SUPPLEMENTARY MATERIALS

SM1: The applied TM process

Overall, the applied topic modeling process was the follows:

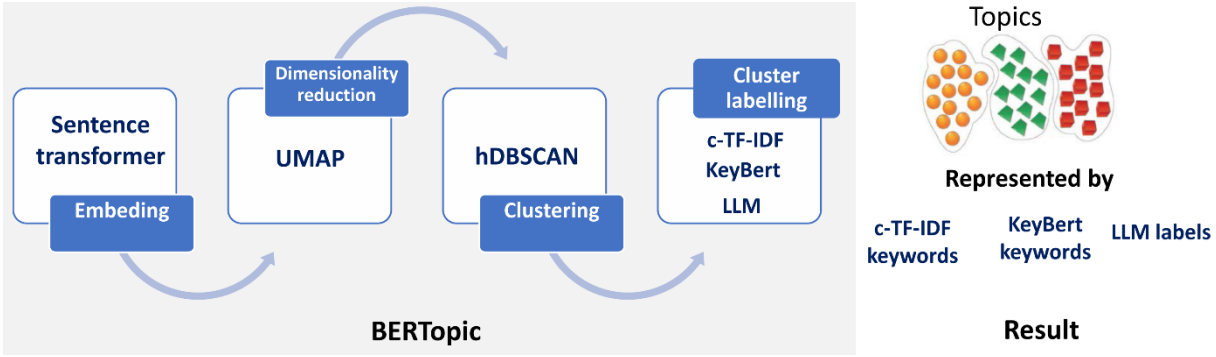
- *First (Embedings)*, the interview texts were converted using the pre-trained sentence transformer language model "all-MiniLM-L12-v2" to a 384-dimensional vector space.
- *Second (Dimension reduction)*, we then decreased these 384 dimensions to optimize the clustering process. We employed the Uniform Manifold Approximation and Projection (UMAP) method, which, based on published research (Grootendorst, 2022; Thompson & Mimno, 2020), better preserves the local and global structure of the corpus than Principal Component Analysis, which is arguably the most widely used technique for dimension reduction. Based on the results of the related literature, in our research, UMAP uses the five nearest neighbours to prefer local structures.
- *Third (Clustering)*, the reduced dimensional embeddings were clustered using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm, which defines clusters of different densities. This algorithm allows noise to be modelled as an outlier, preventing unrelated document parts from being assigned to any cluster and improving topic identification. We set the minimum size of clusters to 45.
- *Fourth (Cluster labelling)*, we used quantized LLM technology, the KeyBERT[1] extraction technique (Grootendors, 2020; Sammet & Krestel, 2023), and the class-based frequency-inverse document frequency approach (c-TF-IDF[2]) to characterize the resulting clusters/topics[3]. The technique of quantized LLM was initially employed to improve the representation of the topics (Grootendorst, 2022). Quantization is crucial in the utilisation of LLMs. It involves reducing the precision of the model's weights by assigning smaller approximations, such as 4-bit or 8-bit values, instead of the original 32-bit floating points. While there may be a slight reduction in accuracy, this approach effectively reduces the memory requirements of the model. This study utilized the pre-trained language model "OpenHermes-2.5-Mistral-7B-GGUF" and the LlamaCPP representation model (Betlen, 2023). A prompt has been established for LLM to utilize when developing topic labels. These labels are derived via clustering comments pertinent to each topic.

[1] KeyBERT is an easy-to-use keyword extraction technique that leverages BERT embeddings to create keywords and keyphrases that are most similar to a document. URL: https://maartengr.github.io/KeyBERT/

[2] The c-TF-IDF value of a term x in a given class c $w_{x,c} = \left\| tf_{x,c} \right\| \cdot \log\left(1 + \frac{A}{f_x}\right)$, where $tf_{x,c}$ is the frequency of word x in class c; f_x frequency of word x in all classes and A is the average number of words in the classes.

[3] As mentioned, the original BERTopic solution proposed by Thompson and Mimno in 2020 was also applied to test reproducibility with a test-test procedure (Potter & Levine-Donnerstein, 1999).

