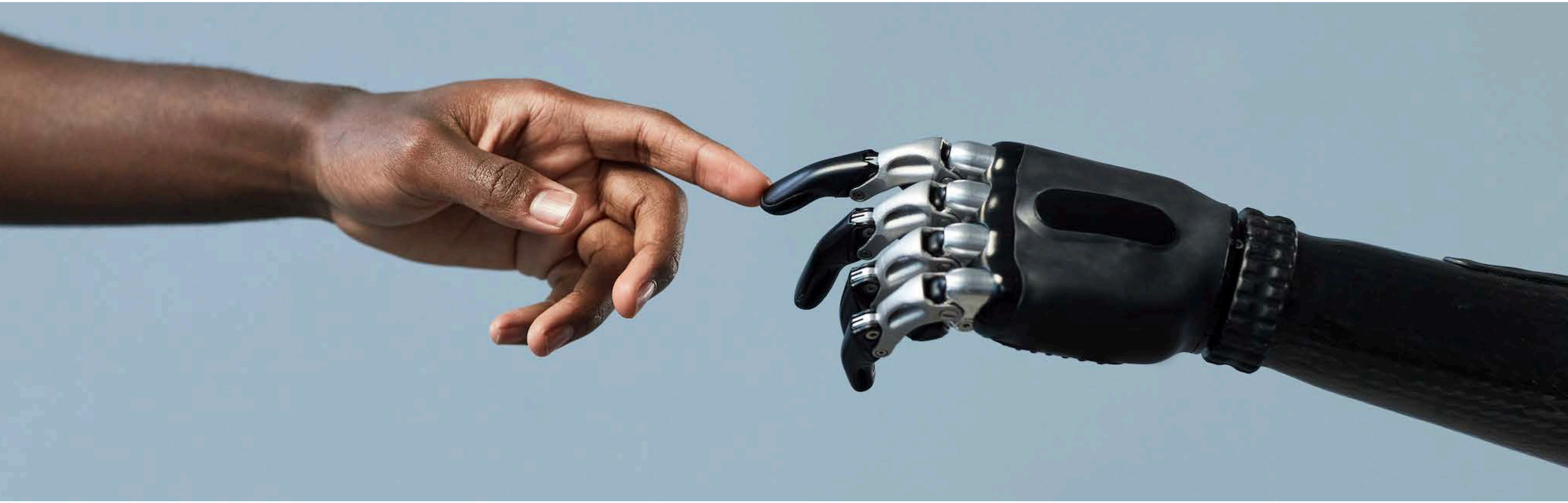


RESPONSIBLE AI

Learn more about Intel's orientation towards the responsible use, design & development of AI capabilities

intel ai

Responsible AI



According to the World Economic Forum, RAI (Responsible Artificial Intelligence) is the practice of designing, building, deploying, operationalizing and monitoring AI systems in a manner that empowers people and businesses and impacts customers and society equitably. RAI should cover both the ethical (social) and technical dimensions of AI, adhering to principles of validity and reliability, safety, fairness, security and resilience, accountability and transparency, explainability and interpretability, and privacy.

The applicability of these principles is context dependent. For example, when AI is used in healthcare, reliability and safety might be of utmost importance, whereas when it is used in hiring processes, fairness could be the key concern.

The challenge lies in finding the right balance between these principles within a given context, which often involves making conscious trade-offs.[1]

Seven Principles of Intel Responsible AI

Intel has identified the following list of seven areas of ethical inquiry that have been integrated into Intel's AI product lifecycle processes. These principles serve as a strong foundation for considering the risks associated with AI products and provide a north star that we at Intel, put into action through our Ethical Impact Assessment process[2].

