FORSEE

Forging Successful AI Applications
for European Economy and Society

D3.4 Gendered perspectives among SME representatives

Project Information
Project Number: 101177579
Project Title: Forging Successful AI Applications for European Economy and Society - FORSEE
Funding Scheme: HORIZON-CL2-2024-TRANSFORMATIONS-01-06
Project Start Date: 1 February 2025

Deliverable Information	
Title	Gendered perspectives among SME representatives
Work Package	WP 3 - Mapping of social expectations: understanding of success from a lifeworld perspective (society and economy)
Lead Beneficiary	NUID UCD
Due Date	31/01/2026
Revision Number:	V1.0
Authors	Alexandros Minotakis, University College Dublin[1] Elizabeth Farries, University College Dublin[2] Loredana Bucseneanu, European DIGITAL SME Alliance Sandra Sieron, WZB Berlin Social Science Center[3]

[1] https://orcid.org/0009-0001-9294-4564
[2] https://orcid.org/0000-0002-6524-2671
[3] https://orcid.org/0009-0005-0988-1259

Funded by
the European Union

	Molly Newell, Think-Thank for Action for Social Change
Dissemination Level	PU - Public
Deliverable Type	Report

Revision History				
Version	Changes implemented by	Revision Date	Reviewed by:	Brief description of changes
V1.0	Alexandros Minotakis, Elizabeth Farries, Loredana Bucseneanu	22/01/2026	temi lasade-anderson , Aphra Kerr	Revised according to internal peer-review.

Approval Procedure			
Version	Deliverable Name	Approved by	Approval Date
<Fx.x>	<.....>	<Name>	<dd/mm/yyyy>

Abbreviations and definitions

AI gender bias: AI systems' reinforcement of existing technical errors together with structural, systematic and societal prejudices in relation to gender.

AI lifecycle: Includes design: dataset collection and selection, model design, development and validation; deployment, and monitoring.

Demographic characteristics: gender, sexuality, race, age, income, residence, and more.

SMEs: Small and Medium Enterprises, defined as organizations with fewer than 250 employees and an annual turnover below EUR 50 million or a balance sheet total of up to €43 million.

Sociotechnical analysis: Analysis that considers the interdependent relationship between social elements (people, culture, rules, values) and technical systems (tools, processes, technology).

Forging Successful AI Applications for European Economy and Society

Mapping of social expectations: understanding of success from a lifeworld perspective

The capabilities of artificial intelligence (AI) are advancing rapidly, yet understanding what constitutes successful AI for society and the conditions that enable its effective deployment remains limited. AI promises economic growth, knowledge creation, and broader societal benefits, but realising this potential depends on developing and integrating applications that are successful not only technologically and economically but also socially, and ethically. AI applications are embedded within complex social contexts, reflecting and shaping aspirations, biases, and inequalities; thus, understanding AI success requires attention to these broader dimensions.

The FORSEE project (*Forging Successful AI Applications for European Economy and Society*) adopts a sociological perspective to examine these dynamics, focusing on how different stakeholders define success and how controversies, and unequal distributions of risks and benefits are articulated and potentially resolved.

This cluster of research papers maps social expectations of AI success across stakeholders, adopting a lifeworld perspective that situates understandings within societal and economic contexts. It comprises four interrelated reports. The first examines **digital small and medium-sized enterprises' (SMEs) success narratives**, identifying recurring themes and operational challenges. The second addresses **civil society organisations (CSOs) perspectives on gendered risks related to AI**, examining potential paths to advocate for gender vulnerable communities. The third investigates **criteria for awards and prizes**, providing an external perspective on standards of AI success. The fourth applies a **gendered lens**, exploring SMEs' perspectives on AI and gender bias. Together, these reports link SMEs' and CSOs' viewpoints with societal concerns, offering a multidimensional understanding of AI success in Europe.

The present report seeks to explore how SMEs understand and respond to gender bias associated with artificial intelligence, including the strategies and measures they implement to address it. The analysis is framed within the wider European Union environment and takes into account the constraints characteristic of SMEs, with the objective of generating practical and implementable policy guidance.

Funded by
the European Union

Table of Contents

Funded by
the European Union

Executive Summary

The goal of the present study is to examine small and medium-sized enterprises' (SMEs) perspectives on AI-related gender bias, as well as the practices and policies they adopt to mitigate it. These perspectives are situated within the broader European Union context and the specific structural and organisational challenges SMEs face, with the aim of informing concrete and actionable policy recommendations.

AI bias generally, and AI gender bias in particular, has emerged as a prominent concern as artificial intelligence systems become increasingly embedded across both public and private spheres. Existing research highlights not only the role of biased datasets, but also the influence of social, systemic and organisational structures and processes, including the lack of diversity in AI development processes, as key factors contributing to biased AI systems. However, the literature remains largely focused on large technology firms, often overlooking SMEs either as developers or users of AI, despite their central role in the European industrial landscape and the distinct constraints under which they operate.

Through 39 semi-structured interviews, the present study addresses this gap by focusing on European digital SMEs[4] and exploring their awareness of AI gender bias, as well as existing mitigation practices and policies. The findings indicate that awareness of gender AI bias among SMEs is generally high, with many recognising the risks of discrimination and reputational harm associated with biased AI systems. At the same time, most SMEs frame bias primarily as a data-related technical issue and report implementing a range of technical safeguards to address it.

Crucially, the data also point to an emerging shift in understanding, suggesting that effective bias mitigation must extend beyond data engineering solutions to encompass inclusive design practices and broader organisational change. Only a limited number of SMEs explicitly connect bias to organisational cultures within the technology sector, which continue to be shaped predominantly by white, male perspectives and experiences. Diversity within engineering teams remains limited, with SMEs facing structural barriers such as the "pipeline problem", i.e. the flow of work and workers around the AI lifecycle, and constrained organisational capacities that restrict meaningful intervention.

Overall, the study reveals a clear divide among SMEs: while some integrate diversity considerations into their core business strategies, others view them as beyond their remit, and a small minority perceive diversity as a barrier to accessing skilled personnel. Based on

[4] SMEs are defined as organizations with fewer than 250 employees and an annual turnover below EUR 50 million or a balance sheet total of up to €43 million. Digital SMEs are enterprises for which digital technologies play a central role in value creation, organisational processes, and competitive positioning. A defining characteristic of these firms is the strategic importance of digital capabilities and skills, which underpin continuous innovation, support scalability, and enable adaptation to rapidly evolving technological and market conditions

these findings, the study highlights the need for more comprehensive policy interventions to address automated bias beyond narrow technical solutions. Furthermore, findings indicate that SMEs, operating under specific constraints and in dependence on Big Tech and their infrastructures, often have limited capabilities to engage meaningfully with the AI systems they deploy or even the ones they develop and integrate in their own products. Therefore, the issue of digital sovereignty emerged through present study and the corresponding need for policymakers to provide institutional and financial support mechanisms tailored to SMEs' limited capacities.

Such measures should aim to enable SMEs to move beyond narrow data-driven fixes towards holistic approaches to AI development and deployment, incorporating ethical by design practices and in-house diversity. Without coordinated policy support, existing structural inequalities risk being reproduced at scale through AI systems, disproportionately affecting SMEs' ability to develop and deploy fair and trustworthy AI.

Section 1: Introduction

As AI systems become embedded across more areas of the economy and public policy, concerns about bias and discrimination have become increasingly prevalent. FORSEE's approach understands AI not as a fixed technological object but as a shifting socio-technical field that is continuously contested. More concretely, this research started with an assumption derived from existing scholarship that controversies surrounding AI coalesce around both recurring tensions as well as new types of controversies. A substantial body of research shows that contemporary AI, particularly systems built through machine learning, does not function as a neutral or objective decision-making tool. Instead, it often reproduces and may even amplify existing social patterns and harmful stereotypes related to gender, race, disability, and socioeconomic status, which can be perceived as systemic discrimination (Ferrara, 2024) against marginalised populations. These patterns indicate that AI bias operates at the nexus of technical and sociotechnical issues. We consider in particular AI gender bias, a concept informed by researchers Buolamwini & Gebru (2018) who has famously articulated bias as not just a technical error, but a structural and systematic issue where AI systems reinforce existing societal prejudices, including bias against gender.

While addressing bias has traditionally been associated with the datasets used to train AI models, since biased data can directly embed discriminatory patterns into system outputs, this represents only part of the problem. Increasingly, scholars emphasise the need to examine the entire AI production process at any stage of the "AI lifecycle" (Chen et al, 2023). This includes design, dataset collection and selection, model design, development and validation (Yilmaz, Yorgancioglu & Koyutürk, 2025); deployment; and monitoring. With this understanding, AI gender bias isn't eradicated by simply adding more gender diverse people and "diverse" datasets. Unsurprisingly, contemporary research agenda has been dominated by analyses of large tech companies and their technology as Small and Medium Enterprises are at a significant disadvantage in their efforts to adopt and develop AI systems (Schönberger, 2023; Ulrich & Frank, 2021). However, the AI landscape is expanding, therefore SMEs play a growing role (Agbaakin, 2025), especially within the European context. SMEs are both users and developers of AI systems, yet their experiences and perspectives remain under-represented in the research literature.

Research on AI adoption increasingly shows that SMEs approach AI with enthusiasm for its potential to increase productivity, reduce costs, and enhance competitiveness. However, it remains unclear how they approach the topic of AI gender bias within their limited capacities. The present study therefore seeks to broaden the scope of research by examining how SMEs understand AI gender bias (as a risk, in datasets), how they perceive its origins, consequences and broader challenges, and the ways in which they attempt to address it within their own organisations (e.g. bringing diversity in-house).

This report is structured as follows. **Section 2** introduces the theoretical framework on AI and bias with a particular focus on gender bias and the different ways it manifests across the AI lifecycle. **Section 3** situates SMEs within the broader landscape of AI adoption and development, identifying challenges that limit their capacity to engage with gender bias in the AI systems they develop and deploy. **Section 4** outlines the methodological approach, the qualitative methods employed, and the demographic profile of participating SMEs. **Section 5** presents the empirical findings, detailing SMEs' understandings of AI gender bias, the strategies mobilised to mitigate it as well as the limits of their efforts. **Section 6** discusses the findings, linking the broader context of the European AI industry to the SMEs' responses. **Section 7** concludes present research, highlights key findings and provides a brief outline for an alternative approach that empowers SMEs to effectively address AI gender bias.

Section 2: AI, Bias and AI Gender Bias

2.1 AI and lifecycle biases

Scholars have long established that the way contemporary AI systems operate poses significant risks through biased deployments that enable discrimination, while often cloaked as "objective" or "neutral" data-driven systems (Xavier, 2025). The concept of bias is not new; it precedes AI development and reflects the tendency to "promote prejudiced results due to erroneous assumptions" (Mavrogiorgos et al., 2024, p.1). However, as AI systems become integrated into more and more social spheres and decision-making processes, the issue of bias is widely discussed and is a significant concern. AI is trained, validated and tested on large datasets - the raw information or material from which AI systems learn to recognize patterns, make predictions, and perform tasks. Consequential to this training is the ability of AI to detect connections and patterns in the datasets that are not explicitly identified by the people who input or label the data. However, dataset bias arises through systematic errors or prejudices that skew the data, leading AI to develop flawed interpretations of the patterns that emerge (see for eg. Yang et al, 2024; Ferrara, 2024). This can lead to poor performance outside of narrowly constructed norms.