Figure SM1. The applied topic modeling process (Source: Authors' work)



SM2: The eleven topics resulted from topic modeling

Table SM2. The 11 topics resulted from the topic modeling. The extracted topic labels (LLM), the most important eight keywords of KeyBert (KeyBert) and c-TF-IDF. (Source: Authors' work)

Topic	LLM[4]	KeyBert[5]	c-TF-IDF[6]
0	Manufacturing firms profit and automation	['suppliers', 'manufacturing company', 'industry', 'manufacturing', 'companies', 'make profit', 'revenue', 'company']	['manufacturing', 'automotive', 'company', 'production', 'companies', 'example', 'car', 'also']
1	Information technology in companies	['informationtechnology manager', 'informationtechnology roles', 'informationtechnology', 'informationtechnology education', 'informationtechnology system']	['informationtechnology', 'need', 'company', 'even', 'good', 'people', 'manager', 'managers']
2	Human aspects of business processes	['give example', 'example', 'like evidence', 'sure', 'let', 'give', 'picture realist', 'see']	['let', 'people', 'see', 'way', 'example', 'give example', 'sure', 'think', 'let give', 'yellow']
3	Industry 4 0 Definition and Implementation	['industry four', 'industry industry', 'industry', 'industry development', 'industry example', 'industry digitalisation', 'want industry', 'industry systems']	['industry', 'would', 'companies', 'industry would', 'could', 'machine', 'data', 'want']
4	Data collection analysis sensors time series databases	['data information', 'data data', 'data', 'information data', 'data everyone', 'big data', 'lot data', 'collect data']	['data', 'sensors', 'information', 'collect', 'time', 'collect data', 'also', 'database']
5	Regulating Depth Standards in Systems	['constraints', 'rules', 'regulation', 'entitlement', 'standards', 'principle must', 'exceptions', 'rules game']	['standard', 'rules', 'regulation', 'handle', 'regulated', 'depth', 'standards', 'know']
6	Data Security Cybersecurity	['cybersecurity also', 'cybersecurity issues', 'cybersecurity', 'security cybersecurity', 'information security', 'security', 'cybersecurity course', 'data security']	['cybersecurity', 'security', 'data', 'virus', 'vulnerability', 'attack', 'network', 'secure']

[4] The LLM labels were obtained using the prompt sent to the pre-trained language model ("OpenHermes-2.5-Mistral-7B-GGUF").

[5] These words were determined using KeyBERT technology based on sentences clustered to specific topics.

[6] These words were determined using a class-based frequency-inverse document frequency approach used on sentences clustered to specific topics.

7	ERP System Implementation	['erp systems', 'erp system', 'systems erp', 'company erp', 'use erp', 'erp implementation', 'implemented erp', 'enterprise management']	['erp', 'erp system', 'system', 'erp systems', 'systems', 'company', 'erp implementation', 'enterprise']
8	Process Improvement	['process', 'processes need', 'process analysis', 'process obviously', 'process lives', 'processes job', 'processes aware']	['process', 'processes', 'comes', 'end', 'lack', 'written', 'going', 'prepared']
9	Supplier issues and procurement	['delivery note', 'receipt', 'delivery', 'receive', 'arrive', 'received', 'delivered', 'invoice']	['supplier', 'order', 'invoice', 'delivery', 'note', 'procurement', 'deliver', 'delivery note']
10	Digital Transition of SMEs	['small companies', 'small company', 'companies also', 'companies', 'group companies', 'multinationals companies', 'marketing companies', 'within company']	['companies', 'company', 'small', 'group', 'small companies', 'medium', 'group companies', 'within']

SM3: Operationalization of the process management maturity model

The process management maturity reference model builds on CMMI. A detailed description of how it conceptualizes and operationalizes process maturity can be found in the literature (ISO 2015a, 2015b, 2015c; CMMI Product Team, 2010; Chrissis et al., 2011). Only a focused summary is provided here. CMMI has three complex abilities with the following measurement attributes:

- The ability to understand, map, and control processes: This captures how control of business areas can be systematically developed. It moves forward from ad hoc to a consciously controlled operation relying on documented processes designed by company-wide, uniform principles, with regular reviews and redesigns. A further important feature of progress is the ongoing extension of processes with a high-level process control ability.

- The systemic process improvement ability starts with a complete lack of process improvement efforts and moves toward a performance-oriented, data-driven analysis, and feedback-based improvement of individual business areas. Gradual extension of this ability is important too.

- The ability to integrate improvement efforts and align them with strategy: It captures characteristics of an enterprise-wide, integrated process improvement ability used to harmonize ongoing efforts and align them with strategy and the overall logic of the company's value creation. In the first development stage, neither performance relationships among individual processes nor strategic alignment are present. At a higher level of development, performance interdependence and trade-offs are considered for more and more processes. In addition, strategic alignment becomes important to integrate improvement efforts further.

