Inferring Political Orientation from Credit Score-Relevant Variables

An Empirical Study on Profiling and Proxy Inference Through Sensitive Attributes

David Schnepp
Department of Business and Management
Brandenburg University of Applied Science
Brandenburg a. d. Havel, Germany
Email: schnepp@th-brandenburg.de

Olga Levina
Department of Business and Management
Brandenburg University of Applied Science
Brandenburg a. d. Havel, Germany
Email: levina@th-brandenburg.de

***Abstract*— The increasing availability of data in everyday life expands the possibilities for profiling individuals across social, economic, and political domains. Even when sensitive attributes are not explicitly collected, they may be inferred from seemingly non-sensitive demographic and socioeconomic information. This paper explores whether political orientation can be predicted from credit-related variables, motivated by concerns that algorithmic systems might infer sensitive attributes from seemingly non-sensitive data. Using data from the European Social Survey (Round 10), voting behavior in the 2021 German federal election is treated as a multiclass classification problem. Several common supervised learning methods, such as logistic regression, support vector machines, k-nearest neighbors, and boosting-based decision trees, are employed. The results indicate that political orientation cannot be reliably predicted in this context. Nonetheless, this does not eliminate the broader risk of proxy-based inference. Moreover, inferability depends not only on the variables themselves, but also on the broader analytical context. The literature review indicates that differences in data availability, feature construction, preprocessing, and model training approaches can significantly influence information gain and predictive inferability. Instead, the findings emphasize the importance of systematic, context-specific assessments of inferability, with significant consequences for data protection and AI regulation.**

Keywords - Machine Learning; Profiling; Inference of sensitive attributes; Proxy discrimination.

I. Introduction

The increasing deployment of automated, data-driven systems across domains such as finance, public administration, and political processes fundamentally reshapes how personal information is processed and interpreted [1][2]. It also challenges the assumption that data used in such systems can be clearly categorized as sensitive or non-sensitive [3]. Data is not neutral [4]. Additional information can be derived from aggregating, combining, and analyzing the collected data, often beyond the original purpose of data collection [1][5].

This raises a central societal concern. Sensitive personal attributes do not need to be explicitly collected to be effectively used. Yeom, Datta, and Fredrikson, for example, demonstrated that combinations of ostensibly non-sensitive variables in predictive policing-related datasets can function as powerful proxies for racial composition and reinforce discriminatory outcomes in algorithmic decision-making. In their analysis, a proxy constructed from 58 seemingly unrelated features exhibited a particularly strong association with race [6]. They can be reconstructed as so-called proxy variables from data that were originally considered non-sensitive, giving rise to new forms of indirect discrimination. They enable the inference of characteristics such as political orientation, ethnicity, or religion. As a result, individuals may be subject to differential treatment based on attributes that were neither disclosed nor directly processed [7].

The protection concern thus shifts from the question of which data is collected to what information can be inferred from it. This shift has profound implications for data protection and fairness, as it undermines traditional regulatory approaches that focus primarily on input data categories. Algorithmic inference of sensitive attributes therefore represents not only a technical challenge but a structural risk for discrimination in increasingly data-driven societies [5][7][8].

The research question addressed in this study is thus whether the political orientation of individuals, a notably safeguarded sensitive characteristic, can be inferred from non-sensitive, credit-relevant attributes using standard supervised learning models.

Finding an affirmative response would suggest that based on non-sensitive, credit-relevant data an AI system can potentially leverage individual political preferences. Regulatory frameworks such as the EU AI Act (AIA) and the General Data Protection Regulation (GDPR) seek to establish safeguards for sensitive personal data, particularly by restricting their collection and use in automated decision-making contexts [9][10]. However, these protections become less straightforward when sensitive attributes are not directly processed but can be inferred from non-sensitive data. The European Data Protection Board addresses this challenge by clarifying that profiling, defined as the automated processing of personal data to evaluate or predict individual characteristics, also encompasses situations in which sensitive information is inferred. While this interpretation broadens the scope of data protection law, it still raises practical and conceptual challenges regarding the effective regulation and enforcement of such inference-based processing [5].

