



BUSINESS MODEL RED FLAGS UPDATE:

BRINGING AN AI & DATA LENS TO BUSINESS MODEL RISKS TO PEOPLE

SUMMARY

The past five years have witnessed remarkable growth in the development and deployment of Artificial Intelligence (AI), particularly since the emergence of generative AI systems. While the widespread uptake of AI technologies has the potential to drive innovation, productivity gains, scientific advancement, and economic growth, it also presents significant potential risks for people, especially workers, young people and other vulnerable communities.

Many of the human rights risks presented by AI are not new. However, the rapid pace of AI development and adoption has **heightened these risks** and provided **new examples of how they manifest in practice** both within and outside the technology sector. We have also seen the emergence of new risks of harm to people associated with the growing role of AI across business activities and value chains in an ever-increasing number of industries.

With support from the Generation Foundation, Shift has been updating and expanding its Business Model Red Flags tool, a resource which helps companies, investors and regulators identify how business models can put human rights at risk. In 2025 Shift published an additional [thematic lens on climate change](#), which added further information to the Red Flags about how features of a company's business model can interact with climate change to create risks for people. The latest update introduces an "AI and Data lens." The original version of the Red Flags, published in 2021, touched upon impacts associated with the collection and use of data. The updated version places additional focus on how the incorporation of AI technologies, including generative AI, in the business's value proposition, value chain or revenue model, can deepen existing or create new risks.

This update takes place at a time of increasing regulatory expectations on companies to assess and address sustainability impacts connected to their business models and strategies. The European Union's Corporate Sustainability Reporting Directive (CSRD) and Corporate Sustainability Due Diligence Directive (CSDDD) both include requirements for companies to consider how their business strategies, business models, purchasing practices and design choices interact with human rights and environmental impacts and related business risks.

INTRODUCTION

The AI landscape has undergone a significant transformation since Shift published the Business Model Red Flags in 2021. The original version identified several ways in which business model features linked to overall digitalization¹ of the economy were already giving rise to adverse impacts on people's human rights. The first release of the tool included dedicated red flags focused on automation, algorithmic decision-making, platform economies, and both the collection and monetization of data.

Following the release of large language models and the emergence of generative AI tools, the use of AI across the economy, in our workplaces and in our everyday lives, has expanded dramatically. This has led to a rapid growth in the development and deployment of AI in both the “tech” and “non-tech” sectors, which has intensified existing and created new risks to people. This introduces novel considerations about the governance of AI as it becomes increasingly integrated into value creation in the private sector.

At the same time, these technologies can and should be harnessed for positive outcomes and to tackle some of the world's greatest shared challenges. AI tools are being used to accelerate scientific and medical research, including in areas such as [drug discovery](#), and may help improve the efficiency and resilience of [energy systems and climate modelling](#). They also have the potential to enhance productivity and broaden access to information and services.

However, **actual or forecasted benefits like these do not off-set adverse impacts on people**. Moreover, failure to manage those adverse impacts may itself lead to a loss of social license for AI systems, resulting in a chilling effect on legitimate innovation, and creating legal, financial and reputational risks for AI developers, users and investors.

MAKING THE BUSINESS MODEL CONNECTION

Across both the technology sector and more traditional industries, a common pattern emerges: AI often increases the scale, speed and opacity of activities and decisions, some of which affect people's lives, resulting in negative human rights outcomes.

This update of the Business Model Red Flags explores how the **development, deployment and integration of AI technologies is intensifying - and in some cases reshaping - the ways in which existing business model features can give rise to risks to people**.

Because AI and digitalization impact all aspects of the economy, 14 of Shift's original business model red flags have been updated (see Annex). The examples below illustrate how some of these dynamics are playing out both within the technology sector and across a growing range of traditionally “non-tech” sectors.

SHIFT'S BUSINESS MODEL RED FLAGS

First published in 2021, Shift's Business Model Red Flags supports the identification of human rights risks inherent in features of business models – that is, elements of a company's value proposition, value chain and revenue model that give rise to risk of impacts on people. The Red Flags have been used or referenced by a wide range of actors, including investors, lenders, companies, UN bodies, foundations, industry associations and civil society organizations. The tool contains examples of due diligence questions and mitigation measures, reflecting that business models needn't necessarily be abandoned, but that risks need to be addressed, and important questions asked of and by the company's leadership.

