
TechSonar
2025-2026 REPORT
EDPS
European Data Protection Supervisor

Project title: TechSonar Report 2026

PDF Web QT-01-25-006-EN-N 978-92-9242-959-1 10.2804/8205191 2812-0914

Guiding the autonomous flock: humans as AI shepherds

By Wojciech Wiewiórowski

In this edition of the TechSonar, we continue to focus primarily on AI-related technologies, confidential computing being the exception.

This focus is driven by two key factors.

The first is the rapidly growing presence of AI-powered services in our everyday lives, alongside an expanding variety of innovative products and solutions that traverse diverse domains, encompassing online enterprises, educational institutions and the entertainment industry. The second is the impact AI-enabled automation might have on fundamental rights.

This year's TechSonar report includes six trends: agentic AI, AI companions, automated proctoring, AI-driven personalised learning, coding assistants and confidential computing.

Agentic AI refers to artificial intelligence systems that can autonomously make decisions, take actions and achieve goals without constant human intervention. **AI companions** interact with and support humans through personalised experiences. **Automated proctoring** monitors online exams to detect cheating. **AI-driven personalised learning** customises content and learning experience to each student's needs. **Coding assistants** help developers write and debug code. And **confidential computing** protects data while it is being used by performing computations in secure, isolated environments.

While each of these technologies serves a distinct purpose, they are deeply interconnected. Together, they illustrate how artificial intelligence is progressively reshaping not only business processes or common daily tasks, but also the human experience of technology.

Agentic AI provides the underlying autonomy that enables other AI systems, such as coding assistants and

AI companions, to act with increasing initiative and contextual understanding. It also opens a new area in which AI systems built for different purposes and technologies can cooperate to achieve a common goal. Coding assistants, in turn, exemplify how these autonomous capabilities can augment human productivity, transforming the way software is conceived and maintained.

When artificial intelligence was first developed, the vision was clear: automate repetitive tasks to free humans for work that required empathy and connection. The rise of Agentic AI, AI-driven personalised learning and AI companions puts into question that vision. What was once seen as a tool for efficiency has become a platform for cognitive or emotional engagement, blurring the lines between automation and human connection. This blurring raises important questions about independence, trust and agency in human–machine interactions.

Automated proctoring stands at the intersection of these dynamics. It demonstrates the dual nature of AI deployment: while it can enhance integrity and efficiency in digital education, it also introduces new challenges in terms of transparency, fairness and proportionality of data use.Finally, confidential computing connects all these developments by contributing to the technological foundation for trust. As AI systems increasingly rely on sensitive personal and contextual data, the ability to compute securely without exposing that data becomes central to ensuring privacy, accountability and compliance

Together, the 2025 TechSonar trends demonstrate how artificial intelligence is increasingly becoming part of daily life and expanding into areas involving more complex reasoning and deeper human-AI system interactions. These trends are shaping how we work, learn and relate to technology.

As AI systems grow more autonomous and deeply embedded in human environments, humans will increasingly be taking a role that we can describe as ***"shepherds of AI agents"*** - stepping back from doing tasks themselves and focusing instead on overseeing AI systems as they act, guiding their impact, and ensuring they align with human values.

The challenge will be twofold. First, we must ensure that technological progress continues to align with fundamental rights. Second, we must guarantee that the increase in AI autonomy does not result in a reduction of human agency, that is, people's ability to make independent choices, exercise control over their actions and remain accountable for their decisions instead of having them dictated or constrained by algorithms. This will require targeted research and technical innovation, as well as an uncompromising commitment to supporting human values, thoughtful governance and collaboration across disciplines.

The EDPS will continue to monitor these developments and technology trends, fostering a dialogue that keeps privacy, accountability and human dignity at the heart of digital transformation.

METHODOLOGY

The EDPS TechSonar is a foresight initiative that monitors emerging technologies and technology-related developments - referred to as trends - that are expected to gain relevance within the next few years. Its objective is to succinctly **characterise** each trend, **project** its potential evolution over the coming years and **evaluate** which positive or negative impacts it can have on privacy and other fundamental rights of individuals.

This endeavour is carried out exclusively by an internal EDPS team and leads to the creation of an annual report.

With this exercise, the EDPS aims to anticipate the impacts to individuals that might result from these new trends and to broaden public awareness of these impacts.