Intel RAI



RESPECT HUMAN RIGHTS



ADVANCE SECURITY, SAFETY & RELIABILITY



DESIGN FOR PRIVACY



ENABLE HUMAN OVERSIGHT



PROMOTE EQUITY & INCLUSION



ENABLE TRANSPARENCY & EXPLAINABILITY



PROTECT THE ENVIRONMENT

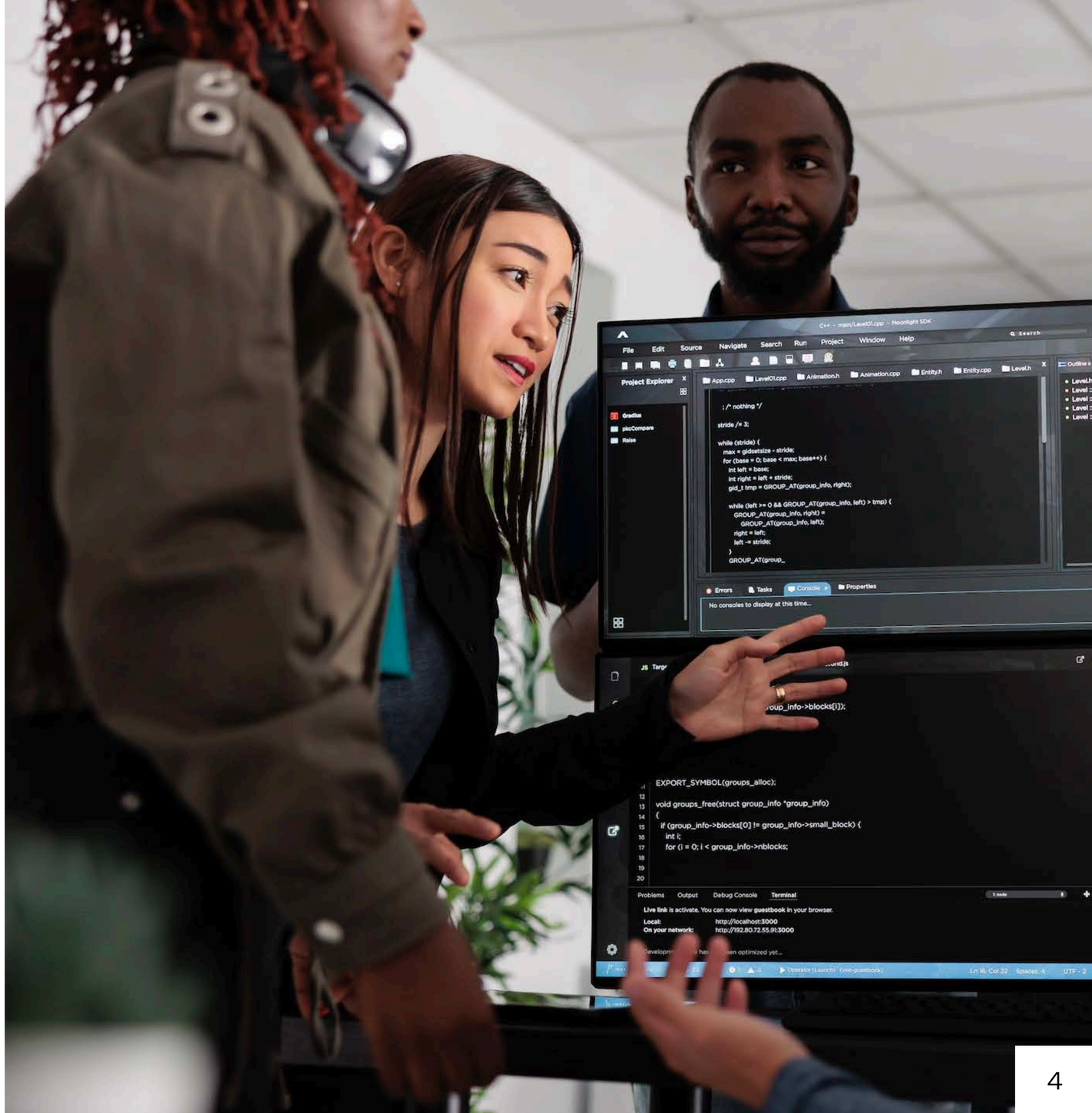
Human Rights

As we enter 2024, now is the time to not only ensure that AI systems are rights-respecting by design, but also to guarantee that those who are impacted by these technologies are meaningfully involved in decision-making on how AI technology should be regulated.

Whilst AI developments do present new opportunities and benefits, we must not ignore the documented dangers posed by AI tools when they are used as a means of societal control, mass surveillance and discrimination. All too often, AI systems are trained on massive amounts of private and public data–data which reflects societal injustices, often leading to biased outcomes and exacerbating inequalities[3].

At Intel, we approach this in alignment with our Global Human Rights Principles and Approach and relevant international frameworks such as the United Nations Guiding Principles on Business and Human Rights and Organization for Economic Cooperation and Development Guidelines for Multinational Enterprises on Responsible Business Conduct[2].

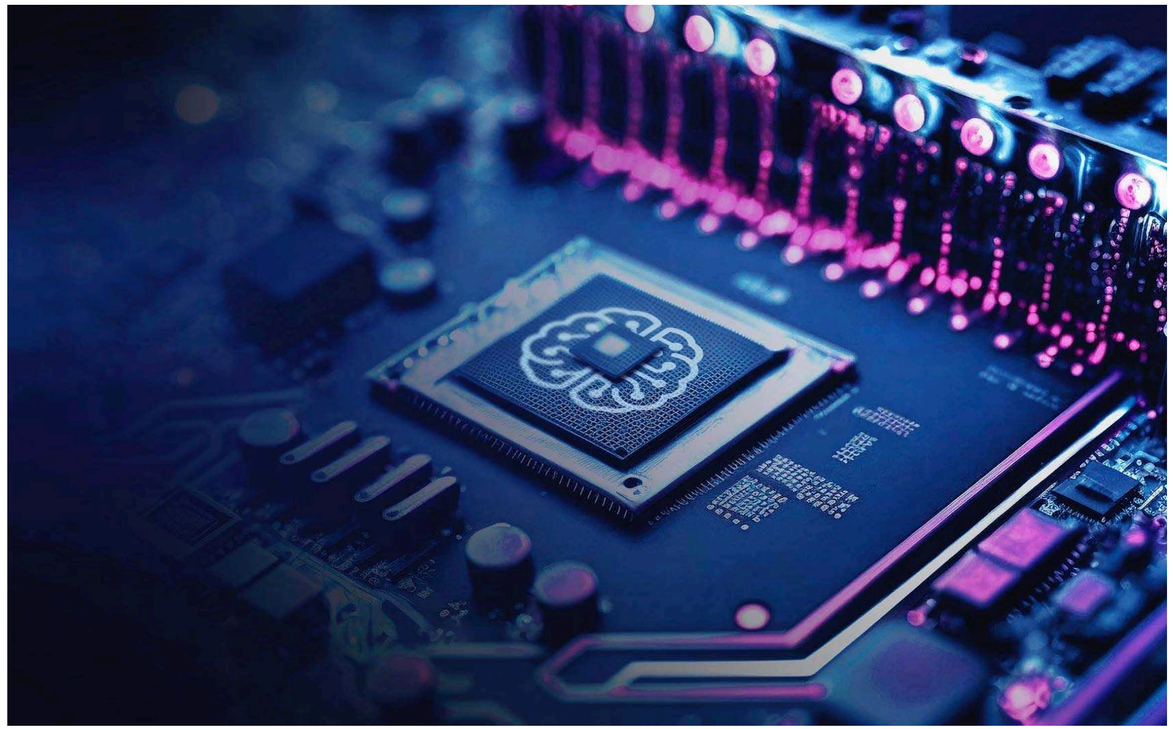
We have also started working with our supply chain responsibility organization to accelerate progress in addressing the human rights of global AI data enrichment workers - those who perform tasks such as data annotation, data cleaning, and conduct a human review of algorithmic outputs[4].



