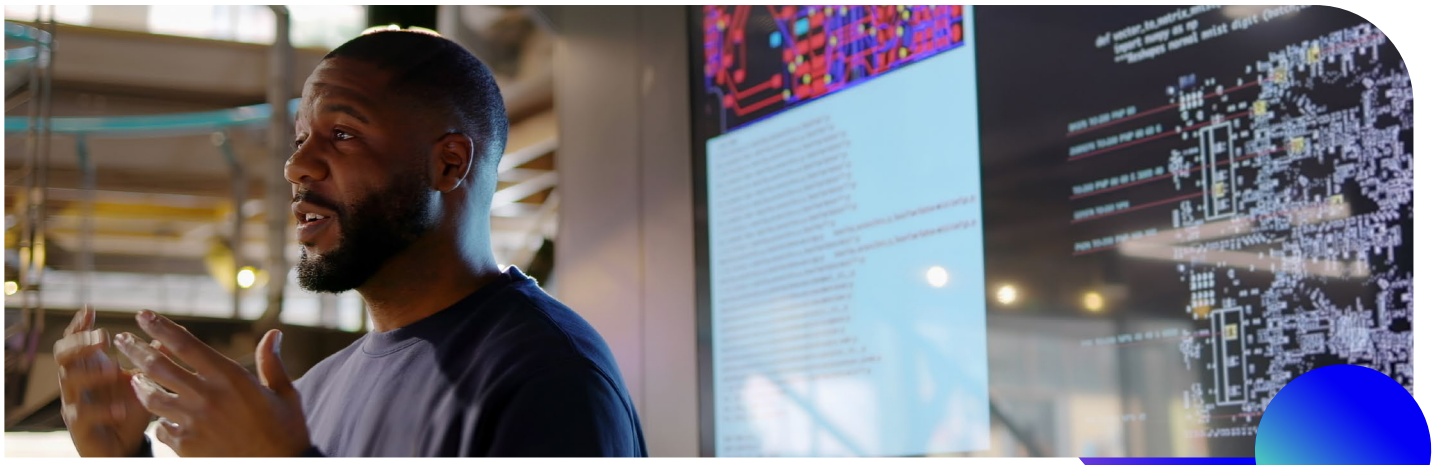

Addressing Artificial Intelligence Concerns: Recommendations from Gen

As our society faces new and escalating challenges to protect safety, freedom and trust in an AI-enabled digital environment, Gen found it imperative to address these challenges through Artificial Intelligence (AI) usage and policy recommendations. With the goal of helping digital citizens and organizations successfully navigate the new digital world, these recommendations are not intended as an exhaustive list but rather as a focused set of actionable key principles aimed at ensuring security and trust. This effort comes at a time when AI-generated threats and risks are developing and increasing, as illustrated in the accompanying paper "The AI Revolution of Good and Bad."

Recommendations regarding AI for citizens

The current AI revolution bears vast consequences that, as of yet, are only partially understood. All citizens of our digital world are and will continue to be, affected. Individuals will increasingly face the ever-growing challenge of navigating the digital world freely and safely. The principles below aim to help them effectively address this challenge:

- **Exercise caution** when encountering all digital information and consider any available context before deciding its trustworthiness. It is exceptionally easy to publish any claim on the internet and create the false impression of it being well-established.
- **Be cautious** of deepfakes and do not rely on your own ability to consistently recognize one. The technical quality of some deepfakes has already reached the point of unrecognizability. Be aware that easily recognized deepfakes still exist, and their visible imperfections, such as inconsistent lighting or blurring in face-swapped videos, may mislead users into feeling overconfident in their ability to recognize all deepfakes across the board
- **Stay cautious** about publicly sharing your own private information. Misuse of personal information by unauthorized third parties can open the door to very convincing personalized scams, fraud or other harm, especially if the data is correlated across time and various services.
- **Only use intelligent chatbots** and generative models from providers who offer clear legal guarantees against the misuse of private data.
- **When communicating digitally** with someone whose identity is not reliably known, do not automatically assume they are a real person, no matter how genuine and authentic the communication feels.