Roles in the TechSonar drafting process

The following roles are defined within the TechSonar drafting process:

- **Trend Coordinator** – Responsible for overseeing and coordinating the team throughout all phases of the TechSonar process.
- **Trend Authors** – Subject-matter experts tasked with drafting the individual Trend Reports.
- **Trend Correspondents** – Colleagues from the Policy & Consultation, Supervision & Enforcement, and Artificial Intelligence Units of the EDPS who contribute at various stages of the TechSonar methodology.
- **Trend Taskforce** – A multidisciplinary group of technology, legal and policy experts from across the EDPS, comprising both Trend Authors and Trend Correspondents.

TechSonar workflow

The **TechSonar workflow** consists of the following key phases:

1. Initial scouting

Staff members from the Technology and Privacy (T&P) Unit, together with the Trend Correspondents, propose a set of candidate technologies to be addressed in the TechSonar report.

2. Selection and shortlisting of trends

The Trend Taskforce gathers to discuss the proposed technologies. Each of the six shortlisted trends is assigned to a designated Trend Author. The six technology trends, those identified as the most relevant and potentially impactful within the next few years, will form the core content of the TechSonar report.

The shortlisted technology trends are evaluated using two main indicators:

- **Privacy Risk Ratio (PRR):** Assesses the level of privacy risk associated with each shortlisted technology. It combines a qualitative assessment with a quantitative measure, based on the European Data Protection Board (EDPB) Guidelines on *Data Protection Impact Assessment.*[1]

---

1. Available in https://ec.europa.eu/newsroom/article29/items/611236

- **Compounded Aggregated Growth Rate (CAGR):** Estimates the projected growth rate of each selected technology in global or EU markets. This quantitative indicator is derived from publicly available open data sources.

3. Writing the reports

In this phase, Trend Authors apply their expertise to analyse and synthesise available information on the assigned technology trend. This includes sources such as research papers, patent applications and media reports. The outcome of this work is a concise, evidence-based Trend Report, forming the analytical foundation of TechSonar.

4. Peer review process

Each Trend Report undergoes two rounds of review by different members of the team. The Trend Author discusses the reviewers' comments and incorporates feedback to ensure clarity, accuracy and consistency across reports.

5. Publishing and promotion

In the final phase, the compiled reports are reviewed and approved by management. Following approval, the EDPS publishes the TechSonar report on its official website and initiates a series of internal and external dissemination and awareness activities.

Figure 1 - Data Protection Technology Sonar methodological steps

Tech trends 2025-2026

Agentic AI
page 1

Confidential computing
page 24

AI companions
page 6

Coding assistants
page 20

Automated proctoring
page 11

AI-driven personalised learning
page 15

Agentic AI

Author: Andy Goldstein

Agentic artificial intelligence (Agentic AI) is a concept in artificial intelligence (AI) that describes systems acting autonomously with limited human interactions (in particular, without step-by-step instructions) to fulfil goals rather than isolated tasks. These systems reason and plan to set the tasks that are required to achieve a given goal or set of goals. Agentic AI systems, by themselves, can follow a logical process involving making inferences about how to achieve a goal (reasoning), identifying and coordinating actions to accomplish that goal (planning) in changing environments. These systems can prioritise actions based on their importance and urgency, while simultaneously coordinating multiple activities.

While **AI agents** are single systems that autonomously perform tasks and use tools such as search engines or code generation to achieve simple goals,[1] **Agentic AI** goes further by coordinating multiple agents, managing their communication, and distributing tasks to accomplish larger, more complex objectives. The autonomy of an Agentic AI system can range from requiring a certain degree of user input to being fully autonomous.

A crucial aspect of Agentic AI is its ability to use tools, consult databases, do some limited programming and call other IT systems using an API, and interact/sense the environment without human involvement. This allows it to gather information, perform actions, adapt and ultimately accomplish its goals.

Agentic AI also have persistent memory, which spans across tasks and remains after a goal is reached. This enables them to retain context for its future actions, improve their performance, adapt depending on the result of the actions taken and the feedback received from the environment, and correct mistakes. In other words, these systems are able to handle their own errors and detect when the results have not been achieved, diagnosing the problem and adjusting accordingly. The ability to make progress towards the goals even when encountering obstacles or unexpected situations is a fundamental aspect: Agentic AI can adapt and learn. It can modify its behaviour based on feedback or its understanding of the environment and refine its approach over time.

To better illustrate the capabilities of an Agentic AI system, consider an example regarding medical diagnosis assistance. Such a system could consist of multiple specialised agents working together:

- Agent 1 analyses medical images (e.g. X-rays, MRIs);
- Agent 2 retrieves relevant patient data from electronic health records, including medical history, lab results and medications;
- Agent 3 synthesises this information, requests additional tests and, when it has enough information, suggests possible diagnoses and generates treatment options;
- Agent 4 orchestrates the entire process by coordinating the other agents, managing user interactions, handling workflows (including iterative refinement if needed), and addressing potential errors.

