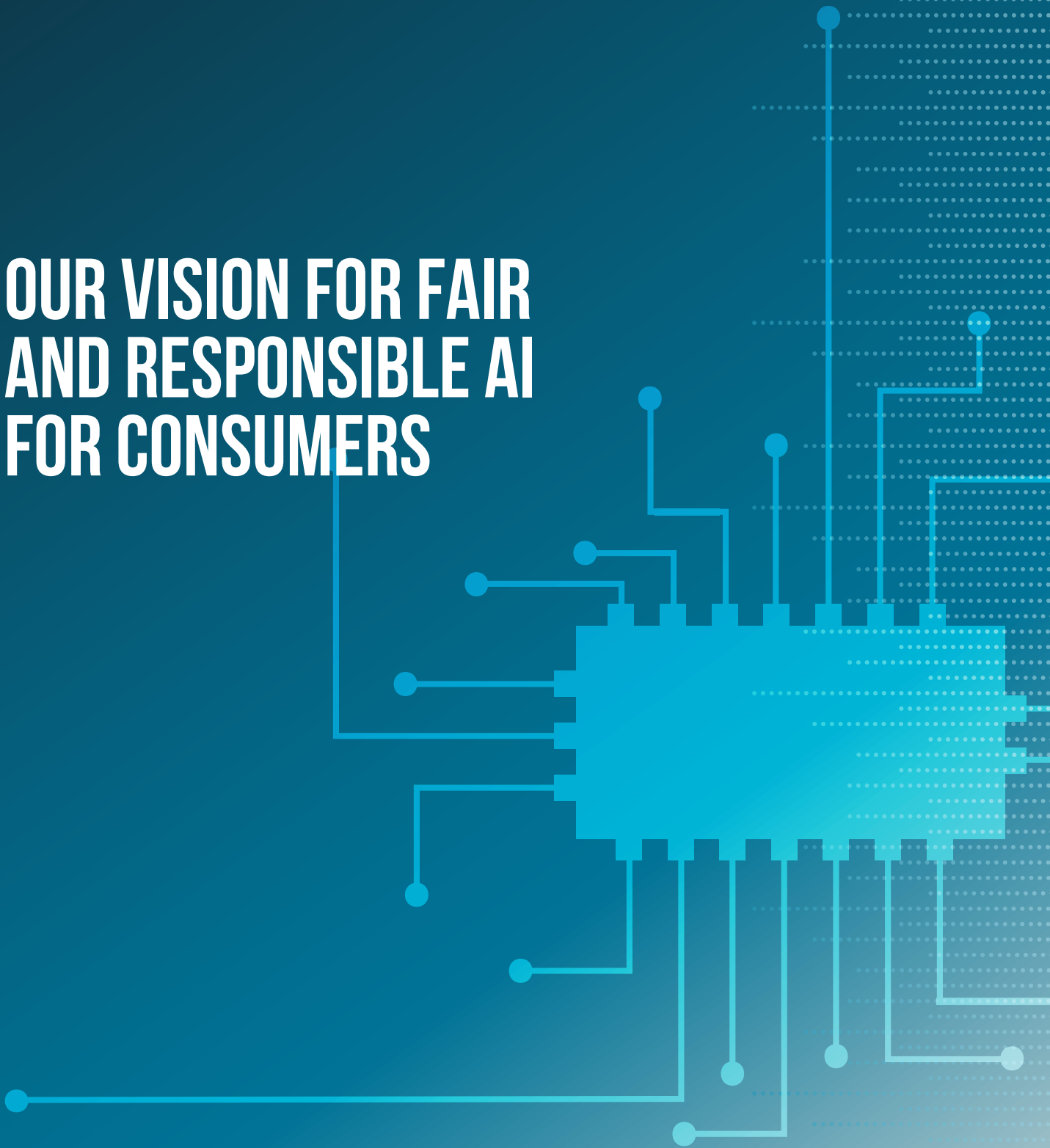
CONSUMERS INTERNATIONAL

COMING TOGETHER FOR CHANGE

# OUR VISION FOR FAIR AND RESPONSIBLE AI FOR CONSUMERS

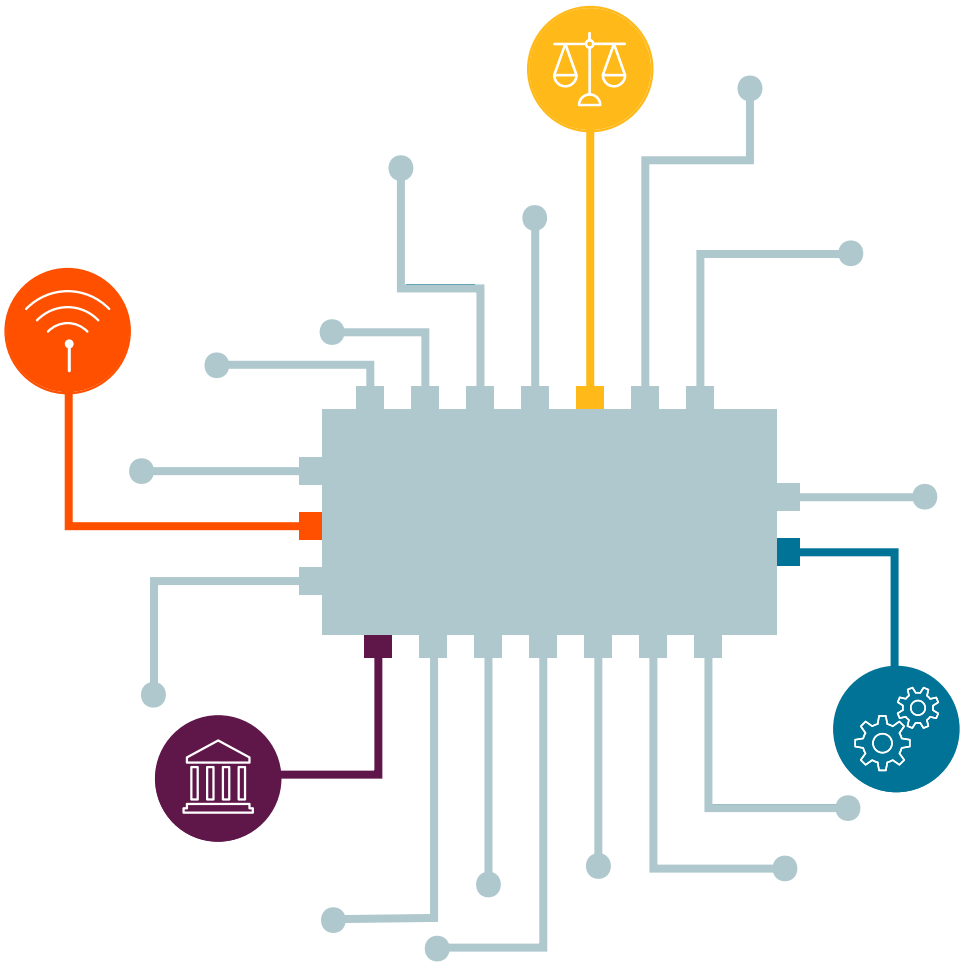


# OUR VISION

Fairness and responsibility in the development and deployment of artificial intelligence (AI) is possible. But it must be more than an aspiration – it is essential.

Convening for World Consumer Rights Day 2024, and as the global voice of consumers to international decision-makers for over 60 years, Consumers International has outlined four priority areas for achieving fair and responsible AI. These drive toward a vision of **digital markets that are truly open and accessible, support high benchmarks for consumer protection, feature inclusive and representative governance frameworks, and maintain the guarantee of redress and representation for consumers.**

We now call on governments to take steps to ensure these priorities are developed and upheld, and provide a set of actions necessary to achieving fair and responsible AI for consumers that must be taken by developers and deployers of commercial generative AI.



# ARTIFICIAL INTELLIGENCE IS AT AN INFLECTION POINT FOR SOCIETY

The past few years have witnessed significant changes in the development and deployment of artificial intelligence. The rollout of generative AI models represents a paradigm shift in the capabilities of AI and a major transition from task-specific applications to versatile tools with diverse use cases.

