The human cogs in the AI machine: Experiences of data annotation and content moderation workers in the BPO sector in India and Kenya

Uma Rani*†[1] and Morgan Williams*[1]

[1]International Labour Organization – Switzerland

Abstract

Artificial intelligence (AI) has gained a lot of attention in recent years to handle complex issues autonomously, without human intervention; this hype obscures the significant human labour that underpins AI lifecycle. This perception of autonomy is misguided given the two interrelated types of labour that are crucial for the functioning of AI systems at each stage of the AI lifecycle: algorithmic labour and data labour. Algorithmic labour consists of highly skilled and well-compensated professionals who develop, test, and maintain the algorithms. Data labour, on the other hand, contributes to creating, processing and refining datasets that feed into AI systems. Although poor quality data can have severe consequences for the reliability and effectiveness of AI systems, these data workers are often undervalued and hidden from sight of the end user. This paper examines these workers' roles to disentangle how algorithmic and data labour contribute to different stages of the AI lifecycle across four popular AI use cases: large language models, autonomous cars, content moderation, and AI medical diagnostics. Despite the essential nature of their work, data labourers do not have the same degree of job security and recognition that the algorithmic labourers have access to. To explore their working conditions, this paper draws on insights from surveys conducted with data workers in Kenya and India, which demonstrate that power imbalances embedded in the work organisation and lack of transparency contribute to diminished well-being and lack of fair compensation, regardless of whether they hold advanced university degrees. In addition, the pressure for fast, cheap data collection often leads to exploitative labour practices in developing countries, subjecting workers to long hours, low pay, intense pressure, and inadequate social protections. These practices perpetuate inequalities and impact data quality, introducing biases into AI systems and undermining their reliability. The observed differences in the working conditions between algorithmic and data labourers point to the need for a more sustainable and equitable AI ecosystem that acknowledges both types of labour.

Keywords: AI supply chains, algorithmic labour, data labour, artificial intelligence, microtask, bpo, working conditions

*Speaker

†Corresponding author: amara@ilo.org