¹Digitalization refers to the use of digital technologies and data as well as interconnection that results in new or changes to existing activities. Examples of digital technologies include artificial intelligence (AI), cloud computing, blockchain, and biometrics. Source: [OECD](#)

(I) THE DEVELOPMENT AND DEPLOYMENT OF AI IN THE TECHNOLOGY SECTOR

EVOLVING REGULATORY LANDSCAPE

Regulatory expectations on companies are continuing to expand worldwide. On business models and strategies specifically, the EU's Corporate Sustainability Due Diligence Directive (CSDDD) requires companies to make necessary modifications to their own business plan, overall strategies and operations, including purchasing practices, design and distribution practices, to address impacts on people arriving as a result of their operations, products or services². Meanwhile, under the European Sustainability Reporting Standards³, companies are asked to disclose information about the key elements of their general strategy and business model that relate to, or affect, the topics associated with its material impacts, risks and opportunities.⁴ Whilst continuing to evolve, these standards and regulatory developments underscore the increasing imperative of companies understanding the connections across business models, digitalization and human rights, and taking appropriate action.

The development and deployment of AI is intensifying several business model-related risks already present within the technology sector. In particular, the race to develop increasingly capable AI systems is increasing pressure to innovate and deploy products at speed, sometimes before adequate safeguards, oversight mechanisms, or avenues for remedy are in place. At the same time, the substantial infrastructure required to support the rapid uptake of AI technologies is accelerating investment in data centers and other digital infrastructure, with implications for workers, communities and the environment. These developments are creating new pressures on land, water and energy resources. In addition, the technology ecosystem of multiple private sector actors now mirrors the level of complexity seen in the global value chains of other industries. This, again as we have seen in other industries, comes with challenges for even well-meaning companies to build and use leverage with other entities to address impacts on people, as well as for affected people to hold AI developers and deployers accountable for harms.

By way of illustration Shift's new version of the Business Model Red Flags includes the following updates:

- Business models that place a premium on **speed** can **create incentives to bring products and services to market before adequate safeguards are in place** (see [Red Flag 11](#)). Competition among technology companies developing and deploying increasingly powerful AI systems has accelerated product development cycles, which can **undermine adequate attention to governance, oversight, safety and responsible testing**. This can have implications for the health and safety of AI value chain workers, including data annotators and content moderators. Developing and deploying projects at speed can also limit stakeholders' ability to organize and influence, query or seek remedy for adverse human rights impacts, including those on downstream consumers, such as risks to the rights to privacy or non-discrimination
- Further, **platforms**, such as social media, are increasingly AI-enabled, which is increasing the speed, scale and personalization of **harmful content, harassment, misinformation** and other unsafe online conduct (see [Red Flag 6](#)).
- Business models in the tech sector can also put stress, often cumulative, on the **availability of natural resources** (see [Red Flag 13](#)). The explosive growth of AI is spurring enormous demand for computational resources. Digital infrastructure, such as **data centers**, is intensifying pressure on and competition for land, water and energy resources with implications for local communities.

²The EU's Corporate Sustainability Due Diligence Directive, Article 10.2 calls on companies to take appropriate measures to prevent potential adverse impacts where relevant including (d) making "necessary modifications of, or improvements to, the company's own business plan, overall strategies and operations, including purchasing practices, design and distribution practices". And Article 11.3 discusses bringing actual adverse impacts to an end and also requires companies to take appropriate measures, where relevant, including (d) making "necessary modifications of, or improvements to, the company's own business plan, overall strategies and operations, including purchasing practices, design and distribution practices."

³The European Sustainability Reporting Standards (ESRS) operationalize the reporting requirements of the EU's Corporate Sustainability Reporting Directive (CSRD), setting out detailed disclosures that companies must make on sustainability impacts, risks, and opportunities.

⁴Draft ESRS 2 General Disclosures (November 2025), SBM-1 "The undertaking shall disclose ... information about the key elements of its general strategy and business model that relate to, or affect, the topics associated with its material impacts, risks and opportunities".