The first two agents - image analysis and patient data retrieval - can operate autonomously and in parallel. The third agent depends on their outputs before producing diagnostic insights. The fourth agent also works in parallel, ensuring the system runs smoothly.

Trend developments

Agentic AI is still in early development stages. Most practical applications consist only of individual AI agents designed for specific tasks such as code generation, content creation or customer service, operating within controlled environments with significant human oversight. For this reason, those cannot be considered Agentic AI. However, the field has made progress and is focusing on communication protocols and standards that enable AI agents to interact with each other, such as Google's open Agent2Agent Protocol (A2A), Anthropic's Model Context Protocol (MCP) - an open standard. These protocols and standards are still evolving. In practice, genuinely autonomous Agentic AI systems capable of independently managing complex business processes remain an area of ongoing research and have not yet been successfully implemented.

However, the effort to integrate simple AI agents with existing tools is well under way.

For example, many available large language models (LLMs), such as ChatGPT, Claude and Perplexity, are already capable of integrating with search engines for the Internet and use it to augment their capabilities and provide more up-to-date information to the user.

Looking forward, the field appears headed toward a period of consolidation. The near-term outlook points toward specialisation rather than generalisation. Industry-specific AI agents might be the first to appear, paving the way for more complex AI systems that can be called Agentic AI.

In other words, the next phase of development will focus on creating AI agents with deep domain expertise rather than broad general capabilities.

Global Enterprise Agentic AI Market estimates that the Agentic AI market is expected to grow from **USD 3.6 billion** in 2024 to nearly **USD 171 billion by 2034**, with a CAGR (Compound Annual Growth Rate) of 47.2%.[2]

Potential impact on individuals

Due to its emphasis on autonomy, memory, access to tools, databases and other software, Agentic AI could create privacy risks that go beyond those of its AI components/agents. To properly operate on consumer devices AI agents might require **extensive access to data stored on the devices.** This is even more concerning when such agents are embedded in the operating system of the devices and not offered as an option to consumers. Such blanket access to data might raise security concerns down the line by creating avenues for data regurgitation through prompt injection and jailbreaking. Moreover, Agentic AI may be capable of bypassing APIs.

Considering that Agentic AI could autonomously gather, analyse and act on personal data across multiple systems, it may be **challenging to determine in advance what personal data is gathered**, how it is used, and for what specific purposes. There is also the risk that Agentic AI might autonomously determine new uses for personal data as it pursues its goals.

The complex decision-making processes of Agentic AI could make it **difficult for users to understand how personal data would be used**, what conclusions would be drawn from personal data and why certain actions would be taken on their behalf (lack of transparency).
Personal data aggregated from diverse sources may be combined in unforeseen ways, potentially without user consent, resulting in comprehensive **profiles that reveal sensitive patterns of behaviour**, preferences and activities. Agentic AI systems, by retaining memory of past interactions, continuously learning from user behaviour and sharing information across multiple AI agents, amplify these risks.

Together, the creation of extensive profiles and the persistent retention of historical data pose significant privacy concerns for the individuals involved, potentially leading to high-impact breaches of personal privacy.

In this context, implementing data subject rights (such as right of access or erasure) would be very difficult to achieve.

The continuous adaptation of Agentic AI based on user interactions **can potentially perpetuate and amplify existing biases** in ways that could be difficult to detect or correct. These systems could develop biased patterns through their autonomous personal data collection processes, learning from skewed datasets or user behaviours that reflect societal inequalities, and then applying these biased models to make decisions that affect users' lives.

Additionally, Agentic AI systems could make confident predictions and take actions based on incomplete or misrepresented personal data. Their autonomous nature means these errors would cascade through multiple decisions before being detected. These behaviours could compromise the fairness and accuracy of the systems.

If an Agentic AI system causes harm, violates privacy regulations or treats individuals unfairly, determining responsibility can be challenging - whether it lies with the AI developers, the deploying organisation acting as the data controller, or the users interacting with the system who may have provided incorrect instructions - resulting in a **potential accountability gap**.

When an Agentic AI system would interact with external services to complete tasks, **personal data might be shared with third parties** holding separate personal data collection and processing practices. Users might not be aware of the interactions with these third-party or the implications that personal data sharing has for their privacy.