Human Oversight

Lets look at the approach taken in the EU AI Act - The goal of human oversight is to prevent or minimize risks to health, safety, or fundamental rights that may arise from using these systems. The oversight measures should match the risks and context of the AI system's use.

The AI system should be provided in a way that allows the overseer to understand its capabilities and limitations, detect and address issues, avoid over-reliance on the system, interpret its output, decide not to use it, or stop its operation. For certain high-risk AI systems, any action or decision based on the system's identification must be verified by at least two competent individuals.[5]



MINIMISE RISKS OF GENERATIVE AI

OpenAI's ChatGPT* created a frenzy in the generative AI landscape by achieving record-breaking growth. It attracted over one million users within its first week. While big tech companies like Google, Microsoft, and Facebook are competing in the Large Language Model (LLM) race, small startups are also making headway. A growing problem seems to be balancing the secrecy needed for competitive advantage with the transparency needed for safety.

LLMs like ChatGPT have demonstrated impressive capabilities, the emergence of generative AI models raises concerns about potential harmful effects. Whilst the ongoing debates over large vs. small models and open vs. closed systems are important, we must recognize that performance and accuracy are not the only considerations.

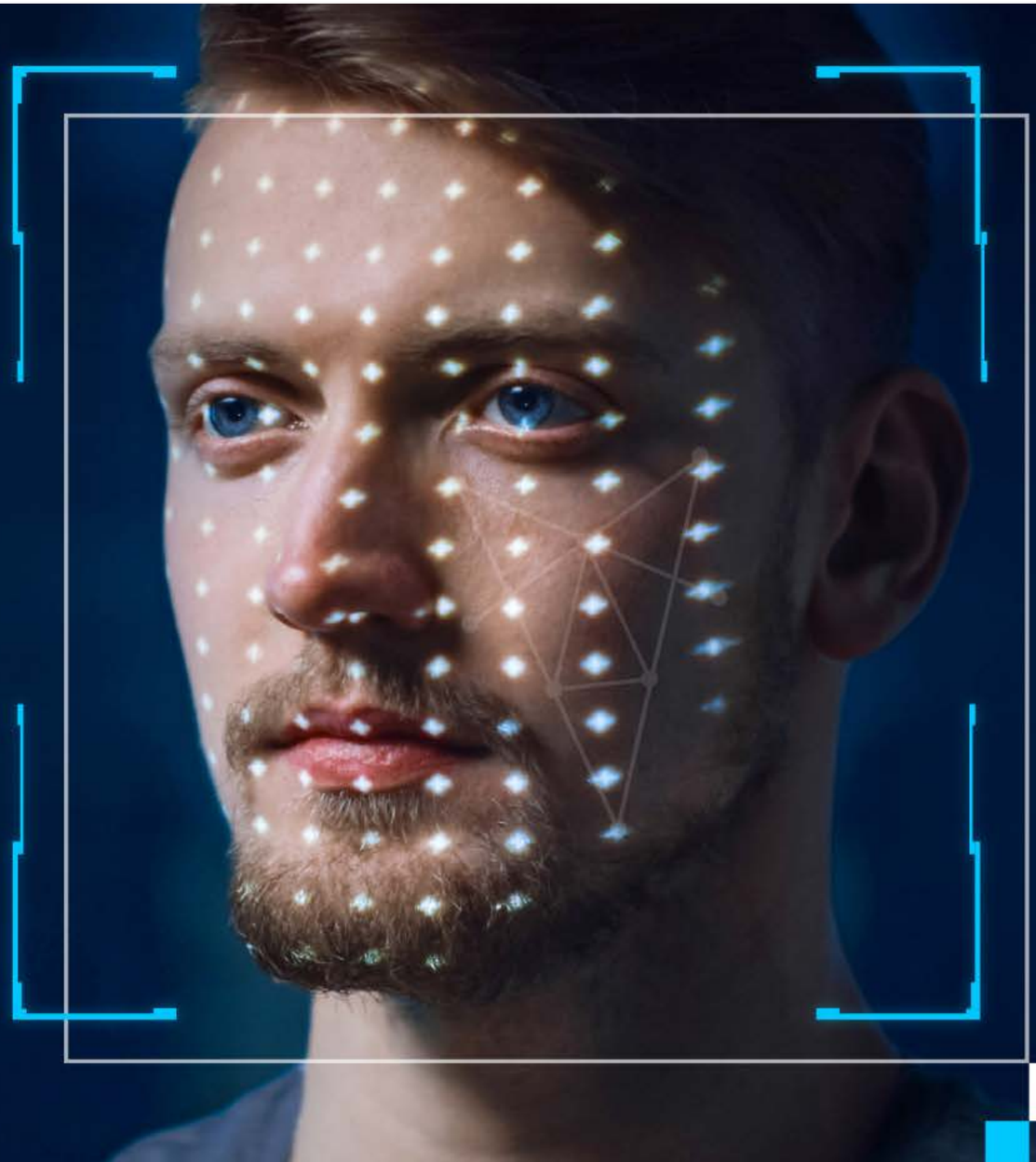
Factors such as fairness, explainability, sustainability, privacy, etc. must also be considered. Upholding responsible AI practices will ultimately determine the societal value of AI[6].