Crenshaw (1989) has written seminally of intersecting discriminations between race and gender and, indeed, recent AI studies have shown that poor AI performances can be amplified in particular at intersections of marginalised identities rather than being reducible to a single demographic axis (Gohar & Cheng, 2023). In social realms, AI is trained on large datasets and can identify patterns across a range of demographic characteristics: gender, sexuality, race, age, income, residence, and more. In many cases, dataset performance declines in intersectional scenarios targeting simultaneously gender and other demographic characteristics (Noble, 2018). For example, LLM-based resume scoring systems have been shown to systematically favor white-associated names while disadvantaging black male candidates, even when all applicants had identical qualifications (Wilson & Caliskan, 2024).

AI gender bias is not a function of flawed datasets alone and so can not be solved simply with more diverse or larger datasets. The systemic literature reveals how categories of bias can emerge at any stage of the "AI lifecycle" (Chen et al, 2023), which includes design: dataset collection and selection, model design, development and validation (Yilmaz, Yorgancioglu & Koyutürk, 2025); however, it also occurs in deployment and monitoring and through opportunities for redress. Bias can be introduced at every stage of this cycle and can interact cumulatively to affect final outcomes. Bias surfacing via these lifecycle stages can overrepresent, underrepresent, and misrepresent findings or influence decisions. While most frameworks evaluate bias in models rather than datasets, it is important that datasets are also a focus of importance because model bias can often be traced back to dataset shortcomings (Dai et al., 2025). Similarly, it is important to evaluate how AI is deployed in real world settings, how it impacts diverse populations, how it is monitored and restrained, and what opportunities for redress and change exist and how different populations can be deployed.

2.2 Systemic discriminations as a function of AI bias

Established and critical findings for bias and discrimination in AI applications have been summarised by Ferrara (2024) in key public policy areas including healthcare, employment and hiring practices, criminal justice, migration, warfare, credit scoring, and generative AI models. Prominent examples in public policy discourse increasingly reflect these risks. For example, social welfare AIs are promoted by governments as tools of efficiency. However, civil society critique these systems for relying on flawed datasets within underfunded and inadequate social protection systems, ultimately resulting in harmful automation of benefit denial (Banerji & Satija, 2025). A widely cited private-sector case is Amazon's recruitment AI, which through its selected datasets trained on male-dominated work forces, built an AI model that when deployed prioritised resumes reflecting historically male hiring patterns, systematically excluding women (Buolamwini, 2019; Dastin, 2018). Furthermore, while sometimes unintentional, biased deployment can reinforce biased policy. In Hungary, biometric surveillance technologies such as facial recognition are now deployed to monitor peaceful LGBTQ+ demonstrations, including Budapest Pride (Hungarian Civil Liberties Union et al., 2025). In such scenarios, discriminatory results can be recorded as vindication of new AI technologies: if social welfare allowances are reduced, if a pool of applicants are surfaced, if a policing system identifies LGBTQ+ protest content for sanction, then this can be characterised as an AI "success" despite the bias and discrimination propelling systemic exclusion. Eubanks (2018, p. 7) describes this phenomenon as a "feedback loop of injustice," in which AI systems not only reproduce but reinforce structural oppression under the guise of neutrality or precision.

2.3 The limits of techno-solutionist responses to bias

In line with techno-solutionist narratives (Lindgren & Dignum, 2023), technological solutions are often proposed for problems that are perceived to be technological. Dataset bias mitigation strategies supported by published research include interventions in the preprocessing, processing, post processing stages. Preprocessing solutions include data training set balancing (costly) and augmentation (synthetic) (Yang et al, 2024). Processing techniques have included causal model-based dataset creation (González-Sendino et al., 2024) or multi-attribute bias mitigation (Duong & Conrad, 2024). Post processing techniques include the validation of dataset representativeness and performance through audits, which are prescribed regularly and for model updates (Chadha, 2024). These purely technical measures however tend to omit what Birhane (2021) and Noble (2018) have emphasised: social and structural conditions that create bias are not merely an error but rather a function of social and cultural processes that amplify structural inequalities. Such sociotechnical analysis, i.e. considering interdependent relationship between social elements (people, culture, rules, values) and technical systems (tools, processes, technology), surfaces the reality that by focusing on technical bias metrics alone, larger facets of exclusionary design are missed, including the power, context, and relational harms that AI performs in everyday life and the way that systems harm social groups through exclusionary and unjust outcomes.

The point of inflection in this debate is how intersectionally biased results in the AI lifecycle discriminate against marginalised populations. Bias does not result from targeting of individuals although, expectedly, they experience discrimination in a personal manner. Rather, people "are targeted as members of social groups, not as individuals" (Eubanks, 2018, p.6). Therefore, as preeminent scholars have pointed out (see Benjamin, 2019; Noble, 2018; Buolamwini & Gebru, 2018; D'Ignazio & Klein, 2020; and Birhane, 2021), bias in AI can lead to social harms for members of social groups according to categories including gender, but also sexuality, and race.

2.4 Sociotechnical drivers for bias: Stakeholder roles

The explanation for biased and discriminatory AI lies not just in technical explanations but at the nexus of technical and sociotechnical issues. Here we categorise biases within the AI lifecycle including dataset, design and development choices, together with deployment and monitoring decisions. Such choices carry assumptions about which data points matter, and which do not. These assumptions can embed cultural norms and privilege certain outputs or conceptualisations over others. For example, a 2025 peer-reviewed study found that popular AI text-to-image models generate from their selected datasets leadership images that significantly overrepresent men and white individuals relative to real-world demographics in U.S. hospitals (Gisselbaek et al, 2025). Despite actual leadership distributions showing substantial female and minority representation, the AI outputs skewed heavily toward men and white leaders, reflecting design assumptions embedded in the dataset facilitating a model that leadership roles are predominantly male and white. Evidently, defining what is normal and what is an outlier relates to design decisions and reflects developers' focus and omissions.

Discrimination also results in intersectional ways when stakeholders enforce biased AI systems outputs. For example, Wilson et al. (2025) showed that when humans acted on biased AI recommendations in hiring simulations, their selection patterns mirrored the AI's intersectional biases in discriminatory ways, even when the participants were aware of potential AI bias. Conversely, when they made screening decisions without AI guidance, choices were relatively balanced across groups. Such findings support classic research on automation bias, showing how people often over-rely on algorithmic recommendations (Mosier & Skitka, 1996; Parasuraman & Riley, 1997). Automation bias belies the understanding that the integration of human oversight, judgment, or intervention at key stages of an AI system's lifecycle is sufficient to detect, prevent, and mitigate bias that the system might otherwise perpetuate (the "human-in-the-loop" solution).

Indeed, human monitoring can actually advocate calibrating systems to heighten biased outcomes in favour of achieving specific results. A 2025 investigation into the UK's national facial recognition deployment revealed that police forces adjusted its system's confidence threshold to re-introduce a more biased version of the system because it yielded more "useful" suspects, despite increased misidentification rates for black and Asian people and women compared with white men (Wilding & Boffey, 2025). Adjusting algorithmic thresholds to actively reinforce bias rather than mitigate it, is vulnerable to organisational priorities in a manner that overrides ethical considerations and confirms that the problem of AI bias is not a purely technical failure.

AI systems are thus shaped and driven by sociotechnical choices that often reflect narrow assumptions about what - and who - matters. AI lifecycle decisions that are made without diverse perspectives can embed cultural norms and disproportionately privilege dominant identities to the exclusion of other demographics. These exclusions are acute at the intersections of gender, race and other social identifiers (Noble, 2018). Bias does not stop at

Funded by
the European Union

deployment or monitoring phases; rather, it is compounded when human actors uncritically enforce AI bias, thereby reinforcing intersectional discrimination (Wilson et al., 2025; Wilding & Boffey, 2025).

2.5 In house diversity solutions to sociotechnical bias

Recognising that AI lifecycle decisions are not neutral technical processes but can reflect bias and discrimination that are shaped by social, political, and institutional values of discrete actors, it becomes important to include a diverse range of stakeholders. AI lifecycle diversity supports "ethical by design" and "fairness" concepts. Ethical by design is intended as a proactive approach to embedding moral principles into technology from the start of the lifecycle to ensure fairness, as well as respect for human rights, dignity, and democratic values throughout (Gurzawska, 2024). Fairness in the AI context means treating people equitably, thereby avoiding unjustified discrimination, and ensuring outcomes don't disproportionately harm certain groups (Information Commissioner's Office, 2023). Diverse lived experiences, disciplinary knowledge, and positionalities are essential throughout the AI lifecycle to identify and counteract bias and to promote ethical and fair design principles, in ways that might otherwise remain invisible to homogenous developer teams.

In this context, diversity within AI development teams across demographics including gender, together with disciplinary diversity, is often proposed as a partial solution to the problem of sociotechnical bias (Shams, Zowghi & Bano, 2025). Demographic diversity in teams may serve to more readily identify harmful defaults or non-inclusive norms in datasets, interfaces, or use cases, making "visible and types of being objectified while other types are erased" (Birhane, 2021). Teams with lived experience across intersectional identities are more likely to raise questions of accessibility, representation, or surveillance harm stemming from AI processes or to challenge exclusionary defaults. Even so, as the experience of Timit Gebru has illustrated, if they have no redress or power, then diversity can also only go so far. The experience of SMEs is particularly relevant here in so far as they inherit or use larger organisation technologies and infrastructures. As discussed further in 3.1, this arguably adds another layer of challenge to the benefits that diverse teams might serve.

Section 3. The perspective of SMEs

One set of diverse perspectives that warrant further engagement are small and medium-sized enterprises acting as important stakeholders in EU AI policy conversations. 99% of European businesses are small and medium-sized enterprises, defined by the European Commission as having fewer than 250 employees and an annual turnover below EUR 50 million (European Commission, 2023; Schwaeke et al., 2024:1300). These firms represent the work of 24.3 million entrepreneurs and provide employment for over 85 million Europeans (European Commission, 2023). Given their centrality to the European economy, it is critical to address under-examined roles, experiences and practices of SME actors in the AI lifecycle. Research on AI adoption increasingly shows that SMEs approach AI with enthusiasm for its potential to increase productivity, reduce costs, and enhance competitiveness (Le Dinh et al., 2025; Sánchez et al., 2025). Conversely, literature suggests that SMEs may lack the digital maturity, data quality standards, and governance structures necessary to adapt to the fast pace of AI change generally (Quenum, Vallée and Ertz, 2025), let alone identify or manage algorithmic bias and discrimination issues effectively. Broader work in AI fairness and AI highlights that such limitations can amplify existing sociotechnical discriminations embedded in data and automated decision-making systems (Barocas & Selbst, 2016; Mehrabi et al., 2021).

Regulation constitutes a distinct challenge in its own right. While large technology firms often operate within relatively light regulatory environments and possess the influence to shape or circumvent rules to their advantage (Couldry & Mejias, 2019), SMEs frequently express concerns about the high compliance costs and confusion arising from overlapping or contradictory regulatory frameworks (Watney & Auer, 2021).

