To cite this version :

Bruno COGNET, Jean-Philippe PERNOT, Louis RIVEST, Christophe DANJOU - Systematic comparison of digital maturity assessment models - Journal of Industrial and Production Engineering - Vol. 40, n°7, p.519-537 - 2023

Taylor & Francis
Taylor & Francis Group

Journal of Industrial and Production Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjci21

Systematic comparison of digital maturity assessment models

Bruno Cognet, Jean-Philippe Pernot, Louis Rivest & Christophe Danjou

To cite this article: Bruno Cognet, Jean-Philippe Pernot, Louis Rivest & Christophe Danjou (2023) Systematic comparison of digital maturity assessment models, Journal of Industrial and Production Engineering, 40:7, 519-537, DOI: 10.1080/21681015.2023.2242340

To link to this article: https://doi.org/10.1080/21681015.2023.2242340

JOURNAL OF INDUSTRIAL AND PRODUCTION ENGINEERING
2023, VOL. 40, NO. 7, 519–537
https://doi.org/10.1080/21681015.2023.2242340

ARTICLE

OPEN ACCESS Check for updates

Systematic comparison of digital maturity assessment models

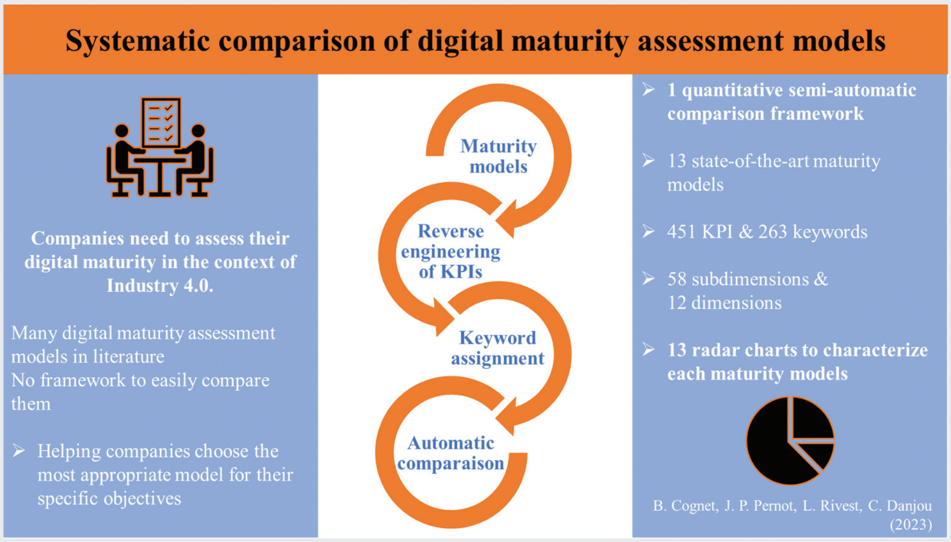
Bruno Cognet[a], Jean-Philippe Pernot[b], Louis Rivest[a] and Christophe Danjou[c]

[a]Systems Engineering Department, ETS, Montreal, Quebec, Canada; [b]Arts et Métiers Institute of Technology, HESAM Université, Aix-en-Provence, France; [c]Mathematics and Industrial Engineering Department, Polytechnique Montréal, Montreal, Quebec, Canada

ABSTRACT
Assessing the digital maturity of companies is essential to prepare for digital transformation in the context of Industry 4.0. Several digital maturity assessment models have emerged in the past few years to support this evaluation. One obstacle for companies is the impossibility of easily comparing themselves to one another quantitatively or qualitatively. This paper introduces a new way to compare digital maturity models through a quantitative framework that is compatible with a wide variety of models. Comparisons are performed in the space of the keywords used to characterize key performance indicators (KPIs) that are reverse engineered from the models. The matches are encoded in a keyword matrix that is used to automatically compute the match level of KPI pairs. The framework has been validated on 13 state-of-the-art maturity models whose analysis resulted in the identification of 451 KPIs characterized using 263 keywords structured according to 12 dimensions and 58 subdimensions.

ARTICLE HISTORY
Received 19 March 2023
Revised 25 June 2023
Accepted 12 July 2023

KEYWORDS
Digital transformation; maturity models; systematic comparison framework; key performance indicators; matching level; model coverage indicators



1. Introduction

The hype around the so-called 4th Industrial Revolution has accentuated the trend towards digitally transforming industrial activities. Governments look to help companies succeed with their digital transformation as quickly as possible. While this generates significant investments from companies, the wrong choice or misuse of new technologies can lead to significant losses [1]. It is therefore crucial to both identify the indicators that most impact companies' development and be able to assess them in order to accurately characterize the level of digital maturity of companies [2]. Rossmann [3] defines a company's digital maturity as its ability to acquire and use so-called digital technologies to improve its overall business. Today, it has become essential for governments to try to assess companies' digital maturity through self-assessment questionnaires or audits. The goal is to compare companies to identify the best-performing ones and determine the best practices for going digital.

A digital maturity assessment model is a tool that identifies best practices and helps to determine a company's level of digital maturity [4]. It provides a foundation to guide a company in its transformation from its current state to its targeted one and thus increase its digital maturity level [4]. Several digital maturity assessment models have emerged in recent years to support this evaluation and enable the benchmarking of companies. Selecting one to use supposes comparing them.

CONTACT Bruno Cognet bruno.cognet.1@ens.etsmtl.ca Systems Engineering Department, ETS, Montreal, Quebec, Canada
Supplemental data for this article can be accessed online at https://doi.org/10.1080/21681015.2023.2242340.

However, comparing available models quantitatively and qualitatively to identify their features is no trivial task.

This paper introduces a new framework to compare the wide variety of digital maturity models that are available. The proposed comparison approach is made possible by a paradigm shift – the concept of moving comparisons into a common space where each maturity model can be efficiently represented. First, key performance indicators (KPIs) are reverse engineered from the various models. Next, keywords are established to characterize each of these KPIs and grouped together. Finally, comparisons are performed directly in the space of these keywords. Matches between keywords are encoded in a keyword matrix that is used to automatically compute the match level of KPI pairs. As a result, the framework can process several hundred thousand comparisons quickly and efficiently. The first two steps are performed manually once for each maturity model to be compared, while keyword comparison is performed automatically. The proposed framework is thus considered semi-automatic.

The matching matrices identified make it possible to characterize the similarities and differences between the maturity models being evaluated. This framework also makes it possible to compute coverage graphs that clearly illustrate the extent to which the maturity models cover the dimensions and subdimensions from the keyword matrix. These quantitative results help to compare the models in order to have a better understanding of their strengths and weaknesses, their specificities, and the concepts they do not cover.

This paper's contribution is threefold: (i) a generic approach that supports the comparison of maturity models (in this case, digital maturity models) using keywords that are assigned to KPIs and that constitute a common space for the comparison of all models; (ii) a semi-automatic comparison framework that computes the match level of KPI pairs that have been reverse engineered from the maturity models and characterized by keywords; and (iii) coverage indicators that quantitatively characterize the extent to which the maturity models (a) cover the dimensions and subdimensions from the keyword matrix and (b) overlap one another, to highlight their similarities and their specific evaluation focuses.

The main idea behind the semi-automatic framework may also be interesting for solving similar comparison questions that pertain to other heterogeneous types of data and can take advantage of the paradigm shift that exploits a common space composed of, for instance, keywords. The objective of this paper is to quantitatively compare existing digital maturity models by defining a common reference space for the models. Using this reference will make it possible to determine the extent to which each model covers the identified evaluation criteria. The approach has been validated on 13 state-of-the-art digital maturity models whose analysis resulted in the identification of 451 KPIs characterized using 263 keywords structured according to 12 dimensions and 58 subdimensions. The framework is systematic and can be complemented by the analysis of new maturity models that may emerge in the future. Its convergence has also been demonstrated.

The rest of the paper is organized as follows. Section 2 reviews the works related to the assessment of digital maturity. The semi-automatic framework is then introduced in Section 3, along with details about its constituent parts. Section 4 presents the results obtained by applying the proposed framework to 13 state-of-the-art digital maturity models. Section 5 offers a discussion, while Section 6 concludes the paper.

2. Literature review

2.1. Digital maturity models and assessment tools

In general, maturity refers to a complete, perfect, or ready state, and consequently to the ability of someone or something to reach their best possible level [5]. In the context of Industry 4.0, an industrial organization needs to first assess its current situation and then establish a digital transformation strategy that will guide its implementation of new technologies. Westerman et al. [6] explain that to become a "Digital Master," a company must distinguish itself not only by its ability to invest in new technologies, but also by its ability to manage the digital transformation of its business. Hizam-Hanafiah et al. [7] note that digital maturity models can also help to assess the success of a digital transformation effort. Westerman et al. [6] show that companies that master both digital and leadership capabilities generally have higher profits than their competitors do. However, not all approaches are suitable for conducting a comprehensive assessment of digital maturity [2].

Several models have emerged to support the assessment of digital maturity, each of which has been designed to provide an indication of a company's digital maturity level based on a target state that is considered the most mature [4]. Repeated use of these models helps to track a company's progress in its digital transformation. The literature points to a nuance between "maturity assessment" models and "readiness assessment" models. On one hand, a maturity assessment model helps an individual or entity to reach a higher level of maturity by following a step-by-step continuous improvement process [8]. Thus, it makes it possible to determine a company's current digital maturity level and to identify the steps it has to take to increase its level of digital maturity [9]. On the other hand, a readiness assessment model examines a company's ability to engage in an organizational transformation [10]. This type of assessment determines its "level of preparedness" [11] based on the conditions and resources that are needed to achieve a goal. In their report, Lichtblau et al. [12] consider Industry 4.0 readiness to represent a company's willingness and capacity to implement the ideas behind Industry 4.0. Finally, both maturity and readiness assessment models aim to describe a company's state at a given point in time. However, their difference lies in their respective objectives. A maturity assessment model looks at an organization's ability to operate/use something, while a readiness assessment model considers an organization's ability to implement/deploy something. The comparison framework proposed in this paper can easily be applied to both types of models. Hence, only the term "maturity model" is used hereinafter to refer to these two types of assessment.

Maturity models generally come from two types of sources [13]: scientific models and models developed by consulting firms. It is also possible to organize digital maturity models into two groups: self-assessment tools and audit tools. With self-assessment tools (e.g. PricewaterhouseCoopers [14]), a company can assess itself and get an idea of the state of its transformation towards the "Industry 4.0" concept. Other tools require the involvement of experts to perform an in-depth audit (for example, the "Audit industrie 4.0" proposed by the Ministère de l'Économie et de l'Innovation du Québec [15]). Following these audits, a detailed report is prepared in which investment paths or modifications related to the company's activities are proposed. Today, multiple public and private organizations offer to perform audits to measure the state of a company's digital transformation in order to precisely guide its future investments [3]. However, as this article will show, these assessment methods are incomplete and cover only some of a company's needs related to the 4th Industrial Revolution [16].