Human oversight is key to minimising the risk of Generative AI. Reason being, these models learn from vast amounts of data available on the internet, which can contain biases present in society and may inadvertently perpetuate and amplify these biases. LLMs can be manipulated to generate or spread misinformation, phishing emails, or social engineering attacks. Malicious actors can intentionally train models with biased or false information, leading to the dissemination of misleading content on a large scale. Such models can be used to create convincing deepfake video and audio content. For example, an AI-generated hoax of an explosion at the Pentagon went viral.

LLMs can often have "hallucinations" and generate inaccurate information, which can be particularly problematic in industries like healthcare, where models can influence diagnostic and therapeutic decisions and potentially harm patients. Even though AI hallucinations are a known phenomenon, people continue using LLMs and uncritically accepting their pronouncements[6].

Enable Transparency

To promote responsible development and use throughout the AI value chain, developers should strive to provide comprehensive information, including things like: recommended uses, potential harms, how AI systems were trained and tested, training sets and the results of bias testing[2].



One example is how we have productized FakeCatcher, a technology that can detect fake videos with a 96% accuracy rate. In fact, Intel's deepfake detection platform is the world's first real-time deepfake detector that returns results in milliseconds[7].

Security, Safety & Reliability

Intel prioritizes security, safety, resistance to tampering, and reliability in the development of AI products. We strive to limit the application of Intel AI products to their intended use. Intel utilizes "security by design" development principles consistent with our Security First Pledge and Cybersecurity Public Policy in addition to the "safety by design" development principles consistent with our commitment to Product Quality and Reliability[2].

Artificial Intelligence and big data have the potential to revolutionize communities by unlocking insights and efficiencies that were previously hard to even imagine. This powerful pairing can, however, prove a nightmare from a security standpoint. The future of minimizing these risks starts with hardware and travels up to the cloud.

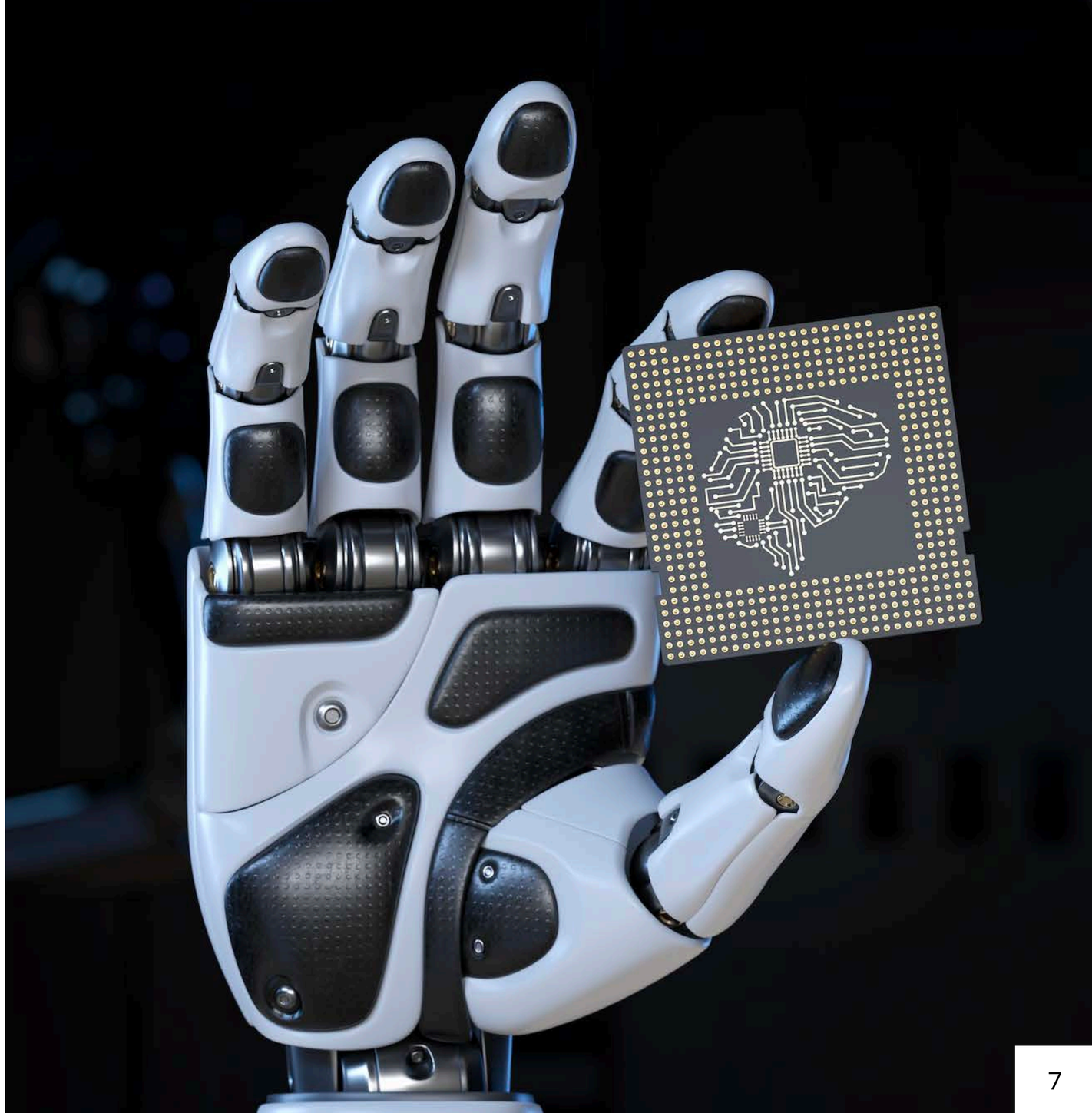
Confidential computing enhances data security and privacy. Intel's vision for AI is to democratize accessibility & provide industry specific solutions[8].

