2006]. This cycle consists of the tasks Service Conception, Service Planning, and Service Integration [Meiren and Liestmann, 2002]. The starting point of the *fourth cycle (Production System Development)* is the conceptual design of the Production System. Here the four aspects are Workflow Planning, Work Equipment Planning, Workplace Planning, and Production Logistics which are to be considered in an integrated fashion.

In addition to the technical tasks structured in the four cycles, PC also includes other processes and activities. These, mainly project-specific activities, are summarized in the *Cross-Sectional Processes*. Examples of this are risk management and configuration management [ISO, 2002].

3 Related Work

The potential as well as the application of AI in PC has been studied in a variety of existing literature reviews. In order to provide an overview of these reviews we categorize them according to their depths with regard to PC and AI, respectively. The categorization is defined in Table 1 and the corresponding result is shown in Table 2. For the assessment of the PC depth we distinguish between publications that focus on *Selected Activities*, *Selected Processes* consisting of several activities, a *Selected Cycle*, and publications that consider several cycles (category *Cross-Cycle*). The category *Not in Focus* subsumes all publications that only indirectly address PC and lay their main focus on other topics (e.g., Industry 4.0).

A similar categorization is also applied to define the AI depth of the publications. *Selected Approaches* denotes the most specific category and *Cross-Branch* the most general category. In the following, we provide a short summary of the main findings based on our categorization. A detailed description of the identified studies can be downloaded from the *AI Marketplace* web page [1].

Our categorization reveals that a high proportion of previous works focuses on specific PC activities, rather than the complete field of PC. For instance, Serna M. *et al.* [2019] investigate AI applications for the automation of software test, whereas Xiu and Wan [2013] focus on the utilization of AI in CAD design. Existing studies with a broader scope are still restricted to specific PC cycles. For example, hu Li *et al.* [2017] focus on the utilization of AI in intelligent manufacturing but do not consider strategic product planning and only partially address product development. In turn, studies with a broader PC scope have often a rather narrow scope with regard to AI. For example, Burggräf *et al.* [2020] consider all processes involved in product development but restrict the AI scope to knowledge-based systems. In summary, although existing literature studies provide valuable information on the potential of AI in different stages of PC, none of them consider the entire spectrum of PC processes as well as AI approaches.

[1] plattform.ki-marktplatz.com/Group/Welcome/Downloads

Category	Definition	Examples
Product Creation Focus		
Selected Activities	Focus on a few selected PC activities.	Requirements Elicitation, Architecture Modelling
Selected Processes	Focus on a few selected PC processes without restrictions on concrete activities in these processes.	Requirements Engineering, Architecture Development
Selected PC Cycle	Focus on a specific PC cycle without restrictions on any of the involved processes.	Strategic Product Planning, Product Development
Cross-Cycle	Focus on PC in general without restrictions on a concrete PC domain.	Whole product lifecycle is considered
Not in Focus	Focus lies not on PC. It is only addressed in passing.	Focus lies on industrial applications in general
Artificial Intelligence Focus		
Selected Approaches	Focus on a few selected approaches (models or algorithms) .	Evaluation of different parameter settings of a few selected models
Selected AI Tasks	Focus on a few selected AI tasks without restriction to certain approaches that address these tasks.	Named Entity Recognition, Sentiment Analysis, Object Detection
Selected AI Branches	Focus on a set of similar AI tasks without restriction to concrete AI tasks.	Natural Language Processing, Knowledge Engineering, AI Vision
Cross-Branch	No restriction to specific AI branches.	All AI branches are considered
Not in Focus	Focus lies not on AI. It is only addressed in passing.	Focus lies on support for engineers in general

Table 1: Categories for the classification of surveys on AI in PC.

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="5">Product Creation Focus</th></tr>
<tr><th>Not in Focus</th><th>Selected Activities</th><th>Selected Processes</th><th>Selected PC Cycle</th><th>Cross-Cycle</th></tr>
<tr><td rowspan="5">AI Focus</td><td>Not in Focus</td><td></td><td colspan="4">[DL11], [GM15], [MSA+17], [PME+18], [PVB+20], [PW11]</td></tr>
<tr><td>Selected Approaches</td><td rowspan="4">[LZ19], [S14], [SPP+19]</td><td></td><td></td><td>[MP19], [PLO19]</td><td></td></tr>
<tr><td>Selected AI Tasks</td><td>[LPM+19], [RS16]</td><td>[JMZ11], [WZL+18]</td><td>[KN20]</td><td>[AB12], [OKÖ+20]</td></tr>
<tr><td>Selected AI Branches</td><td>[BDB+15], [CKW+20], [JJT+19], [QQ18], [SS13], [XSW+17], [ZMB+20], [ZZW15]</td><td>[KKW+18], [TKB+18]</td><td></td><td>[BWW20], [LKZ+18], [RZL+18], [SHB19], [VBD+11], [WZL+12]</td></tr>
<tr><td>Cross-Branch</td><td>[K17], [SAS19], [XW13]</td><td></td><td>[FZZ+20], [HJ20], [LHY+17]</td><td>[our]</td></tr>
</table>

Table 2: Categorization of existing literature reviews on AI in PC.