Only 13 of the 19 digital maturity models that were published prior to 2020 provide sufficient information to be processed by the semi-automatic comparison framework introduced in this paper. For example, the maturity model developed by Schumacher et al. [4] is not fully detailed in the literature, which prevents us from processing it using our data analysis-based comparison technique. The complete list of the 13 models we were able to use is provided in Section 4.1. Each model consists of a set of questions and answers that are generally grouped into dimensions and subdimensions to structure the model around central concepts. The models may introduce some nuances, such as "items" [4], "factors" [17] and "variables" [2], but they all aim to characterize a company's processes, technologies, and organizational structure. In this article, the concepts are equated to KPIs that underlie the questions/answers. In this way, a KPI makes it possible to assess a company's situation at a given moment in time in order to guide its decision-making and help it achieve an objective. From these KPIs, it is possible to define best practices. Even though digital maturity models all follow a somewhat similar structure, they do not all evaluate the same aspects. They may focus on the human aspect of the fourth revolution, or on a more technical dimension related to technology. For example, the IMPULS model [12] focuses more on the evaluation of technology, while the PwC model [18] evaluates business strategy as well as sales and technology performance. Nevertheless, it is possible to identify three categories of aspects that recur regularly: human resources, technology, and company organization [19]. Ideally, the aggregation of the different aspects that are considered by the various available models should provide a more comprehensive assessment of a company's digital maturity on all fronts.

Finally, most self-assessment tools provide a score at the end of the questionnaire that corresponds to the maturity level of the company being assessed. This enables companies to know their strengths and weaknesses for each of the dimensions and subdimensions considered. Unfortunately, the scoring system is not the same for all models. Schumacher et al. [4] use a formula that is based on companies' responses and weights that have been predefined by experts for each KPI. Some models simply report an unweighted average or sum [20]. Thus, as is mentioned earlier, it is difficult to compare models since they do not all return the same types of results and analysis.

2.2. Methods for comparing existing maturity models

Akdil et al. [21] present four maturity models and compare them using the following criteria: maturity levels, dimensions, scope of the study, and type of assessment proposed. Although these results are highly qualitative, the authors synthesize them in a table that they use to create their own maturity model. Ambrosio da Silva et al. [22] present a high-level comparison of 11 maturity models and discuss their evaluation dimensions and maturity scales. Axmann and Harmoko [23] compare three assessment tools (IMPULS, the University of Warwick, and PwC) and use them as a basis for their own proposal intended for small and medium-sized enterprises (SMEs). Their comparison is based on a SWOT (strengths, weaknesses, opportunities, and threats) analysis of each tool and is thus qualitative and high-level. Hizam-Hanafiah et al. [7] propose a systematic literature review of 30 maturity models to identify the most commonly used maturity assessment dimensions by grouping the models' 158 dimensions into 6 broad categories.

Castelo-Branco et al. [24] propose a multi-step methodology to develop a new relevant maturity model. First, they compare six existing maturity models and obtain constructs by analyzing the broad themes and indicators used in them. Then, they identify dimensions from the first analysis, and finally, they conduct interviews with experts to end up with 30 measurement indicators that form the basis of a new maturity assessment tool.

Mittal et al. [8] compare 15 maturity models using a five-step methodology. They first characterize the manufacturing SMEs in different countries. Next, they perform a literature review using the terms "smart manufacturing" and "roadmap" to identify existing models. In their third and fourth steps, they lead discussions to compare the models and assess their limitations with respect to what SMEs require. In their last step, they define the adjustments that need to be made to the models to help manufacturing SMEs to better manage their digital transformation.

Schwer et al. [2] use an existing framework – the "ArchiMate 3.0" tool developed by The Open Group [25] – to compare digital maturity models. This framework consists of six dimensions (a Strategy layer, a Business layer, an Application layer, a Technology layer, a Physical layer, and an Implementation and Migration layer). The authors follow the seven-step method proposed by Fink [26]. First, they define their research questions. Then, they select databases, such as Scopus, and determine the terms to be used in their queries – in this case, words related to digital maturity. An initial selection of articles is made based on the titles and abstracts. A second selection step eliminates the articles that do not contain KPIs for digital transformation. They then identify the KPIs related to digital maturity that are mentioned in the remaining articles. Finally, in the last step, they synthesize the results of their search, but this step is not detailed enough and appears to have been implemented manually. They classify the digital maturity indicators in the "ArchiMate 3.0" tool, which makes it possible to know the scope of the maturity models considered by observing their coverage of the dimensions.

Dikhanbayeva et al. [27] propose a model comparison methodology that is based on the design principles proposed by Hermann et al. [28]. They evaluated 18 maturity models' coverage of the design principles, rating it as "high", "moderate", "low" or "basic". They first defined which criteria must be encompassed in each rating level. After conducting a thorough analysis, the authors concluded with a comparison of the maturity models' design principle coverage.

2.3. Synthesis

Although digital maturity models share common characteristics, such as the use of questions/answers, dimensions, and subdimensions to structure their assessment, they are very heterogeneity in terms of how they work, which makes it difficult to compare them. Consider the IMPULS [29] and PwC [14] models, for example, it is difficult to compare the models at the dimension level. The models each have six dimensions (IMPULS: Strategy and organization; Smart factory; Smart operations; Smart products; Data-driven services; Employees/PwC: Business Models, Product & Service Portfolio; Market & Customer Access; Value Chains & Processes; IT Architecture; Compliance, Legal, Risk, Security & Tax; Organization & Culture). Despite the fact that these models (IMPULS, PwC and the others) have similar constructs, it is therefore necessary to focus on the questions they consider to determine their evaluation criteria in order to compare them.

Some attempts have been made to develop comparison approaches, but few operate at a sufficient level of detail to allow for accurate quantitative comparison. They focus on macro-criteria and do not consider the KPIs underlying the questions/answers. Moreover, it is not possible to know what the models do and do not cover, how they relate to one another, or whether they partially overlap. The methods are also not sufficiently automated, which prevents them from being used at a fine level or when many KPIs must be compared. Thus, they are not fully scalable, and they can make the incorporation of a new maturity model complex.

These conclusions justify the need to develop a new comparison framework, which is detailed in the next section and requires a paradigm shift.

3. Methods

The literature review revealed that a wide range of models exist to assess the digital maturity of companies. There is also much heterogeneity in how these approaches work, using either questionnaires or KPIs with different dimensions and subdimensions, and their scoring and ranking mechanisms, etc. This section describes a framework that has been designed to systematically compare different maturity models. First, the proposed framework permits the semi-automatic computation of matching

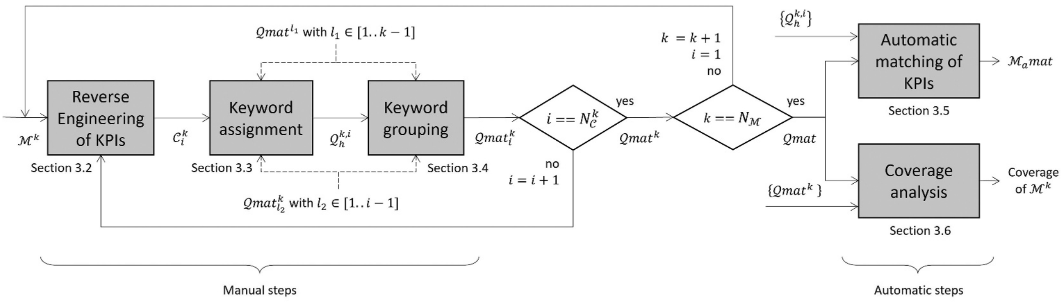


Figure 1. Overview of the semi-automatic comparison framework.

levels, which are also called levels of similarity, between the KPIs that underpin each maturity model. It also enables the computation of coverage indicators characterizing the extent to which the maturity models cover the dimensions and subdimensions associated with the keywords. The comparison results can then be used to design new maturity models that synthesize KPIs from multiple models, but this possible application of the results is not explored in this paper.

3.1. Overall framework

The guiding principle behind the proposed framework is to move to a common space in which comparison is easier and can be automated. To do this, the various maturity models studied are first transformed into a list of KPIs that they are supposed to evaluate. Each KPI is then characterized by a set of keywords that can subsequently be automatically compared with one another to produce matching matrices and coverage analysis charts.

The semi-automatic comparison framework is illustrated in Figure 1. It consists of the three manual steps that must be completed for each maturity model $\mathcal{M}^k$ (with $k \in [1..N_{\mathcal{M}}]$ and $N_{\mathcal{M}}$ being the number of models taking part to the comparison) one after the other following the main loop on the index k:

- **Reverse engineering of KPIs** (Section 3.2): This step makes it possible to consider the large heterogeneity in the way questions and answers are formulated within state-of-the-art digital maturity assessment approaches. A given model $\mathcal{M}^k$ is associated with a set of KPIs, each of which is denoted by $\mathcal{C}_i^k$, with $i \in [1..N_{\mathcal{C}}^k]$, for a total of $N_{\mathcal{C}} = \sum_{k=1}^{N_M} N_{\mathcal{C}}^k$ KPIs.
- **Keyword assignment** (Section 3.3): This step moves comparison to a common space in which each KPI $\mathcal{C}_i^k$ is characterized by one or more keywords, each of which are denoted by $\mathcal{Q}_h^{k,i}$, with $h \in \left[1..N_{\mathcal{Q}}^{k,i}\right]$. Assigned keywords are chosen from a list of keywords identified during the initial analysis of all the maturity models. This list is enriched as the loops on k and i grow.
- **Keyword grouping and dimension/subdimension identification** (Section 3.4): This step makes it possible to group together keywords that have similar meanings so that the KPIs that share keywords will belong to the same match group. The keywords associated with a KPI $\mathcal{C}_i^k$ are grouped in a matrix $\mathcal{Q}mat_i^k$ whose rows correspond to different notions and whose columns are used to store the keywords characterizing said notions. During this step, subdimensions are identified and used to label the rows, and thus the different notions. Rows are also put under the umbrella of specific dimensions that group together several subdimensions. When all the KPIs associated with a maturity model $\mathcal{M}^k$ have been treated ($i == N_{\mathcal{C}}^k$), the matrix $\mathcal{Q}mat^k$ is created, gathering the intermediate matrices. When all the maturity models $\mathcal{M}^k$ have been treated ($k == N_M$), the final keyword matrix $\mathcal{Q}mat$ is produced together with its underlying dimensions and subdimensions.

These nested loops (the manual step) iterate on the maturity models being compared and the KPIs. The intermediate matrices are thus reused as inputs of subsequent iterations (the dashed lines in Figure 1).

Both the final keyword matrix and the final list of keywords $\left\{\mathcal{Q}_h^{k,i}\right\}$ assigned to the KPIs are used to automatically compute the matching matrix $\mathcal{M}_a mat$, i.e. the matrix that gathers the matching levels of the KPIs considered (Section 3.5). They are also used to analyze the maturity models' coverage of the dimensions and subdimensions that emerged from the keyword grouping step (Section 3.6). Furthermore, the approach's convergence can be tracked across various statistical indicators captured throughout the process.

