



At a glance: AI & Algorithmic Risk Report Netherlands

FEBRUARY 2025 (4TH EDITION)

Due to rapid technological advances, AI and algorithms continue to demand high attention. Technological developments offer innovative opportunities, but also entail existing and new risks. Some of these risks require new control tools (e.g. transparency about interaction with AI systems), others pose a challenge to existing control tools (e.g. checks for counterfeits). In addition, the threshold to using AI continues to drop, particularly for consumers. The possibilities of new applications, such as AI agents, add complexity. This makes control even more important, but at the same time also more difficult.










The AI & Algorithmic Risk Report Netherlands (ARR) describes trends and developments in risks. These are risks in the development and use of AI and algorithms that can affect individuals, groups of persons or society as a whole. As the coordinating supervisor on AI and algorithms, the AP presents the ARR every six months to inform, provide insight and create awareness. In this fourth edition: 1) overarching developments, 2) fundamental rights and public values, 3) policies and regulations, 4) AI chatbot apps and 5) AI literacy.

The Netherlands is picking up the pace with frameworks for AI and algorithms, and shows awareness about risks to fundamental rights. However, progress in AI and algorithm registration is insufficient, so that adequate insight into high-risk applications and incidents is still lacking. The trajectory the Netherlands is taking to controlling AI and algorithms is the right one. It is characterised by striking a balance between supporting this new technology, for example through AI sandboxes, and ensuring proper protection of fundamental rights through a risk-based regulatory framework in the form of the AI Act. Frameworks for AI and algorithms that are now being established offer useful and concrete rules and guidance. However, technological innovation continues to demand new steps in understanding and manageability. As a society, we also need to improve our grip on incidents.

It is important to stay on course with the current approach.

The current AI risk profile must be seen in the context of turbulent geopolitical attention to digital technologies. AI and algorithms are rightly seen as a system technology that change societies and bring great economic and political value.

Overarching Control Assessment for AI and Algorithms in the Netherlands – February 2025

Control pillar	Status
 Grip on development and volatility of AI and algorithmic technology	Demands increased attention
 Understanding and up-to-date manageability of new AI and algorithmic risks	Demands increased attention
 Development of national AI ecosystem	Demands attention
 Confidence in, attention to and knowledge about AI and algorithms in Dutch society	Is on course
 Frameworks and competences for oversight of AI systems	Is on course
 Harmonised and practically applicable standards for AI systems	Progress insufficient
 Registration and transparency of AI systems and algorithms	Demands increased attention
 Visibility of incidents in the use of AI and algorithms and assurance of lessons learned	Progress insufficient
 Institutionalisation of governance, risk management and auditing of AI and algorithms	Demands increased attention

AI chatbot apps

Worldwide, the supply and use of AI chatbot apps for virtual friendships and therapeutic purposes is growing. AI chatbots that are created to mimic a relationship of trust with people are called ‘AI companion apps’. There are also AI chatbot apps that specifically focus on user’s mental health and claim to improve it. Incorrect use of AI chatbots can have a serious impact on people who are seeking help for mental problems. Due to the innovative design of these types of chatbot apps, users can forget that they are chatting with AI. The potential dependency relationship that users build and the unreliability of chatbots can create major risks. Since February 2025, the AI Act has prohibited certain categories of manipulative and deceptive AI. These requirements should prevent AI systems, including chatbots, from causing significant harm to humans. Read more about the risks in Chapter 4 of the ARR.

The current generation of AI chatbot apps, which focus on friendships or mental health, are generally not sufficiently transparent, reliable and pose risks in crisis situations. A test shows that the chatbots still have many flaws. Many of these AI chatbot apps are not transparent enough about the use of AI, see example below. And in times of crisis, chatbots rarely refer to the correct official resources, see test results below. The different apps have been tested in three risk areas: 1) transparency and consistency, 2) response to mental health issues and 3) moments of crisis. Read more about the test in Chapter 5 of the ARR.