4 Methodology

Our meta-review is driven by the well-established guidelines proposed by Kitchenham [2004] and Kuhrmann *et al.* [2017]. In accordance with these guidelines, we describe the research questions (Section 4.2), search strategy (Section 4.3), the review process (Section 4.4), and data extraction process of our work (Section 4.5). Beforehand, however, we discuss the preparation in detail and consider different potential of validity threats in the following subsection.

4.1 Meta-Discussion on the Preparation

This section discusses our measures to assure validity of our review as well as the threats to validity. We structure this discussion according to the definitions of common threats to validity provided by Zhou *et al.* [2016].

Construct Validity: Our research questions as well as the used search string were defined in the scope of several workshops with experts from the AI and the PC domain and are

Research Question	Motivation
RQ1: Which use cases are addressed by AI in Product Creation?	We want to find out which use cases or applications are realized by means of AI in product creation, such as market analysis, knowledge based engineering, etc.
RQ2: Which AI approaches are utilized in Product Creation?	We want to find out which AI approaches (algorithms, methods, models etc.) are utilized in the product creation domain.

Table 3: Research questions.

documented in this paper in order to assure reproducibility. However, choosing other research questions or search string might reveal other relevant publications that were not considered in our review.

Internal Validity: In order to assure internal validity, we discussed all steps and selection criteria with the involved researches and assured that each selected publication was reviewed by at least two researchers. In case of discrepancies, more colleagues were consulted for decision making. However, a different team might come to other conclusions which might lead to a deviation in the extracted publications.

External Validity: In order to get a good overview of relevant publications, we searched in the top 3 digital libraries suggested by Kuhrmann *et al.* [2017] as well as in Scopus, the biggest digital library with peer-reviewed publications. However, there are also a variety of other libraries that we did not take into account. Considering these libraries might reveal relevant publications. Moreover, we do not consider papers published before 2010 or in another language than English or German. Considering these publications might lead to deviations in the results.

Conclusion Validity: The data extraction and evaluation procedures were aligned among the whole team consisting of eight researchers from the PC and the AI domain. The concrete data extraction and evaluation steps were conducted by at least two researchers who are familiar with the respective concepts. The results were reviewed and discussed by the whole team. However, there is still a threat for community bias since all researchers work at closely collaborating research institutes.

4.2 Research Questions

After the discussion on the basis we build the literature research on, we start with the presentation of the procedure by posing our research questions. We define two research questions driving our meta-review which are depicted in Table 3 as well as their motivations. These questions are derived from our research goals of identifying the state of the art of AI in PC as well as discovering potentials for further research.

4.3 Review Protocol and Relevant Literature Identification

The review protocol provides a road map for the review, with a sequence of methodological steps intended to reduce researcher bias. A review protocol was developed that defines the search strategy, study selection mechanism, and data extraction criteria. The detailed procedure and the results are shown in Figure 2. This review protocol was elaborated both by AI as well as PC researchers in several iterations based on our research questions.

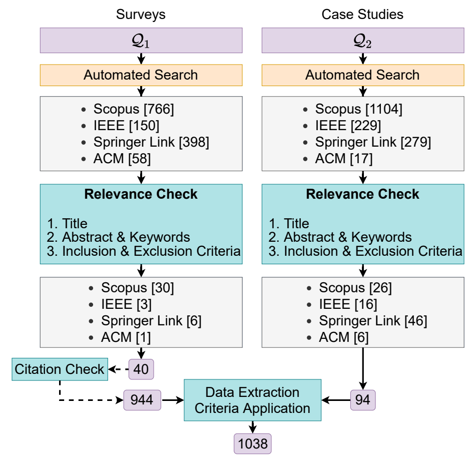

Figure 2: Illustration of the iterative review procedure. The numbers depicted in square braces indicate the number of remaining contributions from each literature database under consideration.

(Automated) Search Strategy The search strategy describes the process to identify literature in the relevant context that is useful in answering the research questions. In the following, we consider the different key points, from the selection of literature databases, to the setting and combination of keywords and a corresponding publication time horizon, to a (manual) selection process within a multi-step procedure.

We conducted an automated search on the online digital libraries *ACM*, *IEEE Xplore*, *Scopus*, and *SpringerLink*, which are often referred to as the top 4 standard libraries (e.g., [Kuhrmann *et al.*, 2017]).

In order to construct search queries at the intersection of AI and PC, we first fixed keywords for both fields, which are depicted in Table 4. Although it is sufficiently clear that these keywords do not cover every niche of AI or PC, they were chosen carefully in order to provide a holistic overview over the most relevant topics in each field. Furthermore, the keywords were tested in several trial runs.