3.2. Reverse engineering of KPIs

The reverse engineering step is designed to transform the questions and answers that are used to evaluate companies' digital maturity into a set of KPIs. This is necessary because of the considerable heterogeneity that exists in how digital maturity is evaluated [30].

During this step, four experts discussed and agreed on the meaning of the questions and answers used in the maturity models and then worked to formulate corresponding KPIs. Several KPIs could emerge from a single question. Some questions were short and covered a single concept, while others were more complex and covered several notions that are best separated from one another in preparation for the comparison step.

Moreover, the experts were asked to focus only on explicit information and to avoid possible interpretations. The objective was to reduce subjective bias due to interpretations that could shift the meaning of the KPI away from that of the original question. Finally, the experts had to ensure that the formalized KPIs were "self-contained" and included all the concepts necessary to understand them without having to refer to the context in which a question was asked. This last aspect makes it possible to focus on the KPIs, which become references, when moving on to the next steps of the framework and disregard the maturity models' questions and answers.

3.3. Keyword assignment

This step occurs once the experts have formalized the KPIs associated with a maturity model $\mathcal{M}^k$. For each KPI C_i^k, the experts identify and assign a list of $N_\mathcal{Q}^{k,i}$ keywords, which is denoted by $\mathcal{Q}_h^{k,i}$, with $h \in \left[1..N_\mathcal{Q}^{k,i}\right]$. This is a very important step, as the assigned keywords are later used to automatically compare the KPIs in light of the groups of keywords resulting from the next step (Section 3.4).

During keyword assignment, the experts have two options. They can use a keyword that has already been introduced to characterize a previously processed KPI (dashed lines in Figure 1), or they can define a new keyword if none of the previously introduced ones are suitable.

Special care must be taken to not introduce notions that are not explicit in the KPIs. To avoid this error, the experts can focus on using the important words from each KPI as keywords.

3.4. Keyword grouping with respect to dimensions and subdimensions

This step groups the keywords $\mathcal{Q}_h^{k,i}$ based on the notions they cover in order to better categorize them and start identifying similarities. The keywords are thus gathered in intermediate $\mathcal{Q}mat_i^k$ matrices. At the end of the process, once all the maturity models have been analyzed, the intermediate matrices are merged into a single $\mathcal{Q}mat$ matrix that is used for comparison (Section 3.5) and contains a total of $N_\mathcal{Q}$ keywords.

The $\mathcal{Q}mat$ matrix is designed to group together keywords that match. Each row in the matrix contains keywords that share the same notion. Thus, all the keywords within a row have close or similar meanings. It is a sparse matrix, as each notion can be characterized by many or few keywords, with rows having some columns that are more filled and others that are less filled. The experts have two options each time they assign a new keyword to a KPI. They can place it in an existing row, meaning the keyword expresses a variant of a previously identified notion, or they can add a new row to the matrix, which mean that the keyword expresses a notion that has not yet been identified by a KPI. It is important to note that each keyword has a specific meaning and appears only once in the matrix.

When integrating a new maturity model or checking the keyword matrix before the synthesis step, the experts may have to change the position of a keyword or modify its formulation. When this situation arises, the experts must ensure that the new position of the keyword in the $\mathcal{Q}mat$ matrix still corresponds to its use in the various KPIs. The objective is to verify that the modified keyword still correctly characterizes all the KPIs with which it is associated. Fortunately, the ad-hoc search tool developed in this paper speeds up this cross-verification step (Section 4).

In the literature, each maturity model is structured with dimensions and sometimes subdimensions. Thus, once all the keywords are placed in $\mathcal{Q}mat$, the experts name each row. Each row covers a notion; this structure corresponds to the concept of the subdimensions of a maturity model. The experts then reorganize the subdimensions into groups of rows that form dimensions, which the experts also name. Subsequently, the new dimensions are denoted by $\mathcal{D}_m$ (with $m \in [1..N_\mathcal{D}]$ and $N_\mathcal{D}$ being the number of dimensions). The new subdimensions are denoted by $\mathcal{SD}_j^m$ (with $j \in \left[1..N_{\mathcal{SD}}^k\right]$ and $N_{\mathcal{SD}}^k$ being the number of subdimensions in dimension m). Dimensions and subdimensions are used during the coverage analysis to better understand how thoroughly the various maturity models cover the underlying notions and concepts (Section 3.6).

3.5. Automatic matching of KPIs through keywords

Once the KPIs have been characterized by keywords, which themselves have been grouped in the keyword matrix $\mathcal{Q}mat$, the automatic matching of KPIs can take place in the common space of the keywords.

The idea is to automatically quantitatively evaluate the extent to which two KPIs, $\mathcal{C}^1_{i_1}$ and $\mathcal{C}^2_{i_2}$, of two maturity models, $\mathcal{M}^1$ and $\mathcal{M}^2$, match, i.e. how much they share similar notions. To do this, each keyword of $\mathcal{C}^1_{i_1}$ is compared with all the keywords of $\mathcal{C}^2_{i_2}$, and vice versa. Thus, for a keyword h_1 of $\mathcal{C}^1_{i_1}$, the function $\mathcal{F}_{\mathcal{Q}}$ evaluates possible matches with keywords assigned to $\mathcal{C}^2_{i_2}$:

$$\mathcal{F}_{\mathcal{Q}}\left(\mathcal{Q}^{i_1,i_2}_{h_1}\right) = \begin{cases} 0 \text{ if } h_1 \text{ is in the same row as another keyword already analyzed for } \mathcal{C}^1_{i_1} \\ 1 \text{ if } h_1 \text{ is used in } \mathcal{C}^2_{i_2} \\ 1 \text{ if } h_1 \text{ is in the same row as a keyword of } \mathcal{C}^2_{i_2} \\ 0 \text{ if no match is possible} \end{cases} \quad (1)$$

with $i_1 \in \left[1..N^1_{\mathcal{C}}\right]$ and $i_2 \in \left[1..N^2_{\mathcal{C}}\right]$.

This function $\mathcal{F}_{\mathcal{Q}}$ is repeated for all keywords of KPIs $\mathcal{C}^1_{i_1}$ and $\mathcal{C}^2_{i_2}$. Next, the number of distinct keywords $N^{k,i}_{\mathcal{Q},distinct}$belonging to each KPI $\mathcal{C}^k_i$ is counted. For instance, the number is equal to 1 if the KPI considered is assigned two keywords that are located in the same row of $\mathcal{Q}mat$. Finally, the level of match of KPIs $\mathcal{C}^1_{i_1}$ and $\mathcal{C}^2_{i_2}$ is obtained as follows:

$$\mathcal{F}_{\mathcal{M}_a}\left(\mathcal{C}^1_{i_1}, \mathcal{C}^2_{i_2}\right) = \frac{\sum_{h_1=1}^{N^{1,i_1}_{\mathcal{Q}}} \mathcal{F}_{\mathcal{Q}}\left(\mathcal{Q}^{i_1,i_2}_{h_1}\right) + \sum_{h_2=1}^{N^{2,i_2}_{\mathcal{Q}}} \mathcal{F}_{\mathcal{Q}}\left(\mathcal{Q}^{i_2,i_1}_{h_2}\right)}{N^{1,i_1}_{\mathcal{Q},distinct} + N^{2,i_2}_{\mathcal{Q},distinct}} * 10 \quad (2)$$

with $i_1 \in \left[1..N^1_{\mathcal{C}}\right]$ and $i_2 \in \left[1..N^2_{\mathcal{C}}\right]$.

The function $\mathcal{F}_{\mathcal{M}_a}$ returns a score between 0 and 10 for each pair of KPIs, with 10 being a perfect match. The results are rounded to the nearest unit and inserted into the matching matrix $\mathcal{M}_a mat$. When a score differs from 0, it means that the two KPIs have some notions in common.

As is demonstrated in Section 4, a good comparison of the maturity models considered can be obtained by analyzing the output matrix. Comparison makes it possible to evaluate the extent to which the models match, which KPIs overlap, which KPIs are not covered by the models, etc. These results are particularly interesting for the development of a new maturity model, as they make it possible to identify KPIs that can be merged or synthesized into more general ones. How the synthesis of existing maturity models can be performed is not part of this paper.

3.6. *Coverage analysis*

The term coverage indicator is introduced to characterize the extent to which a maturity model covers the dimensions formalized by the experts during the keyword grouping step (Section 3.4). Thus, to evaluate how much a model $\mathcal{M}^k$ covers a dimension $\mathcal{D}_m$ (with $m \in [1..N_{\mathcal{D}}]$ and $N_{\mathcal{D}}$ being the number of dimensions), it is necessary to count how many of its subdimensions $\mathcal{SD}^m_j$ (with $j \in \left[1..N^m_{\mathcal{SD}}\right]$ and $N^m_{\mathcal{SD}}$ being the number of subdimensions in $\mathcal{D}_m$) are covered. A subdimension is considered covered when at least one of its keywords is used by a KPI $\mathcal{C}^k_i$ of $\mathcal{M}^k$. The coverage indicator denoting the extent to which a model $\mathcal{M}^k$ covers a dimension $\mathcal{D}_m$ is computed as follows:

$$\mathcal{X}^{k,m}_{Coverage,\mathcal{D}} = \frac{\sum_{l=1}^{N^m_{\mathcal{SD}}} \delta^k_{m,j}}{N^m_{\mathcal{SD}}}, \forall k \in [1..N_{\mathcal{M}}] \text{ and } \forall m \in [1..N_{\mathcal{D}}] \quad (3)$$

with

$$\delta^k_{m,j} = \begin{cases} 1 & \text{if } \exists\, i \in \left[1..N^k_{\mathcal{C}}\right] / \exists\, h \in \left[1..N^{k,i}_{\mathcal{Q}}\right] / \mathcal{Q}^{k,i}_h \in \mathcal{SD}^m_j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The coverage indicator is computed for each maturity model and for each dimension. It makes it possible to characterize the strengths and weaknesses of the various models by highlighting how much the models cover a dimension $\mathcal{D}_m$. It also enables the relative comparison of the models with each other. As is demonstrated in Section 4, the results can be gathered in a "radar" type graph, which makes it possible to observe the maturity models' coverage of the dimensions $\mathcal{D}_m$. Experts can see which model is the most important for each new dimension.

4. Results

4.1. *Digital maturity models studied*

The proposed framework was tested on 13 state-of-the-art digital maturity models that were developed before 2020 and are presented in Table 1. Some were developed by academics, whereas others have been proposed by consulting firms either as a self-evaluation tool (marked with a star * in Table 1) or as a tool to support an audit. The models were selected considering their availability (as we needed to have access to the questions/answers used for assessment), their citation level (with models used in the state-of-the-art of other papers, and cited), and the fact that they have been proposed by experts and validated on different test cases. Nevertheless, our proposed framework can be applied to any new model that may be proposed in the coming years. Note that from now acronyms are used in place of the index i to better relate to the corresponding maturity model. For instance, the IMPULS maturity model is referred to as $\mathcal{M}^{\mathcal{IMP}}$.