INTEL® SGX

For instance, Intel® SGX is a security feature that helps protect data in use through application isolation technology. It creates secure enclaves that are fully encrypted, allowing the application to communicate directly with the CPU to encrypt and decrypt data while bypassing any intermediary layers. This protects the data from being examined or used by other code, including potentially malicious operating systems or hypervisors[8].

PREDICTION GUARD

Another example is Prediction Guard, a startup company operating at the forefront of AI integration & a strong partnership with Intel. Prediction Guard is realizing greater speed, security, scalability, and reliability, as well as reduced costs, using Intel® Gaudi® 2 AI accelerator processors on the Intel® Tiber™ Developer Cloud. The technology is providing a 2x throughput increase for some large language models, for example, allowing the company to support clients using LLMs for a variety of innovative GenAI applications[9].



Promote Equity & Inclusivity

Our path to create a more inclusive and responsible workplace, industry and world is embedded in our purpose. Diversity and inclusion are instrumental in driving innovation and delivering strong business growth[10].

Whether building stand-alone products or working with customers and partners to bring new AI capabilities into the world, Intel is committed to building inclusivity into every step in the value chain. Intel strives to look at all aspects of AI with an inclusion lens, from the diverse backgrounds of developers in accordance with our Diversity and Inclusion policy, to datasets, to models, to intended and unintended uses. We take action to mitigate potential biases and communicate with stakeholders to make it easier for them to do the same[2].

BRINGING AI EVERYWHERE

Intel believes that responsibly bringing "AI Everywhere" is key to the collective advancement of business and society. This belief is the foundation of Intel's digital readiness programming, working to provide access to AI skills to everyone, regardless of location, ethnicity, gender or background.

We were proud to expand our AI for Youth and Workforce programs to include curriculum around applied ethics and environmental sustainability. Additionally, at Intel's third-annual AI Global Impact Festival, winners' projects went through an ethics audit inspired by Intel's multidisciplinary process[11].

Intel uses industry collaboration, co-engineering, and open-source contributions to accelerate software innovation. We also deliver a steady stream of open-source code and optimizations for projects across virtually every platform and usage model[12].



Protect The Environment



AI can serve as a sustainability technology: offering an unprecedented opportunity to accelerate innovation toward net zero and mitigate the impacts of climate change. We have to consider the threat that increased demand for AI compute will have on the simultaneous increase in energy requirements. By applying AI thoughtfully, you can expedite progress toward an organization's decarbonization goals and open new opportunities for greater efficiency and growth[13].

REDUCE ENERGY & WASTE

Apply AI to your existing processes to identify and eliminate resource inefficiencies. This will create greater visibility into your resource consumption and asset utilization, helping not only identify areas for improvement but also automatically detect anomalies to help prolong the life of equipment, minimize waste, and more.

GAIN INSIGHTS TO OPTIMIZE OPERATIONS

Discover opportunities to increase operational efficiency and ensure you're making the most of your available resources. Apply automation and predictive decision-making to increase product quality, right-size engagement across your value chain, and accelerate continuous process improvement reducing waste.

CREATE A TECHNOLOGY POSITIVE TRANSFORMATION

Drive new, sustainable value, finding resiliency and differentiation through AI. Increase your organization's climate resilience and unveil opportunities for new, environmentally-aware offerings that can help your customers in their sustainable transformation[13].