- The energy consumption fueling data center infrastructure expansion is positioning the industry to become a significant contributor to global greenhouse gas (GHG) emissions, unless there are substantial energy efficiency gains and residual energy demand is met through low-carbon sources. The emergence of a significant new driver of global greenhouse gas emissions at the precise moment that we are meant to be reducing emissions, risks jeopardizing global climate targets and brings potentially far-reaching consequences for people and communities worldwide, especially **those most vulnerable to physical climate impacts** (see [Red Flag 25](#)).
- Another update made to this resource connects to the human rights due diligence implications for so-called “low leverage” value chain relationships (see [Red Flag 15](#)). The development and deployment of advanced AI systems increasingly rely on complex networks of commercial relationships involving cloud providers, data suppliers, developers, deployers and enterprise customers. While these arrangements can accelerate innovation and facilitate specialization, they may **reduce visibility of impacts on people and create incentives for actors to view responsibility for identifying, preventing and addressing those impacts as residing elsewhere in the value chain**.
- At the same time, the substantial infrastructure, data and compute requirements associated with advanced AI systems is **concentrating influence** among a relatively small number of actors, increasing dependency on a limited set of providers and potentially reducing visibility into, and leverage over, risks to people across the value chain (see [Red Flag 15](#)).

As reflected in our resource, understanding i) how the AI value chain (or eco-system) is structured, ii) where influence sits, and ii) how parties to contracts attempt to allocate responsibility for impacts (and how this (mis) aligns with their responsibility under the international standards), are considerations that companies and their financiers must incorporate into their due diligence.

(II) THE INCREASING INTEGRATION OF AI-ENABLED PRODUCTS AND SERVICES ACROSS MULTIPLE SECTORS OF THE ECONOMY

AI is increasingly being embedded into products, services and business processes across a broad range of industries. AI-enabled systems are being used to make or support decisions affecting access to employment, finance, healthcare and public services; to curate and personalize online experiences; to manage workers and allocate tasks; and to develop products with dual-use applications. The result is that AI is changing the nature of many existing business models and, by extension, the related risks to people.

For example, many private sector industries (alongside public entities) now increasingly rely on **algorithmic decision-making** to allocate opportunities, assess risk, and determine access to products and services. For example, AI-enabled hiring, lending and insurance systems may replicate or amplify historical biases in underlying data, affecting individuals’ access to jobs, credit and other opportunities. This can amplify the potential for biased, discriminatory, or otherwise harmful outcomes affecting individuals and communities (see [Red Flag 5](#)).

Our update addresses the growth of **AI-enabled dual-use applications** (see [Red Flag 9](#)) such as surveillance, profiling and targeting. For example, AI-enabled decision-support systems used by the military may accelerate targeting processes in ways that undermine meaningful human control, increase both the scale and speed of operations and heighten **risks to civilians**. As with traditional dual-use technologies, there are significant risks to people when these AI-enabled products are sold to and/or deployed by end users that do not respect international human rights norms or international humanitarian law. Further, the increasing sophistication, scalability, and autonomy of these technologies can heighten the severity of impacts on affected individuals and communities, particularly where **meaningful human oversight over decisions affecting life, liberty, privacy, and security is reduced** (see [Red Flag 9](#)).

Rapid integration of AI across diverse industry sectors is also increasing the volume, granularity, and commercial value of data collected about individuals. As organizations seek to collect, combine and analyze **personal and behavioral data to train AI systems**, they may do so in ways that **undermine privacy**, exceed the purposes for which data was originally provided, or increase risks of **discrimination, surveillance, profiling** and other adverse human rights impacts (see [Red Flag 16](#)). For example, organizations may combine data from multiple sources to infer characteristics about individuals that were never explicitly disclosed. These insights may then be used to shape decisions affecting employment, pricing or access to services, creating risks of **discrimination, manipulation and surveillance**. Privacy and data protection regulations, including the EU's General Data Protection Regulation (GDPR), have established important safeguards governing the collection, use and sharing of personal data. Nevertheless, AI systems increase both the demand for and value of data, companies and investors need to remain attentive to how data practices may give rise to risks to people, including where legal compliance alone may not fully address broader human rights concerns.