Overall, market, funding, and regulatory challenges effectively constrain the practical capacity of SMEs regarding AI adoption. In 2025, 18.9% of SMEs employing more than 10 people in the EU deployed internally some form of AI technology, though this does mark a significant increase from the 7.1% of SMEs using AI in 2021. Adoption is uneven throughout the European Union, with Nordic and Western countries enjoying higher adoption than countries in Eastern Europe (Eurostat, 2025). At the same time, the development of AI by SMEs, while even more limited than adoption, is growing, as they increasingly engage in the creation, customisation, or training of AI models tailored to company-specific workflows, data assets, and strategic goals (Agbaakin, 2025; Dinh et al., 2025). SMEs with higher digital maturity tend to develop their own AI solutions rather than deploy off-the-shelf products to better align with organisational workflows, data structures and control, or competitive positioning.

3.1 SMEs and AI bias

As SMEs typically operate with limited resources, their AI adoption is often guided by expectations of efficiency, cost savings, and competitive advantage rather than formal evaluation of ethical considerations or fairness criteria (Le Dinh et al., 2025). Within the same context, SMEs must navigate their dependence on Big Tech, whose infrastructures they rely on for either the deployment or development of AI (Schwaeke et al., 2025; see also FORSEE Deliverable 3.1 for a detailed account of SMEs' perspectives on this topic). Together, these factors shape SMEs' engagement with AI gender bias, contributing to the predominance of a utilitarian perspective and a diminished capacity for action.

The utilitarian orientation influences how SMEs interpret tradeoffs between fairness and performance, often prioritizing immediate business value. The literature on SME digital readiness further shows that many smaller firms approach AI through a pragmatic rather than principled lens, leading to low awareness of potential bias or discrimination risks embedded in algorithmic outputs (Sánchez et al., 2025). This contributes to a context in which discrimination concerns, including gender bias, are often secondary in SMEs' evaluation of AI systems. They remain largely unexamined unless it produces an immediate, visible business problem.

As mentioned in section 2, a central theme in research on AI bias concerns the role of datasets in generating biased model behavior (Suresh & Guttag, 2019). SMEs are especially vulnerable to such issues because their internal datasets tend to be small, fragmented, and generated through informal processes. These data limitations not only increase the likelihood of biased predictions but also reduce SMEs' capacity to detect or correct these biases. Furthermore, SMEs AI IT infrastructure is low (Schwaeke et al., 2025) making them reliant on external AI vendors to integrate AI into their operations. Most smaller firms lack the capacity to build models in-house and therefore adopt ready-made AI technologies that have limited transparency into how these systems are trained on datasets or validated. Prior research on algorithmic bias in domains such as credit scoring and hiring shows that opaque commercial models can contain embedded disadvantages that disproportionately harm certain types of users according to demographic characteristics (Hurley & Adebayo, 2016; Fuster et al., 2022). SMEs effectively "inherit" these external biases without the technical means to audit or modify vendor systems. This dynamic underscores a broader structural vulnerability: SMEs' fairness risks are partly a product of their dependence on technologies from larger companies that they cannot fully interrogate.

Compounding these risks is the dynamic of human-AI interaction within SMEs. Problems of automation bias described above are reported to surface in recent SME studies showing similar patterns. Employees with limited AI experience may accept model outputs with minimal critical scrutiny in routine tasks such as hiring, lead scoring or inventory management (referred to by Sánchez et al., 2025 as limited absorptive capacity). As SMEs increasingly invest in and integrate these systems, the lack of formal oversight mechanisms leaves them with few safeguards to detect biased outcomes or challenge discriminatory recommendations.

Regulatory literature further highlights the tensions between fairness expectations and SME capacity. Emerging frameworks such as the EU AI Act impose transparency, documentation, and risk-assessment requirements, including mandated detection of gender discrimination. Literature consistently emphasizes that smaller firms often lack the financial, technical, and governance capacity to implement such measures effectively (Jarvers, Ullstein, & Grossklags, 2025). AI systems deployed by resource-constrained SMEs may remain insufficiently monitored or audited, thereby increasing the risk of discriminatory outcomes. Large firms only may be able to achieve compliance (Fair Tech Policy Lab, 2023). As a result, SMEs may unknowingly overlook gender-related risks.

Overall, the literature portrays SMEs as operating within multidimensional challenges shaped by resource limitations, data quality issues, vendor dependence, limited governance capacity, and constrained employee oversight. These factors suggest an environment in which AI-related biases are more likely to occur. It requires considering SME perspectives in the context of the structural, organizational, societal forces that shape their engagement with AI technologies, either as developers or as deployers.

Section 4: Methodology

The preceding section describes how digital SMEs' perspectives on AI are of particular importance because SMEs are increasingly adopting AI across the lifecycle, yet do so under distinct structural constraints that shape how bias and discrimination emerge or are evaluated or responded to in practice. SMEs often inherit opaque biases from third-party AI systems and prioritise competitiveness over fairness evaluation. This makes them a critical site where sociotechnical risks are likely to materialise but remain under-examined. Engaging SME perspectives therefore provides essential insight into AI-related risks as well as the challenges that shape the European AI industry, revealing practical gaps between fairness goals and bias mitigation strategies, contextualised within an understanding of real-world organisational capacity.

We ask in this study *How do SMEs perceive AI bias and gender bias in particular and what measures do they take to address it?* Based on our literature review, our research aimed to engage with SMEs that either use or develop AI by examining:

a) How do SMEs themselves understand and interpret AI bias in broad terms?
b) What strategies or practices do they currently employ to mitigate AI gender bias within their enterprises?
c) What challenges limit their capacity to identify, prevent, or address AI gender bias in practice?

To meet our research objectives, ethics clearance was obtained from University College Dublin (UCD). The requirements of this clearance included the protection of participants' anonymity and the principles of data minimisation and secure data storage, which are also key components of the FORSEE Data Management Plan. An agreement was established with FORSEE partner, European DIGITAL SME Alliance, to act as the sole contact point for approaching potential interviewees. Prior to each interview, participants were provided an information sheet, outlining research objectives, data management principles and potential risks as well as a letter of consent. Said letter confirms that participants have received and understood the information sheet, are aware of the purpose of the research, how their data will be collected, stored and used, and understand their rights, including confidentiality and the right to withdraw at any time without consequence

Through this process, between June and December 2025, 39 semi-structured, in-depth online interviews were conducted with European SMEs that either deploy or develop AI systems, averaging 40 minutes each. This method was selected for its ability to capture participants' experiences, perceptions, and interpretations, making it particularly suitable for exploring how organisations conceptualise sociotechnical problems such as AI gender bias. An interview guide was developed to assist interviewers throughout the process, while ensuring sufficient flexibility for participants to elaborate on their perspectives. An indicative list of the interview questions is provided in Appendix 1. The wording and focus of questions were adapted to the role of each participant, with distinct questions for SMEs that use AI systems and for those that develop them. A list of the demographic characteristics of our sample is provided at the end of the present section.

To recruit participants, the European DIGITAL SME Alliance facilitated access to relevant interviewees in their expansive network, who met the study criteria and coordinated the scheduling of interviews. Participants selected for the interviews were representatives of SMEs with direct experience in relation to artificial intelligence. The sample included SMEs that deploy AI-based solutions to support specific business operations, as well as SMEs that develop AI products or services as part of their core activities. As is shown in table 1, the interviewed SMEs further exhibited substantial heterogeneity in terms of business models, spanning business-to-business (B2B), business-to-consumer (B2C), software-as-a-service (SaaS), business-to-government (B2G), and research-and-development partnership arrangements. With respect to technological configurations, many firms relied on existing digital infrastructures, such as external cloud services and large language models, which they adapted and customised to develop proprietary solutions, while a smaller subset pursued fully in-house AI development; others primarily deployed off-the-shelf AI systems. Across these diverse organisational and technological profiles, the SMEs addressed a broad range of application domains, including legal and regulatory compliance, recruitment and human resources, marketing and communication, software development, energy efficiency, insurance, health, manufacturing, robotics, and related sectors. No exclusion criteria were applied with respect to the type, scope, or technical characteristics of the AI systems used, in order to capture a broad range of experiences and perspectives across different application contexts.

Interviews were transcribed and analysed using the NVivo 15 qualitative analysis software. A deductive content analysis approach was employed, whereby a set of predefined labels, derived from the study's conceptual framework and existing literature on AI bias, guided the coding process. These labels ensured that key themes, such as conceptualisations of bias, mitigation strategies and organisational constraints, were consistently examined across the dataset.

Finally, to contextualize the interview data, we compiled a demographic table (Table 1) capturing key characteristics of all participants, including job title, business model, sector, company headquarter country, whether firms act primarily as developers or deployers of AI, the extent of in-house AI development versus reliance on existing architectures, and the implementation of diversity hiring policies.

Table 1

Demographic profile of SME interview participants

ID	Job title	Business model	Sector	Company headquarter	Developers or deployers	In-house or on top of existing architecture	Hiring policies
1	Co-Founder	R&D partnerships	ICT	Greece	Developers	Strategic collaborations with external partners	No formal diversity policy
2	Founder/CEO	B2B	HR Tech	Italy	Developers	Building on top of existing architecture	No formal diversity policy; Mostly referral based (people already known to the founder)
3	Founder	B2B, B2C	ICT	Hungary	Developers	Mostly in-house	Geographical diversity, no formal gender quotas
4	Founder/ CEO	B2G, R&D partnerships	ICT	The Netherlands	Developers	In-house	No formal diversity policy
5	Co-Founder/ CEO	B2C2B2B	Building performance optimization	Italy	Developers	In-house	No formal diversity policy
6	Co-Founder/ CEO	B2B, SaaS	Regulatory compliance	The Netherlands	Developers	Relies on an external cloud service	No formal diversity policy
7	Co-Founder /CEO	D2B SaaS	Regulatory compliance	Italy	Developers	Relies on an external cloud service	No formal diversity policy
8	Co-Founder/ CEO	B2B	Aerospace, defense	Norway	Developers	In-house	Formal HR gender equality plan
9	Director	B2B, B2G	Neurotechnology applied neuroscience	Germany	Developers	Relies on an external cloud service	D&I formalised in internal policies
10	Senior executive and technical lead	B2B	Software development	Romania	Developers	Relies on an external cloud service	No formal diversity policy
11	Founder/ CEO	B2B, SaaS	Predictive sales	Ireland	Developers	Relies on an external cloud service	No formal diversity policy
12	Head of Research and EU Projects	B2G, B2B	Software development	Italy	Developers	Relies on an external cloud service; strategic collaborations with external partners	Recruitment prioritizes diversity and inclusion
13	Founder/CEO	B2B	IT consulting	Germany	Deployers	Integrates feature from global LLMs	Recruitment prioritizes diversity and inclusion
14	Founder/CEO	B2B, R&D Partnerships	IT	Italy	Developers	Relies on an external cloud	Recruitment prioritizes diversity and inclusion