Table 1. Digital maturity models compared.

k	Digital maturity model ($\mathcal{M}^k$)	Number of questions	Number of KPIs (N_C^k)	Acronym
1	Industry 4.0 Readiness Online Self-Check for Businesses [29]*	19	25	IMP
2	Industry 4.0 – Enabling Digital Operations Self-Assessment [14]*	33	33	PWC
3	ADN 4.0 – Autodiagnostic [15]*	50	50	ADN
4	Digitalomètre [31]*	25	25	BPI
5	Adoption Maturity Model (AMM) [32]	35 (in the form of KPIs)	39	SCR
6	The Digital Maturity Model 4.0 [20]	28	28	FOR
7	Maturity and Readiness Model for Industry 4.0 Strategy [21]	68	69	AKD
8	Évaluer la maturité numérique de votre entreprise [33]*	33	33	MED
9	Transformation digitale [34]*	18	18	CEG
10	Industry 4 self-assessment tool [35]*	37	39	WAR
11	Gestion et gouvernance des technologies numériques (MACH 1.6) [36]	49	51	MAC
12	Assessment Tool [37]	16 (in the form of KPIs)	16	SIN
13	SMART DIAG' - Diagnostic numérique [38]*	25	25	AGE

Each maturity model $\mathcal{M}^k$ has a different number of questions, and, as is explained in Section 3.2, the reverse engineering of KPIs is not a one-to-one process, which explains why a model's number of KPIs may differ from its number of questions. Only questions addressing the assessment of digital maturity were considered. Some approaches, such as the IMPULS [29] and PwC [14] models, include additional questions (regarding, for example, the number of employees, turnover, the field of activity) that are used to rank the companies with respect to one another so that, for instance, companies of a similar size can compare their level of digital maturity. These questions have no influence on the evaluation results and do not relate to digital maturity per se, and so they have been disregarded.

While some models, such as those of Scremin et al. [32] and the Singapore Economic Development Board (2017) [37], assess digital maturity directly through KPIs, their KPIs have been reformulated and all maturity models have undergone the reverse engineering step to arrive at a standardized set of KPIs that can then take part in the comparison process.

Finally, it is important to know that some maturity models are not independent and may explicitly refer to other models. For instance, the model proposed by Akdil et al. [21] uses some of the KPIs from the IMPULS [29] and University of Warwick [35] models, and thus the reverse engineering and keyword assignment steps have been performed just once.

To support the different steps of the proposed semi-automatic framework, several tools were implemented in Microsoft Excel by means of Visual Basic for Applications macros. The functionalities developed make it possible to quickly instantiate, search, access, and analyze all the information conveyed throughout the comparison process. For example, it is possible to highlight KPIs that use specific keywords specified by the experts, to detect when two keywords used to characterize a KPI are very similar and could therefore be consolidated, or to determine the number of times a keyword is used.

In addition, all the semi-automatic comparison framework's matrices and formulae have been implemented in such a way as to facilitate automatic calculations and the creation of graphs. The framework's flexibility will support the integration of new maturity models that may emerge in the next few years.

4.2. *KPI lists and associated keywords*

Four experts worked on reverse engineering the KPIs and assigning keywords to each KPI. These two steps are the longest parts of the whole process. Several iterations were required to ensure that the reverse engineered KPIs were not over-interpreted from the questions and answers and that they were "self-contained" (Section 3.2).

In the end, 13 lists of KPIs were obtained for a total of 451 KPIs, including few repetitions for the KPIs of maturity models that are explicitly based on other models. The distribution of KPIs among the maturity models is shown in Table 4.1. In most cases, each question leads to a single KPI. The KPIs written for the IMPULS [29] and PwC [14] models were published in full in [30]. Table 2 presents two of the IMPULS [29] model's KPIs, together with their assigned keywords.

Table 2. Examples of keywords for two of the IMPULS [29] model's KPIs.

i	KPIs (C_i^k)	Keywords ($\mathcal{Q}_h^{k,i}$)
IMP 4	Level of financial investment in the implementation of Industry 4.0 in different company sectors in the next 5 years	future investment, i4.0 implementation
IMP 5	Level of financial investment in the implementation of Industry 4.0 in different company sectors in the past 2 years	past investment, i4.0 implementation

This step gave rise to many exchanges and discussions. Overall, the number of keywords assigned to each KPI ranges from 1 to 8, with an average of 3 keywords per KPI (standard deviation of 1). More specifically, about 85% of the 451 KPIs are characterized by 2 to 4 keywords, and less than 30% of them are characterized by only 1 or 2 keywords. Less than 10% of them are characterized by a single keyword. The experts sought to find a balance between using too many keywords, with the risk of finding false-positive matches in the automatic comparison step, and using too few keywords, with the risk of missing some matches.

4.3. Keyword matrix and identification of dimensions and subdimensions

The methodology detailed in Section 3.3 resulted in 263 keywords belonging to the 451 KPIs being grouped in the keyword matrix $\mathcal{Q}mat$, as shown in Table 3. Each row of the matrix represents a

Table 3. Distribution of the 263 keywords in the keyword matrix $\mathcal{Q}mat$.

Dimension ($\mathcal{D}_m$)	Subdimension $\left(\mathcal{SD}_j^m\right)$	Keywords
Business, strategy & governance	Business model & strategy	business model, business operations, business outcomes, business requirements, commercial activity, economic benefits, revenues, strategic plan, strategic vision, value creation
	Digital awareness	i4.0 awareness, i4.0 knowledge, i4.0 openness
	Digital strategy	digital strategy, digital strategy communication, digital strategy review, external audit, i4.0 strategy, i4.0 strategy indicators
	Digital plan & roadmap	digital plan, digital plan review, digital transformation, roadmap
	Digital leadership	Chief Digital Officer, digital leader, digitalization requirements, future requirements, i4.0 requirements
	Implementation & deployment	business resources allocation, decision-making process, digital activities governance, i4.0 implementation, i4.0 projects, implementation status, IT steering process, last digital initiative, organization's i4.0 resources, project management, selection process
	Societal factors	energetic consumption, environmental benefits, social benefits
	Performance management	KPI, dashboard, decision-making monitoring, forecast monitoring, performance management, time-to-market monitoring
	Digital risk management	mitigation plan, risk management
	IP management	contracting models, external IP, IP protection
Smart portfolio & customer service	Organization's portfolio	organization's portfolio, organization's products
	Smart products & services	add-on functionalities, data-driven services, digital features, digital product portfolio, smart product, smart service, tracking
	Customization	product customization, service customization
	Configuration tools	product configuration tools, service configuration tools
	Dynamic pricing	dynamic pricing
	After-sales services	after-sales, customers assistance, products guarantee
Human resources	Management modes	collective intelligence, corporate culture, employee efficiency monitoring, employee workload monitoring, management modes, manager-employee interactions, right to disconnect
	Managers' involvement	managers' involvement, managers' support
	Hiring	hiring process, human resources
	Working modes	Internet, intranet, remote working, resource mobility
	Change management	adaptation capacity, change management
	Skills & expertise	employees' skills, i4.0 expertise, i4.0 expertise organization, i4.0 organization's capabilities, managers' expertise
	Training	digital training, employees' training, skills acquisition
Digital security & compliance	Digital security	access control, cybersecurity, data protection, encryption, IT security, password management
	Digital compliance	compliance management, digital charter, digital compliance policy
Finance	Investment	future investment, i4.0 investment, past investment
	Funding	funding sources
	Finance	accounting tool, cash flow, cost analysis, financial management, profitability analysis
Innovation & knowledge management	Innovation management	innovation management
	Technological watch	benchmarking, competitive intelligence, technological watch
	Knowledge	knowledge capitalization, knowledge management
	Traceability	order fulfillment rate traceability, traceability means

(*Continued*)

Table 3. (Continued).

Dimension ($\mathcal{D}_m$)	Subdimension $\left(\mathcal{SD}_j^m\right)$	Keywords
IT & software tools	IT technologies	additive manufacturing (AM), digital technologies, Internet of things (IoT), IT technologies, machine-to-machine (M2M), new technologies, RFID
	Infrastructure	equipment infrastructure, facility equipment, IT architecture, IT organization, leading system
	IT integration	communication functionality, interoperability functionality, IT integration, tool interfacing
	Cloud services	cloud computing, cloud services, cloud-based software
	Software tools	computer-aided design (CAD), computer-aided engineering (CAE), computer-aided manufacturing (CAM), computer-aided production management (CAPM), computer-aided maintenance management systems (CMMS), customer relationship management (CRM), electronic document management (EDM), enterprise resource planning (ERP), human resource information system (HRIS), human resource management systems (HRMS), enterprise application integration (EAI), manufacturing execution systems (MES), software tools, software tools' use
	Upgradability & maintenance	maintenance, upgradability
Digital data	Digital twin	digital factory, digital twin
	Data as an asset	customer data, customer insight, digital assets, digital data, financial data, marketing data, massive data, operations data, product data, production data, usage phase data
	Data acquisition & processing	data aggregation, data analysis, data architecture, data collection, data gathering, data interpretation, data processing, data quality check, data storage, data structuring
	Data usage	data-based decision-making, data monitoring, data usage
	Data sharing	data communication, data sharing, data sharing with partners, external data sharing, external information sharing, internal data sharing, internal information sharing
	Data recovery	data backup, data recovery
Smart manufacturing	Autonomous workpiece	autonomous workpiece, self-guiding capacities
	Production digitalization	autonomous response, closed-loop systems, control functionality, manufacturing processes, production, production automation, production control, production equipment digitalization, production monitoring, production processes, quality control checks, self-diagnosing, system uptime monitoring, working instruction digitalization
	Agility	agility, dynamic response, flexibility, just-in-time production, lead times, production changes, real-time response
	Inventory	inventory digitalization, inventory management
Internal value chain digitalization	Internal activities integration	internal activities integration, internal processes, internal value chain digitalization, organization's sectors
	Internal collaboration	internal collaboration, internal exchanges
	Lifecycle simulation	product development phase, product lifecycle phases, product lifecycle simulation, production phase, simulation phase
	Development process	product development process, product improvement process, service development process
External value chain digitalization	External activities integration	customer integration, end-to-end integration, external activities integration, external processes, external value chain digitalization, partners' integration
	Customer experience	client interactions, customer experience, customer feedback, customer interactions
	External collaboration	external collaboration, external exchanges
	Supply chain & logistics	logistics, supply chain
Sales & marketing	Selling & purchasing	billing tool, ordering process, quoting process, sales channels, sales forces, sales tools, selling
	Communication & marketing	communication channels, digital footprint, marketing channels, marketing tools, promotion campaign, responsiveness, social media, social networks, trend analysis, website

subdimension, and the grouping of several subdimensions corresponds to a dimension. As a result, the 263 keywords are distributed over 12 dimensions $\mathcal{D}_m$ (with $m \in [1..N_{\mathcal{D}}]$ and $N_{\mathcal{D}} = 12$) and 58 subdimensions $\mathcal{SD}_j^m$ (with $j \in \left[1..N_{\mathcal{SD}}^m\right]$ and $\sum_{m=1}^{N_{\mathcal{D}}} N_{\mathcal{SD}}^m = 58$).