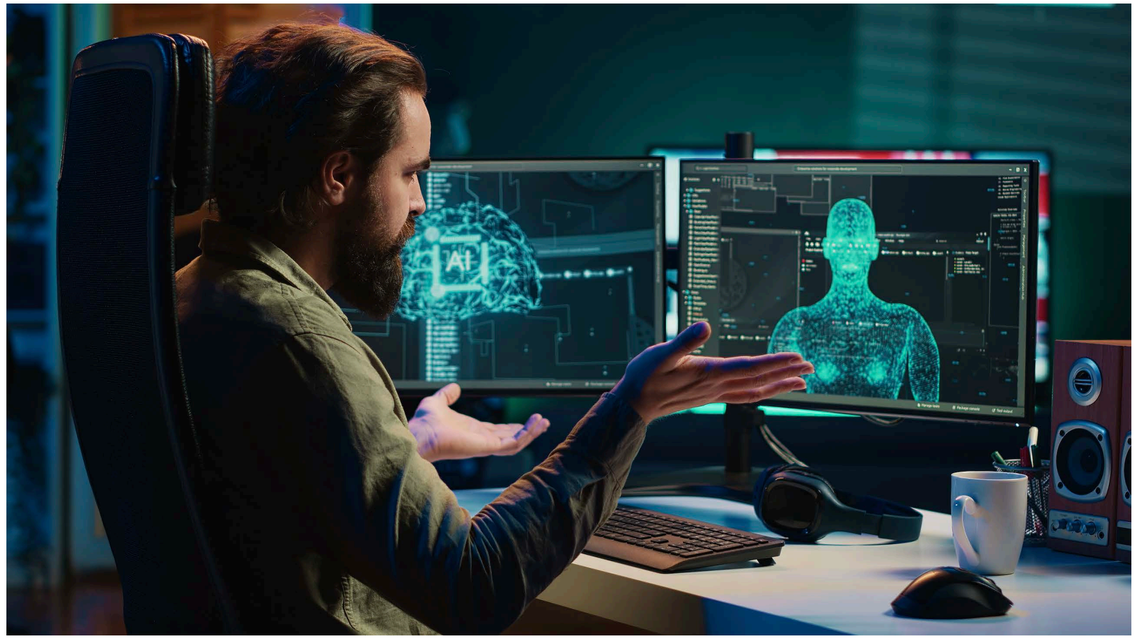
Design for Privacy

AI applications utilize large amounts of data, so respecting and safeguarding privacy and data rights throughout the lifecycle must be prioritized. Consistent with Intel's Privacy Notice, Intel supports privacy rights by designing our technology with those rights in mind, including being transparent about the need for any personal data collection, allowing user choice and control; and, to design, develop and deploy our products with appropriate guardrails to protect personal data[2].

Microsoft

Intel and Microsoft joined forces to establish the Confidential Computing Consortium (CCC), which includes cloud providers, hardware vendors, and solution providers. The group's goal is to establish industry standards for terminology and technology and educate organizations on protecting data in use, both in the cloud and at the edge.

Encryption and Transport Layer Security (TLS) have been available for years, but processing data remained a gap until we introduced new hardware, Drury says. Intel has partnered with Microsoft to deploy this hardware in Azure*, baking encryption keys into the CPU for added security. You can also bring your own customer-managed keys to unlock data during processing. This revolutionizes cloud computing, making it more secure for sensitive data usually kept on-prem. Azure has a high bar for data protection, building on the standard set by the Confidential Computing Consortium[8].



Regulating AI

Countries around the world are launching governance protocols through either legislation or via issued guidance. As our understanding of AI and its capabilities improve, guidance is a moving target with reference to a rapidly changing technology landscape and difficult to lock down specifics today for what might change in the very near future.

THE 2024 EU AI ACT

The EU AI Act establishes a comprehensive regulatory framework for responsible AI development, empowering the European Commission to continually review and update the law with the goal of keeping up with the latest innovations."

The EU uses a risk based approach for categorising AI systems such as in the graphic to the right[15].

EU AI ACT RISK CLASSIFICATION



Organisations can prepare for this legislation using the following 4 steps as recommended by PwC

1 | IMPACT ASSESSMENT	2 | GAP ANALYSIS	3 | OPTIMIZE PROCESSES	4 | IMPLEMENT CHANGES
Ascertain whether AI systems are affected by the regulation.	Identify areas that need adjustments & changes & develop a roadmap.	Drive compliance within organisation across value chain.	Deliver all changes & regularly revisit to review & evaluate.