Proponents of generative AI argue that will significantly change how people interact with technology, and compare it to major innovations like the internet, smartphones and electricity. Their belief is that these models have expanded the horizons of what AI can achieve and point to individual instances of success as evidence. For example, a study by BenevolentAI[1] used generative AI to identify potential treatments for amyotrophic lateral sclerosis (ALS). Results like these have led to high expectations for the step change that generative AI will[2] bring to society, accompanied by equally high estimates of its monetary value.

Sceptics underline that generative AI is built on statistical models[3], and outputs are probability based calculations that attempt to reproduce patterns in those models. Unlike humans, generative AI has no concept of truth or lived experience of the world. Critics caution against the narrative of technological determinism[4] and question whether we should even want generative AI to be applied in such a wide range of fields.

Even as these contrasting viewpoints play out, concerns have emerged quickly and clearly. Due to the computational resources involved in building a generative AI model, developers represent a highly concentrated market. The field is dominated by a few major players, raising questions about representation, accountability and the equitable distribution of benefits.

Regulatory frameworks and accountability mechanisms lag the development of the technology. Governments worldwide are grappling with varying approaches. The United States (US), European Union (EU) and China have each adopted different stances on regulation and policy design. The EU's proposed AI Act[5] emphasises fundamental rights and ethical considerations, while China prioritises economic growth and national security in its AI regulations[6]. The US, in contrast, relies more heavily on industry self-regulation[7]. Urgent calls for action have also been made at the global level, including by the United Nations Secretary-General. International frameworks are in development, including the Council of Europe's Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law[8].

The proliferation of various industry initiatives or coalitions have been hard to track; they range from widely recognised – such as the Partnership on AI[9] – to the niche, like the Association for Computing Machinery's Public Policy Council[10]. In parallel, some businesses have advocated for and formulated their own AI principles and imposed self-constraints that they police voluntarily. For example, Google's

1 https://blog.petrieflom.law.harvard.edu/2023/03/20/how-artificial-intelligence-is-revolutionizing-drug-discovery/
2 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier
3 https://www.heinz.cmu.edu/media/2023/July/generative-ai-is-a-math-problem-left-unchecked-it-could-be-a-real-problem
4 https://link.springer.com/article/10.1007/s43681-022-00148-6
5 https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
6 https://eastasiaforum.org/2023/09/27/the-future-of-ai-policy-in-china/
7 https://www.washingtonpost.com/technology/2023/12/06/ai-policy-eu-act-us-regulation/
8 https://www.coe.int/en/web/artificial-intelligence/cai
9 https://partnershiponai.org/paper/making-ai-inclusive-4-guiding-principles-for-ethical-engagement/
10 https://www.acm.org/binaries/content/assets/public-policy/principles-generative-ai.pdf

AI Principles[11] claim to craft AI in a socially beneficial manner, while Microsoft has refrained from creating or deploying facial recognition technology until robust safeguards are established.

The effectiveness of self-regulation is questioned by civil society and other experts. Concerns are that such measures lack enforcement mechanisms and transparency, and that they encourage companies to engage in 'regulator shopping'[12], leading to a race to the bottom for legislation protecting consumers. Others emphasize the need for robust regulation[13] alongside industry initiatives to ensure responsible development and deployment of generative AI[14].

# FRAGMENTED GOVERNANCE EXPOSES CONSUMERS TO MORE RISK

The speed and scale at which AI is developing, combined with a fragmented governance approach, makes protecting the rights of consumers paramount. The expansion of generative AI applications has widened the pool of consumers that can benefit from the technology, but also exposed more people to risks.

For generative AI, multiple examples of consumer harm have already emerged. These range from the almost comical – like the airline chatbot that invented a refund policy[15] and communicated it to a consumer seeking advice – to the intensely personal – such as the use of generative AI to produce deepfake non-consensual sexual imagery[16]. Demonstrations of how generative AI has been applied to compromise websites, create content reinforcing gender and racial stereotypes, as well as many other problems, form the basis of hundreds of documentations[17].

No technology comes without risk, but consumers must be able to trust that risks they are exposed to have been properly scrutinised and mitigated. Trust in digital environments relies on several factors, including the accuracy and verifiability of information people find there. The implications of low trust extend beyond mere inconvenience: consumers may base decisions on misinformation or skewed perspectives, potentially leading to adverse outcomes, manipulation or exploitation, as well as mental or physical harm. Trustworthiness is also dependent on transparency[18] as a mechanism to highlight how and why flaws exist in systemic design, which is needed to uphold accountability for harm.