In addition to these content-related keywords, we also add keywords regarding the type of each publication, i.e. we focus on surveys and on case studies. By this, we can reduce the amount of search results to a proper quantity on a feasible level. Thus, the search strings used for the automated search are given by $\mathcal{Q}_1$ and $\mathcal{Q}_2$ respectively.

$$\mathcal{Q}_1 := \text{PC} \times \text{AI} \times \text{Survey}$$
$$\mathcal{Q}_2 := \text{PC} \times \text{AI} \times \text{Case Study}$$

Overall we end up with combinations of keywords from the domain of PC combined with AI. The resulting list is in turn combined with keywords indicating a Survey or a Case Study which we consider separately. In total we obtained 513 search queries for Surveys and 513 search queries for Case Studies. The search process is illustrated in Figure 2. There are 1372 unique survey articles identified by the automated process and 1629 case studies. Note that articles have been identified multiple times which we deduplicated afterwards. These publications are the basis for the Review Process we consider in the following section.

Category	Keywords
AI	Artificial Intelligence, Machine Learning , Deep Learning, Neural Network, Computer Vision, Image Processing, Knowledge Representation, Data Mining, Natural Language Processing, NLP, Computational Intelligence, Advanced Data Analytics, Data Science, Big Data
PC	Product Creation, Product Planning, Product Development, Product Engineering, Service Development, Service Engineering, Production System Development, Production System Engineering, Product-Lifecycle-Management
Survey	Review, Survey, Literature Study
Case Study	Case Study, Use Case, Application

Table 4: The individual sets of keywords used for constructing search queries.

4.4 Review Process

The Review Process defines the step of reviewing the identified literature within a multiple-stage process. The goal is to elicit the most relevant publications with respect to the research questions defined in Section 4.2. Note that not all publications published before 2010 could be excluded due to technical constraints within the automated selection process. The same applies to the underlying language of the publications (cf. Table 5). Several publications were also excluded here after the query process. The majority of the publications, however, were processed through a content analysis and accordingly transferred to the next process step or excluded.

This review procedure was carried out by assessing the relevance of each publication in an increasing level of detail. In the first step, relevance was judged based solely on the title, then on abstract/ keywords, and finally, the remaining inclusion and exclusion criteria were applied. The judgement of the (content-based) inclusion and exclusion criteria (cf. Table 5) was carried out according to the 4 eyes principle. Each paper was reviewed independently by 2 domain experts and marked accordingly. An evaluation was carried out quantitatively, so that the evaluation could be averaged. However, it was also marked with a comment, e.g., whether the paper was generally in the engineering domain and not specifically related to the PC domain which helped to exclude the paper from further analysis.

The result of this process step is a list of 40 surveys and 94 Case Studies which are both in the domain of AI and PC (Figure 2). We then specifically screened the surveys. We were able to identify a total of 944 publications of AI applications in the PC area. These, together with the 94 case studies, result in 1038 identified applications together with the corresponding publications, which we examine in more detail in the following section. The procedure for extracting the individual criteria of the applications is explained in the individual subsections.

4.5 Extraction Process

In the last subsection, we described the (sub-)process whose results include 1040 publications. Each application can be seen as a combination of a problem in the PC domain and an AI solution that addresses this problem. Note that an application may also entail several AI solutions, e.g., when applying several AI methods to a problem within the same publication. With a view to answering the research questions, we classified the applications in the PC domain and in the AI domain: The PC tasks were classified based on the 4 cycles (cf. Section 2). For the classification from the AI domain, the terms used were extracted and then grouped into different categories, (e.g., Supervised / Unsupervised Learning). Furthermore, it is determined whether the approach is "data-centric" and the degree of pervasiveness of the solution is captured, i.e., the maturity of the solution is determined. The following section describes the underlying definitions and presents the results of the analysis.

Inclusion Criteria	Exclusion Criteria
Title, keyword list, and abstract make explicit that the paper is related to the utilization of AI in PC	Contribution was published before 2010
Title, keyword list, and abstract make explicit that the paper discusses data structures in product creation	Contribution is neither written in English nor German
Title, keyword list, and abstract make explicit that the paper discusses data quality in product creation	Contribution does not address the utilization of AI in PC

Table 5: Inclusion and exclusion criteria the meta-review was driven by.

5 Results

In the following, we present the results of the meta-review with respect to the initially defined research questions depicted in Table 3. First, we classify the term AI into different areas (Section 5.1). Furthermore, we define the maturity level of the different solutions as well as distinguish whether they are data-centric approaches or others (Section 5.1). Subsequently, we elaborate the research questions RQ1 and RQ2 based on the results from the literature review and present the results found (Sections 5.2 and 5.3). In the next section, we put the results into a context and discuss potential rationales (Section 6).

5.1 Working Definition and Classification for AI

Solution Maturity

In order to asses the adoption of AI in PC, we evaluated the identified case studies with respect to their level of maturity. To this end, we define three maturity levels which are presented in Table 6.

