Towards a Model for Supporting Cloud-based Information Extraction from MICE Content

Gladys Kouakam[1,*], Nico Lindemann[1,2] and Philippe Tamla[1,2]

[1]University in Hagen, Universitätsstr. 1, 58097 Hagen, Germany
[2]FTK e.V. Research Institute for Telecommunications and Cooperation, Wandweg 3, 44149 Dortmund, Germany

Abstract

The Meetings, Incentives, Conferences and Exhibitions (MICE) industry generates large volumes of heterogeneous and predominantly unstructured data (e.g., submissions, programs, feedback, and social media interactions). As a digitally evolving ecosystem, these data streams contain valuable information about Attendee Engagement and Satisfaction as well as Speaker Performance and Reputation, but also increase information overload and hinder data-driven decision-making for Professional Congress Organizers (PCOs). This position paper analyzes the state of the art in lifecycle-oriented information extraction and knowledge management for MICE and identifies key gaps, including the lack of integrated, lifecycle-oriented pipelines for multi-source data collection and preparation, insufficient alignment between extraction outputs, taxonomies, and KPI-oriented knowledge, and limited operationalization through information access and retrieval. In addition, existing approaches often lack unified, cloud-based lifecycle integration. To address these challenges, we introduce *IELC4AS2P*, an Information Extraction Lifecycle (IELC)-based reference model for Attendee Satisfaction and Speaker Performance. The model integrates (i) a use case perspective and (ii) a workflow model connecting Data Collection and Preparation, Information Extraction, Information and Knowledge Organization, and Information Access and Retrieval. As a position paper, this work provides an initial modeling foundation for lifecycle-oriented information extraction in evolving MICE ecosystems.

Keywords

Information Extraction, Information Retrieval, Cloud Computing, Conference Management System, Topic Modeling, Named Entity Recognition, Sentiment Analysis, Information and Knowledge Management

1. Introduction

The **Meetings, Incentives, Conferences and Exhibitions (MICE)** industry supports professional knowledge exchange, collaboration, and networking across business, academia, and public institutions [1, 2]. Recent developments driven by digitization, the COVID-19 pandemic, and climate change have led to the generation of large volumes of heterogeneous and predominantly unstructured data, including scientific submissions, programs, surveys, and social media interactions [3, 4, 5]. In this context, the MICE domain can be conceptualized as a digital ecosystem in which multiple actors, platforms, and data streams interact dynamically across the event lifecycle. While these data streams contain valuable insights into **Attendee Engagement (AE)**, **Attendee Satisfaction (AS)**, **Attendee Interest (AtI)**, **Speaker Performance (SP)**, and **Speaker Reputation (SR)**, their volume and unstructured nature contribute to **Information Overload (IOL)** and hinder data-driven decision-making for **Professional Congress Organizers (PCOs)**.

Existing research projects such as RecomRatio [6], AI4H3 [7], and FIT4NER [8] demonstrate the effectiveness of cloud-based **Information Extraction (IE)** and knowledge management using techniques such as **Named Entity Recognition (NER)** and **Natural Language Processing (NLP)**. These systems successfully extract structured knowledge from domain-specific corpora, particularly in biomedical and technical contexts. *However, these approaches primarily operate on relatively homogeneous and*

TRI-IVIS - AVI 2026, Workshop on Trustable, Reproducible, and Intelligent Information Visualization Systems, Co-located with the 18th International Conference on Advanced Visual Interfaces AVI 2026, June 08–12, 2022, Venice, Italy

*Corresponding author.
gladys.kouakam@fernuni-hagen.de (G. Kouakam); n.lindemann@globit.com (N. Lindemann); philippe.tamla@fernuni-hagen.de (P. Tamla)
0009-0006-0563-9155 (G. Kouakam); 0009-0006-8232-6240 (N. Lindemann); 0000-0002-0786-4253 (P. Tamla)

well-structured datasets and do not address the challenges of heterogeneous, multi-source, and dynamically evolving ecosystems such as the MICE domain [9, 10, 11].