Figure SM3. The operationalization of the process management maturity model of CMMI (Source: Authors' work based on the literature referred to above)

<table>
<tr><th colspan="3">Process Control</th><th colspan="2">Process Improvement</th><th colspan="2">Process Improvement Integration</th><th rowspan="2">Maturity levels</th></tr>
<tr><th>No or ad-hoc process control</th><th>Extent of process control</th><th>Process documentation practices</th><th>Measuring process performance</th><th>Data-driven improvement efforts</th><th>Analysing interrelationships among process performances</th><th>System-level alignment of process improvement efforts</th></tr>
<tr><td rowspan="4">Not ad-hoc</td><td>Large-scale</td><td rowspan="2">Highly developed</td><td rowspan="2">Large-scale</td><td rowspan="2">Large-scale</td><td rowspan="2">Large-scale</td><td>Completely aligned</td><td>Optimizing</td></tr>
<tr><td rowspan="3">Medium-scale</td><td rowspan="3">Partially aligned</td><td>Quantitatively managed</td></tr>
<tr><td rowspan="2">Moderately developed</td><td rowspan="2">Medium-scale</td><td rowspan="2">Medium-scale</td><td>Medium-scale</td><td>Defined</td></tr>
<tr><td rowspan="2">None or small scale</td><td rowspan="2">Not aligned</td><td>Managed</td></tr>
<tr><td>Ad-hoc</td><td>Small-scale</td><td>Underdeveloped</td><td>None or small scale</td><td>None or small scale</td><td>Initial</td></tr>
</table>

SM4: The coding scheme of the process management maturity benchmark model

Table SM4. The coding scheme of the process management maturity benchmark model (Source: Authors' work)

***Abilities*, main features along which process management maturity is evaluated in the reference model**	***Measurement attributes* of abilities, along which their development can be measured**	***Keywords and expressions looked* for in the relevant topics when coding**
Process control	Does process control exist or not? Are processes mapped? Is the process control ad-hoc? Is the process control conscious? Has process control unified guidelines? The extent of the conscious process control? Are process control documents regularly reviewed? Are changes in process control documents traceable? Can changes in process documents be retrospectively verified?	Process control/governance, conscious, ad-hoc, understand /know the process, define the process, write down the process, process map, process mapping, unified guidelines, documentation, process document, process instruction, process standardization, process review, process changes, traceable
Process improvement	Is the performance of operational areas/functions/processes measured? The extent of this performance measurement? Are the performance data used for process improvement initiatives?	Performance, process performance, measured performance, process analysis/assessment, process improvement/development/reorganization, data-driven initiatives, extent of these initiatives
Integration of improvement efforts	Are relationships among the performance of individual processes considered when initiating improvement projects? Is the system-level performance and the overall value creation of the firm analyzed when deciding on improvement projects? Are improvement projects aligned with strategic objectives?	Trade-offs in performance, interdependences among performance, system-level (systemic) performance, overall value creation logic, effort integration, strategic objectives, strategic considerations

SM5: The IT management maturity reference model

As indicated in Section 3.2.2, the IT management maturity reference model is based on SIMMI (Leyh et al., 2016) and the L&K model (Leem & Kim, 2004; Leem et al., 2008). The reference model integrated the human-based organizational aspects of the L&K model with the technological evaluation dimensions of SIMMI.

Like in the case of the process maturity reference model, here we also have five development stages. These are: (1) basic, (2) cross-departmental, (3) horizontal and vertical digitalization, (4) full, and (5) optimized full digitalization. The level of IT management maturity develops along these stages as a combination of three broad features: (i) the level of integration (both horizontally and vertically), (ii) the level of digital continuity in the value creation, and (iii) along key cross-sectional technology criteria. For example, the first two evaluation dimensions create integration within the company and then with partners using organizational/strategic items (e.g, IT investment and related educational practices facilitating integration) and items related to enterprise systems (e.g., methods of data storage and the extent of IT support of the business). The more developed data storage solutions are, and the more business parts have IT support, the higher and broader the integration will be. The complete list of items is provided, and their relationships are illustrated by Figure SM5 below:

(1) The **level of integration** within the firm and along its value chain (vertical and horizontal integration). Elements capturing this have been taken over from the L&K model (investments, IT education, surveys on IT requirements, and business area support) but complemented from SIMMI by a key I4.0 readiness factor, the method of data aggregation, storage, and communication since this is also considered as an important mean of integration. Specifically, the ability of IT integration includes the following attributes:
    a. *Investment rate of IT assets* (from L&K): Since the capacity of physical IT assets has been proved by the L&K model to be proportional to the level of investments in the company's IT infrastructure.
    b. *Periodic requirement survey* (from L&K): Surveys of whether companies regularly conduct needs assessments of IT end-users for IT purchases, renewals, and upgrades are used to assess the appropriateness of IT usage.
    c. *IT education regularity* (from L&K): IT training within the firms is critical for the actual realization of IT usage.
    d. *Target levels of IT education* (from L&K): It is essential that training is provided to all employees, including managers.
    e. *Method of data storage* (from both L&K, SIMMI): It shows the level of data storage in the context of integration. While the focus on process management capability is on integrated management of enterprise processes, the accompanying storage of the generated data in different databases and sometimes in different formats often limits this integration. In different storage and different formats; in different storage but in compatible format; in shared database; in different database with interface; SOA platform.
    f. *Support of business* parts (from both L&K and SIMMI): for the functional areas of the organization, consider whether IT support is available.

(2) **Digital continuity**, especially in production and product development.
    a. *Digital production support* (from SIMMI), whether and to what extent manufacturing is digitally supported.

b. *Collecting customer data on product usage* (from SIMMI), whether data collection from customers on products is ensured.
c. *Using these data for product development* (from SIMMI).

(3) **Cross-sectional technology criteria** including the development of the IT platform at the company, application of service-oriented applications, and business analytics and IT security.[7]
   a. *IT platform* (from SIMMI): The technological maturity of a company is well characterized by the IT platform on which it operates. Isolated system, PC; client/server solution; vertically integrated system; platform enabling vertical-horizontal integration; SOA, cloud.
   b. *Service-oriented cloud applications* (from SIMMI): The use of modern technology is characterized by whether they have service-oriented cloud applications.
   c. *Applications for business analytics* (from SIMMI): To decide whether a company is state-of-the-art, it is important to know whether business intelligence applications are used.
   d. *IT security* (from SIMMI): A high level of IT security is essential to meet the technological criteria, ensuring the risk-free flow of data and information. The operation of a company in Industry 4.0 is heavily dependent on the use of the Internet, which enables the daily transfer of data from machines that are in operation. Due to the continuous internet-based operations, ensuring IT security is a major challenge in using various IT systems. IT security involves ensuring that all electronically accessible information is properly protected. In addition, it is essential to ensure that IT systems and their services are always available to users and that they function properly.

We note that the degree of a firm's overall IT management maturity along the reference model fits well with the levels of the process maturity model. See the scope or the integration levels defined above (e.g., ad hoc process - isolated systems; optimized, managed, cross-enterprise process system - SOA platform).

[7] We mention here that the capability level of the IT management process itself is also crucial for a company. The evaluation of this IT management process is covered, however, by our process management model (CMMI).

Figure SM5. The operationalization of the IT management maturity reference model based on SIMMI and the L&K model (Source: Authors' work based on the literature referred to above)

Organizational/ Strategical Aspects				Enterprise Systems		Digital Production and Product Development			Cross-sectional Technology Criteria				Maturity levels
Investment rate of IT assets	IT education regularity	Target levels for IT education	Periodic requirements survey	Method of data storage	Support of business parts	Digital production support	Collecting customer data on product usage	Using data for product development	IT platform	Service-oriented cloud applications	Applications for business analytics	IT security	
1% or more	Regular and planned	For leaders and employees	Regular and consciously organized	Internal SOA-architecture	All areas (5)	Company and external product development data is processed digitally.	Data collected	Yes	External SOA integration	Successfully implemented	Successfully implemented	ISO-compliant IT security or better (2FA). Quickly adapts to new hazards and resolves security issues. Value network encryption is optimized.	**Optimized full digitization**
Between 0.2% and 1%				In different databases with an interface	All relevant areas (4-5)	The product development process is digitally supported.					Under implementation	ISO-compliant. Data security. Global inter-enterprise data encryption and authentication.	**Full digitization**
	Occasional	For leaders or employees	Occasional	In shared database		Between process steps, product development information is digitally communicated.		No	Integrated internal SOA platform	Under implementation	Under planning	A approach like COBIT underpins it. Their comprehensive security solution protects data access and encrypts company data transmission.	**Horizontal and vertical digitization**
				In different storage but in compatible format	All relevant areas (3)		Not collected		Client – server environment	Under planning	None	Internal governance and control system component. They are creating first IT security models.	**Cross-departmental digitization**
Under 0.2%	None	None	None	In different storage and different format	Some areas (1-3)	Product development is not digitally supported.			Isolated IT systems	None		Confidentiality, availability, and integrity are not guaranteed.	**Basic digitization**