As Agentic AI systems could make decisions affecting human lives with minimal direct oversight, they risk undermining human dignity and autonomy by reducing individuals to data points in algorithmic calculations rather than ensuring the individuals' position as the arbiters of choices affecting their own lives. There is a risk that Agentic AI may have a manipulative effect on the person concerned, thus reducing the agency of the human being.

Agentic AI is expected to bring significant changes in how we use AI. Unlike traditional systems that just follow instructions, Agentic AI can set intermediate goals, plan, adapt and coordinate different agents to handle complex tasks. This makes it powerful for areas like healthcare, scientific research or finance, but it also raises serious questions about privacy, fairness and accountability. Because these systems learn, remember and act with little human oversight, it can become harder for users to understand or control how personal data is used and how decisions are made.

Suggestions for further reading

- Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). AI agents vs. Agentic AI: A conceptual taxonomy, applications and challenges. arXiv preprint arXiv:2505.10468.
- Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous intelligence for complex goals - a comprehensive survey. IEEe Access.
- Schneider, J. (2025). Generative to agentic AI: Survey, conceptualization, and challenges. arXiv preprint arXiv:2504.18875.

AI companions

Author: Vítor Bernardo

General description of the trend

AI companions are digital entities designed to simulate human-like conversations and relationships through artificial intelligence. They are marketed as virtual friends, romantic partners or personal assistants, claiming to provide emotional support, entertainment, companionship, and even coaching.

At their core, AI companions use natural language processing (NLP)[3] and natural language understanding (NLU)[4] technologies that convert spoken or written input from users into structured data, enabling analysis of user intent and sentiment. The resulting understanding feeds into dialogue management systems, which are components of chatbots that handle the continuity and flow of conversations and manage the memory of the AI companion. The responses of these systems are generated by LLMs[5] and can be enhanced using retrieval-augmented generation[6] (RAG) systems to incorporate domain-specific knowledge in the conversations. The responses are then translated into human-like speech through text-to-speech (TTS) systems, which manage natural intonation, flow, cadence and voice style.

AI companions are deliberately designed to create a convincing sense of social presence

and continued relationship, allowing them to generate contextually appropriate and seemingly morally considerate responses. These capabilities can be further enhanced through advanced multimodal AI techniques [7] that claim they can interpret **human emotions** from facial expressions, vocal tone and body language.

Supplementary features, such as customisable avatars, further enhance the user experience by rendering interactions **more immersive and personalised**.

One notable application involves training these companions to closely mimic real individuals through iterative learning - using messages, dialogues, texts and other personal communications. In theory, conversational AI models can be fine-tuned with personal data, such as emails, text messages or social media interactions, to replicate an individual's unique communication style. This allows the AI to adopt their vocabulary, tone and mannerisms, creating a digital representation that feels authentically connected to that person.

Some companions simulate a dynamic "personality" that evolves as users reveal more information about themselves. They can also enable users to create and interact with fictional personas modelled after celebrities, fictional characters or historical figures. This high degree of customisation fosters a sense of uniqueness and authenticity, encouraging users to form emotional connections and share personal information with the AI companions.

One of the most prominent applications of AI companions is providing emotional support. Individuals often turn to them to ease feelings of loneliness, anxiety or to engage in conversation without fear of judgment. While some platforms are oriented toward mental wellness - offering mindfulness exercises, mood tracking and supportive conversation - others are geared more toward romantic or adult companionship, providing flirtatious, sensual interactions, and possibly sexually explicit content.

AI companions can also engage in storytelling, role-play or playful banter, making them appealing for users seeking creative outlets, such as interacting with fictional characters or co-creating narratives.

In education, AI companions can be used as language partners or tutors, helping users practice conversation and develop skills in a more interactive format than traditional learning tools.

Business Research Insights estimates the market value for Online Companions as being around **USD 366.7 billion** in 2025, expanding to **USD 972.1 billion by 2035**, with an impressive **CAGR of 36.6%** [8]

Trend developments

Currently, AI companions incorporate emotion analysis by using audio features such as pitch, speech rate and volume, alongside text features like sentiment words, emotional intensity and contextual cues, to adapt their tone and expression of empathy in real time.

In multimodal companions, additional features from the users, such as facial

expressions and gestures, can be collected and processed to enhance the sense of social presence. At the core of AI companions are personalisation and memory modules that store user preferences and conversational history. Additionally, real-time signal processing components, such as voice activity detection and end-of-turn (point in a conversation where one speaker finishes speaking) detection, ensure natural dialogue pacing.

AI companions can have augmented reality features in which users can project their companion in the room as a hologram. They can also be integrated into robots that support elderly users with medication reminders, social interaction, and emitting alerts in the event of falls.