11 https://ai.google/responsibility/principles/
12 https://srinstitute.utoronto.ca/news/tech-self-regulation-democratic-oversight
13 https://www.freiheit.org/sub-saharan-africa/stakeholders-call-robust-ai-regulations-africa-amid-concerns-over-ethical
14 https://sloanreview.mit.edu/article/are-rai-programs-prepared-for-third-party-and-generative-ai/
15 https://arstechnica.com/tech-policy/2024/02/air-canada-must-honor-refund-policy-invented-by-airlines-chatbot/
16 https://www.404media.co/inside-the-ai-porn-marketplace-where-everything-and-everyone-is-for-sale/
17 See https://incidentdatabase.ai/apps/discover/?is_incident_report=true&s=generative%20AI and https://www.aiaaic.org/aiaaic-repository
18 https://arxiv.org/abs/2310.12941

There is a growing body of research into the trustworthiness of generative AI. Efforts to gauge the transparency of foundation models19 underpin the principle that trust is dependent on transparency. Some studies have looked at the extent to which large language models (LLMs)[20] invent information that sounds convincing but is false, as well as whether disclosure in human-to-generative AI interactions[21] can improve trust. Others have focused on ways to verify output accuracy and quality[22], or the comprehensiveness of the information cited[23]. There have been many attempts to "jailbreak" models[24] with the intention of identifying vulnerabilities before release. The field is building as fast as the technology.

# THE VALUE OF A CONSUMER LENS ON GENERATIVE AI

Research to provide evidence of harm is useful, and regulators should feel empowered that they are not starting with a blank canvas. Speaking at the Consumers International 2023 Global Congress[25], United States Federal Trade Commissioner Rebecca Kelly Slaughter argued that regulators and enforcement agencies in many countries already have the tools and remedies at their disposal to investigate and uphold and breaches of consumer protection laws. Principles for safe and responsible AI put forward by the Norwegian Consumer Council[26] also underline that consumer law is a logical and practical starting point to "future proof" policy development. The Consumer Policy Research Centre, however, notes[27] that for regulators to feel empowered to enforce, they must have appropriate powers to investigate, restrict and apply remedies to those using AI to harm consumers, even unintentionally.

The United Nations Guidelines for Consumer Protection (UNGCP)[28], housed at the United Nations Conference on Trade and Development (UNCTAD), are another critical foundation that can be adapted and applied to different challenges. They include the idea of a fair market where access to essential goods and services is universal; the need for consumer safety and empowerment; a right to privacy and reliable, free-flowing information; the need for adequate education and awareness; and, critically, clear procedures for dispute resolution, including routes to redress and representation.

19 https://arxiv.org/abs/2310.12941
20 https://arxiv.org/abs/2401.11452
21 https://arxiv.org/abs/2401.12773
22 https://arxiv.org/abs/2402.06414
23 https://arxiv.org/abs/2304.09848
24 https://arxiv.org/abs/2310.06987
25 https://www.consumersinternational.org/consumers-international-global-congress-2023/programme/
26 https://storage02.forbrukerradet.no/media/2023/06/generative-ai-rapport-2023.pdf
27 https://cprc.org.au/report/in-whose-interest/
28 https://unctad.org/topic/competition-and-consumer-protection/un-guidelines-for-consumer-protection

While the guidelines offer a robust framework for designing consumer protections across sectors, they are deliberately broad and are not intended to fully capture the intricacies and complexities specific to a sector. We recognise that to comprehensively tackle the issues of generative AI the guidelines must be complemented by additional, more direct regulation working in tandem.

As the global voice of consumers to international decision-makers for over 60 years, Consumers International has outlined an approach to do so. We have developed four priority areas, which combine the UNGCP with a set of actions needed by developers and deployers of commercial generative AI to protect consumers. We now call on governments to work with Consumers International, UNCTAD and other relevant stakeholders to ensure these priorities are developed and upheld.