To address this question, an empirical investigation was conducted using the European Social Survey (ESS) Round 10. The target variable is voting behavior in the 2021 German

 ISBN: 978-1-68558-382-8

federal election, which results in a multiclass classification problem [11]. Core supervised-learning algorithms were used for processing the non-sensitive attributed: logistic regression, Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and a boosting-based decision tree.

The empirical results consistently show across all model types that no reliable prediction of political orientation is possible. Nevertheless, this analysis demonstrates that systematic approaches to estimate ex ante the likelihood of attribute inference under different conditions are needed in research and practice. The inability to infer political orientation in this setting does not guarantee safety in others; practitioners should not interpret negative results as proof of absence of profiling risk and regulators may need to consider implementing flexible, case-by-case evaluation mechanisms instead of rigid assumptions about inferability.

The contribution of this research is twofold. First, an empirical assessment shows that, in this dataset and under this label definition, predictive performance remains close to simple baselines, suggesting limited inferability from credit-relevant attributes alone. Second, the scope of this finding is clarified: it does not generalize to all contexts. Inferability is shaped by the feature space, label granularity, sample size, class balance, and modeling choices. We therefore recommend systematic, context-specific inferability audits as a prerequisite for claims about the feasibility or safety of inferring sensitive attributes. This targeted evidence can help calibrate regulatory discussions that currently alternate between universal warnings and untested assurances.

The remainder of the paper is structured as follows: First, related work on the inference of sensitive attributes is reviewed. This is followed by a description of the study's methodology. Subsequently, the empirical results are presented and discussed, considering their implications.

II. Data Processing and Sensitive Attributes: State of the Art

Empirical research provides substantial evidence that sensitive personal attributes can be inferred from non-sensitive data with considerable accuracy. A prominent example is the work of Kosinski et al. [12], which demonstrates that digital behavioral traces such as Facebook Likes can predict a wide range of sensitive attributes. Exploiting a dataset of approximately 58,000 individuals (170 likes on average) and applying logistic regression for dichotomous variables and linear regression for numeric variables combined with dimensionality reduction, the study achieved high predictive performance across multiple categories. Both regression models used 10-fold cross-validation and k=100 SVD components. Area Under Curve (AUC) values reached 0.95 for distinguishing between Caucasian and African American individuals, 0.82 for religion (Christianity vs. Islam), and 0.85 for political orientation (Democrat vs. Republican). These results highlight the strong inferential power of seemingly non-sensitive behavioral data. Performance for numeric features was also partly strong. Age was predicted with a Pearson correlation of 0.75, whereas predictions for the number of Facebook friends were less accurate, with a Pearson correlation of 0.47.

Similar findings emerge in financial contexts. In a study by Hassani [13], ethnicity was predicted from standard credit-scoring variables such as income, credit limits, and account balances. The dataset comprised about 400 cases covering African-American (99), Asian (102), and Caucasian (199) individuals. Applying a Random Forest model (750 trees) and a 75/25 train-test split, the study achieved an F1-score of approximately 0.65 on the original dataset, which increased to around 0.70 under modified conditions and up to 0.99 after applying the Synthetic Minority Oversampling Technique (SMOTE) to rebalance the dataset by creating synthetic data. The results indicate that financial variables can act as strong proxies for ethnicity, although the study is limited by its small sample size and reliance on data manipulation techniques.

More recent work further demonstrates the inferability of sensitive attributes across data modalities. Chaturvedi & Chaturvedi [14] show that religion can be predicted with very high accuracy from personal names alone. Employing a large-scale dataset of over 115,000 households from India and an 80-10-10 train-validation-test split, they applied several machine learning models, including dictionary-based approaches, language models, logistic regression, SVM, and Convolutional Neural Network (CNN) sequencing models. Evaluation was performed with F1, precision, and recall. The study achieved strong predictive performance, exceeding that of typical dictionary-based approaches. The best predictive performance for single names was obtained by CNN with F1 of 95.86, while for concatenated names (containing more information) SVM outperformed CNN with F1 of 97.33. The results illustrate that even basic identifiers contain rich latent information that can be exploited for sensitive inference.