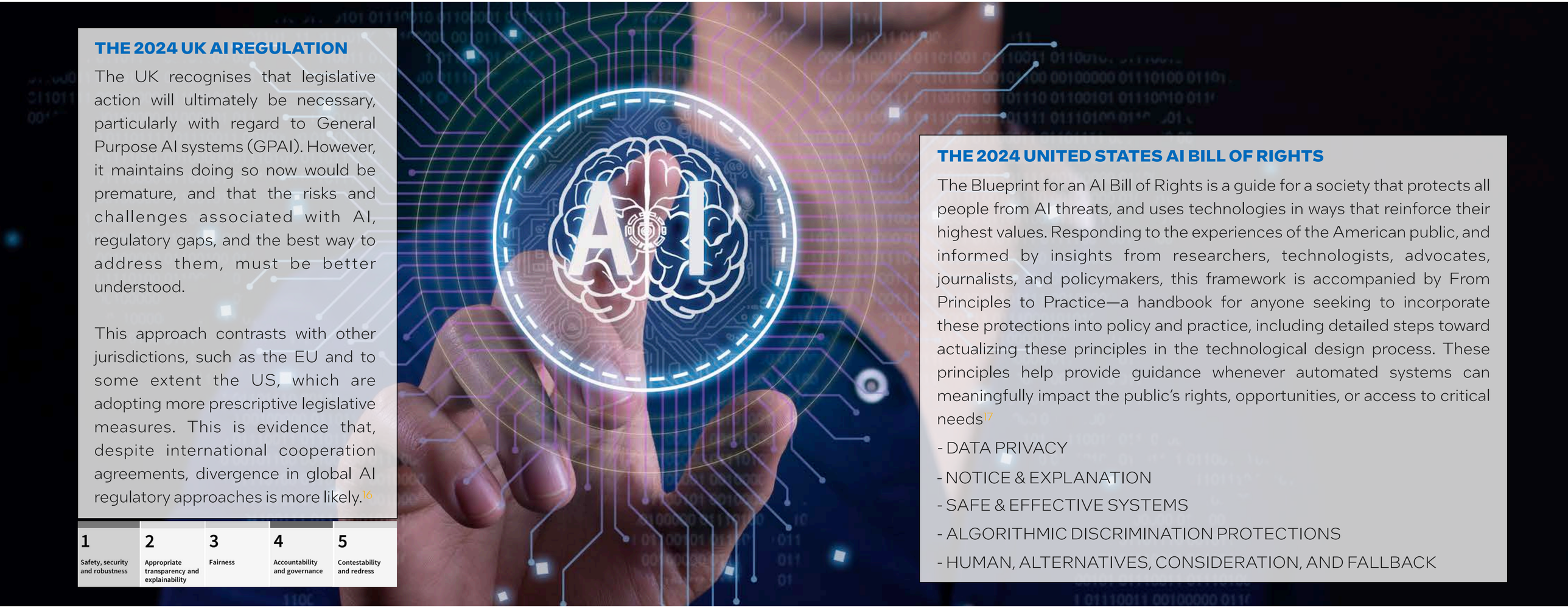
Regulating AI

THE 2024 UK AI REGULATION

The UK recognises that legislative action will ultimately be necessary, particularly with regard to General Purpose AI systems (GPAI). However, it maintains doing so now would be premature, and that the risks and challenges associated with AI, regulatory gaps, and the best way to address them, must be better understood.

This approach contrasts with other jurisdictions, such as the EU and to some extent the US, which are adopting more prescriptive legislative measures. This is evidence that, despite international cooperation agreements, divergence in global AI regulatory approaches is more likely.[16]

1	2	3	4	5
Safety, security and robustness	Appropriate transparency and explainability	Fairness	Accountability and governance	Contestability and redress



THE 2024 UNITED STATES AI BILL OF RIGHTS

The Blueprint for an AI Bill of Rights is a guide for a society that protects all people from AI threats, and uses technologies in ways that reinforce their highest values. Responding to the experiences of the American public, and informed by insights from researchers, technologists, advocates, journalists, and policymakers, this framework is accompanied by From Principles to Practice—a handbook for anyone seeking to incorporate these protections into policy and practice, including detailed steps toward actualizing these principles in the technological design process. These principles help provide guidance whenever automated systems can meaningfully impact the public's rights, opportunities, or access to critical needs[17]

- DATA PRIVACY
- NOTICE & EXPLANATION
- SAFE & EFFECTIVE SYSTEMS
- ALGORITHMIC DISCRIMINATION PROTECTIONS
- HUMAN, ALTERNATIVES, CONSIDERATION, AND FALLBACK

Intel’s commitment to Responsible AI

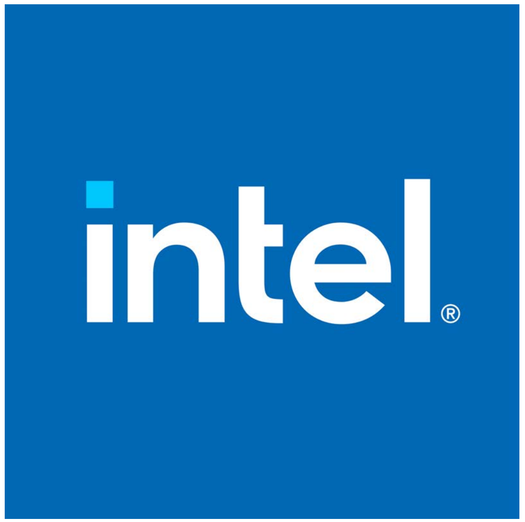


BE BETTER

Intel is committed to advancing AI technology responsibly. We do this by utilizing rigorous, multidisciplinary review processes throughout the development lifecycle, establishing diverse development teams to reduce biases, and collaborating with industry partners to mitigate potentially harmful uses of AI. We are committed to implementing leading processes founded on international standards and industry best practices, including open standards. AI has come a long way but there is still so much more to be discovered, as technology evolves[18].