AI and Machine Learning

We subdivide the AI approaches that we encountered in the discovered literature into two groups: (i) More traditional, knowledge-based approaches, such as heuristic search, planning, or rule-based (expert) systems, and (ii) modern data-centric approaches making use of machine learning (ML) and data mining (DM) methods to extract models and patterns from data. In the recent past, a clear shift in focus from the former to the latter could be observed in the field of AI at large. Reasons for the increasing popularity of data-centric solutions are manifold, including of course the increased availability of data thanks to advances in digitalization combined with improved computational power. Thus, while data and compute ressources are becoming "cheaper" and learning from data more automatized, knowledge-based approaches require the availability of human experts and the

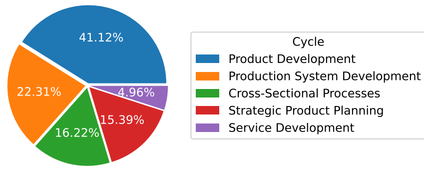

Figure 3: Distribution of PC cycles among all considered solutions.

Maturity	Description	Example
Level 1	The paper discusses the application of AI in product creation on a theoretical level. The discussed approaches are applied on dummy data or not validated at all.	The paper suggests that the discussed approaches might be applied in product creation without verifying this hypothesis.
Level 2	The discussed approaches are applied on real product creation data but in a lab environment.	The approaches are trained and/or tested on data of prior projects but the results are not used in real development projects.
Level 3	The discussed approaches are applied on real data in a real development project.	The approaches are verified in the scope of a pilot project.
Not Provided	The paper does not provide any information on the used data sources.	

Table 6: Overview of maturity levels.

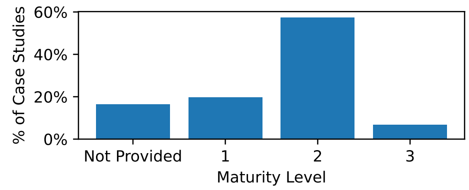

Figure 4: Maturity levels of identified case studies.

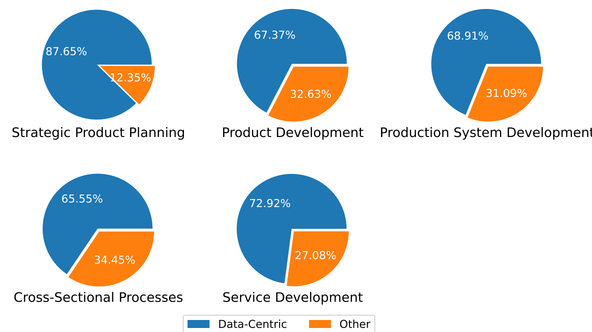

Figure 5: Distributions of *data-centric* solutions among the 4 cycles of PC and the Cross-Sectional Processes.

formalization of domain knowledge, which is often difficult and cumbersome.

Admittedly, the distinction between knowledge-based and data-centric approaches is rather broad, and a more fine-granular categorization of methods might be desirable. For example, in machine learning a distinction is commonly made between supervised, unsupervised, and reinforcement learning, and the former could further be sub-categorized into classification, regression, etc. However, for most papers we sighted, this type of information could simply not be extracted, because the description remained on a rather abstract level and did not provide enough technical details.

5.2 RQ1: Which use cases are addressed by AI in PC?

As discussed in Section 2, the PC domain is divided into 4 cycles. Generally speaking, the first research question addresses the distribution of AI methods within the PC area. In order to answer RQ1, we analyze which of these areas are addressed by the solutions presented in the considered literature. Figure 3 shows the distribution of considered applications among the PC cycles.

We observe, that the majority of considered applications stem from the cycle of Product Development. The remaining shares are referenced in approximately the same proportions in the literature, up to the area of Service Development (with $4.96\,\%$). Thus, a tendency towards Product Development can be observed in our data set. The share of the three categories varies not considerably between $15.39\,\%$ to $22.31\,\%$. These results are discussed in the following section.

5.3 RQ2: Which AI approaches are utilized in PC?

To answer the second research question, we look at the solution approaches used in the papers. These include the maturity level of the data basis used and the project context as well as the solution type used. Recall, we classify the solution type into data-centric and other approaches. We start by presenting the research with reference to the maturity of the project and then consider the solution approach.

The results of our evaluation with respect to the maturity level are visualized in Figure 4. For a large proportion of the case studies, the degree of penetration can be determined; only $16.38\,\%$ of the studies did not provide a note on the data basis used. A small proportion ($19.56\,\%$) published a study at a very early stage of the investigation; the majority of studies ($57.32\,\%$) were able to make an informed statement using real data from the domain under laboratory conditions. The fewest studies ($6.74\,\%$) did so based on an application that could be localized to an industrial standard in a real development project.