This table is the result of lengthy discussions to reach an agreement on the different concepts, what they cover and the proper formulation of the many keywords, subdimensions and dimensions. The matrix evolved throughout the comparison process, and many changes were made when new maturity models, KPIs or keywords were considered. Section 4.5 discusses the convergence of the comparison process, which was made possible by carefully tracking the evolution of the number of keywords and subdimensions. The resulting dimensions and subdimensions partially correspond to the structures proposed by each of the maturity models studied. The new semi-automatic approach was used to study each maturity model's coverage of the dimensions (Section 4.6).

	KPI IMPULS																								
KPI PwC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																			5	3		8			
2																			10	4		7			
3															4										
4																				4	4		5		
5									3								3	4			3	4	3		
6													4							3	3				
7																									
8													4							3	3				
9			5							4	3														
10																									
11									7								3	4		6	9		7		
12									3				3							5	5				
13						5				5		5				5				3	3		4		
14							4	3						5	10										
15	3			3	3	3	3	2		3	3	3	3			3				2	2				
16							4	3	4					5	5										
17													5							7	7		4		
18							4	3			4													5	
19							4	3		5	4			5	5										
20							3	3	7		3						3	4		3	6		3		
21			4															3				3			
22							3	3			3				3							3			
23							4	3			4		5							3	3				
24																									
25																									
26							3	3						4	4				5	3		4			
27						3				3		3	3			3				3	3				
28																	5								
29													3							3	3				
30									3								3	4			3	4	3		
31			7													5								5	5
32																5								5	5
33													4			4				3	3			4	4

Figure 2. Matching matrix $\mathcal{M}_a mat^{IMP,PWC}$ indicating the match scores of pairs of KPIs from the IMPULS and PwC maturity models (a blank indicates no match, and 10 indicates a perfect match).

4.4. Automatic computation of the matching matrices $\mathcal{M}_a mat$

Once the keyword matrix $\mathcal{Q}mat$ is complete, the KPIs of the 13 maturity models can then be compared with each other using the formula introduced in Section 3.5. Figure 2 shows the result of the comparison of the IMPULS and PwC maturity models. It consists of a matrix $\mathcal{M}_a mat^{IMP,PWC}$ of size $(N_\mathcal{C}^{PWC} \times N_\mathcal{C}^{IMP})$ gathering the KPI pairs' match scores on a scale from 0 to 10 (with 0 being left blank and indicating no match, and 10 indicating a perfect match). Scores between 0 and 10 refer to a partial match and are highlighted in yellow in Figure 2.

Overall, the comparison of the $N_\mathcal{M} = 13$ maturity models generated a total of $N_\mathcal{M}(N_\mathcal{M} - 1)/2 = 78$ comparison matrices $\mathcal{M}_a mat^{k_1,k_2}$ that satisfied the following property:

$$\mathcal{M}_a mat^{k_1,k_2} = \mathcal{M}_a mat^{k_2,k_1^{\,T}}, \forall (k_1, k_2 \neq k_1) \in [1..N_\mathcal{M}]^2 \tag{5}$$

Following this process, the $N_\mathcal{C} = 451$ KPIs were compared automatically with one another. The number of comparisons performed can be computed as follows:

$$\sum_{k_1=1}^{N_\mathcal{M}-1} \sum_{k_2=k_1+1}^{N_\mathcal{M}} N_\mathcal{C}^{k_1} \times N_\mathcal{C}^{k_2} \tag{6}$$

Thus, for the maturity models considered, a total 92,540 comparisons were performed automatically, which is a feat that would not have been possible to manage manually. The proposed framework not only performs comparisons automatically, it also assures that any changes made to the assigned keywords or the keyword matrix $\mathcal{Q}mat$ are automatically reflected in the results. This is particularly powerful because it avoids spending time on comparisons that result in no match and also ensures more uniform and equitable treatment of the KPI pairs step after step, thus limiting the risk of errors.

Nevertheless, the results obtained using the semi-automatic framework were validated by comparing some of them with those obtained following the more traditional manual comparison of KPI pairs. More specifically, the semi-automatic framework was tuned based on the comparison of the IMPULS [29] and PwC [14] models. To do this, the experts compared the automatically generated matching matrix to a manually generated matrix that was obtained by manually evaluating the level of matching of $N_C^{PWC} \times N_C^{IMP} = 825$ KPI pairs. During this step, the keyword matrix $\mathcal{Q}mat$ was adjusted and tuned so that the results of the semi-automatic framework fit the results obtained by the experts [30].

The resulting matrices are sparse, with less than a third of the pairs having non-zero match values. If one looks at the rows and columns they can directly identify which KPIs are completely distinct from the others, thus highlighting some aspects that are not treated by either of the maturity models compared. For instance, the KPI IMP 2 does not match any of the PWC model's KPIs, which results in an empty column in the matrix shown in Figure 2. Moreover, a match value equal to 10 indicates a KPI pair that matches perfectly. For instance, IMP 9 and PWC 2, and IMP 15 and PWC 14 are perfect matches. Partial matches – those having a value greater than zero and less than 10 – require a subsequent manual comparison step.

Finally, this understanding of what the models do and do not cover is particularly interesting for the development of new maturity models based on the synthesis of existing ones. However, the development of a new maturity model is not part of this paper. The proposed approach also makes it possible to verify whether a maturity model self-covers. Indeed, comparing a model to itself makes it possible to create a square matching matrix $\mathcal{M}_\dashv mat^{k,k}$ that highlights how much model k's KPIs match one another. A maturity model can be self-covering to allow for cross-checking of the maturity assessment.

4.5. Convergence of the keyword and subdimension lists

The processing of the 13 maturity models considered resulted in 263 keywords, 58 subdimensions and 12 Chronological evolution of the number dimensions, and involved all the stages of the semi-automatic comparison framework. This section studies how the new concepts, i.e. subdimensions, appeared, step by step.

First, Figure 3 shows how the keywords appeared in the matrix $\mathcal{Q}mat$ as the maturity models were processed by the semi-automatic framework. Each of the maturity models has its own vocabulary and areas of interest, which resulted in the addition of new keywords. The x-axis represents the sequence in which the maturity models were analyzed in this study, as they were processed in their chronological order, i.e. IMP first and AGE last. It can be observed that about 60% of the keywords (165 of the 263 keywords) come from the first three models (IMP, PWC and ADN). The other maturity models generated about 10 new keywords each, on average. Thus, the number of keywords stabilizes step by step.

However, the observation of keywords alone does not suffice to validate the convergence of the results. A new keyword can lead to the identification of either a variant of an existing concept or a new concept and thus a new subdimension. Figure 3 shows the evolution of the number of subdimensions by maturity model analyzed. As in Figure 2, the x-axis can be viewed as the model analysis sequence.

The subdimensions graph follows the same evolution as the keywords graph before it. In this case, 90% of the concepts, which became subdimensions, come from the first three maturity models (IMP, PWC and ADN); the remaining 10 maturity models identified only six concepts. After the MED model's analysis, new keywords were added to the $\mathcal{Q}mat$ matrix (Figure 3), but no new concepts were identified (Figure 4).

This convergence of keywords and subdimensions validates the relevance of the framework. The convergence observed indicates that there is significant redundancy between the concepts covered by the maturity models analyzed. Since both graphs aggregate the appearance of new concepts step by step, the results of this cumulative study

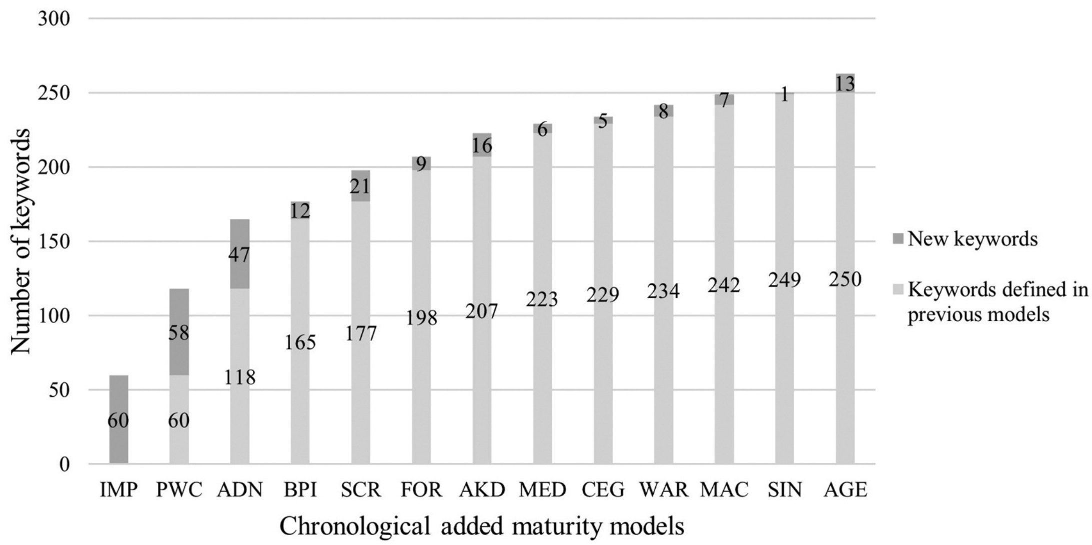


Figure 3. Chronological evolution of the number of keywords.

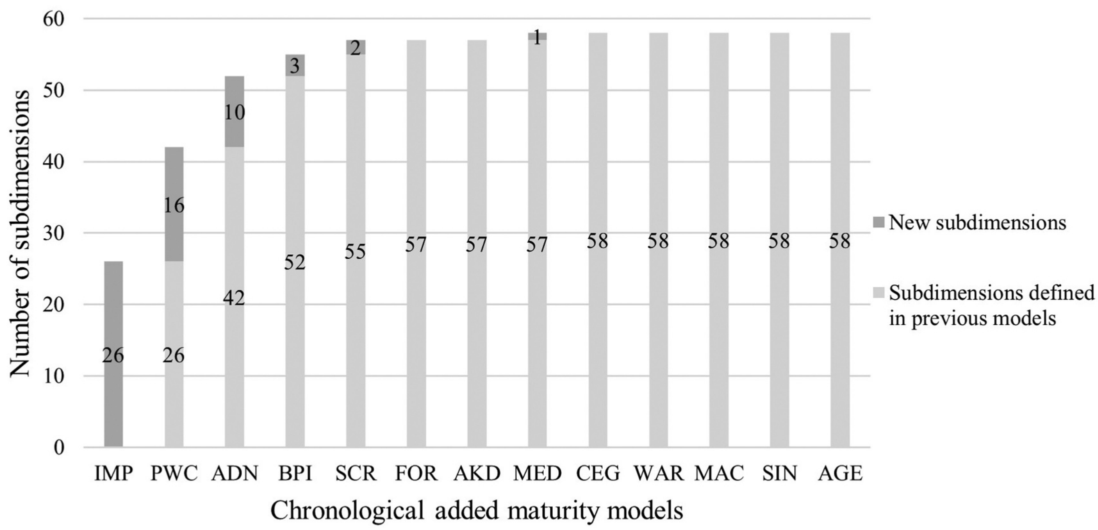


Figure 4. Chronological evolution of the number of subdimensions.