						service; strategic collaborations with external partners	
15	Co-founder/Business Lead	B2B	FinTech	Italy	Developers	Relies on an external cloud service	Recruitment prioritizes diversity and inclusion
16	Head of Innovation Department	B2B	Process automation	Spain	Developers	In-house	No formal diversity policy
17	Co-Founder/ CEO	B2B	Health	Italy	Developers	Relies on an external cloud service	Holds an official gender equality certification
18	Founder and Lead Consultant	B2B	IT consulting and training	Italy	Deployers	N/A	No formal diversity policy
19	Head of Research	B2B	Digital manufacturing	Germany	Developers	Non-strategic use of LLMs	No formal diversity policy
20	Researcher/ data analyst and software developer	B2G, R&D	Software development	Italy	Developers	Relies on an external cloud service; strategic collaborations with external partners	Holds an official gender equality certification
21	Co-Founder	B2B	AI systems management	France	Developers	Relies on an external cloud service; strategic collaborations with external partners	Internal diversity charter
22	Founder/Director	B2B	Human Factors and Responsible AI Consulting	Ireland	Deployers	N/A	N/A (solo entrepreneur)
23	CTO	B2B	Marketing	Portugal	Deployers	N/A	Recruitment prioritizes diversity and inclusion
24	Head of Research and Development and AI project lead	R&D	R&D	Belgium	Developers	In- house	No formal diversity policy
25	Founder/CEO	R&D	Health	Bulgaria	Developers	Relies on an external cloud service	Recruitment prioritizes diversity and inclusion
26	CEO	R&D, B2B, B2C	Robotics	France	Deployers	N/A	No formal diversity policy
27	Co-Founder/ CCO	B2B, B2G	Supply chain traceability	Belgium	Developers	In-house	Recruitment prioritizes diversity and inclusion
28	Co-Founder/ Managing director	B2B, B2C	Health	Greece	Deployers	N/A	No formal diversity policy
29	Founder/ Managing Director	B2B, B2C	Consulting on sustainable AI	France	Deployers	N/A	Recruitment prioritizes diversity and inclusion
30	Data engineer	B2B	IT consulting, software development	Spain	Deployers	N/A	No formal diversity policy
31	Project Manager and AI Training Coordinator	B2C	Education	Spain	Deployers	N/A	D&I formalised in internal policies

32	Founder/CEO	B2B	Creative media, communication	Croatia	Deployers	Using off-the-shelf solutions	N/A (solo entrepreneur)
33	Founder/CEO	B2B	Industry, hardware innovation	Austria	Developers	In-house	Recruitment prioritizes diversity and inclusion
34	Head of the AI transformation	B2B, B2C	Insurance	Belgium	Developers	Building on top of existing architecture	Diversity, inclusion & fairness in recruitment; Gender balance in career development: No gender pay gap
35	Founder and Director	B2B	Digital transformation and innovation	Spain	Developers	In-house	LGBT-owned startup committed to hiring a diverse workforce
36	Founder	B2B	Agri-food trade	Italy	Deployers	Using off-the-shelf solutions	N/A (solo entrepreneur)
37	Founder	B2B	Digital transformation	The Netherlands	Deployers	Using off-the-shelf solutions	N/A (solo entrepreneur)
38	Founder/CEO	B2B	Software development	Romania	Developers	In-house and building on top of existing architecture	No formal diversity policy
39	CEO	B2B	Defense	France	Developers	In-house and building on top of existing architecture	N/A

Section 5: Findings

5.1 Acknowledgement of bias as a risk factor

This section examines how SME interviewees understand and interpret AI bias in the context of their own development and use of AI systems. It highlights the extent to which participants recognise bias as a sociotechnical risk, as well as the sources to which they attribute it. The findings reveal variation in how SMEs conceptualise responsibility for bias, particularly in relation to data, development practices, and external vendors. Taken together, these perspectives shed light on the practical and organisational factors shaping how AI bias is perceived and addressed within SMEs.

Broadly speaking, SMEs interviewees indicated a significant understanding of bias as a risk factor. More specifically, they had a clear grasp of how contemporary AI systems are capable of reproducing and amplifying preexisting biases, including AI gender biases, and can do so in intersectional ways. Interviewee 27 summarised this understanding:

> AI is a multiplicator of what most people think and it is built definitely with a bias. It can take many forms, it doesn't need to be pro-feminine or pro-masculine, it can also be age bias, as well, which is very strong, I think. It's a multiplicator, and we see it at all levels.

Interviewee 1 identified datasets and development bias as a key area upon which they consult their clients:

> You need systematic consulting over this process to understand what you're already doing now and understand that the machine learning system might end up mimicking your preferences, but this might be legally problematic, because this is your bias, right? There is also implicit bias, inducing features, because, you know, I might not seem as a man if I hide my face, or whatever else but if you see my preferences right, you might imply that I am, at least typically, a man.

Interviewee 20 also identified the role of data and its gendered implication:

> I think the data is important, if it has more, let's say female and male. I don't want to go in the other genders. For now, let's say these two. If you have, for example, 70% of your data male, and the other is female, so the data has caused this bias, because the model has seen more examples from one side than the other.

Biased datasets are the key origin of imbalance in the way AI systems operate according to our interviewees. It is possible that this reflects both their origins as skilled engineers and people investing in tech companies, and their knowledge of contemporary literature and news articles that often focus on the role of datasets. However, interviewee 22, a member of a company consulting businesses on AI, provided a broader analysis, linking societal biases of unbalanced datasets with biased team engineering composition:

> First of all, when we build and train these models there are inherent biases in society, and I'm pretty sure there's enough research backing this. So the data that's used to train these models is skewed between male and female, so almost in the data that these models are trained on and evaluated on. Plus, if you look at the average person working in AI, the actual person building it, there's still a massive skew between male and female, which is another thing. You know, again, you are building this for the whole of humanity, not just the subsection of it. You need diversity there too, whether it's gender or people from different cultures, languages.

This same interviewee was pessimistic about the prospects of SMEs in addressing AI bias when weighing the pace of technological change generally against organisational capacities:

> SMEs are almost inundated with the realities of what's happening. The change is so big, the changes are fast, the change is so expensive. People are worried about their roles, there's a lot of uncertainty in the world, including politics, so a lot of companies are almost fighting for their survival, so I find there's a little bit of looking the other way. That's just my take on it, but I do think there's a bit of looking the other way. It's almost as if people have enough to deal

> with, and if they don't openly voice it, then they don't necessarily have to take responsibility for it, or have to face it, so I genuinely think the biases are there, for sure.

Conversely, some claims were made in relation to their understanding that Big Tech companies will have largely resolved AI bias through fine-tuning as technosolutions. Interviewee 34, a user of AI systems for insurance claims, said:

> In our case, it doesn't really apply, first of all, because we don't send a lot of information about the person to the model. And also, I think bias mitigation has been a big focus for model developers, like Google, OpenAI, etc. And at least to what I've read, they've baked in a lot of fine-tuning to avoid biases in their model.

However, this statement did not resonate with the majority of interviewees who acknowledged bias as an open challenge.

Overall, SME interviewees broadly recognised AI bias as a significant risk, particularly its capacity to reproduce and amplify existing social inequalities such as gender and age bias. Many attributed these issues to imbalanced or skewed datasets, as well as implicit developer biases embedded during the system design phase. Several demonstrated an understanding of intersectional harms and linked bias not just to data, but also to the lack of diversity in engineering teams. Some interviewees expressed concern over SMEs' limited capacity to address bias due to resource constraints and the overwhelming pace of technological change. A minority believed that Big Tech vendors had already mitigated bias through fine-tuning, reflecting a reliance on external solutions. However, most interviewees acknowledged that bias in AI systems remains an unresolved and ongoing challenge.

Figure 1 presents a word cloud summarising the most frequently used terms by interviewees when discussing AI biases in general. All word clouds presented in section 5 were generated through word frequency queries in NVivo 15 and are used as indicative visualisations of recurring themes in participants' accounts.

Figure 1
Interviewees discussing AI bias.

Funded by
the European Union

Figure 2 presents a word cloud summarising the most frequently used terms by interviewees when discussing AI gender biases in particular.

Figure 2

Interviewees discussing gender bias.

In the following subsections, we examine interviewees' described strategies for addressing bias in datasets, along with the challenges they encounter in building diverse engineering teams. The final subsection of the findings considers whether diversity is viewed as a trade-off that limits access to skilled personnel or as an outdated concern, "a thing of the past."

5.2 Addressing bias in datasets

Unsurprisingly, the link between discriminatory outputs of AI systems and biased datasets occupied a prominent position in present research. Our interviews revealed that many interviewees demonstrated awareness of these risks and outlined technical strategies to mitigate them. These included separating training, testing, and validation datasets; sourcing data from diverse origins; implementing data quality control systems; auditing outcomes; and increasing transparency in how data is used and shared. However, other interviewees treated bias as a matter of secondary significance, prioritising technical fidelity to existing datasets and shifting responsibility for fairness and debiasing onto end users rather than embedding it within system design. Overall, interviewees predominantly framed AI bias as a technical problem rooted in data imbalance and engineering limitations, demonstrating awareness of risks but often overlooking broader sociotechnical dynamics that shape exclusionary design. A small number of participants moved beyond this framing by recognising bias as a societal issue and experimenting with user feedback, co-design, and participatory approaches that engage with civil society organisations (CSOs) and members of marginalised communities. While limited in practice so far, these approaches indicate existing "best practices" that, coupled with an overall empowerment of SMEs, could point to a viable path forward.

5.2.1 Techniques for addressing bias

Regarding separating training, testing and validation datasets, interviewee 16 suggests that properly separating data into distinct subsets is essential for building AI models that generalise well and avoid a feedback loop, thus mitigating the emergence of biases:

> Well, we try to use this training to separate data into training, test, and validation sets in order to avoid overfitting and underfitting with our data. We are always using data from machines, and we are not taking photos of people or information from our ongoing projects. So, at least from my point of view, if you need to create an AI algorithm with data, you split the data according to the 70%, 15%, and 15% reference, and you can avoid that.

The 70%/ 15%/ 15% split is a common heuristic in machine learning used to divide a dataset into three subsets for training, testing and validation.

Regarding the importance of representative data, interviewee 3, part of a company developing a generative AI system, mentions the importance of accessing data from different sources:

> Our system is a quite artistic kind of image generator AI. So we contract databases, different kinds of database owners who have a quite different database of art and for cultural differences we use public databases. This is very important to us for diversity.

While the data quality of the databases is uncertain, interviewees understood that diverse and representative datasets are mechanisms for reducing biased outcomes, and that mitigation strategies commonly involve improving data quality and diversity as well.

Interviewee 13 seeks to combine the heuristic with the database in a dual strategy that they view as a potential answer to all biases:

> You need to take two steps. The first one is internal, because we have some tools in order to test our solution and try to understand if there is some bias when we are developing the solution. The second step is more important and is made together with the organization, because when we analyze the data we try to understand if there are some issues related to this data. So before matching any data, you must be aware about what is the source of this data, if there is any backend bias. So a critical thinking approach to data is the most important step before starting any machine learning model or any machine learning solution. For this reason we create a sort of RAG. RAG is a system in which we can control the quality of data, the quality of information, the source of information and source of knowledge, that is the context that we give to the AI. In this way, we can control a good percentage of all issues that can emerge in this kind of project. It's an iterative and incremental approach to solve all kinds of bias, all kinds of problems that could emerge from the first installation.