We now proceed to the consideration of the solutions used. The solutions are divided according to the definition of data-centric and other approaches (for the definition see Section 5.1). This subdivision is further broken down by the individual process steps within the 4-cycle model, which is enriched by the Cross-Sectional Processes. On average, 2/3 of the approaches are data-centric, so the Product Development, Production System Development, and Cross-Sectional Processes areas reflect this average. The areas of Strategic Product Planning and Service Development differ significantly: Strategic Product Planning is almost exclusively ($87.65\,\%$) addressed by a data-centric approach, where Service Development is still about 3/4 of the approaches ($72.92\,\%$). It should be noted that this study is based exclusively on AI approaches and the relative statistics are within this context.

6 Discussion

In this section we revisit the various results from the previous section and put them into context: Different attempts to explain the observations are presented and checked for plausibility. We also draw conclusions from these and derive recommendations for the project implementation as well as for potential follow-up studies. Besides the suggested research directives, we also plan a more detailed analysis of the identified applications with regard to the addressed PC problems as well as the utilized AI algorithms for future work.

Use Cases in PC for AI First we consider the results regarding the first research question, i.e. the number of solutions distributed over the PC areas. We observed, that most of the proposed AI approaches are solutions to problems from the area of Product Development and the smallest share is dedicated solely to the Service Development area.

The reasons for these observations can be multifaceted: On the one hand, the field of product development in general could present a particularly large number of problems that can be handled with above-average success using AI methods. On the other hand, there could also be an increasing number of problems that are particularly difficult to solve, e.g., (NP-hard) combinatorial optimization problems that entail a large number of heuristics for special subproblems. Based on the available data collection, no conclusive statement can be formulated here. From our experience, this observation is valid and leads to the assumption that this is a comparatively large area, i.e., it results in an above-average number of problems. This aspect could be investigated in more detail in a follow-up study: The standard model of PC could be further broken down to concrete problem definitions. An assignment of the papers to the concrete problems would show which individual problems are frequently considered (many solutions for few problems) or whether there are a large number of problems (in each case few solutions for many problems).

Furthermore, we observe a reduced representation in the area of service development. The interpretation of this observation is comparatively obvious: On the one hand, this area often involves the manual execution of a process, and on the other hand, the area of service development is closely related to product development. Intuitively, a provider can offer a service based on the know-how of the product - the development of a service to a product without in-depth knowledge is much more difficult. In addition, publications rarely consider this point in isolation, as the development of a service represents the extended arm of the product.

Maturity We assessed the maturity of the approaches in terms of their degree of practical application. Here, we observed that the majority of approaches was evaluated on real-world data e.g., from prior projects, but *not* in the scope of a dedicated development project, i.e. they do not fulfill an industrial standard. This suggests that, in general, real-life PE problems can be addressed by means of AI. However, there seems to be a lack of research on the adoption of AI in industrial practice. Also, a high number of analyzed case studies does not provide sufficient information on the used data and verification strategies in order to evaluate the maturity level.

From our perspective, this shows that there is currently a lack of understanding of the importance of accurately describing the underlying data set that forms the basis for learned models. Comparability between different approaches and their solutions can only be made on the basis of a comparability of the different data sets. Furthermore, a common view of the data, e.g., in the form of benchmarks or specific problem instances in the form of landmarks would be supportive of comparability. A coordination process with the goal of being able to compare and classify different approaches in this area is suggested at this point.

Data-Centric Methods During the period of 2010-2020, there was a clear tendency towards the use of data-centric methods in PC in general and also in each of its cycles. This is in agreement with our intuition, as during the considered time period, the fields of Machine Learning, Data Mining, and Data Science have received a considerable amount of attention both in theory and practice.

Among the four cycles of PC, the cycle of Strategic Product Planning has been shown to be most prone to data-centric solutions. We have attributed this to the tasks within Strategic Product Planning which mainly deals with the identification of market potentials based on observations, i.e. in this cycle a lot of approaches are concerned with analyzing user and review data [Hou and Jiao, 2020]. Another aspect that makes Strategic Product Planning more amenable to data-centric solutions might be the utilized data formats. A majority of processes in this cycle utilize textual and numerical data which can be directly processed by existing data-centric solutions. In contrast to this, the other cycles often use complex, PC-specific data formats (e.g., geometrical models, graph-based structures). Processing these data requires PC-specific preprocessing steps which increase the development effort for data-centric solutions. Additionally, there is not as much potential for the use of optimization or search algorithms in Strategic Product Planning as in the other areas. In Product or Production System Development, the engineers are often faced with high-dimensional optimization problems, e.g., when designing analog integrated circuits using evolutionary algorithms [Tlelo-Cuau *et al.*, 2010].

Addressing the topic of data-centric approaches in the context of AI-based solution approaches has identified potential for further research questions: The use of data to create a solution is often unavoidable; Deep Learning methods in particular are served by the use of large amounts of data (Big Data). How has the amount of data underlying the solution changed over the past few years? We suspect that the creation of Big Data has increased precisely due to digital transformation efforts and increasing harmonization and linking of data within an industrial context. This highlights once again that an explicit description of the underlying data in the form of meta-data is important. Currently, we see this in a large part of the case studies we have examined, where the size is not explicitly stated in whatever form.

