

Week 4 SGTA

1. How can epistemic fragmentation contribute to hermeneutical injustice in the context of algorithmic profiling, according to Milano and Prunkl (2025)?*

Epistemic fragmentation involves the isolation of individuals through personalised ML algorithms designed to account for the preferences, traits and experiences of the individual. Milano and Prunkl (2025) argue that epistemic fragmentation used in digital systems from online shopping, social media, search results and advertisements contributes to hermeneutical injustice. Milano and Prunkl (2025) utilise the example of online job listings, referring to findings that setting the user's gender to 'female' on Google's AdSettings page leads to fewer advertisements for high paying jobs (Datta et al., 2015). Users who are subjected to this kind of algorithmic profiling often do so in isolation due to epistemic fragmentation, and subsequently lack the epistemic awareness to share and compare their experiences with others.

The use of epistemic fragmentation in ML algorithms and personalised systems effectively 'silos' individuals, making them less likely to question whether what they're experiencing is normal or common, and as a result less likely to consult others in their circles. This contributes to hermeneutical injustice by inhibiting epistemic awareness, preventing the formation, uptake and application of epistemic resources, and crippling epistemic infrastructure that would otherwise be present.

2. Kay et al. (2024) argue that generative AI systems such as chatbots based on Large Language Models and text-to-image generators can contribute to what they call 'amplified testimonial injustice.' In a first step, briefly summarise their account of amplified testimonial injustice. In a second step, critically discuss this account. Do you agree with it? If yes, why? If not, why not?*

'Amplified Testimonial Injustice' as discussed by Kay et al. (2024) describes the tendency for generative AI systems, by learning from large datasets containing existing societal biases, reinforce testimonial injustices against marginalised groups and subsequently undermine their credibility. This is a result of existing social biases in datasets, as well as the inherently disproportionate representation of more socially dominant groups when compared to marginalised ones. As a result, this creates a feedback loop wherein underrepresented perspectives are further suppressed and dominant ones amplified.

I agree with this perspective because it exposes a fundamental flaw in current generative AI systems: their tendency to reinforce and amplify pre-existing biases in their training data. This not only amplifies testimonial injustices but also fuels increasing skepticism about the reliability of AI outputs. Issues such as fabricated content and misinformation further challenge the credibility of generative AI models. It is essential to adopt robust bias mitigation strategies and enhance transparency throughout AI development.