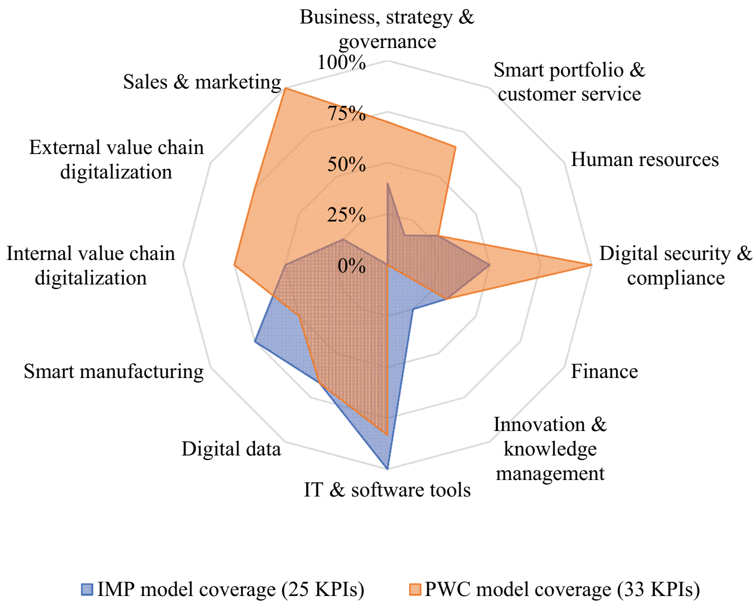


Figure 5. Comparison of the coverage of the IMPULS and PwC models according to the 12 dimensions listed in Table 3.

would be similar if the maturity models were processed in a different order.

4.6. Coverage of dimensions by the digital maturity models

Not all digital maturity models address the same aspects of digital transformation. This section compares the 13 models considered and highlights their specificities. The coverage indicator $\mathcal{X}^{k,m}_{Coverage,\mathcal{D}}$, which denoting the extent to which a model $\mathcal{M}^k$ covers a dimension $\mathcal{D}_m$, can be used for this comparison (Section 3.6). The 12 dimensions mentioned in Figure 5 are those listed in Table 3. For a given dimension, the percentage corresponds to the number of subdimensions that are covered by at least one of the KPIs of the maturity model considered.

A comparison of the IMP and PWC models' coverage is provided in Figure 5. The IMP model focuses on the dimensions "IT & software tools" and "Smart manufacturing". The PWC model focuses on the dimension "IT & software tools" too, but also on the dimensions "Internal value chain digitalization", "External value chain digitalization", "Digital security & compliance" and "Sales & marketing". These maturity models are complementary. Some dimensions (e.g. "Sales & marketing") are not assessed at all by the IMP model, whereas the PWC model covers them extensively. However, if only these two models had been selected for the study, the list of

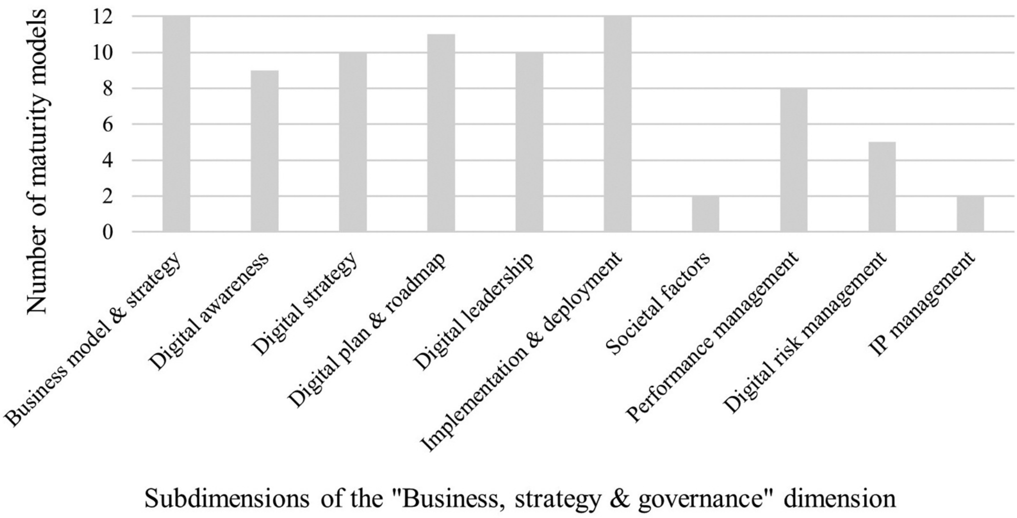


Figure 6. Coverage of the subdimensions of the "Business, strategy & governance" dimension by the 13 maturity models.

KPIs would not completely cover the dimensions "Human resources" or "Innovation & knowledge management". Indeed, some concepts were identified in other models, hence the need to consider the maximum possible number of available digital maturity models. These are therefore two limitations of the IMP and PWC models.

The coverage graphs and specificities of all 13 maturity models are presented in the Appendix. The AKD and WAR models have the greatest coverage, but neither covers all dimensions, which reveals the need to develop a new maturity model that does. While that would be a worthy goal, it is not part of this paper. The CEG model, which uses fewer than 20 KPIs, has the smallest coverage and focuses on just 2 main dimensions: "Human resources" and "Sales & marketing." The radar charts show that most maturity models fully cover at least one dimension. The FOR model is the only one that does not cover any dimension 100%. Instead, it focuses on several dimensions without covering them completely. These results make it possible to know the main scope and specificities of each model.

Radar charts such as those shown in Figure 5 or the Appendix thus allow companies to determine the model(s) that are most appropriate for them based on their development choices. For example, if a company wants to strengthen its IT & software tools, the IMP and PWC models are the ones that best cover this dimension, while the CEG and MED models do not cover it at all. If the priority is Digital security & compliance, the company should choose a model such as PWC or MAC because they cover that dimension 100%. The IMP and FOR models should be disregarded because they either do not cover this dimension or cover it only to a limited extent.

Finally, it is possible to note from the radar charts that all dimensions are covered 100% by at least one maturity model. The fact that each concept is evaluated by multiple models from different points of view and that each of our dimensions/subdimensions is supported by at least two models helps to validate the dimensions and subdimensions identified. The only exception is for "Business, strategy & governance" (see the radar charts in the Appendix), which is the broadest dimension (10 subdimensions). None of the models cover exactly the same subdimensions. Thus, although no single model covers all the subdimensions of the "Business, strategy & governance" dimension, all of its subdimensions are covered by at least two maturity models (Figure 6).

5. Discussion

This work was carried out by several experts during collaborative meetings. Thus, several competencies and fields of expertise, such as Industry 4.0, product lifecycle management (PLM), innovation management, customer needs assessment and production management, were gathered to develop and use this systematic comparison framework.

The literature review showed that digital maturity models propose various levels to evaluate an organization's progress in its digital transformation. Many models propose an overall result corresponding to a level and also give a level for each dimension. For instance, the IMPULS [29] model defines six levels ranging from "Outsider" to "Top performer." The objective of this paper was to develop an approach to compare KPIs from maturity models identified in the literature. The comparison of the KPIs was done automatically using keywords. This made it possible to focus the comparison on the content of the KPIs rather than their formulation. Thus, an initial analysis of the literature showed that the scope of each model was different, i.e. none of the models cover the dimensions in the same way. The semi-automatic comparison framework that was put in place provided a picture of the scope of each model using the common space of the keywords (see the radar charts in the Appendix). An additional step could be conducted to compare the different levels of classification used to assess companies in their digital transformation. This comparison could be conducted between the KPIs that have a link to one another.

The results obtained by answering the maturity assessment questionnaires provide some insights into how to improve one's digital maturity level. In addition, analyzing the proposed answers can help to understand the developments to be achieved. However, these details are not sufficient to define a concrete action plan or digital strategy. The creation of a digital roadmap is complex and should be considered as future work.

Following the proposed framework, 263 keywords were identified to characterize 451 KPIs distributed over 13 digital maturity models. The underlying concepts were structured and organized around 12 dimensions and 58 subdimensions. It is possible to understand from the matching matrices developed the similarities and differences between the maturity models evaluated. The radar charts in the Appendix show which dimensions are and are not covered by which maturity models. These results and could help to develop a new maturity model that would synthesize and aggregate all the notions. Indeed, some of the many pairs of KPIs that were compared automatically match perfectly (10/10) and could therefore be merged. Approaches to simplify the full list of 451 KPIs are part of projected future work.

The proposed framework has the advantage that new maturity models (those that have not been considered in this study as well as future models) can easily be integrated in it. Using this approach would make it possible to obtain a picture of the coverage of these models and to detect the emergence of new dimensions or subdimensions.

The proposed approach makes it possible to identify the differences and similarities between digital maturity models in order to have a good picture of their coverage of dimensions (defined from the existing models). End users can use the radar charts in the Appendix to help them decide which model best meets their needs.

As future work and to help companies in their digital transformation, an approach should be developed to compile all the dimensions identified in one model. This model would contain all the KPIs currently present in the literature. Thus, end users would be able to consider only the dimensions necessary for their digital transformation.

This semi-automatic comparison framework has some limitations. As mentioned above, it requires the participation of several experts to reverse engineer the KPIs, and assign and classify the keywords. These manual steps are time-consuming and involve some subjectivity. One way to improve this method would be to automate the manual steps using artificial intelligence (e.g. text mining methods).

Finally, this study does not determine whether any KPIs are omitted from the digital assessment. The objective of this work was not to identify potential failings of existing digital maturity models, but rather to compare the models. It would be useful to carry out a new study with industrial partners and to develop new maturity models that consider failings.

6. Concluding remarks

This paper introduced a new automatic quantitative approach to compare state-of-the-art digital maturity models that is capable of being used with a wide variety of heterogeneous models. It avoids tedious data processing and is able to rapidly process several hundred thousand comparisons. This performance has been made possible by a paradigm shift that is based on the idea of moving comparisons into a common space where each maturity model can be efficiently represented. In this case, the comparisons were performed directly in the space of the keywords used to characterize the KPIs that were reverse-engineered from the digital maturity models considered. Matches were encoded in a keyword matrix that was then used to automatically compute the level of matching of KPI pairs. This idea could be used in other contexts and to compare other objects by instrumenting traditionally manual comparisons and making them more quantitative and automatic.