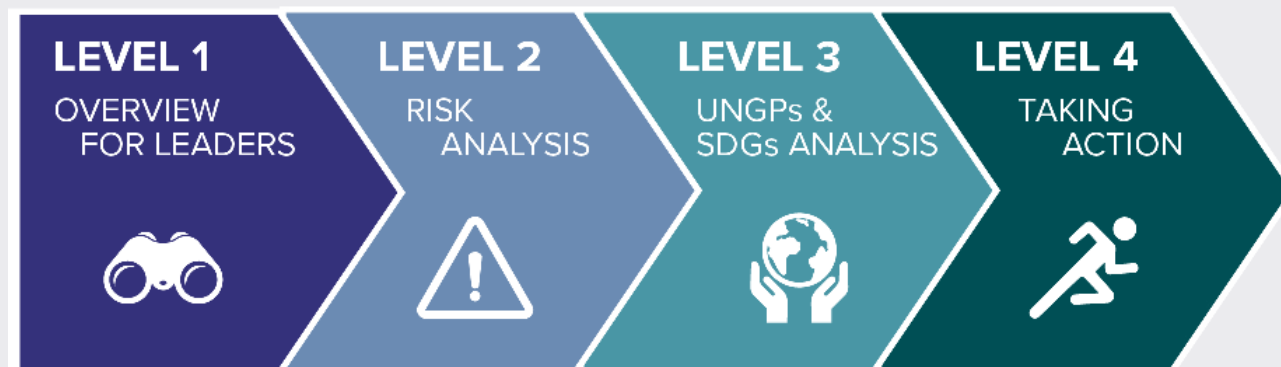
AI systems are also increasingly being used to **allocate tasks, monitor performance, determine compensation, and manage access to work** within the logistics and services sectors. These so-called algorithmic management systems can reduce workers' visibility into how decisions and disciplinary actions are arrived at while negatively affecting the predictability of work, earnings, and scheduling. The updated Red Flags include examples of cases where AI management systems have put individuals' mental and physical health at risk while also making it more difficult to understand, challenge, or seek remedy for adverse impacts (see [Red Flag 2](#) and [Red Flag 17](#)).

The uptake and adoption of AI technologies across sectors is also reshaping labor markets more broadly. For example, the adoption of AI may **automate** tasks previously performed by workers, alter skills requirements across occupations and reduce demand for some categories of labor, creating transition risks for affected workers and communities. While this may create new opportunities, it also risks **accelerating job displacement and labor market disruption** if workers are not adequately supported to retrain and upskill as part of the digital transition (see [Red Flag 21](#)).

As AI-enabled products and services become increasingly integrated across the economy, greater attention to the business model features that give rise to adverse impacts will be essential to understand and address **root causes**, and, ultimately, to improve outcomes for people.

A TOOL FOR ACTION: HOW TO USE THIS UPDATE

The updated Red Flags retain the features of the original Business Model Red Flags. The flags are organized around three features of a business model: The Value Proposition, The Value Chain and the Cost Structure and Revenue Model. Each Red Flag is supported by a guidance document organized into four levels:



As part of this update, users can explore, at each of the four levels, how digitalization interacts with dominant or emerging business models. This includes an analysis of how people are at risk, examples of where these issues have manifested as financially material risks for business, due diligence lines of inquiry, examples of companies that have made efforts to mitigate or prevent these risks, and the inclusion of business model and climate-specific resources.

Red Flag 6: Providing online platforms with potential for online and offline harm

- **AI & Data Lens:** AI-driven recommendation, ranking and personalization systems have become the core operating logic of many online platforms, optimizing content delivery for wide engagement at speed and scale. For example, recommendation systems may prioritize content that maximizes user engagement, even where doing so increases users' exposure to harmful, misleading or polarizing content. The same systems can be used to micro-target individuals based on data collected and inferred characteristics, including characteristics that might make them particularly vulnerable to potentially harmful online activities.

AI also enables the rapid creation and broad dissemination of content, some of which can be harmful (e.g. non-consensual intimate imagery, deepfakes, disinformation). The inherent opacity of these systems and how they are deployed can limit the ability of users, regulators and affected communities to understand how content is curated and amplified, to challenge the choices made, and to deliver effective remedy for victims when harm occurs.

- **Risks to people:** Online platform users, particularly more vulnerable groups such as women, children, adolescents or those with mental health challenges, face heightened risks. These risks include exposure to content that contributes to mental health harms (including body image issues, eating disorders, self-harm and suicidal ideation), exploitation, harassment, and discrimination, as well as to disinformation that can undermine participation in public life.
- **Risks to business:** In March 2026, a New Mexico jury found Meta liable under the state's Unfair

Practices Act for misleading consumers about the safety of Facebook and Instagram and for endangering children, ordering it to pay \$375 million in civil penalties. "The evidence presented at trial ... established that Meta's design features enabled pedophiles and predators to engage in child sexual exploitation on Meta's platforms." A second-phase trial began in May 2026, with the state seeking approximately \$3.7 billion in abatement costs and court-ordered changes including the redesign of Meta's recommendation algorithms (so they no longer prioritize constant engagement), effective age verification, and the appointment of an independent monitor.