MAKE BETTER

We believe that technology should enrich the life of every person on Earth. AI can create global change, provide us with powerful tools, and enable a responsible, inclusive, and sustainable future. We are harnessing the power of AI to tackle critical global challenges like pandemics, natural disasters, and global public health. We are developing AI capabilities and solutions to amplify human potential, enhance inclusion and improve accessibility for people with disabilities. We hope that by being able to provide AI access to all, we can help build a better stronger tomorrow[19].



Forward-Looking Statements
Intel's Business Outlook and other statements in this release that refer to future plans and expectations are forward-looking statements that involve a number of risks and uncertainties. Words such as "accelerate," "achieve," "adjust," "allow," "anticipates," "believes," "committed," "continues," "could," "deliver," "estimates," "expand," "expects," "focus," "forecast," "future," "goals," "grow," "guidance," "improve," "increasing," "manage," "may," "on-track," "opportunity," "outlook," "plan," "positioned," "potential," "progress," "ramp," "refocus," "regain," "sharpen," "should," "support," "will," "would," and variations of such words and similar expressions are intended to identify such forward-looking statements. Statements that refer to or are based on estimates, forecasts, projections, uncertain events or assumptions, including statements relating to Intel's strategy and its anticipated benefits, including our February 2022 Investor Day financial model, Smart Capital strategy, and updates to our reporting structure; manufacturing expansion, financing, and investment plans, including the impacts of plans such as our announced investments in the U.S. and abroad; plans, customers, and goals related to Intel's foundry business; projected costs and yield trends; supply expectations, including regarding industry shortages, constraints, limitations, pricing and sufficiency of future supply; pending transactions, including the pending acquisition of Tower Semiconductor Ltd. and the wind-down of our Intel® Optane™ memory business; the proposed initial public offering of Mobileye; total addressable market (TAM) and market opportunity; business plans and financial expectations; future macroeconomic and geopolitical conditions; future legislation, including any expectations regarding anticipated financial and other benefits or incentives thereunder; tax- and accounting-related expectations; future responses to and effects of the COVID-19 pandemic, including manufacturing, transportation, and operational restrictions or disruptions, such as the recent port shutdowns in China; future products, technology, and services, and the expected regulation, availability, production, and benefits of such products, technology, and services, including product ramps, manufacturing goals, plans, timelines, and future progress, future process nodes and technologies including Intel 20A, RibbonFET, and PowerVia, process performance parity and leadership expectations, future product architectures, Alder Lake, Meteor Lake, Rapid Lake, Sapphire Rapids, and future GPU and IPU products; future business, social, and environmental performance, goals, measures, and strategies; expectations regarding customers, including with respect to designs, wins, orders, and partnerships; projections regarding competitors; and anticipated trends in our businesses or the markets relevant to them, including with respect to future demand and industry growth, also identify forward-looking statements. All forward-looking statements included in this release are based on management's expectations as of the date of this release and, except as required by law, Intel disclaims any obligation to update these forward-looking statements to reflect future events or circumstances. Unless specifically indicated otherwise, the forward-looking statements in this release do not reflect the potential impact of any divestitures, mergers, acquisitions, or other business combinations that have not been completed as of the date of this presentation. Forward-looking statements involve many risks and uncertainties that could cause actual results to differ materially from those expressed or implied in such statements. Intel presently considers the following to be among the important factors that can cause actual results to differ materially from the company's expectations •Demand for Intel's products is highly variable and can differ from expectations due to factors including changes in business and economic conditions; customer confidence or income levels, and the levels of customer capital spending; the introduction, availability, and market acceptance of Intel's products, products used together with Intel products, and competitors' products; competitive and pricing pressures, including actions taken by competitors; supply constraints and other disruptions affecting customers; changes in customer order patterns or forecasts including order cancellations; changes in customer needs and emerging technology trends; and changes in the level of inventory and computing capacity at customers •Intel's results can vary significantly from expectations based on capacity utilization; variations in inventory valuation, including variations related to the timing of qualifying products for sale; changes in revenue levels; segment product mix; the timing and execution of the manufacturing ramp and associated costs; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources, including as a result of ongoing industry shortages of components and substrates; suppliers extending lead times, experiencing capacity constraints, limiting or canceling supply, allocating supply to other customers including competitors, delaying or canceling deliveries or increasing prices, or other supply chain issues; product manufacturing quality/yields; and changes in capital requirements and investment plans. Variations in results can also be caused by the timing of Intel product introductions and related expenses, including marketing programs and Intel's ability to respond quickly to technological developments and to introduce new products or incorporate new features into existing products, as well as decisions to exit product lines or businesses, which can result in restructuring and asset impairment charges.

For more complete information about performance and benchmark results, visit www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Performance varies by use, configuration and other factors. Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Sources: 1. WEF 2. Intel RAI 3. Amnesty 4. Intel CSR [Pg 51] 5. EU Act 6. Intel Gen AI 7. Intel FakeCatcher 8. Intel Security 9. Intel Prediction Guard 10. Intel D&I 11. Intel RAI [B] 12. Intel Overview 13. Intel Environment 15. EU AI Act 16. Deloitte 17. White House 18. Intel [A] 19. Intel [B]
Image Credits: ENVATO

intel ai