AI discrimination also manifests during deployment. Automated decision-making systems used in insurance, welfare, or immigration disproportionately harm already marginalised groups. In the Netherlands, the Dutch Tax, and Customs Administration used algorithms to flag potential childcare benefit fraud. Factors such as "foreign-sounding names" and "dual nationality" were used as indicators, leading to thousands of racialised low- and middle-income families being wrongly accused and forced to repay benefits (Heikkila, 2022). Similar patterns are evident in AI-driven hiring tools (Sánchez-Monedero et al., 2020) and border surveillance technologies. Metcalfe et al. (2023) highlight how EU investments in border tech subject migrants to invasive AI scrutiny, while EU citizens are protected from such monitoring.

To categorise these impacts, Redden (2022) developed a taxonomy of AI-related harms: exploitation, discrimination, surveillance, manipulation, loss of privacy, exclusion and injustice. These harms stem from and reflect structural inequality. Rather than neutral or objective tools, AI systems often replicate the very power imbalances embedded in the societies that create them.

Exploited Labour, Extracted Resources, and Environmental Impact

The question of the labour that goes into machine learning and similar AI techniques has been covered extensively in earlier sections. The way in which digital labour is divided between developed and developing countries, between the global North and the global South (or global majority countries) reflects global inequalities and power asymmetries. As Casilli (2017), Gray and Suri (2019), Irani (2022), and others have shown, low skilled digital work is outsourced to countries of the global South, while well paid, highly skilled engineering work tends to take place in countries of the global North. It is further important to point out that AI applications rely on a material infrastructure, which in turn requires intensive extraction of resources to satisfy demand. Brevini (2021) argues that to understand AI's environmental impact we should include the material resources that go into the hardware that support these systems, the energy consumed, and the e-waste created.

In particular, critical minerals, essential to the production of high-tech, include rare earth minerals (a group of some 15 elements), lithium, cobalt, as well as gold and tungsten. As these resources are only found in some parts of the world, important geopolitical tensions arise. China is the world's biggest supplier of most of the rare earths leading to Western fears of dominance and supply restrictions (Kalantzakos, 2020). Other minerals, specifically tin, tantalum (coltan), tungsten, and gold, are associated with conflict, especially in Central Africa and particularly in the Democratic Republic of Congo (DRC) (Vogel, 2022). These conflict minerals which are used in the production of digital goods, including mobile phones, are extracted with the use of forced labour, have a negative environmental impact, and are traded with weapons. Despite regulations to track supply chains, Vogel (2022) shows that these measures have led to a de facto embargo on minerals from areas such as Congo, which has severely impacted artisanal miners who rely on mining for their livelihoods. Instead of reducing

violence, these policies have pushed many miners into deeper poverty and as a result pushed them into the very armed groups these policies aimed to dismantle.

Adding to these issues, the environmental footprint of AI applications is entirely at odds with the need for green policies. Brevini (2020) argues that the rhetoric around AI obscures the materiality of these systems and their reliance on a material infrastructure with an astonishing carbon footprint. In a widely cited study, Strubell et al. (2019) estimated that training a large language model emits 626,155 lbs of carbon dioxide, which is almost five times the lifetime emissions of an average car (126,000 lbs). Adding to this, the costs of data centres and cloud computing which are necessary to support AI applications, the picture that emerges is one of excessive environmental impact: according to Brevini (2023), data centres use more energy that the whole national consumption of some countries. In Ireland, which hosts a large number of data centres, their energy consumption represented 21% of all electricity consumption in Ireland in 2023. This exceeded electricity consumption in urban homes which was 18% in 2023 (Gain, 2024). The rate of energy consumption for AI is likely to increase, since Google and other companies have already indicated that they plan to increase their data centres' capacity; Google in particular reported that its carbon emissions had grown by nearly 50% compared to 2019 (Gain, 2024), and this is almost entirely due to AI needs.

Given the serious environmental impact of AI, it is important to identify mitigation strategies that reduce or eliminate energy consumption and waste. But it is equally important to ask whether it is worth pursuing some of these applications, especially if they are associated with harm. The question of AI harms and fairness has led to proposals for the creation of an ethical framework for the development of AI applications. This approach is discussed next.

Addressing AI Harms: Ethics or Politics?

The negative impact, actual and potential, of AI systems has driven advances in ethics as a potential means for preventing and redressing harms and ensuring fairness and trustworthiness. In 2018, when advances in AI and machine learning were evident, the European Commission established a high-level expert group on artificial intelligence, tasked with providing an overall ethical framework for the development of AI applications. In parallel, corporations that develop AI systems began establishing ethics departments to internally check and steer the production of these systems in an ethical manner. Finally, academic work on AI ethics has sought to develop normative approaches on how AI ought to be developed in order to ensure fairness. Yet, despite these efforts, critics have argued that these ethical approaches remain superficial (Birhane et al., 2022), and that corporate efforts are engaged in 'ethics-washing' rather than actually dealing with the ethical implications of the AI they develop. It may be therefore that we need a political rather than an ethical approach to create a binding framework for AI.

The European Commission considers that AI technologies are part of the future, and is looking to ensure that Europe produces ethical, secure and cutting-edge AI. To help