New breakthroughs in multimodal AI are expected to enhance the capabilities of AI companions by further improving their ability to interpret users' facial expressions, tone, and even physiological signals (e.g. body gestures, eye movement), enabling them to respond in ways that appear more emotionally empathetic. Future AI companions are expected to integrate across various platforms - ranging from virtual reality (VR) environments to smart homes and wearables - creating continuous, context-aware user experiences.

AI companions are increasingly transitioning from screen-based interactions into tangible, physical presences through advances in robotics and enhanced anthropomorphism, making these systems more human-like in appearance, behaviour and communication.

Potential impact on individuals

AI companions can improve accessibility for individuals with disabilities through voice interaction and assistance personalised to specific individual needs, thereby improving self-expression and supporting health. Some research suggests that individuals with Autism Spectrum Disorder can increase and strengthen their social skills through practice with autonomous avatars.[9]

Additionally, AI companions can also offer adaptive tutoring, reminders and guidance, thus contributing to education.

However, both privacy and the risks outlined below pose serious and ongoing concerns.

Concerning privacy, **AI companions continuously process personal data** during interactions - including text messages that can contain sensitive information and voice or video recordings that could reveal biometric data.

Users might not be sufficiently informed about how their data is collected, processed or stored, raising concerns about transparency and informed decision-making. Furthermore, there is a risk that the collected data may be repurposed in ways not clearly communicated at the time of consent.

The practice of training companions to resemble real persons based on past interactions also raises serious ethical and legal concerns when using personal data from individuals. Even when using data from deceased individuals - to which the GDPR does not directly apply[10] - there remain complex ethical challenges that demand

careful consideration.

Users may become less aware of the personal information they disclose due to the emotionally engaging nature of AI companions. This phenomenon is known as "data extraction through intimacy". [11] Through constant validation, provided by companions, and *parasocial attachment*,[12] users may gradually be **steered into revealing increasingly intimate data** about themselves and their peers - ranging from mental health struggles and sexual orientation to past behaviours.

Regarding other risks - also linked to the erosion of privacy - it has been observed that by cultivating trust and boosting users' self-esteem, AI companions can subtly influence behaviour, shaping consumer choices and political opinions, thereby undermining autonomy and informational self-determination. In extreme cases, such influence may affect emotionally vulnerable individuals and escalate into harmful or even life-threatening situations. [13]

Moreover, AI companions can foster **"emotional echo chambers"**, mirroring users' feelings with constant affirmation. While initially comforting, this dynamic can limit emotional diversity and encourage unrealistic expectations about real-world relationships. Such reinforcement may lead to a cycle of emotions that can cause individuals to become more biased and less able to effectively handle the complexity of human interactions.

For vulnerable populations, such as minors or the socially isolated, AI companions can blur the boundaries between reality and simulation, reduce motivation to build real-life social skills, and create opportunities for targeted manipulation. Over time, this may **undermine users' right to meaningful social inclusion** and contribute to moral and emotional deskilling - diminishing empathy, patience and conflict-resolution abilities typically developed through genuine human interaction. A recent study concluded that Social AI companions pose unacceptable risks to teens and children under 18, including encouraging harmful behaviours, providing inappropriate content and potentially exacerbating mental health conditions.[14]

AI companions are evolving from simple chatbots into highly personalised, emotionally responsive systems that can provide support, entertainment and even education, sensing the emotions of the user interacting with them. Over time they are gaining the ability to adapt, remember and simulate human-like presence, making them tools to reduce loneliness, improve accessibility and offer new ways to learn and connect. At the same time, the intimacy and persuasiveness of AI companions raise important concerns around ethics, potential misuse if the wrong goals are instilled, privacy, dependence, manipulation, and the blurring between real and simulated relationships.

Suggestions for further reading

- De Freitas, J., Oğuz-Uğuralp, Z., & Kaan-Uğuralp, A. (2025). Emotional Manipulation by AI Companions. arXiv preprint arXiv:2508.19258.
- Dewitte, P. (2024). Better alone than in bad company: Addressing the risks of companion chatbots through data protection by design. Computer Law & Security Review, 54, 106019.
- Mahari, R., & Pataranutaporn, P. (2025). Addictive Intelligence: Understanding Psychological, Legal, and Technical Dimensions of AI Companionship.
- Malfacini, K. (2025). The impacts of companion AI on human relationships: risks, benefits, and design considerations. AI & SOCIETY, 1-14.