## PRIORITY AREAS FOR FAIR AND RESPONSIBLE GENERATIVE AI

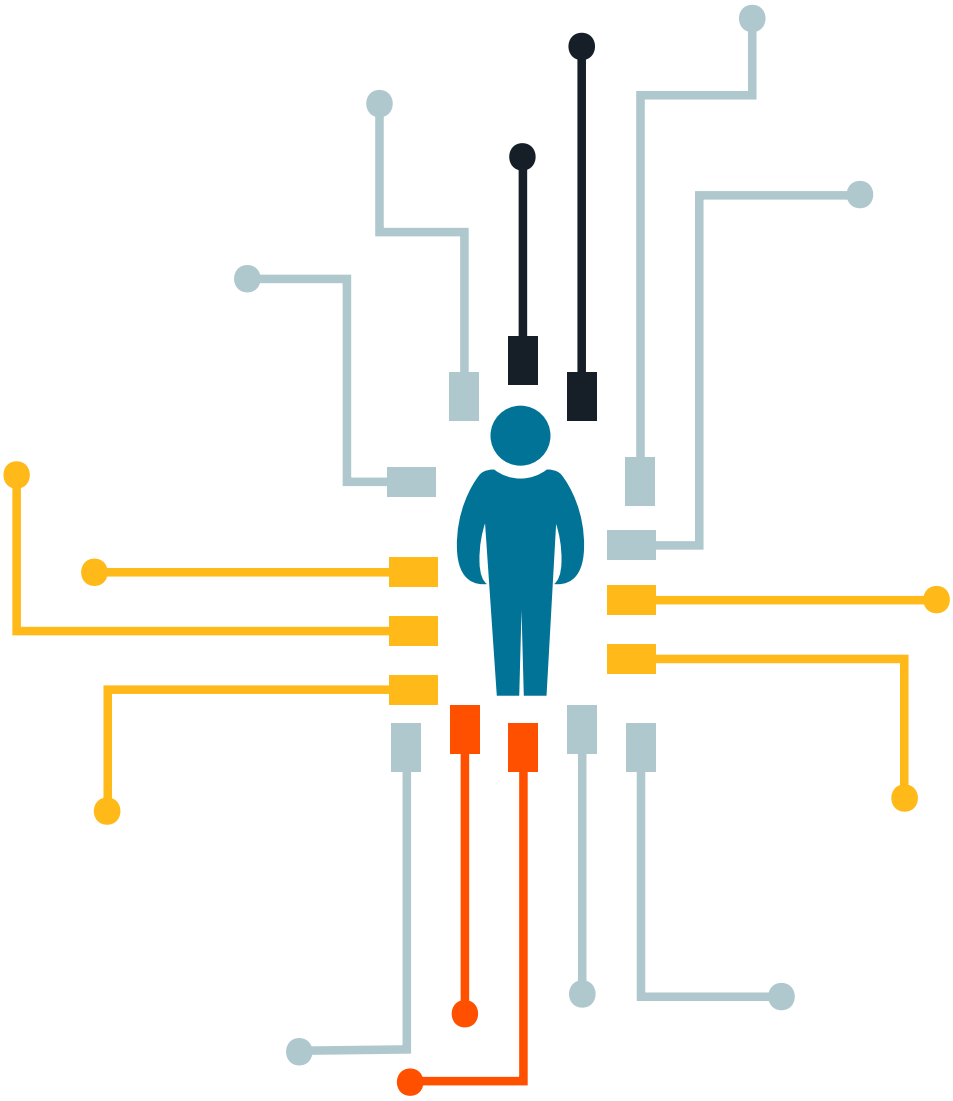
1. **Transform digital markets to make them open and accessible for all** – through expanded and culturally relevant access points; privacy respecting uses of data; affordable and meaningful connectivity; and a healthy and sustainable information ecosystem.
2. **Establish and maintain high benchmarks for consumer protection** – through the creation and application of stringent and globally consistent procedures that safeguard people from harm; and continuous and independent monitoring and evaluation of trustworthiness and transparency of commercial developers and deployers of AI.
3. **Develop inclusive and representative governance frameworks** – by disclosing protocols for training data and model design; including trusted consumer advocates in their development; and devoting adequate resources to maintaining them.
4. **Guarantee that redress and representation is available, respected and enforced** – by establishing clear and transparent processes to report harms; ensuring that rights to appeal are meaningful and fair; sharing information with consumer protection authorities when risks are identified; and building representation mechanisms into systems where consumers are impacted.