Beyond explicitly sensitive targets, related work also demonstrates the predictive power of behavioral financial data for complex personal characteristics. Kim et al. [15] applied large-scale open-banking transaction data comprising approximately 100,000 individuals and 180 million transactions to model financial vulnerability. Combining open-banking data with typical credit-scoring approaches opens wide potential for credit risk valuation. Applying machine learning models such as Random Forest, logistic regression, and XGBoost, the study achieves an accuracy of around 0.77 and an F1-score of approximately 0.62 in identifying people with disabilities. They used an 80/20 train-test split with hyperparameter tuning via grid search and 10-fold cross-validation.

The aforementioned studies underscore that the efficacy of sensitive attribute reconstruction is fundamentally contingent on the alignment between the model architecture and the underlying data structure. This architecture-performance gap is particularly evident in Chaturvedi & Chaturvedi [14], where CNNs excelled at extracting features from sequential name patterns (F1: 95.86), whereas SVMs demonstrated superior performance on high-dimensional, concatenated datasets (F1: 97.33). Such variations indicate that additional data only enhances predictive performance if the model has sufficient complexity to process the increased information density. The application of resampling strategies, such as SMOTE in

Hassani [13], demonstrates that addressing class imbalances can drastically improve model performance, increasing F1 Scores from 0.65 to 0.99 by enabling algorithms to learn patterns in minority classes that would otherwise be obscured by statistical noise. The inferential power of these models is highly domain-specific and depends on the quality of proxy variables. The selective predictive power observed in Kosinski et al. [12] illustrates this limitation. While Facebook Likes serve as a high-performance proxy for ethnicity (0.95 AUC), they are less informative for quantitative social metrics such as friend count (0.4 correlation).

The existing research suggests that the transition from non-sensitive input to sensitive inference is governed by a rigorous methodological framework involving individual feature pre-processing (i.e., dimensionality reduction), cross-validation, and systematic hyperparameter tuning. However, the entire prediction process is highly individualized, based on the dataset and (target) sensitive variables, which collectively challenge current technical and regulatory safeguards [12][13][14][15].

III. RESEARCH DESIGN

The research question of this study is whether individuals' political orientation, operationalized via voting behavior as a particularly sensitive attribute, can be reconstructed using established supervised learning models based solely on legitimate, credit-relevant features. These models were selected as established approaches in applied machine learning for structured data, without relying on highly complex model architectures.

To address this question, an empirical investigation was conducted using the European Social Survey (ESS) Round 10. The ESS is a social science dataset that includes demographic and socioeconomic information, as well as respondents' political orientations, such as their party affiliation and views on specific political issues, e.g., the government's role in reducing income inequality [16]. For the analysis, 15 features relevant to real creditworthiness assessments were selected from the dataset, as shown in Table I.

TABLE I. OVERVIEW OF SELECTED FETAURES AND DATA TYPES.

ESS-Variable	Credit-Scoring related feature	Data Type
prtvfde2	Voting behavior (target var)	Nominal
gndr	Gender	Nominal
agea	Age	Metric
domicil	Domicile / housing location	Ordinal
rshpsts	Relationship status	Nominal
hhmmb	Household size	Metric
hhcd	Children living in household	Metric
wrkctra	Contract type	Nominal
educde1	General education level	Ordinal
edubde2	Vocational education level	Ordinal
hinctnta	Household income	Ordinal
hincsrca	Main source of income	Nominal
nacer2	Industry sector	Nominal
isco08	Occupation	Nominal
ctzcntr	Citizenship	Nominal
brncntr	Country of birth	Nominal

The selection of input features followed a structured three-step approach. First, relevant variables were identified based on established criteria from the literature on creditworthiness assessments and complemented by practical data requirements derived from two real-world credit application forms. These criteria were then systematically mapped onto the ESS dataset by identifying where possible and valid approximations. This applies, for instance, to credit scoring categories such as address or the number of children in a household. The *address* category was mapped to *domicil*, whereas the number of children living in the household was determined by aggregating the variables for *persons in household* in combination with the *relationship to respondent* indicators (*rshipa2–rshipa15)* [16]. Voting behavior in the 2021 German federal election was used as the target variable, resulting in a multiclass classification problem. The ESS variable prtvfde2 was restricted to seven valid party classes: *1 = CDU/CSU (15.9%), 2 = SPD (20.2%), 3 = Die Linke (3.9%), 4 = Bündnis 90/Die Grünen (15.4%), 5 = FDP (9.8%), 6 = AfD (3.8%), and 7 = Andere (4.9%).* Responses without substantive informational content, including 66 = not applicable, *77 = refusal, 88 = don't know,* and *99 = no answer*, were recoded as NaN and excluded from the modeling dataset. Since supervised classification requires a valid target label, only observations with values between 1 and 7 were retained for model training.