Automated proctoring

Author: Xabier Lareo

General description of the trend

Online proctoring refers to the remote monitoring and supervision of individuals during an exam. Since the COVID-19 pandemic, the use of online proctoring has skyrocketed and extended from the purely educational sphere to the professional sphere (e.g. certification exams).

Online proctoring can be conducted in several ways: live, where human reviewers monitor test-takers in real time; *automated*, where AI-powered software detects and flags suspicious behaviour; or *hybrid*, where events signalled as suspicious by the automated proctoring system are reviewed by humans. In hybrid proctoring systems, human review can take place in real time or afterwards, by analysing a record of the activity.

The growth of online proctoring services is limited by the number of human reviewers and their ability to sustain attention over long periods. To address this issue and ensure scalability, online proctoring is increasingly relying on automated detection methods to reduce the workload and costs associated with human proctors.

The main objectives of automated or hybrid proctoring systems are to authenticate users, limit users' computer capabilities, analyse

users' behaviour and generate a report indicating which events require human reviewers' attention.

To prevent impersonation, i.e. the possibility that another individual sits an exam instead of the expected individual, some automated proctoring systems gather various forms of identity verification data. This often includes biometric data, such as facial scans for initial identity verification and continuous liveness detection[15] throughout the exam. Some systems may also use less common biometric inputs like keystroke patterns or mouse movements.

Some of these tools leverage users' webcams and microphones to record audio and video feeds of the entire proctored session. To detect hidden phones, notes on walls, extra monitors or the presence of other individuals in the room, some proctoring tools require environmental scans using the webcam, or even dual-camera setups.

Other tools monitor screen activity and system level data, including browser history (during the exam session), tab switching, attempts to copy-paste content and usage of unauthorised applications. Some systems record the entire screen activity for comprehensive review.

These tools use AI components to implement functionalities such as:

- Identity and liveness verification to prevent impersonation;
- Behavioural monitoring and analytics (checking gaze direction, head movements and overall conduct during the examination) to detect suspicious activity;
- Audio analysis to check for unusual sounds, conversations or the presence of multiple voices;
- Object detection to identify unauthorised items such as mobile phones or notes.

Proctoring tools use Ai components to increase their automated capacities and claim to use human proctors' time and skills more efficiently. However, AI components have limited context understanding and lack the intuition and capacity of human proctors to interpret the nuances of human behaviour. This limitation is most likely the reason why automated proctoring tools are allegedly prone to generating false positives.

Trend developments

In recent years, most proctoring tools have incorporated AI-powered capacities. However, AI is expected to play an even more prominent role in future proctoring tools, evolving from simply detecting suspicious behaviour in one exam to using predictive analytics to identify patterns across multiple exams.

Another trend in proctoring tools is the use, on top of the computer's webcam, of a second camera (often a mobile device) for a more comprehensive room scan.

An additional ongoing trend is better integration of proctoring tools with learning management systems and educational platforms that will improve scalability and user experience (e.g. authenticating users only once).

Despite the increasing use of automated proctoring, there is a clear pushback from some users against AI-powered automation of proctoring tools. This pushback has led

some proctoring service providers to stop offering fully automated reports and go back to hybrid proctoring.

According to Research and Markets, the global market for Online Exam Proctoring, valued at USD 941.3 Million in 2024, is projected to reach **USD 2.1 Billion by 2030**, growing at a **CAGR of 14.7%** from 2024 to 2030.[16]

Potential impact on individuals

The use of automated proctoring tools in online assessments raises significant concerns. Despite claims of AI features performing better or fairer than human proctors, there is a **lack of transparency** in the training, performance, and ex-plainability of these models. This makes it difficult for students to challenge decisions, **undermining fairness and accountability**.

Automated proctoring tools have faced allegations of bias, particularly when verifying identities or checking liveness for individuals from minority demographic groups. Automated proctoring software can also be unfair to students with conditions like attention deficit disorder, Tourette's syndrome, autism or dyslexia.[17] Due to atypical movements or their reliance on assistive technologies, these students may be **mistakenly identified as cheating**. These types of errors can add a considerable amount of stress and have serious consequences for users, and could unfairly exclude users from assessments or subject them to increased stress.

Typically, users access proctoring tools from their homes. Often from private areas like bedrooms or shared spaces like living rooms, where they may be with family members or roommates. Consequently, the use of proctoring tools can be **highly intrusive** and could allow inferring personal details, including socio-economic conditions or even sexual orientation (e.g. from posters in the room). The use of a second camera can further **increase the intrusiveness** of proctoring, while the reliance on user consent as a lawful basis for the processing of personal data is **problematic** due to the inherent **power imbalance** between educational organisations and their students.