- **Due diligence lines of inquiry:** How does the company seek to prevent the posting and spread of harmful content and/or misinformation? Does it engage in content moderation, and how does it strike a balance between preventing harmful content and potential censorship?
- What are the processes the company has in place to engage with civil society and other experts to remain aware of the potential impacts on people of their platforms, and to explore any dilemmas that may arise in seeking to mitigate those risks?
- If the platform(s) include AI-enabled functionality, how does the company ensure that AI solutions are developed and deployed with human rights protections in mind, such as human oversight, transparency in decision making and privacy protections?
- Do affected users and groups have meaningful access to information about how content is curated, recommended or removed, and to effective remedy where harms occur?

Investors and lenders can use this tool to:

- build the business model analysis capacity of relationship and portfolio managers
- screen portfolios to identify high-risk segments and prioritize investees for deeper due diligence
- design and strengthen portfolio company engagement
- build more resilient portfolios by recognizing and addressing the ways in which business model-related risks to people interact with digitalization-related risks
- inform double materiality assessments and sustainability reporting on relevant topics

HOW FINANCIAL INSTITUTIONS HAVE USED RED FLAGS

Investors and lenders have used the Red Flags alongside other tools developed by Shift to drive better evaluation and engagement on human rights. For example, FIs have: incorporated the Red Flags into their social risk identification tools; screened their portfolios by mapping the Red Flags against industry classification codes; mapped the red flags and potential mitigation measures against commercially available data; leveraged the tool's due diligence lines of inquiry to design engagement strategies; and identified root causes of portfolio controversies to proactively identify similar risks across the portfolio. See Examples of Investor and Lender Application of Red Flags [here](#).

Companies can use this tool to:

- inform double materiality assessments and sustainability reporting on human rights topics
- identify opportunities for business model, product and service innovation
- support more resilient and strategic decision-making

Regulators, auditors, analysts and civil society organizations can use this tool to:

- identify business model features that will predictably put people at risk
- engage meaningfully with companies and investors on the interactions between digitalization and business model-related risks to people
- assess company efforts to address risks related to the interaction between business model features and AI/digitalization

With generous funding from the Generation Foundation, Shift has been working to update and expand its popular Business Model Red Flags tool. In September 2025, Shift released 13 updated and 1 new Red Flag that have an important climate-related dimension. This latest tranche of 14 Red Flags explores the ways in which business model features interact with AI – including through the collection and use of data - to create or intensify risks to workers and communities. There are also 5 red flags to which both the Climate Lens and AI & Data Lens have been added where business model features intersect with both the digital and climate transition.

To help financial institutions address the human rights performance of their clients and investees, including how human rights risks manifest in business model features, Shift regularly convenes and facilitates social sustainability clinics. In the months ahead, Shift will convene clinics on [inequality](#) and the [nexus between climate, nature and people](#), as well as [training for measuring human rights performance](#).