Interviewee 13 does not connect the importance of other biases in the AI lifecycle to dataset bias.

Interviewee 25 stressed the need to audit the outcomes in order to identify biases early on in the process:

> Yes, we have in place a framework to develop the systems. This is in our benefit, because for instance, you have to audit the outcomes of Generative AI to get insights on how this outcome was generated.

This is recognition of auditing as post processing techniques to validate dataset representativeness and performance.

Interestingly enough, interviewee 4 described as bias mitigating strategies both a data review process and a commitment to transparency, which was not echoed by other interviewees in our sample:

> We have dedicated people that we call data analysts and data engineers. And what they're doing is manually checking and cleaning data. And one of the tasks that they have is to look also into these biases. So, we are very aware of that. And that's also why we want to make our AI as transparent as possible. So, the building blocks of the AI, how we created it. We have it also in Github, so everybody can see, what, let's say, the principles of how we developed it are. And we're also open for feedback and comments on that.

This reflects open access approaches, allowing others to to review the dataset validation processes.

Figure 3 presents a word cloud summarising the most frequently used terms by interviewees when discussing bias in datasets.

Figure 3
Interviewees discussing bias in datasets.

5.2.2 Bias as a second-order consideration

In addition to addressing bias methodologically, there were cases where interviewees approached biases as a matter of secondary significance. For example, interviewee 17 mentions that bias will be captured and reproduced in the design stage. However, they provide certain tools, and say responsibility lies with the user to deploy those tools for reasons of fairness or debiasing.

> In the current business that we are doing, we try to reproduce the statistical properties of the dataset of our customers using this tool that we have developed. This includes biases that data might have. So, if our customer dataset is biased, for some reason, because it is collected in a biased way, or whatever other reason, we will reproduce those biases. This is by design in our system. We do have tools for our customers to identify and remove those if they want. Because from the real dataset we generate a synthetic one, we have control over the dataset. For example, if the original dataset is unbalanced, say, it's 70% male and 30% female, we can correct that in the synthetic data if the customer wants. Anyhow, fairness is not a primary concern of our system, because the primary concern of our system is fidelity with respect to the original data, and then the fairness and unbiasing of the dataset is left to the customer.

Funded by
the European Union

Therefore, bias is not a priority for this interviewee's own company, instead they provide fidelity to the dataset.

In underplaying the importance of bias in data, only interviewee 28 demonstrated a narrow understanding of bias:

> I don't think there is any such risk of bias on our part, because all the data we enter, apart from the first and last names, is demographic data, age, gender, there's nothing to cause discrimination.

This is consistent with the understanding that some SMEs resource constraints preclude governance structures to identify or manage algorithmic bias and discrimination issues effectively.

5.2.3 Technical orientations towards bias problems

Overall, interviewee' responses indicate a technical orientation to addressing AI bias. This reflects both certain limitations in their approach but also an adequate understanding of the risks inherent in biased datasets. The absence of a sociotechnical analysis combining social and technical systems cannot adequately address the larger facets of exclusionary design. Notably two interviewees, sought to engage with broader approaches to data related biases. Interviewee 18 was among the few who identified gender bias as a societal concern that cannot be addressed solely in technical terms:

> Because in AI, bias can come from various different places. Your dataset might be biased. For example, your dataset contains a gender bias. You have more data about men than you have for women. It's just a bias in the data, and sometimes it's very hard to correct it. You might do some data engineering to compensate for class imbalance, but at the end of the day, your dataset will be imbalanced.

Interviewee 20 also suggested an approach based on interaction with system users as a pathway to achieving diversity. They received user feedback that darker skinned persons were finding them misrepresented in systems photos:

> To understand if it's an individual problem, or it's a total bias, we did a survey. We tried to study by research, by asking the feedback of a lot of people to see if they can also see this or not. So we tried to use subjects from different facial characteristics and different genders. We discovered that there is some bias. Because of the collection of the feedback, they say that this is not okay, for example, this is better, so they need more identity features in this specific application.

In addition to the survey, interviewee 20 designed co-design features to solve the problem of bias:

> To solve this, one idea is to have an interactive application that you give and take: you give a result and take feedback. So, in this case, you understand where you need to go for this specific group of people. Because you can't say "OK, in my opinion, it's correct, so everyone else should agree with me". It's not like this. Especially in this kind of application, where it's very subjective.

Overall, interviewees displayed a largely technical orientation toward the problem of data bias, relying on standard practices such as dataset partitioning, quality control tools, and outcome audits.

5.3 Challenges in bringing diversity in-house

Overall, interviewees recognised the importance of establishing diverse engineering teams, including ones with gender balance, within their own companies. This trend was significantly more prominent across interviewees that develop AI systems. However, in contrast with addressing bias in datasets, interviewees find it significantly more challenging to establish diverse engineering teams and achieve gender balance within their own companies in practice. As shown in table 1, the majority did not have a hiring policy in place to foster diverse hiring. Despite this, SMEs that did have a hiring policy were content with the results and understood it as part of their business plan that needs to be consistently implemented. As interviewee 13 notes:

> We are 50-50, male and female. It's something that is happening because we are focused to be representative regarding diversity, regarding inclusion, it's not something that comes natural due to the kind of organisation we are.

Similarly, interviewee 12 noted that:

> I think maybe half of the people here are women and in technical roles. So yeah, we have a very strong policy regarding that.

On the other hand, for some of our interviewees this issue was resolved organically in the process of developing their company. For example interviewee 25 mentioned that:

> Actually, we are a gender-diverse company, because we have 60% ladies in our company. So, they are the decision makers in our small team. Of course, gender diversity is more than that, and, sometimes we are going beyond expectations within this point, when we're talking about gender diversity, really. I mean, for me, it doesn't matter if I'm working with LGBTQ or if I'm working with a lady or if I'm working with a man. The important thing is how we are creating a team of understanding and knowledge, right?

Similarly interviewee 27 notes that:

> We don't have it as a hiring policy, but now we have a bit more women. A year ago, we were very balanced. Now we have a bit more women on the admin side, and then on the developers, it's quite balanced. It's a pro-feminine company, if I can say so.

These quotes suggest that even in the absence of formal hiring policies, a section of our interviewees acknowledge the importance of diversity and are willing to highlight their achievements in this area. For some participants, diversity is seen as a goal worth celebrating, possibly reflecting an implicit awareness that tech companies are typically male-dominated and that deviating from this norm is noteworthy. However, this remains a part of the picture and is relevant to “pipeline” issues discussed in the next section.

5.3.1 Persistence of male-dominated environments and the “pipeline” problem

Nonetheless, most of the SMES in the sample reflected a mostly male environment, typical of many companies in the tech sector in general (Eurostat, 2025) and in the AI industry in particular (Pal, Lazzaroni, & Mendoza, 2024). Interviewees from male-skewed companies have undergone the process to acquire a gender equality certification but in practice they reported an inability to achieve gender balance in the workplace. As interviewees 17 mentioned:

> Well, we actually have a certification for gender equality. We have just obtained it. But to be fair, our team, the tech team, is very unbalanced in terms of gender ratio. It's actually quite hard, at least here in Italy, to hire women in tech. We have a policy in place for gender equality. If we have two candidates that have the same abilities, we do prefer the female candidate, especially because we are very, unbalanced right now. The tech side is 90% men.

Interviewee 20, having also a gender equality certificate, did not consider gender as a criterion when hiring:

> We have a certificate in gender equality. But, in general, as employees, we don't see a difference between a female or a male... it's not a criteria to distinguish people in our company. It is more about creativity or the natural preferences in terms of work for each one of us, because in our company, we come from very different areas of study. Everybody brings a particular background on the table, and this is probably the only main differentiation between us.

The above excerpts are indicative of two trends identified in our interviews: the first (as exemplified by interviewee 17) concerns the broader challenges of recruiting women in the tech sector, while the second (as exemplified by interviewee 20) reflects a preference for skill over gender balance and an underlying assumption that enhancing diversity may come at the expense of technical competence; therefore, there is a tradeoff at play.

Proceeding to engage with the first point, it was reiterated by SMEs that have an active strategy to ensure gender balance that they find it increasingly difficult to implement it. According to interviewee 21:

> We are pleased to have different genders in our company. We have a document, we write a charter, but most important for us is that we want to have diversity in our company because

> it's where we learn a lot. But I don't think it's a charter which changes things. It's the mindset of people working on the subject. What is difficult is that in tech, there's fewer women than men, and it's quite difficult to find them.

They consider diversity as an asset, while noting that overall the opportunities to hire women are low in the tech sector. In a similar note, interviewee 22, running a consultancy company, tries, among others, to help their clients with diversity goals and clearly states that "on the tech side and the AI startup side, the gender bias is skewed towards males". Interviewee 29 clearly identifies it as a "pipeline issue" as fewer women are studying engineering:

> For gender, we have a problem. There are few women in this process, mostly men. But this is not easy in France because the engineering schools have only 15% of girls. So it's a big topic, not very easy to deal with.

Echoing this narrative with reference to a different country, interviewee 14 mentioned that:

> We only have one woman in our team. I think it's a real problem in universities in Italy. A class will have 200 people with only 2,3,4 women in it. After that, the women tend to go into UX, UI, graphics, something frontend. So something completely different from developing solutions. So for us, it's very difficult to find good women developers, not because we are men and we don't want to, but because it's very difficult to find.

The other side of the "pipeline problem" is not only that fewer women pursue certain studies, but also that far fewer show interest in specific sectors in the job market. This is reflected in the following quote by interviewee 15:

> There are about 40 people in our company. Unfortunately, the reality is that we only have 6 women out of the 40 people we are, and only 2 of them are in the tech department. What we do is an AI in financial markets where the reality is that most of the people are men. Therefore, the combination of AI and financial markets makes the problem squared, somehow. So, if we consider the number of applications of men and women, we have more men than women.

The "pipeline/sector" narrative was common among many of our interviewees that have hiring policies in place or in some way aspire to achieve gender balance in their engineering teams. In this case, the resource constraints interact with the broader limitation of European educational systems and the marginalisation of women across different economic sectors. When discussing the "pipeline problem", SMEs tend to frame the lack of diversity as a wider societal issue. However, once they recognise the scale and structural roots of the problem, they often feel powerless, perceiving it as an insurmountable challenge that cannot realistically be addressed by a single SME.

Figure 4 presents a word cloud summarising the most frequently used terms by interviewees when discussing hiring policy within their own organisations.

Figure 4
Interviewees discussing hiring policy.

Furthermore, as the next section will elaborate, a certain section of interviewees not only identified challenges in implementing diversity policies but considered diversity itself as an ambivalent or illegitimate objective.

5.3.2 Diversity as an ambivalent or illegitimate objective

For some interviewees, diversity within the company was not recognised as a legitimate goal in itself, but rather was treated as an open question to be evaluated in terms of the immediate benefits it might bring in the future. For example, according to interviewee 24:

> It's true that our company is mostly made of white men, between 30 and 40 years old, so probably there is a huge bias there which we haven't really addressed, but, I mean we are happy to continue that research provided that, it can be somehow beneficial for us in terms of our visibility or financially, essentially.