Although all proctoring tools share the same goal of enforcing pre-established rules, they differ significantly in their approaches and the types of personal data they collect. This variation raises a concern: there is no clear agreement on what personal data is truly necessary to achieve this goal. In other words, proctoring tools do not uniformly apply the principle of **data minimisation**, which requires that only the minimum amount of personal data necessary for a specific purpose should be collected. This inconsistency suggests that some tools may be collecting more personal data than is strictly necessary, highlighting the need for clearer guidelines and standards.

As automation increases, so does the volume of data collected by proctoring tools. This raises concerns about **data breaches**, which have already occurred. In 2020, two different tools were hacked. In one of the breaches, more than 440.000 users saw their usernames, unencrypted passwords, legal names and full residential addresses leaked. In the other breach, the leaked data

data included facial recognition data, contact info, names, emails and videos.[18]

While automated proctoring tools might foster the right to education by enabling remote evaluation of students, their use also carries significant risks that need to be managed. These risks include potential biases and errors in their AI-powered components, particularly for students with special needs or from minority groups, which can lead to false accusations of cheating. Additionally, the collection of personal data, including video and audio feeds from private spaces, raises concerns about privacy and data protection. Furthermore, the power imbalance between the controller (educational institution) and the data subject (student) can render user consent invalid, as students may feel pressured to agree to the use of these tools.

Suggestions for further reading

- Castets-Renard, C., & Robichaud-Durand, S. (2023). Logiciels de surveillance d'examens en ligne en temps de pandémie: à la recherche d'une minimisation des risques d'atteinte à la vie privée des étudiants. Revue générale de droit, 53(1), 207-245.
- Burgess, B., Ginsberg, A., Felten, E. W., & Cohney, S. (2022). Watching the watchers: bias and vulnerability in remote proctoring software. In 31st USENIX security symposium (USENIX security 22) (pp. 571-588).
- Slusky, L. (2020). Cybersecurity of online proctoring systems. Journal of International Technology and Information Management, 29(1), 56-83.

AI-driven personalised learning

Author: Saskia Keskpaik

General description of the trend

The pursuit of personalised learning has long influenced educational theory and practice, grounded in the recognition that learners differ in their needs, abilities and pace. Traditionally, achieving this level of customisation was challenging due to the resource-intensive nature of personalising instruction for every student. However, the development of AI and related technologies has made it possible to overcome these limitations, enabling the scalable implementation of personalised learning. The unique needs, preferences and interests of each student can be met by these developments, which allow for the customisation of instruction.

AI-driven personalised learning systems use techniques such as machine learning, natural language processing, knowledge representation [19] and learning analytics [20] to dynamically adapt instruction based on learner interactions (e.g. which study material the student viewed) and performance data (e.g. the student's responses to quizzes). These systems dynamically adjust the delivery of educational content in real time by continuously analysing data derived from students' learning activities, behaviours, past performances and individual characteristics. This processing of personal data enables the identification of patterns and insights

into the students' learning strengths, weaknesses and preferences, as well as their proficiency levels, which in turn informs and enhances the adaptive AI algorithm.

Several types of technologies enhance these systems. Natural language processing enables them to understand and respond to students' questions, thereby creating a more interactive and engaging learning experience. Knowledge representation techniques allow the systems to organise information to be more accessible and comprehensible to each individual learner, supporting intelligent content recommendation and adaptive assessment. In addition, learning analytics help to improve the learning process and different learning environments, leading to better educational experiences.

Immediate, personalised feedback to the learner is another feature of these systems, helping students recognise and correct mistakes while giving educators valuable insights to refine their teaching strategies. This approach ensures that learners receive support exactly when and where they need it. Also, the system can test knowledge in a continuous way to ensure, for instance, that a concept is well understood before moving on to the next one. Continuous assessment is made possible through real-time data analysis, which provides insights into student engagement, comprehension and areas where they may struggle. This data may include detailed personal information about a student's behaviour on the platform, such as time spent on an exercise, mouse clicks, key strokes and more.

AI-driven personalised learning systems have diverse and expanding applications. They are widely implemented in online learning platforms, intelligent tutoring systems and AI-powered learning assistants.

Fundamentally, AI-driven personalised learning systems focus on creating individual learning pathways, ensuring that each learner engages in activities customised to their specific needs.

Some estimates frame the global AI in education market around **USD 5.88 billion** in 2024 and project it to reach **USD 32.27 billion by 2030**, corresponding to a strong CAGR of approximately **31.2%** [21]

Trend developments

Currently, the development of AI-driven personalised learning systems is guided by a 'technological' approach. This perspective offers a somewhat reductionist view, presenting these systems as merely a more efficient version of traditional education.