The analytical approach is based on the Cross-Industry Standard Process for Data Mining framework, a process model for structuring machine learning projects [17]. All data processing were implemented in Python within Jupyter Notebook environments using pandas for data manipulation and scikit-learn for machine learning workflows. Four supervised learning algorithms were used: logistic regression, Support Vector Machines (SVMs), k-Nearest Neighbors (KNN), and a boosting-based decision tree (HistGradientBoostingClassifier). Data pre-processing included the encoding of categorical and ordinal variables as well as the handling of missing values through variable-specific imputation strategies. For logistic regression, SVM, and KNN, categorical variables were transformed using one-hot encoding to obtain a numerical representation suitable for model training. Numerical predictors were subsequently scaled according to model-specific requirements. StandardScaler was applied for logistic regression and SVM, whereas MinMaxScaler was used for KNN due to its distance-based nature. For the boosting-based decision tree model, categorical variables were retained as categorical data types where applicable, allowing them to be processed without one-hot encoding. All preprocessing operations were integrated directly into the machine learning pipelines to ensure a standardized structure of the input matrices and to prevent data leakage during cross-validation. Particular attention was given to the stratified train (0.8)-test (0.2) split to enable evaluation on independent data and to control for variance and bias. Due to preprocessing procedures, the final analytical sample used for model training comprised n=6,443 observations.

For hyperparameter optimization, GridSearchCV was used to perform a structured, exhaustive search across the entire defined parameter space. Systematic hyperparameter variations were evaluated. For SVM, the grid search explored variations in C, gamma, and different kernel types, while for k-NN it evaluated adjustments to n_neighbors, distance metrics, and weighting schemes. Each configuration was tested in experimental setups, allowing the performance of all parameter combinations to be assessed under consistent conditions and enabling the identification of the optimal model configuration. To obtain robust and distributionally faithful estimates of generalization performance, stratified 5-fold cross-validation was used. Preliminary experiments with stratified 10-fold cross-validation did not yield meaningful performance improvements, while substantially increasing computational complexity. The combination of GridSearch and StratifiedKFold ensures that model variants are optimized independently from the final test data. To support reproducibility, additional details on missing-value handling, imputation strategies, hyperparameter grids, the specific algorithm, and the encoding of isco08 and nacer2 variables are available from the authors upon request.

Model performance was assessed using common classification metrics, particularly accuracy, F1 score, precision, sensitivity, and specificity. The no-information rate, defined as the accuracy achieved by always predicting the most frequent class, served as a reference point. This methodological framework enables a nuanced evaluation, especially in the presence of imbalanced class distributions typically associated with electoral choices.

IV. RESULTS

The empirical results consistently show that no reliable prediction of political orientation is possible across all model types. Across all models, accuracy ranged from 0.246 to 0.316 and thus only marginally exceeded the no-information rate of 0.272 (Fig. 1).

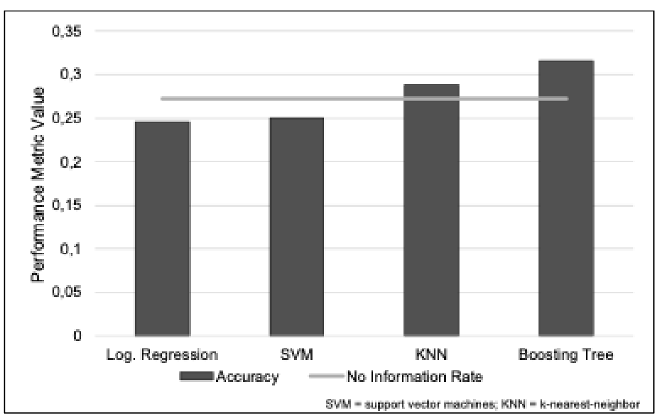

Figure 1. Comparison of Classification Models Accuracy and No Information Rate.