ANNEX: UPDATED BUSINESS MODEL RED FLAGS: AI & DATA LENS

The business' commercial success substantially depends upon:		How does the AI & Data lens interact with business models to intensify or create new risks to people and business
RELATED TO THE VALUE PROPOSITION		
2	High-speed delivery that places pressure on warehouse workers and logistics workers in the "last mile"	Digital and algorithmic management systems are increasingly being used in last-mile logistics to allocate work, optimize routes, monitor performance, and evaluate workers at scale. While these technologies can improve efficiency, they may also increase work intensity, reduce worker autonomy, expand worker surveillance, and heighten risks to workers' health, safety, privacy, and access to effective grievance mechanisms.
5	Using or providing algorithmic systems that make predictions, recommendations or decisions that can affect people's access to rights and resources	As companies digitalize their operations and integrate AI across business functions, algorithmic systems increasingly intermediate access to essential resources and determine who receives credit, employment, healthcare, housing, and public services in ways that can create or amplify existing adverse impacts on people at scale.
6	Providing online platforms for individuals to interact where use of the platform can lead to human rights harms	As platforms integrate AI-driven content curation and monetize user data to maximize engagement, the risks of harm to people are intensified and scaled, amplifying risks of harassment, exploitation, and discrimination while eroding users' ability to understand or contest the systems shaping their experience.
8	Offering products that when used at high levels can affect peoples' health, particularly when targeted at vulnerable populations	The design and marketing of digital products and services, such as social media or online gaming/gambling, are increasingly underpinned by AI systems and behavioral analytics optimized to drive higher levels of use. The design and marketing choices at play in these optimizations can lead to harm to people, particularly those who are most vulnerable, such as children, adolescents and people grappling with mental health challenges.
9	Offering products or services with the potential for harmful forms of use, especially by end-users that are inclined to pursue those uses	The increasing accessibility of AI and autonomous weapons technologies amplifies risks, as dual-use capabilities can reach state and non-state actors inclined to deploy them in ways that violate international humanitarian law, undermine human control over the use of force, and cause civilian harm at a scale and speed that forecloses meaningful human oversight or accountability.
10	Products that harm when used as intended	Advances in AI are increasing the feasibility and deployment of weapons systems that remove meaningful human judgment from decisions about the use of force. Where products are designed to operate without sufficient human control, understanding or predictability, risks to civilians may become inseparable from the intended functionality of the product itself.
RELATED TO THE VALUE CHAIN		
11	Speed in developing products or services or delivering projects, with risks to health and safety	- Across various sectors, the rapid emergence of AI technologies has sped up their practical application in various AI-enabled products and services, which are being developed and deployed at a remarkable pace. This rapid deployment has left, limited, if any, time for effective human rights due diligence, including meaningful stakeholder engagement. Where speed in product innovation and deployment is prioritized over effective safeguards, risks to people increase as AI technologies and AI-enabled solutions are developed and released without: (i) adequate assessment of and mitigations for their impacts on rights (e.g., rights to health, privacy, and non-discrimination), or (ii) sufficient oversight and remedy measures in place. - The rapid surge in data center construction, driven by AI demand, is intensifying health and safety risks for workers, as compressed schedules, skilled labor shortages, and inadequate training heighten exposure to hazards.

ANNEX: UPDATED BUSINESS MODEL RED FLAGS: AI & DATA LENS

13	Depleting natural resources or public goods such that it undermines access to health	AI growth is driving a steep rise in demand for the data centers that fuel it, as well as for the land on which these data centers are situated (and for the considerable energy and water they consume). This can intensify competition for land, energy and water resources, creating risks for communities whose access to these resources may be reduced or whose environments may be adversely affected.
15	Structuring business relationships in ways that limit influence to address risk to people	Complex AI ecosystems can create assumptions that responsibility for impacts on people sits elsewhere in the value chain, while growing dependence on a small number of providers of data, compute and infrastructure concentrates influence and can reduce visibility into, and leverage over, risks to people across the value chain.
16	Collecting, holding or monetizing data about customers or users in ways that lead to adverse impacts on privacy, and other human rights.	Increased digitalization and use of Internet of Things, AI, facial recognition and other digital solutions, lead to a substantial increase in the creation and use of personal data, which increases the risk of that data being used in ways that impact on people's rights, including their rights to privacy and non-discrimination, as well as a range of civil and political rights.
25	Activities, products and/or services that significantly contribute to cumulative greenhouse gas emissions and the resulting physical climate change impacts that negatively affect people's rights	The rapid expansion of AI infrastructure is driving significant increases in energy demand and land use for data centers, which depending on the energy mix, can accelerate greenhouse gas emissions and contribute to land conversion that undermines natural carbon sequestration.

RELATED TO THE COST STRUCTURE & REVENUE MODEL

17	Labor relationships that are structured to avoid costs that come with formal employment arrangements	Digitalization is further intensifying the impacts being felt by workers in the gig economy. For example, pay and access to work are increasingly determined by algorithmic systems that often lack human oversight, transparency and due process making it more difficult to understand or challenge decisions affecting workers' livelihoods. At the same time, data-driven surveillance and algorithmic management increase pressure on workers and worker insecurity, while weakening bargaining power.
21	Rapid digitalization of processes and key functions such that planning or support for upskilling or redeployment of displaced employees is challenging to achieve	If not managed appropriately, the simultaneous economic transformations of decarbonization and digitalization have the potential to expand the range of impacts on people. For example, in response to "green skills" shortages employers may choose to bypass workforce training and upskilling by deploying digital and AI solutions.
22	Incentive structures that are designed to maximize sales or minimize costs, but which also create risk to the health and welfare of customers or patients.	AI-driven healthcare or financial platforms may advise patients/customers or make recommendations or decisions on their behalf. Commercial incentives embedded into these platforms may lead to advice that prioritizes revenue generation, cost reduction or efficiency over customer or patient needs.