Furthermore, a certain ambivalence was observed regarding diversity itself. Interviewees, even when asked explicitly about gender diversity in their engineering teams proceeded to emphasise diversity of nationality in their company, effectively bypassing gender as a criterion. Interviewee 24 presents this juxtaposition in a clear manner:

> You know, we have Italian, Turkish, Chinese, Iranians, Congolese, Spanish, Belgian, for sure. We don't really put gender...I don't consider that yet as a KPI, or as a rule when I interview people.

Similarly, interviewee 3, when asked about diversity related hiring policies, they replied that:

> The developers are from different countries now and one of our managing directors is a woman. The other manager is male. The developers are from 2 countries, Hungary and Romania, for now. In general, we don't make differences.

By deploying this kind of response, interviewees avoided discussing the gender composition of their engineering team. However, interviewee 8 mentioned that:

> Being a small company, we hire the people that are good for the team. So we have both. You know, females and males. So we are about 25% women. Not too bad. If you have any female CVs, send one: we are hiring people. We also have a lot of people from various countries if you look into that aspect. We have people from Germany, Sweden, India, Russia, Vietnam.

Within this narrative, gender diversity is, explicitly or implicitly seen as a burden, especially when correlated with company size. The resource limitations of SMEs are reiterated, echoing literature on how SMEs prioritise cost saving and efficiency over fairness criteria (Le Dinh et al., 2025). Interviewee 2 relates company size with a limited capability to address all possible angles of diversity, while also expressing satisfaction with team composition:

> Obviously, diversity of thought and experience is key. But since we are a small team, it's not always possible to look at all angles for diversity. As far as the tech team is concerned, this is where we have been able to find the technology experts from a very small group, so if you consider both advisors and the core team, we've got a good diversity. But if you're just looking at a core tech team, we probably are not as diverse in all aspects, but we have one aspect of diversity. The team lives in different countries. We are not from one country.

However, it should be noted that the most of the interviewees also offer a remote option for their employees. Expectedly, a hybrid working environment will attract workers from different countries and this is not uncommon in the tech sector in general. In that sense, presenting "diversity of nationality” is at least a very partial answer to diverse teams generally, and to gender concerns specifically. The above quotes indicate a feeling of uneasiness in discussing gender diversity that warrants further examination.

However, this uneasiness was seldom expressed as explicit indifference towards diversity. Only one interviewee (29) stated that diversity training and gender considerations are “a thing of the (recent) past”.

> We had many training sessions on diversity, gender, etc. And now, when I talk with companies about what they do, they do less and less training about this and, it's not now on the topics. It was, but it's not anymore. Maybe it's just because we have all these problems, politics and economics etc.

While data from our sample do not indicate a blatant disregard for gender diversity, future research could corroborate or refute this statement. We have to consider the possibility that interviewee 29 is echoing a broader shift in the wider geo-political climate, already reflected in the removal of DEI mandates from company policies and can be expected to also influence SMEs; further research is necessary on this topic.

5.3.3 Broader strategies for achieving diversity

The findings thus far have highlighted the challenges of bringing diversity within SMEs and the ways diversity is sometimes overlooked as an objective for a company operating under specific resource constraints.

However, this does not preclude the possibility of SMEs' capability of developing alternative strategies to ensure diversity. For example, interviewee 31 was confident that SMEs can address the “pipeline issue” by building synergies with CSOs specialised in gender rights. Their experience was the following:

> We have an agreement signed with an important gender NGO working at the national and international level. And we have committed to hiring women if they have the knowledge we are asking for that job position. Because this organisation is working not only with women, but also with women living in hard situations. If they have been suffering some sort of discrimination, something like this. So we have an agreement with them, and we try to hire these women. This is one tool, or one way we are working towards, gender equality. We are aware that in these careers, engineering and IT, there is this difficulty of finding more women than men, but my advice would be maybe to stay in touch with gender NGOs that are working on reskilling, upskilling, training for women, and they are, trying also to support them in finding a job position so they can connect.

Similarly, interviewee 23 paints an optimistic picture, one of successfully changing their engineering team, empowering the company in the process. As they mention:

> We have a hiring policy and the balance is not 50-50, it's 48-52. And 10 years ago, or 5 years ago, it was like 20-80. I know that a man and a woman give different things to the company, and I see that in our teams. So the balance I think it's very good. I have way more women in development than I had on the start. That's nice. And the team keeps being quite balanced. When you have too much for one side or for the other side, it doesn't work very well.

Here balance may be approached in a narrow manner, but overall it is an example of how an SME can gradually and effectively transform their composition.

At this point, it is worth mentioning that gender diversity was approached strictly in terms of a binary (male-female) by the large majority of our interviewees. The exceptions are quite limited. As mentioned before, interviewee 20 hints at the possibility of a spectrum (“the other genders”) but chooses to not engage with it. Interviewee 5 was the only one who referenced the LGBTQ+ community and the possibility for a future expansion of their hiring policy:

> We are 50% male and 50% female in the hiring. So yeah, we really like to have a mixed team and a mixed policy and so we have a high diversity level in our team. We still don't have any disability program in our company, but we are surely about to start it very soon. We'll also consider that as well for the LGBTQ+ community.

Given the substantial constraints SMEs face, the findings reveal a complex and somewhat polarised landscape. On one hand, there are SMEs that are either unable (or in some cases unwilling) to bring diversity within their organisations. On the other hand, there are SMEs that view diversity as integral to their business strategy and actively seek to develop internal policies and partnerships to overcome existing barriers. This divergence will be explored further in the next section.

.

Section 6: Discussion

This study revealed how SMEs understand and respond to bias within AI systems, focusing on both technical and sociotechnical dimensions. Interviewees widely acknowledged bias as a significant risk, especially in relation to datasets, and described common mitigation strategies such as data partitioning, quality audits, and dataset diversity. However, most approaches remained narrowly technical, with limited engagement with structural or intersectional sources of bias. Only a minority of participants advocated broader practices like transparency, open data, or co-design. This technical emphasis was echoed in discussions of team diversity: although many interviewees recognised its value, gender and equity considerations were also deprioritised or dismissed as external problems or outdated goals. The findings suggest that while SME actors are aware of bias risks, systemic limitations including resource constraints and male-dominated organisational cultures hinder diversity efforts needed for ethical-by-design or fair AI orientations.

6.1 Acknowledgement of bias as a risk factor

While the literature on SME digital readiness shows that many smaller firms approach AI through a pragmatic rather than principled lens, our finding showed a relatively high level of awareness amongst SMEs. Interviewees represent SMEs and themselves occupy dual identities as both businesspersons and technology users and developers. This position enables them to conceptualise bias as a risk inherent in AI systems, one they wish to address for professional, commercial and ethical reasons. In particular, due to their familiarity with machine-learning systems as both developers and users, almost all interviewees demonstrated an awareness of the risks embedded in datasets that may replicate pre-existing biases and reproduce stereotypes or discriminatory patterns in AI-generated outputs or AI-assisted decisions.

6.1.2 Techniques for addressing bias (in the datasets)

As the findings indicate, SMEs employ a range of techniques to pre-emptively mitigate dataset bias (e.g. dataset splitting to avoid overfitting, sourcing diverse data, internal bias-testing tools, output auditing, manual data review). This also suggests that the highly publicised cases of AI bias in recent years have had an impact: SMEs, especially those developing AI systems, recognise that exposure to biased systems could severely compromise their reputation and jeopardise their long-term viability. For example, interviewee 16 mentioned that the 70%/ 15%/ 15% split is a common heuristic in machine learning used to divide a dataset into three subsets for training, testing and validation. Further, interviewee 3's understanding is consistent with Ferrara's (2024) understanding that diverse and representative datasets are mechanisms for reducing biased outcomes, and that mitigation strategies commonly involve improving data quality and diversity as well as fairness-aware algorithm design. In particular, interviewee 25's recognition of auditing as post processing techniques to validate dataset representativeness and performance per Chadha (2024) is a significant finding.

6.1.3 Technical orientations towards sociotechnical bias problems

At the same time, SMEs' position within the tech sector is reflected in their largely technical orientation toward the problem of data bias, through their reliance on standard practices such as dataset partitioning, quality control tools, and outcome audits. Such approaches on their own may exclude ethical-by-design or fairness factors. For example, bias can still be introduced in the 70%/ 15%/ 15% debiasing heuristic if any of these subsets are unrepresentative of a broader dataset (e.g. they are skewed across gender or sexuality) (Pagano et al., 2023). Fairness-aware AI systems must therefore ensure that each subset maintains demographic diversity and that outcomes are assessed across subgroups, not just in aggregate. Further, while these approaches demonstrate an awareness of algorithmic risk, few responses are engaged with the broader sociotechnical structures that produce or amplify exclusion through data.

6.1.4 Beyond technical solutions: transparency, open data and co-design

Exceptions include interviewees who linked data bias to structural gender inequality or who incorporated user feedback loops and co-design methods into their systems. For example, interviewee 4 described bias mitigating strategies both as a data review process and a commitment to transparency, which was not echoed by other interviewees in our sample. This reflects open access approaches, allowing others to review the dataset validation processes and is consistent with literature research on ethical AI governance highlighting transparency, fairness, and explainability as core ethical principles embedded throughout the AI lifecycle (Gebru et al, 2021; Ferrara, 2024). Even more noteworthy is interviewee 20 emphasising co-design features to solve the problem of bias. This approach aligns with scholars who argue that co-design and co-production are a tool in a suite of approaches through a fundamental re-architecture of the AI lifecycle (Mushkani et al., 2025). From these authors' perspectives, re-design should center co-production, diversity, equity, inclusion (DEI), and multidisciplinary collaboration. These outliers signal an important shift: that bias mitigation must move beyond data engineering alone and into inclusive design processes. The following subsection examines this further by exploring how stakeholder diversity, especially within SME teams and user groups, is understood as a potential driver for more ethical and fair AI design and the challenges this brings.

6.2 Diversity solutions

6.2.1 Persistence of male-dominated environments and the “pipeline” problem

Overall, interviewees recognised the importance of establishing diverse engineering teams, a trend that was significantly more prominent across interviewees that develop AI systems. Thus, they seem to at least partially align themselves with scholars who emphasise the importance of diversity in AI development teams (Shams, Zowghi and Bano, 2025). However, most of the SMEs in the sample reflected a mostly male environment, typical of many companies in the tech sector in general (European Commission, 2024) and in the AI industry in particular. This is consistent with long-term observations of the tech sector in general (Chang, 2018) and with the particular limitations of SMEs (Fair Tech Policy Lab, 2023).

The predominantly technical orientation to bias in male-dominated development teams leads most interviewees to treat bias chiefly as a data-related problem stemming from training practices of machine learning algorithms. Only some of them connected bias to broader organisational cultures within the tech sector that remain skewed towards white, male perspectives and experiences. As shown in table 1, the majority did not have a hiring policy in place to foster diverse hiring. This is exacerbated by a finding of our research that is often overlooked in contemporary literature: the inadequacies of gender equality certifications. This relates both to a narrowly defined understanding of AI bias (supposedly pertaining exclusively to datasets) but also links back to what has been identified regarding limited capabilities of SMEs for AI governance (Jarvers, Ullstein, & Grossklags, 2025).