F1-scores, precision, and sensitivity remained consistently low, with particularly poor model performance for smaller parties and rare classes. Regardless of whether the F1-score is computed as a macro or weighted average, the results follow a similar pattern. Macro F1-scores ranged from 0.194 to 0.225, while weighted F1-scores ranged from 0.251 to 0.280, indicating overall weak classification performance (Fig. 2).

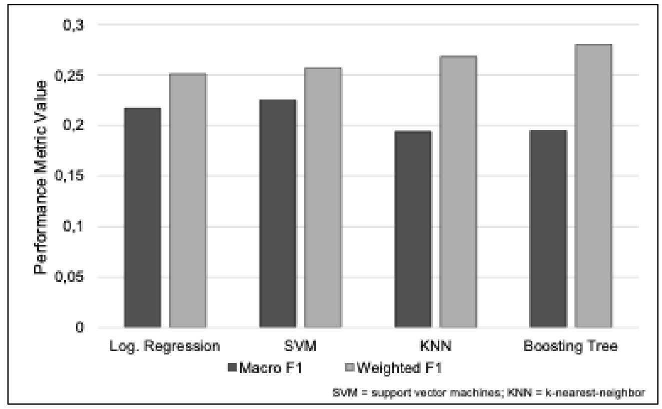

Figure 2. Comparison of Classification Models by F1-Score Metrics.

Macro sensitivity between 0.197 and 0.253 shows that only about one-fifth to one-quarter of actual class instances were correctly identified, and macro precision between 0.222 and 0.234 indicates that fewer than one-quarter of positive predictions were accurate. The models consistently achieved high macro specificity values ranging from 0.870 to 0.874, indicating reliable identification of non-memberships. However, they nonetheless failed to produce meaningful differentiation for positive class assignments (Fig. 3).

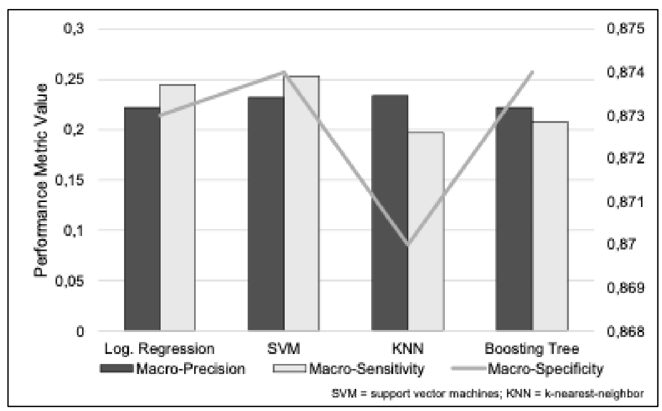

Figure 3. Comparison of Classification Models by Core Performance Metrics.

A permutation-based feature-importance analysis was added to understand the importance of each variable. No feature showed strong predictive relevance; age had the highest relative importance (0.0297 ± 0.0112), followed by general education level (0.0115 ± 0.0081), domicile (0.0109 ± 0.0070), children living in the household (0.0106 ± 0.0030), and household income (0.0099 ± 0.0039). Several variables contributed only marginally or negatively, suggesting that the selected creditworthiness-related features provide limited information for reconstructing political orientation.

V. Discussion

This research addressed the question of whether it is possible to infer voting behavior based on data collected for socio-economic and credit score- related purposes. Popular classification algorithms were used to combine and strucutre the data. The results showed that no reliable prediction of political orientation was possible across the applied model types. The findings of this study therefore indicate that no meaningful inference of political orientation could be drawn from the given setup, suggesting that attribute profiling based on typical features related to creditworthiness is not feasible in this specific context.

However, this result should be interpreted in light of the imbalanced target distribution. Several party classes are sparsely represented, particularly *Die Linke* (3.9%) and *AfD* (3.8%), which may limit the ability of supervised learning models to learn reliable class-specific patterns. Poor performance should therefore not be interpreted solely as evidence of real non-inferability, but also as potentially reflecting methodological constraints resulting from sparse and unevenly distributed target classes. The negative result may therefore be due to several reasons, which have yet to be fully clarified.