This heterogeneity challenges all phases of the **Information Extraction Life Cycle (IELC)** [12]. Current approaches lack scalable integration of heterogeneous sources and do not sufficiently support preprocessing, normalization, and temporal alignment [13, 14]. In addition, the combined use of methods such as NER, **Topic Modeling (TM)**, **Sentiment Analysis (SA)**, and **Predictive Analytics (PAs)** for deriving structured vocabularies and predictive insights remains limited. Furthermore, the transformation of extracted outputs into taxonomies and **Key Performance Indicators (KPI)**-oriented knowledge, as well as their operationalization in **Content and Knowledge Management System (CKMS)** such as **Content and Knowledge Management Ecosystem Portal (KM-EP)**, is insufficiently addressed [15, 16, 17, 18, 19].

This paper addresses these gaps by proposing a lifecycle-oriented, cloud-based **Information Extraction Activities (IEA)** approach that integrates Data Collection, Data Preparation, IE, **Information and Knowledge Organization (IKO)**, and **Information Access and Retrieval (IAR)** into a unified framework to derive **Professional Congress Organizer Knowledge Vocabularies (PCOKV)**, transform them into **Professional Congress Organizer Knowledge Taxonomies (PCOKT)**, and operationalize KPI-oriented knowledge for decision-making.

The research addresses three **Research Questions (RQs)**: (RQ1) how the MICE domain can be characterized as a digital ecosystem to support lifecycle-oriented model for data collection, data preparation, IE, and metrics-based measurement to derive PCOKV and predictive scores for AE, AS, SP, and SR; (RQ2) how a PCOK IKO model can be developed to transform PCOKV into PCOKT and support the derivation of KPIs; and (RQ3) how can a PCOK IAR model enable efficient search, browsing, and access to PCOKT, PCOKV, predictive scores, and KPIs for data-driven decision-making.

To investigate these RQs, this study follows the Information Systems framework of Nunamaker et al. [20], focusing on the Observation and Theory Building phases, while System Development and Evaluation are left for future work. Section 2 analyzes the MICE domain and reviews related work across the IELC phases. Section 3 presents the proposed models, and Section 4 concludes the paper and outlines future work.

2. State of the Art in Science and Technology

This chapter addresses the Observation phase by reviewing the scientific and technological foundations for lifecycle-oriented, cloud-based IEA in the MICE domain. Based on literature and an initial field study of a PCO software ecosystem, it characterizes MICE as a digital ecosystem and analyzes methods for data collection, data preparation, IE, IKO, IAR, and their cloud-based operationalization.

2.1. MICE Ecosystems and Lifecycle-Oriented Models

The MICE industry has evolved into a digitally transformed ecosystem characterized by increasing complexity, technological diversification, and strong interdependencies between stakeholders, platforms, and data flows [21, 5, 11]. Event-related data is generated across preparation, execution, and post-event phases, resulting in heterogeneous landscapes combining structured, semi-structured, and unstructured data. While such environments offer significant potential for data-driven decision-making [22], they require systematic transformation processes from raw data to actionable knowledge, as conceptualized by **Data, Information, Knowledge, Wisdom (DIKW)** [23] and knowledge creation models such as **Socialization, Externalization, Combination, and Internalization (SECI)** [24]. The IELC structures this transformation across Data Collection, Data Preparation, IE, IKO, and IAR [25], yet *current approaches do not provide an integrated, lifecycle-aware framework capable of consistently transforming heterogeneous, dynamically evolving MICE data into structured and decision-relevant knowledge across all phases*, indicating the need for a unified lifecycle model that enables continuous and coherent knowledge generation within such ecosystems.

2.2. PCOK Vocabulary Construction

MICE data collection relies on APIs, databases, and web-based sources, often resulting in fragmented and evolving datasets [9, 26]. Data preparation transforms raw data into analysis-ready corpora through **Extract, Transform, Load (ETL)/Extract, Load, Transform (ELT)** pipelines, cleaning, normalization, deduplication, and integration [27, 14], but remains largely tool-centric and insufficiently aligned with lifecycle requirements and temporal dynamics. On this basis, IE methods derive structured vocabularies, where NER extracts entities such as persons and organizations [28, 29], TM identifies thematic structures including **emerging** and **trending topics**, and neural approaches such as BERTopic improve topic quality in short texts [30], while extensions such as BERTrend incorporate temporal evolution [31]. In addition, SA and PA enable the derivation of indicators related to AE, AS, SP, and SR [17, 32]. However, these methods are typically applied in isolation and lack integration into coherent, lifecycle-aware pipelines, highlighting the *need for a unified approach that orchestrates data collection, preparation, and multi-method IE to enable consistent, temporally aware construction of PCOKV and predictive indicators.*

2.3. PCOK Knowledge Organization and Access

Building on extracted vocabularies, IKO and IAR transform PCOKV into structured and actionable knowledge assets. IKO organizes extracted information into PCOKTs and semantic structures that enable interpretation, comparison, and reuse [33, 16], with approaches such as NetTaxo supporting automated hierarchy construction [34]. It further enables the derivation of KPI-oriented knowledge linking extracted information to decision-relevant metrics such as AE, AS, SP, and SR [35]. In parallel, prior work in Human-Computer Interaction has demonstrated that user interaction data can reveal implicit indicators of system performance and user experience. In particular, the concept of *usability smells* [36] has been introduced as observable patterns in user behavior (e.g., navigation deviations, task failures, or repeated backtracking) that indicate potential usability issues and user confusion. This information is typically derived from interaction logs and visualized to support usability evaluation. However, such approaches remain primarily focused on isolated evaluation contexts and do not formalize this information into structured, reusable knowledge assets or integrate them into taxonomy-driven or KPI-oriented knowledge representations. IAR enables access through search, browsing, and retrieval mechanisms, where **Information Access (IA)** provides structured access to known resources and **Information Retrieval (IR)** supports exploratory search [37, 38, 18], increasingly enhanced by machine learning and Transformer-based models for semantic retrieval [39, 40], as well as faceted search for interactive exploration [41, 42]. Despite these advances, there is a *lack of integrated frameworks that consistently connect vocabularies, taxonomies, behavioral information (such as usability smells), and KPI-oriented knowledge with lifecycle-aware access mechanisms*, indicating the need for unified models that combine semantic organization with efficient, exploratory, and decision-oriented access.

2.4. Cloud Computing

Cloud computing (CC) provides the infrastructural basis for coordinating data collection, preparation, extraction, organization, and retrieval in large-scale systems by enabling scalable processing, elastic resource allocation, and continuous workflows [43, 44]. It further supports orchestration, monitoring, reproducible deployment, and data provenance for traceability and comparability [45, 46], while modern architectural paradigms such as microservices, container orchestration, and serverless computing enable flexible and distributed processing environments [47, 48, 49]. However, current cloud-based solutions often remain tool-centric and do not achieve end-to-end lifecycle integration across all IELC phases, highlighting *the need for cloud-native architectures that unify these phases into a reproducible, continuously evolving system capable of handling heterogeneous and dynamic MICE ecosystems.*

2.5. Summary

The analysis shows that MICE environments constitute heterogeneous and dynamically evolving digital ecosystems requiring lifecycle-oriented approaches for transforming data into knowledge. Existing methods for data collection, data preparation, and IE remain insufficiently integrated for systematic vocabulary construction, while IKO and IAR approaches lack alignment with KPI-oriented knowledge and decision-making, and cloud solutions do not yet provide unified lifecycle integration. These limitations motivate the need for an integrated, lifecycle-oriented, and cloud-based model that consistently connects all IELC phases, from data acquisition to knowledge access, to transform heterogeneous and unstructured MICE data into structured, actionable knowledge, forming the basis for the design and modeling approach presented in the next chapter.

3. Modeling Strategy

This chapter presents the conceptual and technical modeling required to address the previously introduced RQs and the related RCs identified in the state of the art, structured along the phases of the IELC. The chapter develops two complementary modeling views. First, a *use case modeling* specifies the main lifecycle phases, actors, and functional responsibilities, ranging from data collection to data preparation, IE, IKO, and IAR. Second, a *workflow modeling* describes how these phases are connected in an iterative end-to-end process that transforms heterogeneous event data into governed knowledge assets and predictive information. Therefore the name of the reference model is **Information Extraction Lifecycle for Attendee Satisfaction & Speaker Performance (IELC4AS2P)** as it combines a lifecycle-oriented IE with a focus on supporting attendee satisfaction and speaker performance analysis in the MICE domain.

3.1. Use Case Modeling

The use case model structures the IELC4AS2P lifecycle as a set of interconnected, multi-entity interaction processes spanning Data Collection and Preparation, IE, IKO, and IAR. The model explicitly reflects the include relationships between functional components and the continuous transformation of data into knowledge assets. The overall lifecycle is supervised and configured by the Administrator, who manages data ingestion, preprocessing, and analytical workflows. The PCO acts as the primary consumer of the resulting knowledge assets through search and exploration functionalities.

As illustrated in Figure 1, the process begins with Data Collection and Preparation (orange box), where heterogeneous event data is continuously ingested from multiple sources. This phase includes data cleaning, deduplication, and tokenization, producing analysis-ready corpora. The IE phase (green box) builds on this prepared data and includes the coordinated application of multiple analytical methods. Specifically, methods such as NER, TM, and SA contribute to the generation of PCOKVs. In parallel, PA is applied to derive predictive scores, which quantify aspects such as Attendee Engagement, Satisfaction, and Speaker Performance. In the IKO (blue box), the extracted vocabularies and predictive scores are transformed into structured knowledge assets. This includes the construction of PCOKTs, the derivation of KPI-oriented knowledge, and the creation of profile structures for speakers and attendees. These components are explicitly linked through include relationships, ensuring a consistent transformation from extracted data to decision-relevant knowledge. Finally, the IAR phase (purple box) enables PCOs to interact with the system through faceted search and browsing mechanisms. This phase provides integrated access to profiles, events, topics, and interaction data, supporting exploratory analysis and decision-making processes such as session planning and performance evaluation.

Across all phases, the lifecycle is supported by a cloud-based environment that enables scalable processing, orchestration, and continuous updates, ensuring that knowledge assets evolve dynamically with incoming data streams.

3.2. Workflow Model

The IELC4AS2P workflow model describes the end-to-end transformation of heterogeneous MICE data into structured and actionable knowledge assets through a lifecycle-oriented, cloud-based process. As

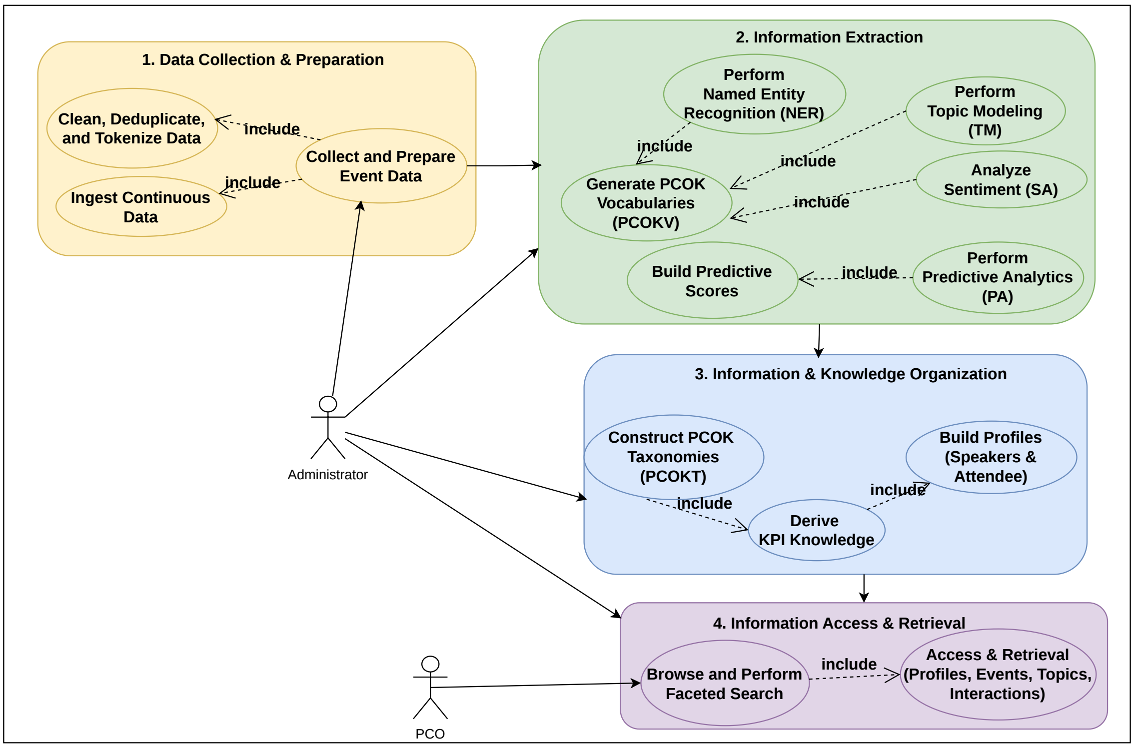

Figure 1: Use Cases for IELC4AS2P

illustrated in Figure 2, the workflow begins with Data and Content Collection, where event-generated data is acquired from multiple sources across the event lifecycle. These include pre-event data (e.g., submissions, participant information, program drafts), in-event data (e.g., session attendance, interaction logs, real-time feedback), and post-event data (e.g., surveys, evaluations, and reports). This results in a heterogeneous mix of structured, semi-structured, and unstructured data.

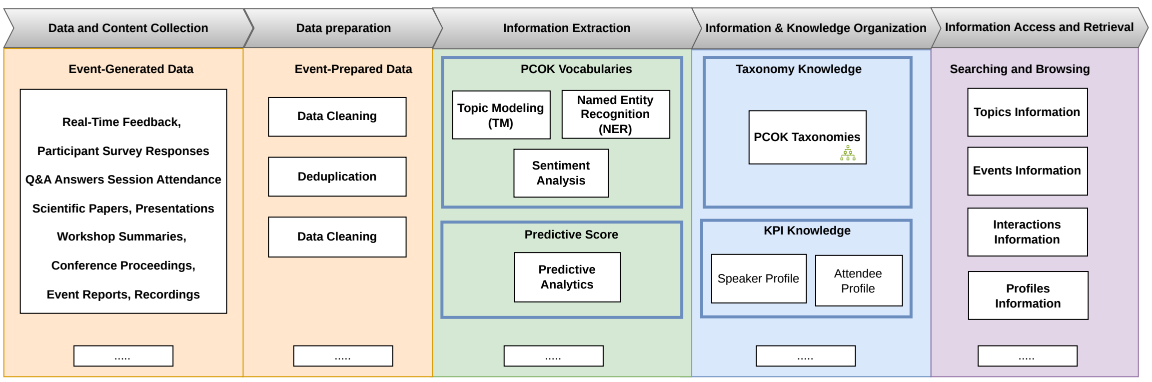

Figure 2: IELC4AS2P Workflow Model

In the Data Preparation phase, raw data is transformed into event-prepared data through systematic preprocessing steps, including data cleaning, deduplication, normalization, and integration. These steps ensure data consistency, comparability, and readiness for downstream analytical processing. The IE phase applies multiple analytical methods in a coordinated manner to derive structured knowledge representations. TM, NER, and SA jointly contribute to the construction of PCOKVs, capturing entities, topics, and sentiment patterns. In parallel, PA is applied to generate predictive scores, representing measurable indicators such as attendee engagement, satisfaction, and speaker performance. In the IKO phase, these outputs are systematically transformed into structured knowledge assets. The PCOKV is organized into PCOKTs, enabling semantic structuring and hierarchical representation. At the same time, predictive scores are integrated into KPI-oriented knowledge, and profile structures for speakers and attendees are constructed, linking extracted information to decision-relevant metrics. Finally, the IAR phase operationalizes these knowledge assets by enabling structured access through searching and browsing mechanisms. This includes access to topics, events, interaction data, and profile information. In particular, faceted search and taxonomy-guided navigation support exploratory analysis and decision-oriented use cases, such as session recommendation, performance monitoring, and event optimization.

Overall, the workflow model reflects a continuous and iterative lifecycle, where newly ingested data is continuously processed and integrated into evolving knowledge structures. The cloud-based infrastructure ensures scalability, orchestration, and reproducibility across all phases, enabling consistent transformation from raw data to actionable knowledge in dynamic MICE ecosystems.

3.3. Summary

This chapter presented the modeling foundations of the proposed *IELC4AS2P* strategy for transforming heterogeneous MICE data into decision-relevant knowledge assets for PCOs. Three complementary views were developed: a use case model, a workflow model, and a cloud-based architecture. The models structure the lifecycle into four phases—*Data Collection & Preparation*, *IE*, *IKO*, and *IAR*—describing the transformation from raw event data to structured vocabularies, predictive scores, taxonomies, and KPI-oriented knowledge, and their subsequent access through search and browsing. As a position paper, this work represents an initial modeling attempt that conceptualizes a lifecycle-oriented approach for MICE ecosystems. Future work will extend this foundation through deeper domain and data understanding, refined modeling, and comprehensive system development and evaluation.

4. Conclusion and Future Work

This position paper addressed the challenge that the MICE domain produces large volumes of heterogeneous, predominantly unstructured data, which increases information overload and limits data-driven decision-making for PCOs. Based on an IELC-oriented analysis, the state of the art was structured along data collection and preparation, information extraction, knowledge organization, and information access and retrieval. This analysis revealed key challenges, including the lack of lifecycle-aware integration, insufficient alignment between vocabularies, taxonomies, and KPI-oriented knowledge, and the absence of unified, cloud-based end-to-end architectures for evolving MICE ecosystems. To address these gaps at the modeling level, the paper proposed the *IELC4AS2P* strategy and introduced a use case model and workflow model that make the transformation from heterogeneous event data to decision-ready knowledge explicit. Future work focuses on three main directions. First, the proposed *IELC4AS2P* models will be instantiated and validated through a prototype implementation using real, edition-based MICE data and multi-source ingestion. Second, the modeling approach will be extended by developing a detailed MVC-based architecture to operationalize the lifecycle phases within a modular, cloud-based system. Third, the approach will be evaluated through technical and user-centered studies, assessing extraction quality, taxonomy and KPI consistency, and the effectiveness of information access and retrieval for decision-oriented use.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT as an AI-assisted writing support tool. ChatGPT was used to review grammar, improve the clarity and structure of the manuscript, suggest revisions for conciseness, and provide feedback on the organization and presentation of the proposed models. The final wording and all scientific content were developed, verified, and approved by the authors. All AI-generated suggestions were critically reviewed, revised where necessary, and incorporated only after verification by the authors, who take full responsibility for the content of this publication.

References

[1] T. Warner, T. rogers , conferences and conventions: a global industry (second edition), butterworth-heinemann, oxford (2008) isbn 978-0-7506-8544-3 400 pp., eur 36.95, gbp 24.99, usd 43.99, Tourism Management - TOURISM MANAGE 30 (2009) 146–147. doi:10.1016/j.tourman.2008.03.010.

[2] M. of Tourism, Ministry of Tourism Government of India NATIONAL STRATEGY FOR MICE INDUSTRY, 2022. URL: https://tourism.gov.in/sites/default/files/2022-09/National%20Strategy%20for%20MICE%20Industry%202022.pdf.

[3] M. S. Wilhelm Bauer, Stefan Rief, Ökosysteme im wandel–zukunftsszenarien für business events im zeitalter grenzenloser kommunikation, in: Future Meeting Space. Fraunhofer-Institut für Arbeitswissenschaft und Organisation IAO [Themenheft], Fraunhofer IAO, 2022.

[4] H. Wee, N. D. Kamarulzaman, M. S. Anas, Mice industry survival: A systematic literature review, International Journal of Academic Research in Business and Social Sciences (2022).

[5] A. Magdy, N. H. Abouelazm, Towards digital transformation technologies application in the mice industry in egypt, Minia Journal of Tourism and Hospitality Research MJTHR 17 (2024) 81–117.

[6] Forschungsinstitut für Telekommunikation und Kooperation e.V. (FTK), RecomRatio, https://www.ftk.de/de/projekte/recomratio, n.d.

[7] FTK, Artificial Intelligence for Hospitals, Healthcare & Humanity (AI4H3), R&D White Paper, FTK e.V. Research Institute for Telecommunications and Cooperation, Dortmund, Germany, 2025.

[8] F. Freund, P. Tamla, F. Wilde, M. Hemmje, Making medical experts fit4ner: Transforming domain knowledge through machine learning-based named entity recognition, International Journal on Natural Language Computing 14 (2025). doi:10.5121/ijnlc.2025.14202.

[9] Y.-H. Lin, C.-H. Chang, H.-M. Chuang, X.-S. Lin, T. Yeh, M.-J. Hong, Cost-effective event mining on the web via event source page discovery and data api construction, IEEE Access (2024).

[10] X. Liu, R. Seevers, Z. Gu, X. Yang, Smart mice: Definitions, foundations and development, in: 2020 7th International Conference on Information Science and Control Engineering (ICISCE), 2020, pp. 1296–1300. doi:10.1109/ICISCE50968.2020.00262.

[11] P. Warintarawej, N. Thongsri, J. Seksan, W. Marchoo, Mice city ecosystems: building collaboration through information systems – a case study from thailand, Cogent Business & Management 12 (2025) 2460627. URL: https://doi.org/10.1080/23311975.2025.2460627. doi:10.1080/23311975.2025.2460627. arXiv:https://doi.org/10.1080/23311975.2025.2460627.

[12] P. Tamla, Towards a Reference Model for an Information Extraction Lifecycle Supporting Data Collection, Data Preparation, Information Extraction, Information and Knowledge Organization, and Information Access and Retrieval, Technical Report, FernUniversität in Hagen, 2025.

[13] C. Bizer, T. Heath, T. Berners-Lee, Linked data: The story so far, International Journal on Semantic Web and Information Systems (IJSWIS) 5 (2009) 1–22. doi:10.4018/978-1-60960-593-3.ch008.

[14] U. Leser, F. Naumann, Informationsintegration, 2006.

[15] M. Gucma, A. Deja, J. Szymonowicz, Environmental solutions for maritime ships: challenges and needs, Production Engineering Archives 29 (2023) 216–224. doi:10.30657/pea.2023.29.25.

[16] O. Medelyan, S. Manion, J. Broekstra, A. Divoli, A.-L. Huang, I. H. Witten, Constructing a focused

taxonomy from a document collection, in: The Semantic Web: Semantics and Big Data: 10th International Conference, ESWC 2013, Montpellier, France, May 26-30, 2013. Proceedings 10, Springer, 2013, pp. 367–381.
[17] A. Thakur, et al., Survey into predictive key performance indicator analysis from data mining perspective, in: IEEE ETFA, 2020.
[18] R. Baeza-Yates, B. Ribeiro-neto, D. Mills, O. Bonn, S. Juan, M. Mexico, C. Taipei, A. Wesley, L. Limited, Modern information retrieval (1999).
[19] J. Paul, Predictive analytics in the automobile industry: Driving efficiency and customer satisfaction (2022).
[20] J. F. Nunamaker Jr, M. Chen, T. D. M. Purdin, Systems Development in Information Systems Research, Journal of Management Information Systems 7 (1990) 89–106. URL: https://doi.org/10.1080/07421222.1990.11517898. doi:10.1080/07421222.1990.11517898, publisher: Routledge _eprint: https://doi.org/10.1080/07421222.1990.11517898.
[21] G. G. C. B. e. V., EUROPE AS A DESTINATION FOR MEETINGS AND CONFERENCES, 2019.
[22] Ž. Erceg, N. Đalić, The importance of digital transformation for business decision-making in organizations., New Economist/Novi Ekonomist 18 (2024).
[23] L.-W. Lung, Y.-R. Wang, Ssd-based innovations for improved construction management, Journal of Engineering Research (2025).
[24] I. Nonaka, H. Takeuchi, The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, New York, 1995.
[25] P. Tamla, Towards a Reference Model for an Information Extraction Lifecycle Supporting Data Collection, Data Preparation, Information Extraction, Information and Knowledge Organization, and Information Access and Retrieval, Technical Report, FernUniversität in Hagen, 2025. URL: https://ub-deposit.fernuni-hagen.de/receive/mir_mods_00002257, available via the institutional repository of FernUniversität in Hagen.
[26] S. Kim, S. Park, M.-R. Sun, J.-H. Lee, A study of smart beacon-based meeting, incentive trip, convention, exhibition and event (mice) services using big data, Procedia Computer Science 91 (2016) 761–768. URL: https://www.sciencedirect.com/science/article/pii/S1877050916312558. doi:https://doi.org/10.1016/j.procs.2016.07.072, promoting Business Analytics and Quantitative Management of Technology: 4th International Conference on Information Technology and Quantitative Management (ITQM 2016).
[27] A. A. A. Fernandes, M. Koehler, N. Konstantinou, P. Pankin, N. W. Paton, R. Sakellariou, Data preparation: A technological perspective and review, SN Comput. Sci. 4 (2023). URL: https://doi.org/10.1007/s42979-023-01828-8. doi:10.1007/s42979-023-01828-8.
[28] N. Chinchor, Muc-7 named entity task definition, in: Proceedings of MUC-7, 1997.
[29] Z. Hu, W. Hou, X. Liu, Deep learning for named entity recognition: A survey, Neural Computing and Applications (2024). doi:10.1007/s00521-024-09646-6.
[30] M. Grootendorst, Bertopic: Neural topic modeling with a class-based tf-idf procedure, arXiv preprint arXiv:2203.05794 (2022).
[31] A. Boutaleb, et al., Bertrend: Neural topic modeling for emerging trends detection, arXiv preprint arXiv:2411.05930 (2024).
[32] V. Kumar, M. Garg, Predictive analytics: a review of trends and techniques, International Journal of Computer Applications 182 (2018) 31–37.
[33] P. Lambe, Organising knowledge: taxonomies, knowledge and organisational effectiveness, Elsevier, 2014.
[34] J. Shang, X. Zhang, L. Liu, S. Li, J. Han, Nettaxo: Automated topic taxonomy construction from text-rich network, in: WWW 2020, 2020, pp. 1908–1919. doi:10.1145/3366423.3380259.
[35] R. S. Kaplan, D. P. Norton, The balanced scorecard: Measures that drive performance, Harvard Business Review 83 (1996) 172–180.
[36] P. Buono, D. Caivano, M. Costabile, G. Desolda, R. Lanzilotti, Towards the detection of ux smells: The support of visualizations, IEEE Access PP (2019) 1–1. doi:10.1109/ACCESS.2019.2961768.
[37] M. A. Hearst, Search user interfaces, Cambridge university press, 2009.

[38] C. D. Manning, Introduction to information retrieval, Syngress Publishing„ 2008.
[39] E. G. Jacques Savoy, Handbook of Natural Language Processing, Taylor and Francis Group, LLC, 2010, pp. 455–484.
[40] K. A. Hambarde, H. Proenca, Information retrieval: recent advances and beyond, IEEE Access (2023).
[41] D. Tunkelang, Faceted Search, Morgan & Claypool / Synthesis Lectures, 2009. doi:10.1007/978-3-031-02262-3.
[42] X. Niu, et al., Understanding faceted search from data science and human factors perspectives, ACM Transactions on Interactive Intelligent Systems (2019).
[43] P. Mell, T. Grance, et al., The nist definition of cloud computing (2011).
[44] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, A view of cloud computing, Communications of the ACM 53 (2010) 50–58. doi:10.1145/1721654.1721672.
[45] A. Sunyaev, Cloud computing, in: A. Sunyaev (Ed.), Internet computing, Springer, Cham, 2020, pp. 195–236. doi:10.1007/978-3-030-34957-8{\textunderscore}7.
[46] M. Nandy, A. Dubey, Investigating topology optimization techniques for enhancing network performance in cloud computing environments, in: 2025 International Conference on Automation and Computation (AUTOCOM), 2025, pp. 1240–1245. doi:10.1109/AUTOCOM64127.2025.10957600.
[47] B. Burns, B. Grant, D. Oppenheimer, E. Brewer, J. Wilkes, Borg, omega, and kubernetes, Communications of the ACM 59 (2016) 50–57. doi:10.1145/2890784.
[48] N. Dragoni, et al., Microservices: Yesterday, today, and tomorrow, Present and Ulterior Software Engineering (2017) 195–216. doi:10.1007/978-3-319-67425-4_12.
[49] I. Baldini, P. Castro, K. Chang, P. Cheng, S. Fink, V. Ishakian, N. Mitchell, V. Muthusamy, R. Rabbah, A. Slominski, P. Suter, Serverless computing: Current trends and open problems, Research Advances in Cloud Computing (2017) 1–20. doi:10.1007/978-981-10-5026-8_1.