6.2.2 Diversity as an ambivalent or illegitimate objective

This absence of sociotechnical understanding is reflected in how interviewees discussed diversity in-house, particularly with regard to their own engineering teams. As this issue came up, a range of additional challenges surfaced. As table 1 indicates, most participants lacked concrete policies for fostering diversity, and some did not consider it a priority. Others invoked diversity of nationality as a proxy for gender diversity. This argument is effectively weakened if we take under consideration the transnational production model of several interviewees, many of whom rely on remote work and maintain an international presence. For some, bias was framed as a societal problem originating in stereotypes embedded within education systems that result in fewer women entering software engineering. Yet even in such cases, interviewees tended to emphasise the "pipeline problem" in a way that implicitly absolved their organisations of responsibility. However, based on our findings, diversity may be less of a "pipeline" issue and more of a recruitment issue. Strategies for gender diverse recruitment, or even policy acknowledging the benefits of gender diverse recruitment were not apparent.

6.2.3 Broader strategies for achieving diversity

Only a handful identified concrete strategies going forward to address pipeline constraints: their suggestions regarding alliances between SMEs and civil society organisations focused on gender equality merit further exploration and may point to a broader, underutilised approach. As FORSEE Report D3.2 indicates, civil society organisations often operate under a broader, sociotechnical understanding of AI gender bias but lack the technical expertise, whereas SMEs develop certain technical capabilities while sometimes lacking a broader perspective; and multiple perspectives has the potential to benefit both. Indeed, broader perspectives must be welcome across the AI lifecycle given vulnerabilities that extend beyond SME influence. AI systems can be calibrated and used to override concerns about bias and discriminatory impacts at the monitoring stage in an AI system's use. Decisions about AI's operational settings, thresholds, and risk tolerances can still **favour actionable as opposed to accurate outcomes.**

6.2.4 Concerning trends for diversity

Finally, although a minority position within our sample, another concerning tendency emerged in our findings. Some interviewees reproduced a "meritocracy" narrative, suggesting that DEI considerations in hiring may restrict their ability to recruit highly skilled personnel. In a similar vein, some participants claimed that diversity goals are increasingly being removed within companies because they are viewed as outdated or no longer relevant. Such responses reflect broader political shifts towards deregulation and the downgrading of gender-equality agendas. While the present data do not indicate that this constitutes a majority trend, its presence is noteworthy and requires further investigation in order to assess its dynamic and potential implications. This is particularly the case given that gendered forms of knowledge and skill are valued and, as Corcoran and Cullen (2020) outline in the volume they edit, are established as leading to downstream impacts on technology. As Harvey (2022) writes in relation to gaming, the barriers women face have been firmly established, including well-documented harassment and material forms of structural discrimination such as gender pay gaps. Despite this, explanations of homogeneity as being due to a 'leaky pipeline' between training and the workforce persist, extending discourse familiar from the history of computing. Harvey (2022) argues that efforts to 'get in' to exclusionary tech spaces based on discourses of feminine lack don't account for how such environments require marginalized people to develop strategies for coping with exclusionary norms to 'stay in'. Indeed, as FORSEE discusses in Report D3.2, AI gender bias, as researchers Buolamwini & Gebru (2018) have famously articulated, is not just a technical error, but a structural and systematic issue where AI systems reinforce existing societal prejudices, including bias against gender. Connected with AI gender bias is discrimination, where bias acts as a cause or mechanism and discrimination becomes the real world consequence of differential treatment along categories of identity including gender. An AI hiring tool with a gender bias that preferentially selects men leads to a discriminatory consequence for applicants who are women. Therefore, a form of AI gender bias at the recruitment level is one that dismisses the importance of diversity across the AI lifecycle.

6.3 Ethical-by-Design and fairness concepts is also a matter of context

Findings indicate that " ethical-by-design" principles enshrined in regulation like the AI Act are not easily introduced in SMEs due to, as a point already highlighted in relevant literature, their limited capacities (Jarvers, Ullstein, & Grossklags, 2025). Interviewees appeared significantly less familiar with concepts of "ethical-by-design" and "fairness" and the connection between a male-skewed environment (Chang, 2018) and biased AI systems was consistently underplayed. This suggests a disconnect between technical mitigation strategies and the broader socio-organisational conditions that shape how technologies are produced and perceived. This gap is important: scholars have long argued that strategic diversity within development teams (Gurzawska, 2024; Shams, Zowghi and Bano, 2025) is not merely a matter of representation but a structural safeguard against reproducing discriminatory assumptions in design, implementation and evaluation. Therefore, fairness as an objective in technological innovation is closely related to ethical-by-design principles.

The limited attention paid by SMEs to these dimensions is shaped less by a lack of willingness and more by the limited frameworks available to them. SMEs operate within a context of infrastructural dependency on large tech firms (Schwaeke et al., 2025), regulatory fragmentation and limited support from public authorities, as it is also demonstrated in FORSEE report 3.1. Within such a landscape, SMEs have neither the agency nor the resources to articulate or pursue more substantive approaches beyond engaging with biased datasets. Their responses therefore reflect the understandings embedded in the legal and procedural environment in which they operate, rather than an absence of concern or ethical commitment.

At the same time, the abovementioned factors are important for situating respondents' perspectives without absolving SMEs of their responsibility to meaningfully engage with AI gender bias. References to "natural preferences in terms of work" (as expressed by interviewee 20) implicitly legitimise male-skewed organisational environments by naturalising the lack of diversity in engineering teams. Similarly, and perhaps more insidiously, the "pipeline" metaphor requires critical interrogation rather than uncritical acceptance. While it may partially reflect existing structural conditions on all educational levels (Varma & Hahn, 2008), it also serves to shift responsibility away from SMEs and their hiring policies. The term "pipeline" implies a given structure ensuring the trajectory of software engineers from education to employment, simultaneously obscuring the mediating factors that shape recruitment. The conclusion of the "pipeline" narrative is that diversity in hiring is pre-determined upstream and lies beyond organisational influence. Instead of focusing linear trajectories, Soe & Yakura (2008), examine societal, occupational, and organizational cultures, and "argue that these cultural contexts form layers that are interdependent and to some extent self-reinforcing" (Soe & Yakura, 2008:184). Therefore, by implicitly reinforcing gender stereotypes, we run the risk of creating a vicious circle where focusing on the "pipeline issues" actually exacerbates the problem and diminishes diversity in hiring.

6.4 Limitations

Several limitations of the present study should be acknowledged. First, the research did not collect gender data on interviewees due to the project's data management plan and commitments to anonymity. While this protects participants, it limits the ability to empirically assess how gendered positionality shaped interview responses. This constraint is significant given evidence that women are underrepresented in technical roles in EU digital and AI-related work (European Commission, 2024) and also in many companies in the tech sector in general and in the AI industry in particular (see section 5.3.1). This "pipeline" and representation gap discussed in our findings can shape organisational perspectives on AI bias feeding our research data and may contribute to the relative absence of sociotechnical or gender-aware analyses in AI development contexts.

Second, interviewees participated as representatives of their organisations rather than as private individuals, which likely influenced how openly they addressed gender-related bias. Several responses suggested a tendency toward socially acceptable or reputationally cautious framing, particularly when discussing fairness, discrimination, or diversity issues. This raises concerns about social desirability bias and corresponding dynamics in methods of qualitative research. Such dynamics may have led interviewees, particularly those in senior or client-facing roles, to downplay structural gender bias in AI systems they develop or deploy.

Third, the study primarily captured perspectives from individuals in technical, managerial, or executive positions, rather than from employees in junior, operational, or marginalised roles within SMEs. Prior research indicates that experiences of bias, exclusion, and discrimination are often more visible to workers positioned lower in organisational hierarchies or outside dominant demographic groups (Newman, et al, 2025). As a result, the findings may underrepresent critical perspectives on gendered harms and everyday discriminatory effects of AI systems in SMEs' contexts. This route of inquiry might also entail conducting additional qualitative future research on subjectivities of women and LGBTQI tech workers and successful empowerment strategies, linking those perspectives critically back to the "pipeline" problem.

Finally, with regard to hiring policies, the present study did not examine whether interviewees used AI-based hiring software within their own organisations. This remains an important aspect that could further illuminate the current landscape of diversity (or lack thereof) within digital SMEs and should be addressed in future research.

Taken together, these limitations suggest that the findings should be interpreted as reflecting organisationally mediated and predominantly male-dominated SME perspectives, rather than a comprehensive account of gendered or other marginalised experiences of AI bias. Future research would benefit from purposive sampling strategies that explicitly incorporate gender-diverse participants, junior staff, as well as mixed-methods approaches that triangulate organisational narratives with demographic data.

Section 7. Conclusion

The present study focused on how European SMEs perceive AI bias, and gender bias in particular, as well as the strategies and constraints that shape their capacity to address it. The findings reveal that SMEs are aware of potential risks that biased AI systems pose and many adopt practical measures to reduce dataset bias where possible. However, their understanding and responses are largely shaped by the limited frameworks, resources, and infrastructures available to them. As a result, SMEs tend to approach bias primarily as a technical problem located in training datasets, rather than as a sociotechnical issue closely linked to workplace diversity, development practices, or broader power dynamics in the AI lifecycle. Only some SMEs connected bias to broader organisational cultures within the tech sector that remain skewed towards white, male perspectives and experiences. However, data outliers signal an important shift: that bias mitigation must move beyond data engineering alone and into inclusive design processes.

Policymakers therefore need to be aware of these concerns and consider how regulatory frameworks, support programmes and capacity-building initiatives might incentivise or enable SMEs to integrate diversity, inclusion and ethical design principles more systematically into their development processes. At the same time, the position of SMEs within a fragmented European AI ecosystem is marked by funding shortages, dependence on external, often non EU providers, and limited organisational capacity. Therefore their agency is significantly constrained. While some SMEs embrace diversity as part of their business plan, others feel immobilised by the scale of the problem and tend to consider it as beyond their remit. These divergent approaches underscore the uneven landscape in which SMEs operate. Therefore, policymakers should support SMEs not only through regulatory compliance frameworks but also by strengthening the wider ecosystem with a particular focus on funding and computational infrastructure, while simultaneously facilitating networking with relevant CSOs, a practice that, albeit limited until now, seems to yield results and foster diversity in SMEs. Therefore, while often not included in these debates, digital sovereignty is crucial in promoting fair, non discriminatory AI use and development.

Overall, while concrete policy recommendations fall beyond the scope of the present research, an outline of an alternative pathway nevertheless emerges. Building on participatory approaches and collaboration with CSOs, as well as with women and LGBTQ+ individuals, SMEs can more effectively address AI-related gender bias, provided they are also empowered by public authorities to reduce their dependency on Big Tech infrastructures. This dual approach both acknowledges SMEs' responsibility to act and frames digital sovereignty as a crucial prerequisite. Ultimately, addressing AI gender bias entails enabling SMEs to move from technical mitigation to more holistic approaches, while increasing the trustworthiness and acceptability of their systems to end users.

References

Agbaakin, O. (2025). Leveraging Artificial Intelligence as a Strategic Growth Catalyst for Small and Medium-sized Enterprises. https://doi.org/10.48550/arXiv.2509.14532

Banerji, L., & Satija, D. (2025, February 7). AI as double speak for austerity. TechPolicy.Press. https://www.techpolicy.press/ai-as-double-speak-for-austerity/

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(3), 671–732.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity Books.