Acknowledgements

This work was supported by the German Federal Ministry of Economic Affairs and Energy (BMWi) within the “Innovationswettbewerb Künstliche Intelligenz”.

References

AI Marketplace. AI Marketplace. https://ki-marktplatz.com/en/, 2021. Online; accessed April 26, 2021.

Hans-Jörg Bullinger and August-Wilhelm Scheer. Service Engineering -— Entwicklung und Gestaltung innovativer Dienstleistungen. In *Service Engineering*, pages 3–18. Springer, 2006.

Peter Burggräf, Johannes Wagner, and Tim Weisser. Knowledge-based problem solving in physical product development – a methodological review. *Expert Systems with Applications*, page 100025, 2020.

J Gausemeier, T Schierbaum, and T Westermann. Interdependencies during the conceptual design of an analytical telemedical device. In *DS 77: Proceedings of the DESIGN 2014 13th International Design Conference*, 2014.

Jürgen Gausemeier, Roman Dumitrescu, Tomas Pfänder, Daniel Steffen, and Frank Thielemann. *Innovationen für die Märkte von morgen: strategische Planung von Produkten, Dienstleistungen und Geschäftsmodellen*. Carl Hanser Verlag GmbH Co KG, 2018.

Liang Hou and Roger J. Jiao. Data-informed inverse design by product usage information: a review, framework and outlook. *Journal of Intelligent Manufacturing*, 31(3):529–552, March 2020.

Bo hu Li, B. Hou, W. Yu, Xiaobing Lu, and Chun-Wei Yang. Applications of artificial intelligence in intelligent manufacturing: a review. *Frontiers of Information Technology & Electronic Engineering*, 18:86–96, 2017.

ISO. ISO/IEC 15288:2002 Systems engineering – System life cycle processes, institution = International Organization for Standardization. Technical report, 2002.

Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26, 2004.

Marco Kuhrmann, Daniel Méndez Fernández, and Maya Daneva. On the pragmatic design of literature studies in software engineering: An experience-based guideline. *Empirical Software Engineering*, 22:2852–2891, 12 2017.

Thomas Meiren and Volker Liestmann. Service Engineering in der Praxis. *Kurzstudie zur Dienstleistungsentwicklung in deutschen Unternehmen, Stuttgart: Fraunhofer-IRB Verlag*, 2002.

Edgar Serna M., Eder Acevedo, and Alexei Serna. Integration of properties of virtual reality, artificial neural networks, and artificial intelligence in the automation of software tests: A review. *Journal of Software: Evolution and Process*, 31:e2159, 2019.

E. Tlelo-Cuau, I. Guerra-Gom, M.A. Duarte-Vil, Luis G. de la Frag, G. Flores-Bec, G. Reyes-Salg, C.A. Reyes-Garc, and G. Rodriguez. Applications of Evolutionary Algorithms in the Design Automation of Analog Integrated Circuits. *Journal of Applied Sciences*, 10(17):1859–1872, August 2010.

Yi Xiu and Zhen-Kai Wan. A survey on pattern-making technologies in garment cad. In *IEEE Conference Anthology*, pages 1–6, 2013.

Xin Zhou, Yuqin Jin, He Zhang, Shanshan Li, and Xin Huang. A map of threats to validity of systematic literature reviews in software engineering. In *2016 23rd Asia-Pacific Software Engineering Conference (APSEC)*, pages 153–160, 2016.

Appendix

The supplementary material contains a comprehensive list of the considered surveys, which are discussed in the related work section of this paper.

