Appendix A. Overview of the search strategy employed

The following provides details on the literature review conducted as part of the construction of the RAVEE, in order to gain a comprehensive understanding of the current approaches related to the problem and ensure novelty of our design. As a suitable means for this Sonnenberg and vom Brocke [1] consider a literature review. To ensure a systematic and grounded approach, we conducted a structured literature review based on the methodology defined by Kitchenham and Charters [2] for a Systematic Literature Review (SLR). First, the selection of our search string "process mining" AND "video data" is motivated, followed by a listing of the databases employed for the literature search, as well as an explanation of their selection. Finally, the search criteria applied to the results obtained from the database search are described.

The search string has been derived from our research question. Since there is hardly any research in our addressed field, we have aimed to keep the search string as broad as possible to gain a comprehensive overview. Hence, we extracted the term "video data" from the research question as this term describes part of the addressed research field. We identified "process mining" as a second term, covering a broader spectrum as the superordinate term "process discovery" from the research question. Further terms and abbreviations such as "PM" for process mining have been tested against the predefined sources but did not enhance the quality of results. Both of the terms have been connected with the AND-operator to form the final search string: "process mining" AND "video data".

Building on the idea of Wolfswinkel et al. [3] to generate a comprehensive list of all possible sources, which contain relevant literature, Ahmed and Letchmunan [4] argue that the data-bases ACM Digital Library, IEEE Xplore, Springer Link, SciVerse Scopus and Web of Science do include all relevant primary literature within the field of IS. By conducting a literature search with the previously defined search string in these databases, we obtained 115 results. These results were tested against different inclusion and exclusion criteria as proposed in Wolfswinkel et al. [3]. An overview of the different selection criteria is given in Table A.1.

As a result of our literature search we found seven different publications, that address the topic of extracting information from video data for the usage in process mining. The relatively small number of selected publications already implies a lack of research in this area. Table A.2 provides a short overview and summary of the approaches identified.

Evidently, all the approaches derived in our SLR provide valuable insights in their respective research area, but they do not identify activities solely based on the available video data without prior process-related input. Instead they do so with the help of previously acquired process knowledge. However, since the identification of process deviations and new process steps is embedded within the core of process mining [13], the objective of this work is to facilitate this by using unsupervised mechanisms that can be deployed in a wide range of application domains.

Criterion	Explanation
I1 - Not a duplicate	For this criterion, a duplicate is defined according to Kofod-Petersen [5]. A publication is seen as a duplicate if it is to be found under the same name from the same author in different sources, or if the same author publicized different studies of approximately same content.
I2 – Published in English	To achieve comparability in our SLR, we only considered English publications
I3 – peer reviewed primary literature	Only primary literature from peer-reviewed journals, conferences or workshops has been considered
I4 - Accessible	The literature must be accessible with the available online accesses
I5 – Clear focus on process mining in title and abstract	Only literature with clear focus on process mining is considered
I6 - Clear focus on video data in title and abstract	Only literature dealing with video data is considered
E1 – No implementation of a method to use video data in process mining	The literature must implement a solution for the usage of video data in process mining to not be excluded. *If the literature incorporates unstructured data only in a superficial or qualitative manner or exclusively relies on manual pre-processing to extract meanings, it is excluded.*

Table A.1: Overview of the applied Exclusion and Inclusion Criteria

Reference	Short description
Abbad Andaloussi et al. 2021 [6]	Exploration of how hybrid process artifacts can support users in practice by conducting, among others, process mining activities on video data obtained through eye-tracking.
Fichtner et al. 2023 [7]	Development of an approach to extract process relevant details from among others video data and enrich existing process models.
Lepsien et al. 2023 [8]	Development of a pipeline to extract process mining compliant event logs from unstructured video data and evaluation on a real-world dataset.
Knoch et al. 2019 [9]	Detection of material picking and placement during manual assembly of manufacturing processes including computer vision capabilities on generated video data and the use of the information thus obtained for process mining activities.
Knoch et al. 2020 [10]	Mining of actual process behavior from manual activities in manufacturing processes by equipping workstations with video cameras and extracting event logs using computer vision capabilities for process mining efforts.
Kratsch et al. 2022 [11]	Development of a reference architecture to facilitate the extraction of process-related information from video data for process mining activities using various computer vision techniques.
Zhou et al. 2023 [12]	Analyzation of group interactions in an educational setting using among others, process mining on video data.

Table A.2: Approaches that use Video Data for Process Mining.

References

[1] C. Sonnenberg, J. Vom Brocke, Evaluation patterns for design science research artefacts, in: European Design Science Symposium, Springer, 2011, pp. 71–83. doi:10.1007/978-3-642-33681-2_7.

[2] B. Kitchenham, S. M. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, Vol. 2, 2007.

[3] J. F. Wolfswinkel, E. Furtmueller, C. P. M. Wilderom, Using grounded theory as a method for rigorously reviewing literature, European Journal of Information Systems 22 (1) (2013) 45–55. doi:10.1057/ejis.2011.51.

[4] M. M. Ahmed, S. Letchmunan, A systematic literature review on challenges in service oriented software engineering, International Journal of Software Engineering and its Applications 9 (6) (2015) 173–186. doi:10.14257/ijseia.2015.9.6.17.

[5] A. Kofod-Petersen, How to do a Structured Literature Review in computer science, 2015. URL https://www.researchgate.net/publication/265158913_How_to_do_a_Structured_Literature_Review_in_computer_science

[6] A. Abbad Andaloussi, F. Zerbato, A. Burattin, T. Slaats, T. T. Hildebrandt, B. Weber, Exploring how users engage with hybrid process artifacts based on declarative process models: a behavioral analysis based on eye-tracking and think-aloud, Software and Systems Modeling 20 (5) (2021) 1437–1464. doi:10.1007/s10270-020-00811-8.

[7] M. Fichtner, S. Jablonski, A specification of how to extract relevant process details to improve process models, in: J. Filipe, M. Śmiałek, A. Brodsky, S. Hammoudi (Eds.), Enterprise Information Systems, Springer Nature Switzerland, Cham, 2023, pp. 391–414.

[8] A. Lepsien, A. Koschmider, W. Kratsch, Analytics pipeline for process mining on video data, in: C. Di Francescomarino, A. Burattin, C. Janiesch, S. Sadiq (Eds.), Business Process Management Forum, Springer Nature Switzerland, Cham, 2023, pp. 196–213.

[9] S. Knoch, N. Herbig, S. Ponpathirkoottam, F. Kosmalla, P. Staudt, P. Fettke, P. Loos, Enhancing process data in manual assembly workflows, in: ICBPM, Springer, Cham, 2019, pp. 269–280. doi:10.1007/978-3-030-11641-5_21.

[10] S. Knoch, N. Herbig, S. Ponpathirkoottam, F. Kosmalla, P. Staudt, D. Porta, P. Fettke, P. Loos, Sensor-based human–process interaction in discrete manufacturing (2020). doi:10.1007/978-3-030-11641-5_35.

[11] W. Kratsch, F. König, M. Röglinger, Shedding light on blind spots – developing a reference architecture to leverage video data for process mining, Decision Support Systems 158 (2022) 113794. doi:10.1016/j.dss.2022.113794.

[12] Q. Zhou, W. Suraworachet, M. Cukurova, Detecting non-verbal speech and gaze behaviours with multimodal data and computer vision to interpret effective collaborative learning interactions, Education and Information Technologies 29 (2024) 1071–1098. doi:10.1007/s10639-023-12315-1.

[13] W. van der Aalst, A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, et al., Process mining manifesto, Springer, Berlin, Heidelberg, 2012, pp. 169–194. doi:10.1007/978-3-642-28108-2_19.