Additionally, prior research demonstrates that reconstructing sensitive attributes from non-sensitive data can be highly effective in other settings. This contrast highlights that such inference processes are not universally generalizable in either direction. Rather, their success critically depends on the underlying data, preprocessing, the presence and strength of proxy variables, and the chosen modeling approach [12][13][14][15]. Ultimately, the current findings do not definitively determine whether political orientation can be derived from creditworthiness-related features.

Furthermore, the findings underscore the methodological sensitivity of inference outcomes. Even within the same empirical setting, variations in modeling choices, such as different architectures or resampling strategies, can yield different results. Although this study did not observe meaningful predictive performance, these dependencies indicate that outcomes are not method-invariant. In addition, the reconstructability of sensitive attributes appears to be attribute-specific, as some characteristics may be more easily inferred than others depending on their relationship to available proxy features. Consequently, the absence of predictive performance in a given setup does not imply the absence of attribute profiling risk, because absolute certainty about the absence of attribute profiling risk, or about non-profiling, cannot be established.

This context dependence has important implications for assessing the risks of attribute profiling. The results suggest that the mere availability of non-sensitive data does not automatically enable the reconstruction of sensitive attributes. At the same time, inferability is the mechanism by which non-sensitive variables may become proxies for sensitive attributes, thereby enabling indirect discrimination. However, the variability observed across studies indicates that such risks cannot be reliably excluded ex ante. Consequently, the feasibility of attribute profiling remains inherently difficult to predict, posing a significant challenge for regulatory frameworks such as the GDPR, which aim to govern and restrict such practices.

VI. Conclusion and futre work

In this paper, we examined whether creditworthiness-related variables can be used to infer political orientation as a sensitive attribute. Using selected socio-demographic and economic features from the ESS dataset, we trained and evaluated several machine learning models to assess the feasibility of attribute profiling in this specific context. The results indicate that no reliable prediction of political orientation could be achieved across the applied models, suggesting that such inference is not feasible under the given conditions. At the same time, these findings should be interpreted cautiously, as the absence of reliable predictive performance does not necessarily imply the absence of inferability, but may also reflect methodological constraints such as class imbalance.

The results of this study suggest several important directions for future research and regulatory consideration. First, the inherent unpredictability of attribute profiling's feasibility calls for a shift away from binary risk assessments toward more probabilistic, context-sensitive evaluation frameworks. Future work should therefore focus on systematically mapping the conditions under which sensitive attribute inference becomes viable, including the roles of data richness, feature correlations, class distributions, and model complexity. In particular, comparative studies across a wider range of modeling approaches, from classical statistical methods to advanced machine learning architectures such as neural networks, would help to better understand the extent to which inference risks depend on methodological choices. Future studies should also examine resampling strategies such as SMOTE to address class imbalance and assess whether sparse target classes affect inferability outcomes. In addition, datasets with larger sample sizes and potentially richer socioeconomic variables, such as the SOEP dataset provided by the *Deutsches Institut für Wirtschaftsforschung* (DIW), may offer further opportunities to investigate proxy-based inferability in contexts related to creditworthiness assessments.

Second, the observed attribute-specific variability highlights the need for more fine-grained analyses of which categories of sensitive information are most vulnerable to indirect inference. This suggests that regulatory approaches such as the GDPR and AIA may benefit from incorporating differentiated risk assessments rather than treating all sensitive attributes uniformly. Developing standardized benchmarks and evaluation protocols could support more consistent assessments of profiling risks across contexts and studies.

Finally, the findings underscore the importance of adopting precautionary principles in both system design and policy. Given that the absence of evidence for predictive performance does not constitute evidence of absence, future research should explore robust auditing methods and risk mitigation strategies that remain effective under uncertainty. This includes investigating privacy-enhancing technologies, model-auditing techniques, and data-minimization practices that can reduce the likelihood of unintended attribute inference, even when such risks are not immediately observable.

REFERENCES

[1] X. Gao et al., "Fairness in machine learning: definition, testing, debugging, and application", *Sci. China Inf. Sci.*, vol. 67, no. 9, p. 191201, Sep. 2024, doi: 10.1007/s11432-023-4060-x.