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns, 2(2)*, 100205. https://doi.org/10.1016/j.patter.2021.100205

Buolamwini, J. & Gebru, T. (2018), Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Buolamwini, J. (25 January 2019). Response: Racial and Gender bias in Amazon Rekognition – Commercial AI System for Analyzing Faces. *Medium.* https://medium.com/@Joy.Buolamwini/Response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced

Chadha, K. S. (2024). Bias and fairness in artificial intelligence: Methods and mitigation strategies. *International Journal for Research Publication and Seminars*, 15(3), 36–49. https://doi.org/10.36676/jrps.v15.i3.1425

Chang, E. (2018). *Brotopia: Breaking up the boys' club of Silicon Valley*. Portfolio/Penguin.

Chen, Y., Clayton, E. W., Novak, L. L., Anders, S., & Malin, B. (2023). Human-centered design to address biases in artificial intelligence. *Journal of Medical Internet Research*, 25, e43251. https://doi.org/10.2196/43251

Corcoran, M., & Cullen, P. (Eds.). (2020). Producing knowledge, reproducing gender: Power, production and practice in contemporary Ireland. *University of Chicago Press.*

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336–349. https://doi.org/10.1177/1527476418796632

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics. *University of Chicago Legal Forum, 1989*(1), 139–167.

Dai, K., Kim, J. J. Y., Džeroski, S., Wicker, J., Dobbie, G., & Dost, K. (2025). *Assessing the risk of discriminatory bias in classification datasets. Machine Learning, 114*(9), Article 204. https://doi.org/10.1007/s10994-025-06843-9

Dastin, J. (2018, October 10). Amazon scrapped 'sexist AI' recruiting tool. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

D'Ignazio, C., & Klein, L. F. (2020). Data feminism. MIT Press.

Duong, M. K., & Conrad, S. (2024). Measuring and mitigating bias for tabular datasets with multiple protected attributes (arXiv:2405.19300). arXiv. https://arxiv.org/abs/2405.19300

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European Commission. (2023). *Entrepreneurship and small and medium-sized enterprises (SMEs).* European Commission. https://single-market-economy.ec.europa.eu/smes_en

European Commission. (2024, August 12). *Women in Digital Scoreboard 2024.* https://digital-strategy.ec.europa.eu/en/news/women-digital-scoreboard-2024

Eurostat. (2025). *Artificial intelligence by size class of enterprise*. European Union. https://doi.org/10.2908/ISOC_EB_AI

Eurostat. (2025, July 8). *Rising share of ICT specialists among employed people*. European Commission.
https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20250708-2

Fair Tech Policy Lab. (2023). *The equity blind spot: How the EU AI Act fails to protect the most vulnerable*. Fair Tech Policy Lab. https://www.fairtechpolicylab.org/post/the-equity-blind-spot-how-the-eu-ai-act-fails-to-protect-the-most-vulnerable

Ferrara, E. (2024). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. Sci, 6(1), 3. https://doi.org/10.3390/sci6010003

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance, 77*(4), 1977–2022. http://dx.doi.org/10.2139/ssrn.3072038

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM, 64*(12), 86–92. https://doi.org/10.1145/3458723

Gisselbaek, M., Berger-Estilita, J. M., Minsart, L., Saxena, S., et al. (2025). *Gender disparities in artificial intelligence-generated images of hospital leadership in the United States. Mayo Clinic Proceedings: Digital Health, 3*(2), 100218. https://doi.org/10.1016/j.mcpdig.2025.100218

Gohar, U., and Cheng L. (2023). A survey on intersectional fairness in machine learning: Notions, mitigation, and challenges. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI, 6619–6627 https://www.ijcai.org/proceedings/2023/0742.pdf

González-Sendino, R., Serrano, E., and Bajo, J. (2024). Mitigating bias in artificial intelligence: Fair data generation via causal models for transparent and explainable decision-making. Future Generation Computer Systems, 155, 384–401. https://doi.org/10.1016/j.future.2024.02.023

Gurzawska, A. (2024, January 25). The power of human-rights centred ethics-by-design. Trilateral Research. https://trilateralresearch.com/emerging-technology/the-power-of-human-rights-centred-ethics-by-design

Information Commissioner's Office. (2023, March 15). What about fairness, bias and discrimination? In Guidance on AI and data protection. ICO. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-fairness-in-ai/what-about-fairness-bias-and-discrimination/

Harvey, A. (2022). Making the grade: Feminine lack, inclusion, and coping strategies in digital games higher education. *New Media & Society, 24*(9), 1986–2002. https://doi.org/10.1177/1461444820986831

Hungarian Civil Liberties Union, Civil Liberties Union for Europe, European Digital Rights, & European Center for Non-Profit Law. (2025, May 6). *Hungary's new biometric surveillance* *laws violate the AI Act*. European Digital Rights. https://edri.org/our-work/hungarys-new-biometric-surveillance-laws-violate-the-ai-act/

Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of big data. *Yale Journal of Law & Technology, 18*(1), 148–216. https://yjolt.org/credit-scoring-era-big-data

Jarvers, S., Ullstein, D., & Grossklags, J. (2025). *Uncertainty as a Primary Barrier for Trustworthy AI Under the EU AI Act: German SME Perspectives*. Proceedings of Fourth European Workshop on Algorithmic Fairness, PMLR 294:281-287

Le Dinh, T., Nguyen, T. H., & Bui, T. (2025). Artificial intelligence in SMEs: Enhancing business performance and innovation. *Sustainability*. MDPI.

Lindgren, S., & Dignum, V. (2023). Beyond AI solutionism: Toward a multi-disciplinary approach to artificial intelligence in society. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 163–172). Edward Elgar Publishing. https://doi.org/10.4337/9781803928562.00019

Mavrogiorgos, K., Kiourtis, A., Mavrogiorgou, A., Menychtas, A., & Kyriazis, D. (2024). Bias in Machine Learning: A Literature Review. Applied Sciences, 14(19), 8860. https://doi.org/10.3390/app14198860

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35.

Mosier, K. L., & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other? In *R. Parasuraman & M. Mouloua (Eds.), Automation and human performance: Theory and applications* (pp. 201–220). Lawrence Erlbaum.

Mushkani, R., Berard, H., Ammar, T., & Chatonnier, C. (2025). Co-producing AI: Toward an augmented, participatory lifecycle. arXiv. https://doi.org/10.48550/arXiv.2508.00138 (preprint)

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Pagano, T. P., Loureiro, R. B., & Lisboa, F. V. N., et al. (2023). Bias and unfairness in machine learning models: A systematic review on datasets, tools, fairness metrics, and identification and mitigation methods. Big Data and Cognitive Computing, 7(1), 15. https://doi.org/10.3390/bdcc7010011

Pal, S., Lazzaroni, R. M., & Mendoza, P. (2024, October 10). *AI's missing link: The gender gap in the talent pool* (Data brief). Interface. https://www.interface-eu.org/publications/ai-gender-gap

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253.

Quenum, G. G. Y., Vallée, S., & Ertz, M. (2025). The digital maturity of small- and medium-sized enterprises in the Saguenay-Lac-Saint-Jean region. Machines, 13(9), 835. https://doi.org/10.3390/machines13090835

Schönberger, M. (2023). Artificial intelligence for small and medium-sized enterprises: Identifying key applications and challenges. *Journal of Business Management*, 21, 89–112. https://doi.org/10.32025/JBM23004

Sánchez, E., Rodríguez, A., & López, S. (2025). Artificial intelligence adoption in SMEs: A survey-based assessment of opportunities and risks. *Applied Sciences, 15*(4), 1123–1144.

Schwaeke, J., Peters, A., Kanbach, D. K., Kraus, S., & Jones, P. (2025). *The new normal: The status quo of AI adoption in SMEs. Journal of Small Business Management, 63(3), 1297–1331. https://doi.org/10.1080/00472778.2024.2379999*

Shams, R. A., Zowghi, D., & Bano, M. (2025). AI and the quest for diversity and inclusion: A systematic literature review. *AI Ethics,* *5*, 411–438. https://doi.org/10.1007/s43681-023-00362-w

Soe, L., & Yakura, E. K. (2008). What's Wrong with the Pipeline? Assumptions about Gender and Culture in IT Work. *Women's Studies*, 37(3), 176–201. https://doi.org/10.1080/00497870801917028

Suresh, H., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning (arXiv:1901.10002). *arXiv*. https://arxiv.org/abs/1901.10002

Ulrich, P., & Frank, V. (2021). Relevance and adoption of AI technologies in German SMEs. Results from survey-based research. *Procedia Computer Science,* 192, 2152–2159. https://doi.org/10.1016/j.procs.2021.08.228

Varma, R., & Hahn, H. (2008). Gender and the pipeline metaphor in computing. *European Journal of Engineering Education*, 33(1), 3–11. https://doi.org/10.1080/03043790701745936

Watney, C., & Auer, D. (2021). Encouraging AI adoption by EU SMEs. *Progressive Policy Institute.* https://www.progressivepolicy.org/wp-content/uploads/2021/03/PPI_Encouraging-AI-adoption-by-EU-SMEs-3.24.21-2.pdf

Wilding, M., & Boffey, D. (2025, December 10). *UK police forces lobbied to use biased facial recognition technology. The Guardian.* https://www.theguardian.com/technology/2025/dec/10/police-facial-recognition-technology-bias

Wilson, K., & Caliskan, A. (2024). Gender, race, and intersectional bias in resume screening via language model retrieval. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES). https://arxiv.org/abs/2407.20371

Wilson, K., Sim, M., Gueorguieva, A.-M., & Caliskan, A. (2025). No thoughts, just AI: Biased LLM hiring recommendations alter human decision making and limit human autonomy. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 8(3), 2692–2704. https://doi.org/10.1609/aies.v8i3.36749

Xavier, B. (2025). Biases within AI: challenging the illusion of neutrality. AI & Society, 40(3), 1545–1546. https://doi.org/10.1007/s00146-024-01985-1

Yang, Y., Lin, M., Zhao, H., Peng, Y., Huang, F., & Lu, Z. (2024). A survey of recent methods for addressing AI fairness and bias in biomedicine. Journal of Biomedical Informatics, 154,104646. https://doi.org/10.1016/j.jbi.2024.104646

Yilmaz, S., Yorgancioglu, K., & Koyutürk, M. (2025). Bias-aware training and evaluation of link prediction algorithms in network biology. *Proceedings of the National Academy of Sciences, 122*(24), e2416646122. https://doi.org/10.1073/pnas.2416646122

Appendix 1 - Interview Guide

Gender and inclusivity interview questions for SME representatives

Could you tell us about your engineering team?
Are there any considerations for diversity and representation within the team, especially in technical roles?
If yes, are there any tools or practices you are using/considering?
Have you considered gender biases in your AI products or services? If so, how?
Are there difficulties in designing AI systems that work well across different groups of people?/ What are the key issues in making AI systems perform reliably for diverse user populations?
How do you assess issues of bias and discrimination that might arise from using AI in your workplace?
(Follow-up) How do you preemptively address sexist bias that may result from using AI in your workplace?