Gen

Recommendations regarding AI for organizations

Technology providers, digital industries, service providers and content providers have the technical power to take strides in improving both safety and freedom in the digital world. Their incentives, however, may be only partially aligned with the interests of users and communities. The principles below aim to help organizations contribute to a safer AI-enabled digital environment:

- Internet ecosystems should benefit from tools to detect potential AI-generated content (invisible watermarks, certified metadata, generated content detection technology, claim verification ecosystems, etc.). Likewise, online environments should utilize additional pervasive digital trust technology acceptance, allowing users to navigate the digital realm with a lower risk of fraud.

- Using AI models in production opens a new type of security and privacy threat to organizations. AI models are opaque, meaning they are hard to explain and test. The technology to properly evaluate AI models and systems to ensure safety and privacy is currently not mature or pervasive enough. In addition, arbitrary third parties that cannot be trusted could amplify existing security and privacy risks and introduce new AI-centric attack vectors. Organizations must contribute to the development and acceptance of new AI safety control tools and policies to prevent known and unforeseen AI misuse and failure scenarios.

- Organizations should enhance transparency regarding AI-technical solutions, open widely used AI-supported ecosystems to public or third-party scrutiny and enable third-party participation in policy control or safety measures in AI solutions.

Recommendations regarding AI for regulators and other governmental bodies

Regulators and governmental bodies face the challenge of finding the right balance between regulation and freedom, safety and fostering business growth. Another concern arises from delayed action, given the speed with which the AI revolution is progressing. The principles below aim to help regulators and other governmental bodies make the right decisions and adopt adequate policies related to AI:

- Support any measure that aims to raise public awareness regarding the issue of content authenticity on the Internet.

- Increase public understanding of the risks associated with sharing personal information publicly.

- Update laws regarding ownership of AI-generated data regarding data used for generative AI training and clarify rules of recycling AI-generated content, including the use of digital avatars and personas.

- Considering the problem of deepfakes and misinformation developed through technology that is also influenced by sociological and psychological factors, provide a combination of technological solutions (fact checking, authentication of content or authors, automatic detection of deepfakes from content using AI) and societal solutions (improved education, improved systems of liability assessment, legal responsibility, identity management on the internet, etc.).

- Support measures that can help raise understanding and awareness of technological aspects like human manipulability through the use of behavioral models.

- Support efforts aimed at achieving societal consensus on the balance and trade-off between free-and-dangerous versus safe-but-restricted modes of operating AI systems.

- Encourage clarity and openness regarding AI providers and their systems.

Gen

Gen, a global leader in Cyber Safety

Gen is a global leader in Cyber Safety, with dual headquarters in Prague, Czech Republic and in Tempe, Arizona. The company marks its presence in over 150 countries, catering to nearly 500 million users worldwide. The Gen portfolio includes comprehensive cybersecurity solutions from a family of trusted brands such as Norton, Avast, LifeLock, Avira, AVG, ReputationDefender, and CCleaner.

Digital Freedom as a key principle

Powering Digital Freedom lives at the heart of everything Gen does. This goes beyond the Company's mission to create solutions that enable people to navigate their digital lives safely, privately, and confidently. It's about empowering both the generations of today and future generations to be able to take advantage of the ease technology offers, worry free. That's why Gen approaches everything we do with the customers and communities we serve in mind. We champion the simplification and safeguarding of customer experiences in the ever-evolving digital landscape, reinforcing our role as a leader in digital security and empowerment.

Read the full report here

If you want more information, please reach-out to:

Kim Allman
Head of Corporate Responsibility,
ESG & Government Affairs
Kim.Allman@GenDigital.com

Transparency Register number: 083146048556-68

United States: 60 E Rio Salado Pkwy STE 1000 Tempe, AZ 85203
Czech Republic: Enterprise Office Center Pikrtova 1737/1A 140 00 Prague 4

Gen | norton | Avast | LifeLock by norton | Avira | AVG | Reputation Defender by norton | CCleaner