Reference	Source
[AB12]	Agard, Bruno; Barajas, Marco (2012): The use of fuzzy logic in product family development: literature review and opportunities. In: J Intell Manuf 23 (5), pp. 1445–1462. DOI: 10.1007/s10845-010-0465-5.
[BDB+15]	Blondet, Gaëtan & Le Duigou, Julien & Boudaoud, Nassim & Eynard, Benoit. (2015). Simulation data management for adaptive design of EXPERIMENTS: a litterature review. Mechanics and Industry. 16. 611. 10.1051/meca/2015041.
[BWW20]	Burggräf, P., Wagner, J., & Weißer, T. (2020). Knowledge-Based Problem Solving in Physical Product Development–A Methodological Review. Expert Systems with Applications: X, 100025.
[CKW+20]	Chan, Kit Yan; Kwong, C. K.; Wongthongtham, Pornpit; Jiang, Huimin; Fung, Chris K.Y.; Abu- Salih, Bilal et al. (2020): Affective design using machine learning: a survey and its prospect of conjoining big data. In: International Journal of Computer Integrated Manufacturing 33 (7), pp. 645–669. DOI: 10.1080/0951192X.2018.1526412.
[DL11]	Du, J., & Liu, J. (2011, September). Knowledge management for product development: A review. In 2011 12th International Conference on Computer-Aided Design and Computer Graphics (pp. 38-43). IEEE.
[FZZ+20]	Feng, Y., Zhao, Y., Zheng, H., Li, Z., & Tan, J. (2020). Data-driven product design toward intelligent manufacturing: A review. International Journal of Advanced Robotic Systems, 17(2)
[GM15]	Lan, Bo. (2014). Knowledge Management for Product Development: A Review. Advanced Materials Research. 1037. 494-498. 10.4028/www.scientific.net/AMR.1037.494.
[HJ20]	Hou, L., & Jiao, R. J. (2020). Data-informed inverse design by product usage information: a review, framework and outlook. Journal of Intelligent Manufacturing, 31(3), 529-552.
[JJT+19]	Jadhav, R., Joshi, pp. D., Thorat, U., & Joshi, A. pp. (2019, March). A Survey on Software Defect Prediction in Cross Project. In 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 1014-1019). IEEE.
[JMZ11]	Yaqiong, L., Man, L. K., & Zhang, W. (2011). Fuzzy theory applied in quality management of distributed manufacturing system: A literature review and classification. Engineering Applications of Artificial Intelligence, 24(2), 266-277.
[K17]	Kumar, pp.P.. (2017). State of The Art-Intense Review on Artificial Intelligence Systems Application in Process Planning and Manufacturing. Engineering Applications of Artificial Intelligence. 65. 294-329. 10.1016/j.engappai.2017.08.005.
[KKW+18]	Kim, Dong-Hyeon & Kim, Thomas & Wang, Xinlin & Kim, Mincheol & Quan, Yingjun & Oh, Jin & Min, Soo-Hong & Kim, Hyungjung & Bhandari, Binayak & Yang, Insoon & Ahn, Sung- Hoon. (2018). Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry. International Journal of Precision Engineering and Manufacturing-Green Technology. 5. 555-568. 10.1007/s40684-018-0057-y.
[KN20]	Khosravani, Mohammad Reza; Nasiri, Sara (2020): Injection molding manufacturing process: review of case-based reasoning applications. In: J Intell Manuf 31 (4), pp. 847–864. DOI: 10.1007/s10845-019-01481-0.

[LHY+17]	Li, Bo-hu; Hou, Bao-cun; Yu, Wen-tao; Lu, Xiao-bing; Yang, Chun-wei (2017): Applications of artificial intelligence in intelligent manufacturing: a review. In: Frontiers Inf Technol Electronic Eng 18 (1), pp. 86–96. DOI: 10.1631/FITEE.1601885.
[LKZ+18]	Lv, Shengping; Kim, Hoyeol; Zheng, Binbin; Jin, Hong (2018): A Review of Data Mining with Big Data towards Its Applications in the Electronics Industry. In: Applied Sciences 8 (4), pp. 582. DOI: 10.3390/app8040582.
[LPM+19]	Lupinetti, Katia; Pernot, Jean-Philippe; Monti, Marina; Giannini, Franca (2019): Content- based CAD assembly model retrieval: Survey and future challenges. In: Computer-Aided Design 113, pp. 62–81. DOI: 10.1016/j.cad.2019.03.005.
[LZ19]	Liao, Xiaofeng; Zhao, Zhiming (2019): Unsupervised Approaches for Textual Semantic Annotation, A Survey. In: ACM Comput. Surv. 52 (4). DOI: 10.1145/3324473.
[MP19]	Kabir, Sohag; Papadopoulos, Yiannis (2019): Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review. In: Safety Science 115, pp. 154– 175. DOI: 10.1016/j.ssci.2019.02.009.
[MSA+17]	McMillan, Alison J.; Swindells, Norman; Archer, Edward; McIlhagger, Alistair; Sung, Anna; Leong, Kelvin; Jones, Rhys (2017): A review of composite product data interoperability and product life-cycle management challenges in the composites industry. In: Advanced Manufacturing: Polymer & Composites Science 3 (4), pp. 130–147. DOI: 10.1080/20550340.2017.1389047.
[OKÖ+20]	Onar, pp. Ç., Kahraman, C., Öztayşi, B., & Boltürk, E. (2020). Fuzzy production systems: A state of the art literature review. Journal of Intelligent & Fuzzy Systems, 38(1), 1071-1081.
[PLO19]	Papadopoulos, Chrissoleon T.; Li, Jingshan; O'Kelly, Michael E.J. (2019): A classification and review of timed Markov models of manufacturing systems. In: Computers & Industrial Engineering 128, pp. 219–244. DOI: 10.1016/j.cie.2018.12.019.
[PME+18]	Preidel, M., Wang, W. M., Exner, K., & Stark, R. (2018). KNOWLEDGE IN ENGINEERING DESIGN: A SYSTEMATIC LITERATURE REVIEW ON ARTIFACTS AND IT SYSTEMS. In DS 92: Proceedings of the DESIGN 2018 15th International Design Conference (pp. 881-892).
[PVB+20]	Pereira Pessôa, Marcus Vinicius; Jauregui Becker, Juan Manuel (2020): Smart design engineering: a literature review of the impact of the 4th industrial revolution on product design and development. In: Res Eng Design 31 (2), pp. 175–195. DOI: 10.1007/s00163-020- 00330-z.
[PW11]	Pirmoradi, Zhila & Wang, Gary. (2011). Recent Advancements in Product Family Design and Platform-Based Product Development: A Literature Review. Proceedings of the ASME Design Engineering Technical Conference. 5. 10.1115/DETC2011-47959.
[QQ18]	Qayyum, Shamaila; Qureshi, Ahsan (2018): A Survey on Machine Learning Based Requirement Prioritization Techniques. In: Proceedings of the 2018 International Conference on Computational Intelligence and Intelligent Systems - CIIS 2018. the 2018 International Conference. Phuket, Thailand, 17.11.2018 - 19.11.2018. New York, New York, USA: ACM Press, pp. 51–55.