[2] Risk Research, "Der EU AI Act in Kreditinstituten [The EU AI Act in Credit Institutions]", Aug. 2024. Accessed: May 09, 2026. [Online]. Available: https://www.risk-research.de/fileadmin/userdaten/docs/PDF-Dateien/Whitepaper/Risk_Research_Whitepaper_7_EU_AI_Act.pdf.

[3] P. Quinn and G. Malgieri, "The difficulty of defining sensitive data—the concept of sensitive data in the EU data protection framework", Ger. Law J., vol. 22, no. 8, pp. 1583–1612, Dec. 2021, doi: 10.1017/glj.2021.79.

[4] F. Elsafoury, "Diskriminierung | Wenn der Schein trügt – Deepfakes und die politische Realität [Discrimination | When appearances are deceptive – deepfakes and political reality]". Accessed: May 09, 2026. [Online]. Available: https://www.bpb.de/lernen/bewegtbild-und-politische-bildung/556762/diskriminierung/.

[5] Artikel-29-Datenschutzgruppe, "Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679 [Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679]", Oct. 2017. Accessed: May 09, 2026. [Online]. Available: https://dsb.gv.at/sites/site0344/media/downloads/wp251rev01_de.pdf.

[6] S. Yeom, A. Datta, and M. Fredrikson, "Hunting for discriminatory proxies in linear regression models", 2018, arXiv. doi: 10.48550/ARXIV.1810.07155.

[7] A. D. Selbst, D. Boyd, S. A. Friedler, S. Venkatasubramanian, and J. Vertesi, "Fairness and abstraction in sociotechnical systems", in Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta GA USA: ACM, Jan. 2019, pp. 59–68, doi: 10.1145/3287560.3287598.

[8] K. Tiwari, N. Sarkar, K. Bisht, and S. Shukla, "A systematic survey on bias and fairness in machine learning", in Data Science and Security, S. Shukla, H. Sayama, K. Tiwari, and J. V. Kureethara, Eds., Lecture Notes in Networks and Systems, vol. 1355. Singapore: Springer Nature Singapore, 2025, pp. 77–91, doi: 10.1007/978-981-96-4883-2_8.

[9] European Data Protection Supervisor, AI Act Regulation (EU) 2024/1689. Luxembourg: Publications Office of the European Union, 2025, doi: 10.2804/4225375.

[10] European Parliament and Council of the European Union, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), May 4, 2016. Accessed: May 09, 2026. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/oj.

[11] European Social Survey European Research Infrastructure (ESS ERIC), "ESS10 - data from Interviewer’s questionnaire, edition 3.1". Sikt - Norwegian Agency for Shared Services in Education and Research, 2023, doi: 10.21338/ESS10INTE03_1.

[12] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior", Proc. Natl. Acad. Sci., vol. 110, no. 15, pp. 5802–5805, Apr. 2013, doi: 10.1073/pnas.1218772110.

[13] B. K. Hassani, "Societal bias reinforcement through machine learning: a credit scoring perspective", AI Ethics, vol. 1, no. 3, pp. 239–247, Aug. 2021, doi: 10.1007/s43681-020-00026-z.

[14] R. Chaturvedi and S. Chaturvedi, "It’s all in the name: a character-based approach to infer religion", Polit. Anal., vol. 32, no. 1, pp. 34–49, Jan. 2024, doi: 10.1017/pan.2023.6.

[15] S. D. Kim, G. Andreeva, and M. Rovatsos, "The double-edged sword of big data and information technology for the disadvantaged: a cautionary tale from open banking", 2023, arXiv, doi: 10.48550/ARXIV.2307.13408.

[16] European Social Survey ERIC (ESS ERIC), "European Social Survey (ESS), Round 10 - 2020", 2022, Sikt - Norwegian Agency for Shared Services in Education and Research, doi: 10.21338/NSD-ESS10-2020.

[17] S. Selle, Data Science Training - Supervised Learning: Ein praktischer Einstieg ins überwachte maschinelle Lernen [A practical introduction to supervised machine learning]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2024, doi: 10.1007/978-3-662-67960-9.