| Priority areas | Actions for developers and deployers |
|---|---|
|  **Transform digital markets to make them open and accessible for all** | • Support the advancement of public interest and truly open technology, the principles behind it and the organisations that are developing it<br>• Provide greater transparency of how a generative AI model is built – covering the human labour and oversight involved in sourcing and training the data, the protections afforded to those workers, and the energy use and environmental impacts of the model<br>• Take immediate and robust steps to identify and mitigate the use of personal information from consumers in a dataset used to build a generative AI<br>• Exclude personal information from datasets used to train generative AI without informed, express and time-bound consumer consent<br>• Implement stringent and globally consistent procedures to protect consumers from hazards to their health and safety, including but not limited to exposure to harmful content, disinformation and the threat of manipulation<br>• Offer increasing and proportionate degrees of transparency, including sharing information with consumer protection authorities when significant risks are identified and support them in accessing the information they need to protect the public interest |
|  **Establish and maintain high benchmarks for consumer protection** | • Rigorously evaluate and ensure the trustworthiness of generative AI meets standardised criteria accepted by trusted consumer advocates, and do not put on the market any generative AI models that do not meet the thresholds<br>• Devote adequate resources to designing, maintaining and upgrading those criteria in collaboration with consumer advocates and others in civil society, and publish and update benchmarks on a regular basis<br>• Provide a mechanism for independent third party testing and verification of trustworthiness<br>• Document known weaknesses and flaws that impact consumers and develop mitigations for these with trusted consumer advocates<br>• Have a robust due process to swiftly remove or suspend accounts or content using generative AI models to deliberately manipulate or harm consumers<br>• Exclude data sourced from data brokers in generative AI models |
|  **Develop inclusive and representative governance frameworks** | • Disclose – through standardised reporting practices – the protocols for selecting, augmenting and curating the data used in generative AI models, especially where these concern vulnerable and disadvantaged consumers, and publish an impact assessment where these consumers are concerned<br>• Develop processes that can be assessed by independent third parties for assessing the risk of a generative AI model to society |
|  **Guarantee that mechanisms for redress and representation are available** | • Provide clear guidelines and routes for consumers to report instances of harm, and use this data to inform and improve consumer protection processes<br>• Ensure that consumers have a clear and transparent process to object, receive an explanation and seek redress when generative AI is used to make a decision that significantly impacts them<br>• Ensure that consumers have the right to collective redress and to be represented by consumer organisations and other civil society groups when exercising their rights |

CONSUMERS INTERNATIONAL
COMING TOGETHER FOR CHANGE

# BUILDING FAIR AND RESPONSIBLE AI – WITH AND FOR CONSUMERS

Achieving fair and responsible AI for consumers will take sustained effort. It is critical that its development involves – and is informed by – the voice of consumers. Work to highlight the consumer experience of AI is underway already; initial insights from Consumers International are available here. We invite others to join our more than 200 members around the world in building fair and responsible AI.



# ACKNOWLEDGEMENTS

This vision for 'Fair and Responsible AI for Consumers', together with the accompanying research report the 'Consumer Experience of Generative AI', was led by Consumers International and its Digital Consumer Rights programme team. Consumers International is very grateful to the Ford Foundation for funding the programme, which supports the Consumers International Digital Task Force. Several members of the task force advised this work.

Gratitude is also due to the 35 Consumers International members that contributed to the exercise with generative AI chatbots, as well as the 100+ organisations that engaged in our World Consumer Rights Day activities.

Consumers International is thankful for the work of its staff in producing this vision:

- Charlotte Broyd, Head of Communications and Membership
- Stefan Hall, Director, Digital Innovation and Impact
- Hollie Hamblett, Policy and Research Specialist
- Helena Leurent, Director General
- Oarabile Mudongo, Digital Rights Specialist
- Grace Ramsay, Global Communications Coordinator
- Javier Ruiz Diaz, Senior Digital Rights Advisor