[RS16]	K. Renganath and M. Suresh, "Supplier selection using fuzzy MCDM techniques: A literature review," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-6
[RZL+18]	Ren, Shan; Zhang, Yingfeng; Liu, Yang; Sakao, Tomohiko; Huisingh, Donald; Almeida, Cecilia M.V.B. (2019): A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges and future research directions. In: Journal of Cleaner Production 210, pp. 1343–1365. DOI: 10.1016/j.jclepro.2018.11.025.
[S14]	Sandkuhl, K. (2014, July). Knowledge reuse: survey of existing techniques and classification approach. In European Business Intelligence Summer School (pp. 126-148). Springer, Cham.
[SAS19]	Serna M, E., Acevedo M, E., & Serna A, A. (2019). Integration of properties of virtual reality, artificial neural networks, and artificial intelligence in the automation of software tests: A review. Journal of Software: Evolution and Process, 31(7)
[SHB19]	Shabestari, Seyed Sina; Herzog, Michael; Bender, Beate (2019): A Survey on the Applications of Machine Learning in the Early Phases of Product Development. In: Proc. Int. Conf. Eng. Des. 1 (1), pp. 2437–2446. DOI: 10.1017/dsi.2019.250.
[SPP+19]	Sala, R., Pezzotta, G., Pirola, F., & Huang, G. Q. (2019). Decision-Support System-based Service Delivery in the Product-Service System Context: Literature Review and Gap Analysis. In 11th CIRP Conference on Industrial Product-Service Systems (Vol. 83, pp. 126- 131)
[SS13]	Singh, A.; Singh, R. (2013 - 2013): Assuring Software Quality using data mining methodology: A literature study. In: 2013 International Conference on Information Systems and Computer Networks. 2013 International Conference on Information Systems and Computer Networks (ISCON 2013). Mathura, 09.03.2013 - 10.03.2013: IEEE, pp. 108– 113.
[TKB+18]	Tale-Yazdi, A., Kattner, N., Becerril, L., & Lindemann, U. (2018, December). A literature review on approaches for the retrospective utilisation of data in engineering change management. In 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) (pp. 612-616). IEEE.
[VBD+11]	Verhagen, Wim J.C.; Bermell-Garcia, Pablo; van Dijk, Reinier E.C.; Curran, Richard (2012): A critical review of Knowledge-Based Engineering: An identification of research challenges. In: Advanced Engineering Informatics 26 (1), pp. 5–15. DOI: 10.1016/j.aei.2011.06.004.
[WZL+12]	Wen, Kunmei; Zeng, Yong; Li, Ruixuan; Lin, Jianqiang (2012): Modeling semantic information in engineering applications: a review. In: Artif Intell Rev 37 (2), pp. 97–117. DOI: 10.1007/s10462-011-9221-2.
[WZL+18]	Wang, Z., Zheng, P., Lee, C. H., & Chen, C. H. (2018). Knowledge representation and reasoning methods in the concept development of product design: a state-of-the-art review.
[XSW+17]	Xu, Y., Sun, Y., Wan, J., Liu, X., & Song, Z. (2017). Industrial big data for fault diagnosis: Taxonomy, review, and applications. IEEE Access, 5, 17368-17380.
[XW13]	Xiu, Yi; Wan, Zhen-Kai (2013 - 2013): A survey on pattern-making technologies in Garment CAD. In: IEEE Conference Anthology. 2013 IEEE Conference Anthology. China, 01.01.2013 - 08.01.2013: IEEE, pp. 1–6.

[ZMB+20]	Zindani, Divya; Maity, Saikat Ranjan; Bhowmik, Sumit (2020): Decision making tools for optimal material selection: A review. In: J. Cent. South Univ. 27 (3), pp. 629–673. DOI: 10.1007/s11771-020-4322-1.
[ZZW15]	Zhang, Xiaoming; Zhao, Chongchong; Wang, Xiang (2015): A survey on knowledge representation in materials science and engineering: An ontological perspective. In: Computers in Industry 73, pp. 8–22. DOI: 10.1016/j.compind.2015.07.005.