This semi-automatic framework also made it possible to compute coverage graphs that clearly illustrate the extent to which a digital maturity model covers the dimensions and subdimensions it considers. These quantitative results help when it comes to comparing models to gain a better understanding of their specificities and the concepts they do not cover. This systematic comparison of digital maturity assessment models will provide a new perspective to companies needing to identify the digital maturity assessment model(s) that best suit their intended evaluation objectives.

The framework still requires several experts to take part in the manual steps – the reverse engineering of KPIs and the attribution of keywords –

that have not yet been automated. However, these steps could be instrumented to a greater degree through syntactic analysis coupled with a machine learning-based keyword estimation method. A database of the maturity models, KPIs and keywords used is already available here and could directly serve to support this smart comparison development process.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Bruno Cognet received his Engineering degree in Mechanical, specialising in design and development from the Institut National des Sciences Appliquées (Lyon, France), in 2020. At the same time, he completed a research Master's degree (MSc A) from 2018 to 2020 at the École de Technologie Supérieure (Montreal, Canada). This research led him to focus on the digital transformation of companies and their evaluation through maturity indicators.

Jean-Philippe Pernot is a Full Professor of Mechanical Engineering at Arts et Métiers Institute of Technology, France. For more than 20 years, he has acquired a solid experience in coordinating, managing and contributing to national and international research projects. He is currently a member of regional and national boards, and he is an expert for the European Commission for the evaluation of HORIZON projects. His research focuses on computer-aided design, geometric modeling, reverse engineering and machine learning to support the digital transformation in the scope of the Industry 4.0.

Louis Rivest became a Professor at the Ecole de Technologie Superieure in Montreal, Canada, after spending a few years in the aerospace industry. He obtained his PhD from the École Polytechnique de Montréal, in 1993, and Bachelor in Mechanical Engineering, in 1988. His research centres on themodels, methods, tools and processes supporting product development. His teaching and research activities thus relate to CAD, PLM and Digital transformation.

Christophe Danjou became a Professor at École Polytechnique de Montréal, Canada, in 2018, after completing a post-doctoral fellowship at École de Technologie Supérieure in Montreal. Previously he received his PhD in Advanced Mechanics in 2015, and an Engineering degree in MechanicalSystems in 2012 from the Université de Technologie de Compiègne, France. His research and teaching focus on the implementation and adoption of Industry 4.0 in factories and enterprises.

ORCID

Jean-Philippe Pernot http://orcid.org/0000-0002-9061-2937
Louis Rivest http://orcid.org/0000-0003-0112-0090
Christophe Danjou http://orcid.org/0000-0002-9575-0087

References

[1] Taisch M, Arena DN, Gorobtcova P, et al. The 2018 World Manufacturing Forum Report, "Recommendations for the Future of Manufacturing." 2018. [Retrieved 20 January 2020]. Available from: https://worldmanufacturing.org/wp-content/uploads/WorldManufacturingFoundation2018-REPORT-Recommendations-for-The-Future-of-Manufacturing.pdf .
[2] Schwer K, Hitz C, Wyss R, et al. Digital maturity variables and their impact on the enterprise architecture layers. Probl Perspect Managproblems And Perspectives In Management. 2018;16(4):141–154. doi: 10.21256/zhaw-3303
[3] Rossmann A Digital maturity: Conceptualization and measurement model. San Francisco, USA; 2018. p. 9. Available from: https://aisel.aisnet.org/icis2018/governance/Presentations/8.
[4] Schumacher A, Erol S, Sihn W. A maturity model for assessing industry 4.0 readiness and maturity of manufacturing enterprises. Procedia CIRP. 2016;52:161–166. doi: 10.1016/j.procir.2016.07.040
[5] Simpson JA, Weiner ESC. The Oxford English dictionary. Oxford, UK: Clarendon Press; 1989.
[6] Westerman G, Bonnet D, McAfee A. Leading digital: turning technology into business transformation. Boston, USA: Harvard Business Press; 2014.
[7] Hizam-Hanafiah M, Soomro MA, Abdullah NL. Industry 4.0 readiness models: a systematic literature review of model dimensions. Inform. 2020;11(7):364. doi: 10.3390/info11070364
[8] Mittal S, Khan MA, Romero D, et al. A critical review of smart manufacturing & industry 4.0 maturity models: implications for Small and Medium-sized Enterprises (SMEs). J Manuf Syst. 2018;49:194–214. doi: 10.1016/j.jmsy.2018.10.005
[9] Wendler R. The maturity of maturity model research: a systematic mapping study. Inf Software Technol. 2012;54(12):1317–1339. doi: 10.1016/j.infsof.2012.07.007
[10] Sassanelli C, Rossi M, Terzi S. Evaluating the Smart Readiness and Maturity of Manufacturing Companies Along the Product Development Process. In: Fortin C, Rivest L Bernard A, editors. Product Lifecycle Management in the Digital Twin Era. Cham, Switzerland: Springer International Publishing; 2019. pp. 72–81. doi:10.1007/978-3-030-42250-9_7
[11] Benedict N, Smithburger P, Donihi A, et al. Blended simulation progress testing for assessment of practice readiness. Am J Pharm Educ. 2017;81(1):81. doi: 10.5688/ajpe81114
[12] Lichtblau K, Stich V, Bertenrath R, et al. IMPULS - Industrie 4.0 readiness. [Internet]. 2015 [cited 2019 Jul 22]. p. 78. Available from: https://industrie40.vdma.org/documents/.
[13] Felch V, Asdecker B, Sucky E Maturity models in the age of industry 4.0 – do the available models correspond to the needs of business practice? Proceedings of the Annual Hawaii International Conference on System Sciences. 2019; doi: 10.24251/hicss.2019.620
[14] PricewaterhouseCoopers. Industry 4.0 - self assessment [Internet]. 2015 [cited 2019 Jul 22]. Available from: https://i4-0-self-assessment.pwc.nl/i40/landing/.

[15] Ministère de l'Économie et de l'Innovation Québec. ADN 4.0 – autodiagnostic [Internet]. Stiq. 2017 [cited 2019 Jul 22]. Available from: https://www.stiq.com/2017/12/14/adn-4-0-autodiagnostic/ .

[16] Oztemel E, Gursev S. Literature review of industry 4.0 and related technologies. J Intell Manuf. 2020;31(1):127–182. doi: 10.1007/s10845-018-1433-8

[17] Samaranayake P, Ramanathan K, Laosirihongthong T Implementing industry 4.0 — a technological readiness perspective. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2017. p. 529–533. doi: 10.1109/ieem.2017.8289947

[18] Geissbauer R, Vedso J, Schrauf S Industry 4.0: Building the digital enterprise. [Internet].2016 [cited 2019 Jul 22]. p. 36. Available from: https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf.

[19] Cognet B, Pernot J-P, Rivest L. Towards a Novel Comparison Framework of Digital Maturity Assessment Models. In: Fortin C, Rivest L Bernard A, editors. Product Lifecycle Management in the Digital Twin Era. Cham, Switzerland: Springer International Publishing; 2019. pp. 58–71. doi:10.1007/978-3-030-42250-9_6

[20] Gill M, VanBoskirk S The digital maturity model 4.0 [Internet]. Cambridge, USA: Forrester; 2016 [cited 2019 Jul 19]. p. 17. Available from: https://forrester.nitro-digital.com/pdf/Forrester-s%20Digital%20Maturity%20Model%204.0.pdf.

[21] Akdil KY, Ustundag A, Cevikcan E. Maturity and readiness model for industry 4.0 strategy. In: Ustundag A Cevikcan E, editors. Industry 40: managing the digital transformation. Cham (Switzerland): Springer International Publishing; 2018. pp. 61–94. doi: 10.1007/978-3-319-57870-5_4.

[22] Ambrosio da Silva I, César Macêdo Barbalho S, Adam T, et al. Industry 4.0 maturity models: A bibliometric study of scientific articles from 2001 to 2018. 26th EurOMA Conference - Operations Adding Value to Society. Helsinki (Finland); 2019. p. 1829–1838.

[23] Axmann B, Harmoko H. Industry 4.0 readiness assessment: comparison of tools and introduction of new tool for SME. Teh glas (Online). 2020;14(2):212–217. doi: 10.31803/tg-20200523195016

[24] Castelo-Branco I, Oliveira T, Simões-Coelho P, et al. Measuring the fourth industrial revolution through the industry 4.0 lens: the relevance of resources, capabilities and the value chain. Comput Ind. 2022;138:103639. doi: 10.1016/j.compind.2022.103639

[25] The Open Group. ArchiMate® 3.1 Specification [Internet]. 2018 [cited 2019 Dec 9]. Available from: https://pubs.opengroup.org/architecture/archimate3-doc/.

[26] Fink A. Conducting research literature reviews: from the internet to paper. 4th ed. USA: University of California; 2004.

[27] Dikhanbayeva D, Shaikholla S, Suleiman Z, et al. Assessment of industry 4.0 maturity models by design principles. Sustain. 2020;12(23):9927. doi: 10.3390/su12239927

[28] Hermann M, Pentek T, Otto B Design principles for industrie 4.0 scenarios. 2016 49th Hawaii International Conference on System Sciences (HICSS). 2016. p. 3928–3937. doi: 10.1109/HICSS.2016.488

[29] IMPULS. Industry 4.0 readiness - online self-check for businesses [Internet]. 2015 [cited 2019 Jul 22]. Available from: https://www.industrie40-readiness.de/?lang=en.

[30] Cognet B, Pernot J-P, Rivest L, et al. Digital maturity models: comparing manual and semi-automatic similarity assessment frameworks. Intern J Prod Lifecycle Manag. 2021;13(4):291–316. doi: 10.1504/ijplm.2021.119524

[31] Bpi France. Evaluez votre maturité digitale avec le « Digitalomètre »! | Bpifrance servir l'avenir [Internet]. 2018 [cited 2019 Jul 22]. Available from: https://www.bpifrance.fr/A-la-une/Dossiers/Bpifrance-accompagne-la-transformation-digitale-des-entreprises/Evaluez-votre-maturite-digitale-avec-le-Digitalometre-!-39136.

[32] Scremin L, Armellini F, Brun A, et al. Towards a framework for assessing the maturity of manufacturing companies in industry 4.0 adoption. Analyzing The Impacts Of Industry 40 In Modern Business Environments. 2018;224–254. doi: 10.4018/978-1-5225-3468-6.ch012

[33] MEDEF France. Evaluez la maturité numérique de votre entreprise! [Internet]. 2016 [cited 2019 Mar 13]. Available from: http://www.diag-numerique.fr/quiz.

[34] Cegos. Transformation digitale : vos équipes sont-elles dans la course? [Internet]. 2017 [cited 2019 Mar 13]. Available from: https://www.cegos.fr/actualites/news/transformation-digitale-vos-equipes-sont-elles-dans-la-course.