SM6: The coding scheme for the IT management benchmark model

Table SM6. The coding scheme for the IT management benchmark model (Source: Authors' work)

***Abilities*, main features along which IT management maturity is evaluated in the reference model**	***Measurement attributes* of abilities, along which their development can be measured**	***Keywords and expressions looked* for in the relevant topics when coding**
Organizational, strategic aspects	Investment rate of IT assets.	increasing the capacity of IT, expanding the IT assets, IT investments, and investment rate
	Do periodic requirement surveys exist?	IT requirements, IT surveys, periodic surveys, regular surveys
	How regular is IT education?	IT education, IT training
	What are the target levels of IT education?	training/education of end-users, training/education of employees, training/education of management
Enterprise systems	What is the method of data storage?	mode/way of data storage, place of data storage, format of databases, compatible format of storage, shared database, interface between databases, integrated database, integrated data storage, SOA
	How are different business areas supported?	support of business, IT support of functions / operational areas
Digital production and product development	Is production digitally supported?	digital manufacturing, digital support of manufacturing, digital support of production/operation, PLM, manufacturing/production digitalization
	Are customer data on product usage collected?	customer data, collecting customer data /information, data on product usage
	Are the above data used for product development?	product usage analysis, product development
Cross-sectional technology criteria	What IT platform does the firm use?	IT platform, integrated platform, isolates IT system, PC, client-server solution, integrated system, ERP, SOA
	Are service-oriented cloud applications used?	cloud, using cloud, service-oriented cloud, SOA
	Are business analytics applied?	state-of-the-art analytics, business analytics, business intelligence, BI
	What is the level of IT security?	data/information security, IT security, cybersecurity, IT risk, data/information protection, basic security (password), privacy, privilege entitlement, eligibility, ISO, virus protection, firewall, COBIT

References

Betlen, A. (2023). llama-cpp-python. In https://github.com/abetlen/llama-cpp-python

Chrissis, M. B., Konrad, M., and Shrum, S. (2011). CMMI for development: guidelines for process integration and product improvement. Pearson Education. ISBN:978-0-321-71150-2

CMMI_Product_Team. (2010). Improving processes for developing better products and services, CMMI RG for Development 10.1184/R1/6572342.v1

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794. https://doi.org/10.48550/arXiv.2203.05794

ISO. (2015a). Information Technology – Pro-cess Assessment – Process Measurement Framework for Assessment of Process Capability. In Information Technology (Vol. ISO/IEC 33020). Geneva, Switzerland: ISO.

ISO. (2015b). Information Technology – Pro-cess Assessment – Requirements for Process Reference, Pro-cess Assessment and Maturity Models. In Information Technology (Vol. ISO/IEC 33004). Geneva, Switzerland: ISO.

ISO. (2015c). Information Technology – Pro-cess Assessment, International Organization for Standard-ization." Geneva, Switzerland. In Information Technology (Vol. ISO/IEC 33000). Geneva, Switzerland: ISO.

Leem, C., and Kim, I. (2004). An integrated evaluation system based on the continuous improvement model of IS performance. Industrial Management & Data Systems, 104(2), 115-128. https://doi.org/10.1108/02635570410522080

Leem, C., Wan Kim, B., Jung Yu, E., and Ho Paek, M. (2008). Information technology maturity stages and enterprise benchmarking: an empirical study. Industrial Management & Data Systems, 108(9), 1200-1218. https://doi.org/10.1108/02635570810914892

Leyh, C., Bley, K., Schäffer, T., and Forstenhäusler, S. (2016). SIMMI 4.0-a maturity model for classifying the enterprise-wide it and software landscape focusing on Industry 4.0. 2016 federated conference on computer science and information systems (fedcsis). http://dx.doi.org/10.15439/2016F478

Potter, W. J., and Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis.