Dialogue-based XAI for predictive policing: a field study

Fabian Beer[1,*,†], Dimitry Mindlin[1,†], Sebastian Kost[2], Isabel Krause[2], Katharina Schwarz[2], Kai Seidensticker[2], Philipp Cimiano[1,†] and Elena Esposito[1,3,†]

[1]*Bielefeld University, Universitätsstraße 25, Bielefeld, 33615, Germany*
[2]*State Office for Criminal Investigations of North Rhine-Westphalia, Düsseldorf, 40221, Germany*
[3]*University of Bologna, Via Zamboni 33, Bologna BO, 40126, Italy*

Abstract

AI systems are increasingly used in predictive policing settings in order to identify areas at higher risk of residential burglaries or other crimes, and thus support the ground work of field police officers that take preventive measures. Explaining predictions of AI systems is crucial in this setting to increase confidence of officers in acting upon these predictions. This paper presents our field study examining a research unit of the German police that develops and employs predictive policing techniques. We explore the use of a revised version of a previously developed dialogue-based XAI tool, adapted to allow police analysts to inquire the reasons for a model's prediction. Based on semi-structured interviews and ethnographic observations, we analyze user responses on the tool's utility and its potential integration into operational workflows. Our preliminary results concern two distinct levels: individual user interactions and broader organizational communication dynamics.

Keywords

Predictive Policing, Dialogue-based xAI, Explainable AI, Organization

1. Introduction

Algorithmic predictive systems have long been used in the field of crime prevention, leading to the development of predictive policing systems that identify potential criminal activities in order to support interventions to prevent them. Systems such as PredPol, PRECOBS, SKALA, HunchLab or KrimPro have been used for years by police departments in the US and Europe and are the object of much attention by public and private observers [1]. From an ethical perspective, predictive policing represents a sensitive application domain, as it potentially threatens civil and privacy rights [2], can lead to discrimination reinforcement [3] and faces challenges related to ensuring legal and moral fairness [4]. In such a sensitive application domain, there is a particularly urgent need for reliable and effective XAI tools that enable a diverse group of stakeholders – including practitioners, regulators, and the broader public – to evaluate and monitor the operation of predictive policing systems to meet their unique explanation needs, but most importantly also to increase confidence of those who act upon the predictions of such systems.

While some research has been conducted on understandability within the field of predictive policing [5, 6, 7, 8], there is a lack of evidence for the acceptability and usefulness of XAI. To address this gap and better understand the needs and requirements of police analysts and officers regarding explainability of predictions, we carried out a field study with the SKALA team from the Criminological Research Unit of the State Office for Criminal Investigations of North Rhine-Westphalia. SKALA develops their own ML models for making predictions about the probability of residential burglaries for individual residential

Late-breaking work, Demos and Doctoral Consortium, colocated with the 3rd World Conference on eXplainable Artificial Intelligence: July 09–11, 2025, Istanbul, Turkey

*Corresponding author.
†These authors contributed equally.
fabian.beer@uni-bielefeld.de (F. Beer); dimitry.mindlin@uni-bielefeld.de (D. Mindlin); Sebastian.Kost@polizei.nrw.de (S. Kost); Isabel.Krause@polizei.nrw.de (I. Krause); Katharina01.Schwarz@polizei.nrw.de (K. Schwarz); Kai.Seidensticker@polizei.nrw.de (K. Seidensticker); philipp.cimiano@uni-bielefeld.de (P. Cimiano); elena.esposito@uni-bielefeld.de,elena.esposito9@unibo.it (E. Esposito)
0009-0002-5719-9738 (F. Beer); 0000-0003-1367-4365 (D. Mindlin); 0000-0001-8385-4025 (K. Schwarz); 0009-0003-3001-6425 (K. Seidensticker); 0000-0002-4771-441X (P. Cimiano); 0000-0002-3075-292X (E. Esposito)

quarters. At the beginning of every week, predictions about the probability of residential burglaries for individual residential quarters are generated by SKALA's ML model and then shared with *regional police agencies* through a mutually accessible platform [9]. Every regional police agency employs designated *analysts* that receive these predictions, "evaluate" and eventually forward them to several operative forces. In the end, the predictions reach *frontline officers* who are expected to follow their indications, e.g. by patrolling the risk areas more intensely. At each of these steps of the communication chain, difficulties in understanding the ML predictions can occur. However, at the current state of the project, SKALA does not rely on any XAI method to explain their predictions. For this reason, we adapted a previously developed dialogue-based XAI solution by Mindlin and Cimiano [10] and performed a field trial with three different police units and four police analysts to understand how they would use the solution, the main interaction flows as well as observed limitations and problems. As the analysts are the first members of the police outside of SKALA that come in contact with the ML predictions, SKALA suggested to test the XAI tool with them and not with patrol officers.

We discuss related work (Section 2), the XAI tool (Section 3.1), the design of our study (Section 3.2) and present parts of our data and preliminary results (Section 3.3).

2. Related Work

Automated or AI-supported decision making often takes place within organizations [11]. Accordingly, for many XAI application domains like healthcare, the legal field or finance, the problem of explainability occurs in reference to AI-systems employed in organizations. Following Miller's [12] suggestion to study explanation as a social and communicative practice, several studies discuss how within these organizations the problem of explainability occurs and is dealt with practically [13, 14, 15]. As for the case of Predictive Policing, Waardenburg et al. [5] argue that in the case of the Dutch police the organization coped with difficulties in understanding opaque ML models through a new social role the authors term "Knowledge Brokers". This role was occupied by intelligence officers who mediated between data scientists and users. Eventually, these knowledge brokers replaced the ML predictions with their own. Egbert et al.[6] discuss two German predictive policing applications, differentiated by their respective degree of opacity, and analyze their consequences on the decision premises of "decision programs", "communication channels" and "persons" [16].

Studying the problem of explainability in the field is one thing. Another is the implementation of XAI tools within organizations. Until now, there has been little evidence of the usefulness and acceptability of XAI in the daily life of the police. The findings of a study using hand-crafted explanations showed that police officers tended to accept recommendations that align with their intuition, but were unlikely to change their minds in case of misalignment, even when provided with explanations [7]. Another study applied an XAI solution based on LIME [17] to explain the results of a text classifier to police officers [8]. Their results suggest that domain experts preferred natural language explanations over visualizations and numerical representations. While insightful, this use case represents a relatively low-stakes scenario focused on document classification rather than high-risk decision-making like risk area prediction.

To combine the research on the problem of explainability in organizations with research on the use of XAI tools in predictive policing, we conducted a field study in the German police to observe how the implementation of an XAI tool might shape the social practices of dealing with the problem of explainability and how, inversely, those already established social practices might shape the usage of the XAI tool. The data collection process and preliminary results are described in the following section.

3. Field Study

In order to prepare our field study with the analysts of three different police units, we ran bi-weekly meetings with SKALA over a period of six months. We obtained informed consent from the analysts, and

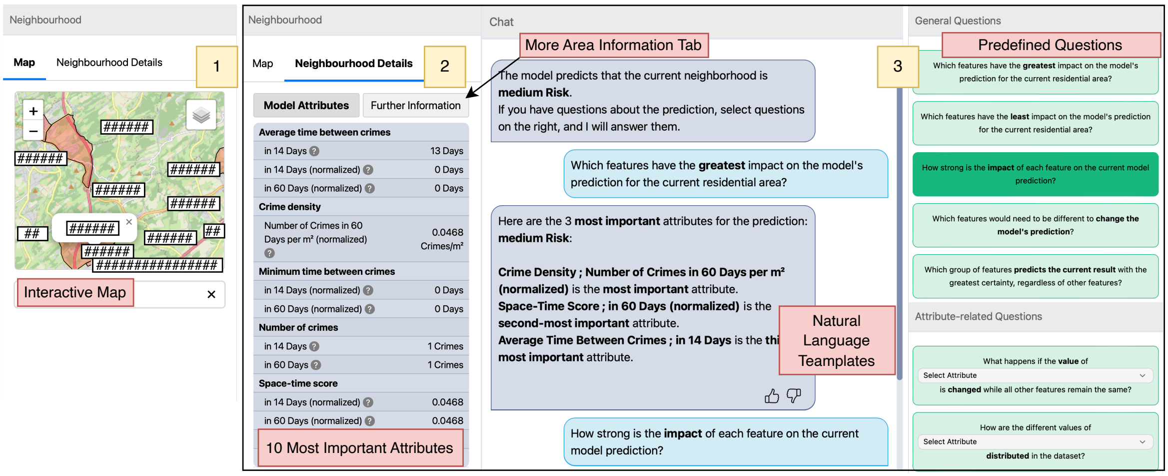

Figure 1: xAI Tool interface used in the field study with key components highlighted by red boxes. Yellow boxes indicate: (1) selection of the residential area, (2) display of neighborhood information, and (3) chat and questions to select. The black border indicates the full User Interface. Anonymised information is marked with #.

a formal approval from the state's Ministry of the Interior.[1] Due to the limited number of participants, we decided against using quantitative measures for this pilot study.

A design choice made in the context of this study was to use a concrete XAI solution that participating police officers could interact with to avoid obtaining only abstract answers on their requirements and needs. We hoped that using a specific and concrete tool would elicit discussions around suitability of current solutions and existing needs would be more specific and tangible. Thus, we used an existing dialogue-based XAI tool developed by Mindlin and Cimiano [10] that was adapted for the study on the basis of feedback received by SKALA in the six-month period.[2] This led to changes in the user interface and result presentation; the updated user interface is shown in Fig. 1.

3.1. Predictive Model and Data

Due to compliance constraints, we did not have access to the random forest model used by SKALA. Instead, we trained a surrogate model using historical data and the original model's labels. As is common in practice, the classes of interest (*risk* and *high risk*) were rare, with the rarest class comprising less than 1% of the data, resulting in a highly imbalanced dataset. Since the highest-risk category is determined through post-processing rules and appears infrequently, we merged both risk classes into a single category. This left us with two final classes: *no risk* and *medium risk*.

To address the dataset's high dimensionality (120 features), we applied univariate feature selection using an ANOVA F-test to retain the top 10 features. While the selected features vary slightly across locations, they generally reflect: i) concentration of offenses within specific time-radius windows, ii) offense count in the same windows, iii) variation in activity patterns over time, and iv) minimal or average time between offenses in those windows.

The system answers questions about feature importance and influence (LIME [17]), counterfactuals (DICE [18]), minimal sufficient explanations (Anchors [19]), individual feature effects (Ceteris Paribus [20]), and feature value distributions.

To train the surrogate models, we experimented with oversampling, undersampling, and class weighting. None of these methods improved F1-scores noticeably. Despite extensive tuning, the models performed modestly on the minority class. Across regions, we observed a trade-off between precision and recall: in some cases, the model was more conservative, identifying only a small portion of true

[1]Two analysts did not consent to be quoted *directly*.
[2]The dialogue-based interface is publicly available in the original paper without our modifications. The system uses model agnostic XAI methods and can be applied to different classification problems or extended to regression problems as well.

risk areas (recall of 0.25) but with higher precision (around 0.63); in others, it captured over half of the actual cases (recall of 0.51) but at the cost of more false positives (precision of 0.22). Macro-average F1-scores ranged between 0.6 and 0.66.

These outcomes reflect the difficulty of learning from sparse positive labels and limited signal quality from the original predictive model's labels. For the purpose of our research—examining the introduction of an XAI tool for the first time into the workflow of analysts—this level is acceptable.

3.2. Field Observations and Interviews with Analysts

One of the authors visited three police departments in North-Rhine-Westphalia in Q1 2025 at the beginning of a week when the analysts receive the predictions of the SKALA system. In total, four police analysts (named A1 - A4 in the following) participated in the study. In a first session, one of the authors with expertise in ethnographic studies [21] observed how the police analysts deal with the predictions of the model. The analysts commented on what they were doing, while the author would sit next to them, at times asking questions and taking field notes. After the analysts evaluated the predictions, we conducted semi-structured expert interviews [22]. The following questions were discussed:

- Could you describe how you process the predictions provided by SKALA?
- When was the last time you encountered difficulties in understanding a prediction?
- We heard from SKALA that they occasionally received requests about the predictions from individual police agencies. If you have ever asked such questions, could you describe how you filed these explanation requests?
- How helpful do you think an application that can answer questions about SKALA'S predictions would be?[3]

A few weeks later, we conducted a second session in which the analysts used the dialogue-based XAI solution. The aim of this "technology probe" [23] was to elicit natural responses of the analysts in their daily working environment. Thus, the analysts were given no instructions other than that they should interact with the tool in whatever way fits their needs. We provided access to the application via a web link.

In these interactions, the interviewer first provided some introductory remarks about the interface, after which the analysts were free to engage with the tool. They started by scrolling through the map, zooming in and out of it, and selecting (familiar) risk areas. After settling for a specific quarter, the analysts would proceed by going through the Model Attributes and Further Information section, reading the informational texts and commenting on the information provided. After this, the analysts started to engage with the questions and their respective answers from top to bottom and eventually discussed the tool in general. These interactions lasted about an hour each.

We then conducted semi-structured interviews focusing on the interaction with the tool, using the following guideline:

- Could you describe your general impression of the tool?
- How does this interaction with the tool differ from asking questions to the SKALA team?
- How did this interaction impact your perception of SKALAs model?
- How can this tool be implemented in your everyday routines?
- How understandable and appropriate were the available questions?
- How understandable and appropriate were the explanations?

The guideline was not strictly followed, deviating whenever the conversational flow required it [24]. We recorded the evaluation of the predictions, the interviews, as well as the tool interaction. The data is fully transcribed, but not openly available, as we did not get consent to share it beyond the project.

[3]All interviews concluded by asking the standard question: Is there anything else you would like to mention?

3.3. Presentation of Interview Data

The analysis of the interviews was performed with ATLAS.ti and Microsoft Excel. For the expert interviews, we applied the conceptual toolbox of sociological system theory, specifically its theory of organization [16], to reconstruct the explanatory communication between SKALA as explainer and the analysts of the police agencies as explainees. We focused on how the communication is conditioned by typical organizational factors like the hierarchical status of the staff, the chosen medium of distribution, the existing communication channels, and the degree of formality/informality.

As for the follow-up interviews regarding the XAI-tool, the answers of the analysts can be summarized briefly as follows. In general, the police analysts had a positive impression of the tool, but also some suggestions for improvement. One positive aspect highlighted is that the tool provides information that was previously inaccessible (e.g. number of crimes, additional socioeconomic information). In terms of suggested improvements, the analysts mentioned that they would like to influence the criteria that are displayed (A4, A3), and would like to receive information about why a certain area was not designated as a prediction area (A1). Two analysts (A3, A4) mentioned that the design should focus on answering simple comprehension questions.

The analysts also identified some limitations of the tool. They highlighted that the tool can answer questions immediately when they arise, but difficulties in understanding still need to be clarified directly with SKALA (A1, A2, A3). One analyst perceived the tool as simple to use, as answers can be obtained with one click (A1). This simplicity was challenged, however, by the need to get to know the tool so that one can understand its explanations (A2). It was also suggested that the tool should include a text field to send open questions to SKALA.

Regarding the question how the interaction with the tool impacted their perception of SKALA's model, two analysts mentioned that the tool enhanced their understanding of the model, revealing in particular the factors used for prediction. Although the analysts had some issues about understanding the tool and its usage, they found that the development was going in the right direction (A1, A2). For one analyst (A4), the tool confirmed the impression that the model used by SKALA is complex and difficult to explain in simple, non-scientific and lay terms. Another analyst (A3) could not relate their conceptualization of how SKALA generates predictions to what they learned from the tool.

As for the adequacy of the predefined questions, two analysts (A1, A2) considered the questions about the most and the least important features, as well as the first attribute-related question as relevant. The query for which group of features predicts the current result with the greatest certainty was deemed unintelligible and its usefulness was questioned. One analyst (A4) concurred that the questions might be relevant to their area of responsibility, highlighting the questions querying feature importance (See Fig. 1). However, the analyst mentioned that this would only be the case if an answer to the question would highlight factors that are considered to be relevant for the operational forces of the police (e.g. the current season) and the statistical aggregates of the model attributes were not considered to be so. Other respondents mentioned that some potentially relevant questions regarding the time of occurrence of the offenses (A3) or the offense code (A4) were not supported by the tool and that open questions are likely to emerge during everyday work (A1, A2).

The explanations were generally regarded as accurate, but also as partially difficult to understand. This difficulty was attributed to the employed "language". Some answers made reference to the difficulty of understanding aggregate statistical features, as can be seen in the following excerpt in which A4 describes an ideal explanation:

> "A4: Well, if down here it was thrown out (.) the connection to the highway is uninteresting. Well, if the attributes that are displayed here (.) if I could understand them
>
> A3: Yes (laughing)
>
> I: So, it's mainly about the attributes?
>
> A4: Yes, exactly. Well, the State Office could probably- 'Yes, okay, I understand, I understand' but when I- I would have to understand, okay (..) this residential quarter is not important

because people with low purchasing power live there, they can't afford watches and that's why they don't get burgled. For me that would be (..) well that's what I could understand."

Lastly, two analysts (A1, A2) mentioned that they would use the tool selectively to either answer questions they receive or to familiarize themselves more thoroughly with the model. Two analysts (A3, A4) stressed that the XAI tool could be helpful for newcomers. That the answers vary from week to week was seen to be an advantage over SKALAs FAQs. For an actual implementation, however, the questions would have to be formulated differently, and the answers would need to be more comprehensible (A3, A4).

These follow-up interviews were fully paraphrased, subsequently coded using in-vivo labels (e.g. 'scientific', 'the Practical'), and eventually compared thematically [22]. Here, two reoccurring themes emerged: 'scientific language' and 'relevance'. Both of these themes pointed towards difficulties in understanding raised by the interaction with the tool itself. Based on these two themes, we searched our material and selected passages in which either difficulties in understanding the XAI tool arose in practice, or were described in the interviews. In this way, a total of 58 text passages were identified. Directly comparing these passages allowed us to conclude that there is a lack of alignment between the conceptualization of the domain by analysts on the one hand, and the XAI tool and the ML model on the other. We specifically identified two dimensions of this misalignment: a mismatch of "language" and a mismatch of relevance structures. First, what the analysts describe as "complicated scientific language", our preliminary results suggest, point to the difference between the categories the analysts use to think about the world and the statistical aggregates processed by the model to generate its predictions. Secondly, the analysts often evaluated the information provided by the XAI tool in terms of relevance for their work by employing a distinction between *questions only of personal interest* and *questions that are relevant for their field of tasks*. Our preliminary results also suggest that the two mismatches are interrelated at least insofar, as the 'scientific' would be attributed to the field of tasks of SKALA. Thus, labeling something as 'too scientific' might also imply a demarcation in the field of tasks and therefore in relevance structures.

By searching for passages of maximum contrast, we identified, for one, that even when the language barrier was not an issue, some analysts struggled in understanding the counterfactual explanations of our tool. Further, the use of the tool exposed analysts to the inherent uncertainty involved in the predictions, which was previously "absorbed" by the model. Here, we refer to the sociological concept of "uncertainty absorption"[16]. As is typical for organizations, decision-making often requires sub-decision to be made and, since decisions have to be made under conditions of incomplete information, persons or units have to assume responsibility for them. This "uncertainty absorption" simplifies the decision-making of others but it requires that only the decision is forwarded to another unit, not the incomplete information on which it is based. As argued by Besio et al. [25], ML models in organizations can take responsibility in this sense of absorbing the uncertainty of a decision. Implementing an XAI-tool, however, enables the analysts to have access to the features that the model's predictions are based on, i.e. to the uncertainty previously absorbed by SKALA.

4. Conclusion

Our field study relied on hands-on sessions with a dialogue-based XAI solution and semi-structured interviews. While our study is limited by the small number of participants and the prototypical status of the tested XAI tool, it could nevertheless provide insights at two levels: the relationship of individual users with the tool and the communication dynamics within the organization.

At the individual level, our preliminary results suggest that analysts struggled with the concepts of machine learning that they could not map straightforwardly into their own world model, labeling them as overly "scientific". Analysts argued that explanations constructing a meaningful narrative with concepts directly mappable to their own conceptualization would be easier to grasp. These preliminary results are in accordance with a study of police employees that suggests that explanations are often too abstract and that usability and usefulness are more valued than interpretability [26].

At the level of the organization, our preliminary results suggest that even if some individual mismatches were resolved, organizations would have to cope with issues of unabsorbed uncertainty in their decision-making processes. This hints at the fact that the introduction of XAI solutions has significant impacts on organizational processes that need to be revisited.

The validity of these preliminary results depends on the extent to which the surrogate model faithfully captures the real model. Regarding the difficulty of understanding model attributes, this is the case, as for non-intuitive counterfactual explanations it may not. We would like to address this distinction in future work. Additionally, many details of this study are not generalizable as it only involves one regional police unit. However, we expect that other law enforcement agencies whose algorithmically supported decision-making process is based on a division of labor will encounter similar questions that arise out of general aspects of organizations. E.g. in what existing "ecology of explanations" [14] will the introduction of an XAI tool intervene, and eventually, how does the tool become a part of this larger ecology?

This study is the result of interdisciplinary cooperation between computer science researchers, sociologists and the Criminological Research Unit of the State Office for Criminal Investigations of North Rhine-Westphalia. This unique collaboration enabled the adaptation of an existing XAI tool to the particular dataset and requirements of SKALA and perform a test of the system 'in vivo' and collecting feedback. We think that such studies are necessary to better understand the 'reality' in which XAI systems are employed, as this reality check can guide the design and further research of XAI Systems.

Acknowledgments

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): TRR 318/1 2021 – 438445824 and by the European Research Council (ERC) under Advanced Research Project PREDICT no. 833749. The studies with adult participants were approved by the Ethics Committee of Paderborn University on January 11, 2021.

Declaration on Generative AI

During the preparation of this work Chat-GPT-4o and Grammarly were used to check grammar and spelling. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] S. Egbert, M. Leese, Criminal Futures: Predictive policing and everyday police work, Taylor & Francis, London / New York, NY, 2021.

[2] E. Schlehahn, P. Aichroth, S. Mann, R. Schreiner, U. Lang, I. D. H. Shepherd, B. L. W. Wong, Benefits and pitfalls of predictive policing, in: Proc. of the European Intelligence and Security Informatics Conference, IEEE, 2015, pp. 145–148. doi:10.1109/EISIC.2015.29.

[3] D. Robinson, L. Koepke, Stuck in a Pattern. Early evidence on "predictive policing" and civil rights, Technical Report, Upturn, 2016.

[4] J. Kleinberg, S. Mullainathan, M. Raghavan, Inherent trade-offs in the fair determination of risk scores, in: Proc. of the 8th Innovations in Theoretical Computer Science Conference (ITCS 2017), volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017, pp. 43:1–43:23. doi:10.4230/LIPIcs.ITCS.2017.43.

[5] L. Waardenburg, M. Huysman, A. Sergeeva, In the land of the blind, the one-eyed man is king: Knowledge brokerage in the age of learning algorithms, Organization Science 33 (2021) 59–82.

[6] S. Egbert, E. Esposito, M. Heimstädt, Vorhersagen und Entscheiden: Predictive Policing in Polizeiorganisationen, Soziale Systeme 26 (2021) 189–216.

[7] F. Selten, M. Robeer, S. Grimmelikhuijsen, 'Just like I thought': Street-level bureaucrats trust AI recommendations if they confirm their professional judgment, Public Administration Review 83 (2023) 263–278. doi:10.1111/puar.13602.
[8] M. David, E. S. Mbabazi, J. Nakatumba-Nabende, G. Marvin, Crime forecasting using interpretable regression techniques, in: Proc. of the 7th International Conference on Trends in Electronics and Informatics (ICOEI), IEEE, 2023, pp. 1405–1411. doi:10.1109/ICOEI56765.2023.10126071.
[9] M. Heimstädt, S. Egbert, Actionable predictions: How designers of algorithmic systems calibrate criminal futures, in: 40th EGOS Colloquium, University of Milano-Biccoca, Italy, 2024.
[10] D. Mindlin, A. S. Robrecht, M. Morasch, P. Cimiano, Measuring user understanding in dialogue-based xai systems, in: Proc. of the European Conference on Artificial Intelligence (ECAI), volume 392, IOS Press, 2024, pp. 1148–1155. doi:10.3233/FAIA240608.
[11] S. Kette, „Computer says no"? Konsequenzen der Algorithmisierung von Entscheidungsprozessen, Soziale Systeme 26 (2021) 160–188. doi:10.1515/sosys-2021-0006.
[12] T. Miller, Explanation in artificial intelligence: Insights from the social sciences, Artificial Intelligence 267 (2019) 1–38. doi:10.1016/j.artint.2018.07.007.
[13] C. Borch, B. H. Min, Toward a sociology of machine learning explainability: Human–machine interaction in deep neural network-based automated trading, Big Data & Society 9 (2022). doi:10.1177/20539517221111361.
[14] M. Gutierrez Lopez, , S. Halford, Explaining machine learning practice: findings from an engaged science and technology studies project, Information, Communication Society 28 (2025) 616–632. doi:10.1080/1369118X.2024.2400130.
[15] S. Lebovitz, H. Lifshitz-Assaf, N. Levina, To engage or not to engage with ai for critical judgments: How professionals deal with opacity when using ai for medical diagnosis, Organization Science 33 (2022) 126–148. doi:10.1287/orsc.2021.1549.
[16] N. Luhmann, Organization and Decision, Cambridge University Press, Cambridge, MA, 2018.
[17] M. T. Ribeiro, S. Singh, C. Guestrin, "why should I trust you?": Explaining the predictions of any classifier, in: KDD, ACM, 2016, pp. 1135–1144. doi:10.1145/2939672.2939778.
[18] R. K. Mothilal, A. Sharma, C. Tan, Explaining machine learning classifiers through diverse counterfactual explanations, in: Proc. of the 2020 Conference on Fairness, Accountability, and Transparency, ACM, 2020, p. 607–617. doi:10.1145/3351095.3372850.
[19] M. T. Ribeiro, S. Singh, C. Guestrin, Anchors: High-precision model-agnostic explanations, in: Proc. of the AAAI Conference on Artificial Intelligence, volume 32, AAAI Press, 2018, pp. 1527–1535. doi:10.1609/aaai.v32i1.11491.
[20] M. Kuźba, E. Baranowska, P. Biecek, pyCeterisParibus: explaining machine learning models with ceteris paribus profiles in Python, Journal of Open Source Software 4 (2019) 1389.
[21] R. M. Emerson, R. I. Fretz, L. L. Shaw, Writing Ethnographic Fieldnotes. Second Edition, Chicago University Press, Chicago, IL / London, 2011.
[22] M. Meuser, U. Nagel, ExpertInneninterviews — vielfach erprobt, wenig bedacht, in: A. Bogner, B. Littig, W. Menz (Eds.), Das Experteninterview: Theorie, Methode, Anwendung, Springer VS, Wiesbaden, 2002, pp. 71–93.
[23] H. Hutchinson, W. Mackay, B. Westerlund, B. B. Bederson, A. Druin, C. Plaisant, M. Beaudouin-Lafon, S. Conversy, H. Evans, H. Hansen, N. Roussel, B. Eiderbäck, Technology probes: inspiring design for and with families, in: Proc. of the SIGCHI Conference on Human Factors in Computing System, ACM, 2003, p. 17–24. doi:10.1145/642611.642616.
[24] C. Hopf, Pseudo-exploration – Thoughts on the techniques of qualitative interviews in social research, Zeitschrift für Soziologie 7 (1978) 97–115. doi:10.1515/zfsoz-1978-0201.
[25] C. Besio, C. Fedtke, M. Grothe-Hammer, A. Karafillidis, A. Pronzini, Verantwortungsvolle Maschinen ohne Verantwortlichkeit? Datenintensive Algorithmen in Organisationen, Soziale Systeme 26 (2022) 129–159. doi:10.1515/sosys-2021-0005.
[26] E. Herrewijnen, M. B. Loerakker, M. Vredenborg, P. W. Woźniak, Requirements and attitudes towards explainable ai in law enforcement, in: Proc. of the 2024 ACM Designing Interactive Systems Conference, ACM, 2024, pp. 995–1009. doi:10.1145/3643834.3661629.