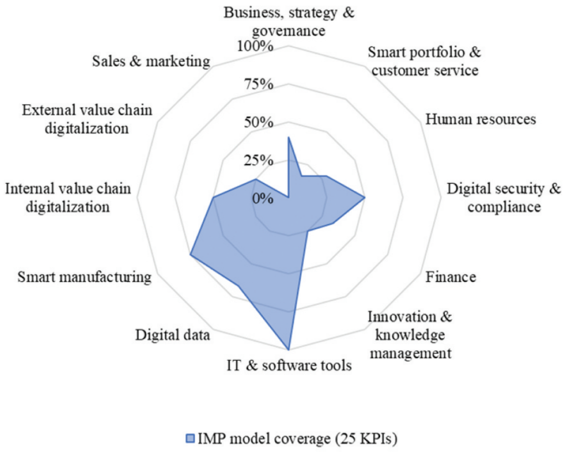
[35] University of Warwick. Industry 4 self-assessment tool [Internet]. 2017 [cited 2020 Jun 6]. Available from: https://warwickwmg.eu.qualtrics.com/jfe/form/SV_7O3ovIWlTCu90uF.

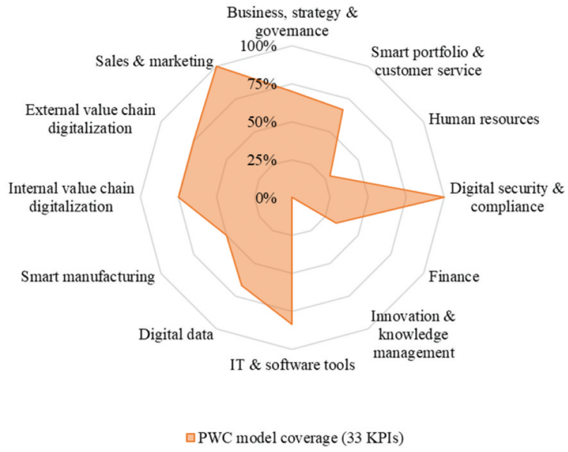
[36] Aéro Montréal. Initiative MACH [Internet]. 2011 [cited 2019 Jul 22]. Available from: https://www.aeromontreal.ca/mach-initiative.html.

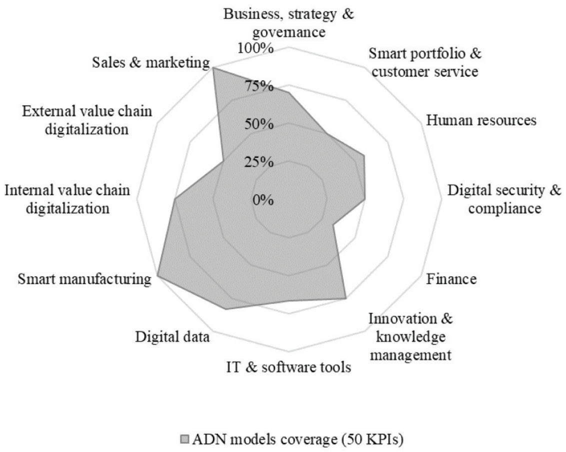
[37] Singapore Economic Development Board. The Singapore smart industry readiness index [Internet]. Singapore: Singapore Economic Development Board; 2017 [cited 2019 Apr 24]. p. 48 Available from: https://www.gov.sg/~/sgpcmedia/media_releases/edb/press_release/P-20171113-1/attachment/The%20Singapore%20Smart%20Industry%20Readiness%20Index%20-%20Whitepaper_final.pdf.

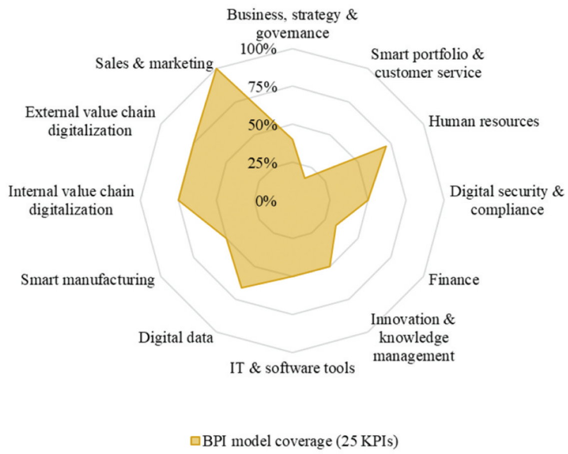
[38] AGEFOS PME. SMART DIAG' - diagnostic numérique [Internet]. 2018 [cited 2019 Jul 22]. Available from: https://www.agefos-smartdiag.com/.

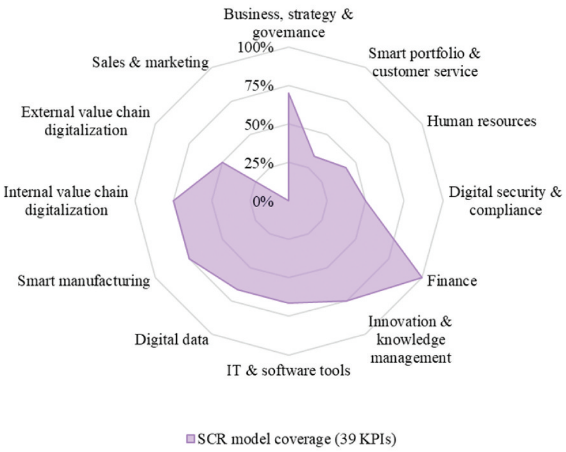
Appendix: Coverage of dimensions by the 13 maturity models compared

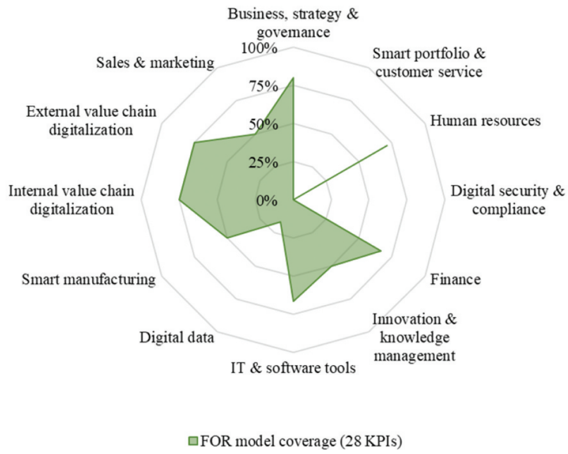


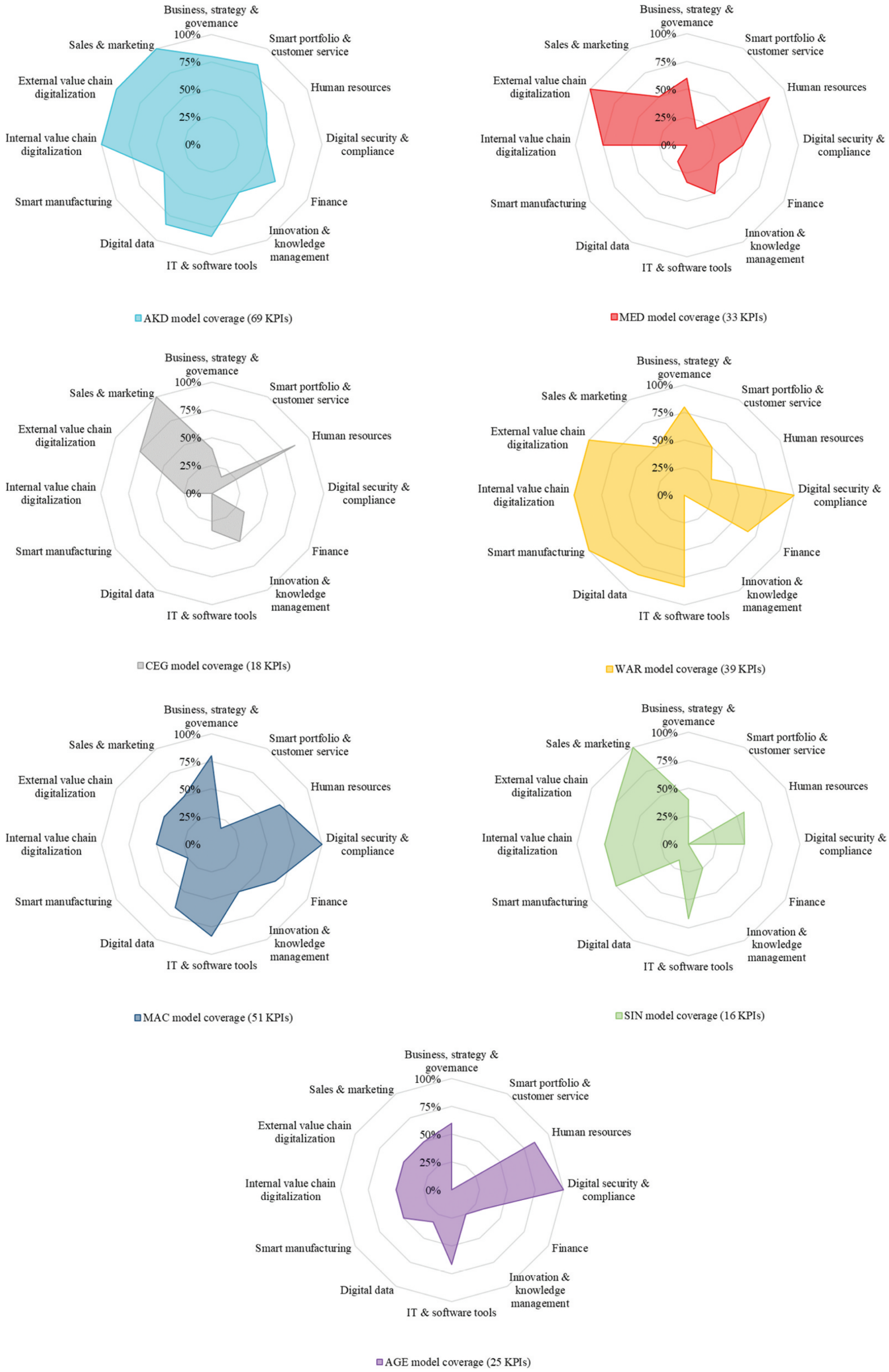











Business, strategy & governance
Smart portfolio & customer service
Human resources
Digital security & compliance
Finance
Innovation & knowledge management
IT & software tools
Digital data
Smart manufacturing
Internal value chain digitalization
External value chain digitalization
Sales & marketing
100%
75%
50%
25%
0%
AKD model coverage (69 KPIs)
MED model coverage (33 KPIs)
CEG model coverage (18 KPIs)
WAR model coverage (39 KPIs)
MAC model coverage (51 KPIs)
SIN model coverage (16 KPIs)
AGE model coverage (25 KPIs)