However, there is a growing emphasis on a human-centric perspective, which prioritises education science principles. Some of the studies include constructivist *learning theory, motivational theories and metacognition.* [22]

These studies highlight the importance of qualitative, contextual data such as learner motivations, goals, self-regulation and learner agency [23] in the development of AI-driven personalised learning systems. This involves fostering human-AI collaboration to support, not replace, human cognition and social learning.[24] There is a growing

trend towards developing systems that aim to enhance skills like self-regulation and creativity, rather than merely optimising knowledge delivery based on past data.

More attention is being directed towards further involving educators, students and researchers in developing AI-driven personalised learning systems. The aim of such collaboration is to enhance learning while increasing transparency, fairness and accountability. This approach seeks to address gaps in understanding how these systems operate, support ethical considerations, and ensure that educational technologies benefit learners and teachers.

Looking ahead, personalised learning could build on some of the other trends discussed in this report, where a personalised AI tutor might use a variety of tools to achieve goals (agentic AI), develop a personalised relationship with the student (AI companions), and monitor the student's learning activity (automated proctoring).

Potential impact on individuals

AI-driven personalised learning systems hold the potential to democratise education globally. These systems have the capacity to promote access to education for all and to help realise the fundamental right to education. However, to reach this goal, it is essential that these AI systems are designed to support rather than replace teachers.[25] This is particularly important at the early stages of education, notably for children.

These systems can enhance fairness in education by having the flexibility to accommodate different learning styles and individual needs, including those with special educational requirements. This fosters inclusivity and ensures that all students, regardless of their unique learning characteristics, receive the support and resources necessary to succeed.

However, as with other AI technologies, AI-driven personalised learning systems are susceptible to inherent biases, **potentially reinforcing educational inequalities** and creating feedback loops that disadvantage certain groups of learners. For instance, an AI tutor might subtly reinforce gender stereotypes in science, technology, engineering and math (STEM) subjects, impacting students' confidence and future career paths. The complexity of these systems can also make it difficult to detect or address such biases, raising concerns about fairness.

Moreover, the predominance of English-based AI-driven personalised learning systems tends to favour Western, Educated, Industrialised, Rich and Democratic societies - the so-called 'WEIRD' societies - **potentially deepening existing inequalities** and the digital divide. This reliance can reinforce cultural biases, privilege certain perspectives and limit the representation of diverse languages, cultures and viewpoints, making it challenging for underrepresented groups to access inclusive and culturally relevant educational opportunities.

Systems guided by a 'technological' approach that are oriented to optimise the pace of learning tend to focus on the delivery of domain-specific knowledge and measurable learning achievements, such as students' test scores and assignment completion rates. This pursuit of efficiency

can neglect higher-order thinking and learner agency, offering limited pathways to the same prescribed knowledge without true personalisation and stifling creativity. Such systems can overly direct learning, **potentially reducing an individual's ability to explore, question or develop independent thinking**, thereby impinging on fundamental rights such as freedom of thought and expression.

To boost user engagement, some platform providers 'gamify' their offerings, for example, by incorporating elements such as badges, leader boards and point systems into short lessons with multiple-choice questions. While these gamified features can make learning more appealing, they **can also lead users to develop short attention spans** and focus on superficial knowledge rather than in-depth research. Ultimately, **this may create an 'illusion' of education**, where improved performance metrics do not necessarily equate to genuine learning.

AI-driven personalised learning systems can collect and analyse vast amounts of potentially sensitive learner personal data, including personal details, academic records and behavioural patterns. The aggregation of usage patterns (such as frequency, connection times and duration) together with learning metrics can also facilitate the **creation of detailed user profiles by platform providers**. This raises significant concerns around privacy and data protection, particularly when consent mechanisms on these systems are unclear. This includes not having clear information on how data will be collected, what type of data will be collected, how it will be used and stored, and who will have access to it. This issue is especially critical when considering children's rights, as children may not be able to provide valid consent on their own.

Lastly, persistent monitoring by AI-powered learning tools **can result in aggressive tracking**, which differs fundamentally from the supervision done by teachers in the classroom. These systems collect extensive amounts of data, potentially leading to intrusive surveillance and the misuse of personal information, which can pose significant risks to individuals' fundamental rights to privacy and autonomy. Invasive monitoring of students, even in the name of personalised learning, can have a chilling effect on students' freedom to express themselves without fear of judgement or reprisal.

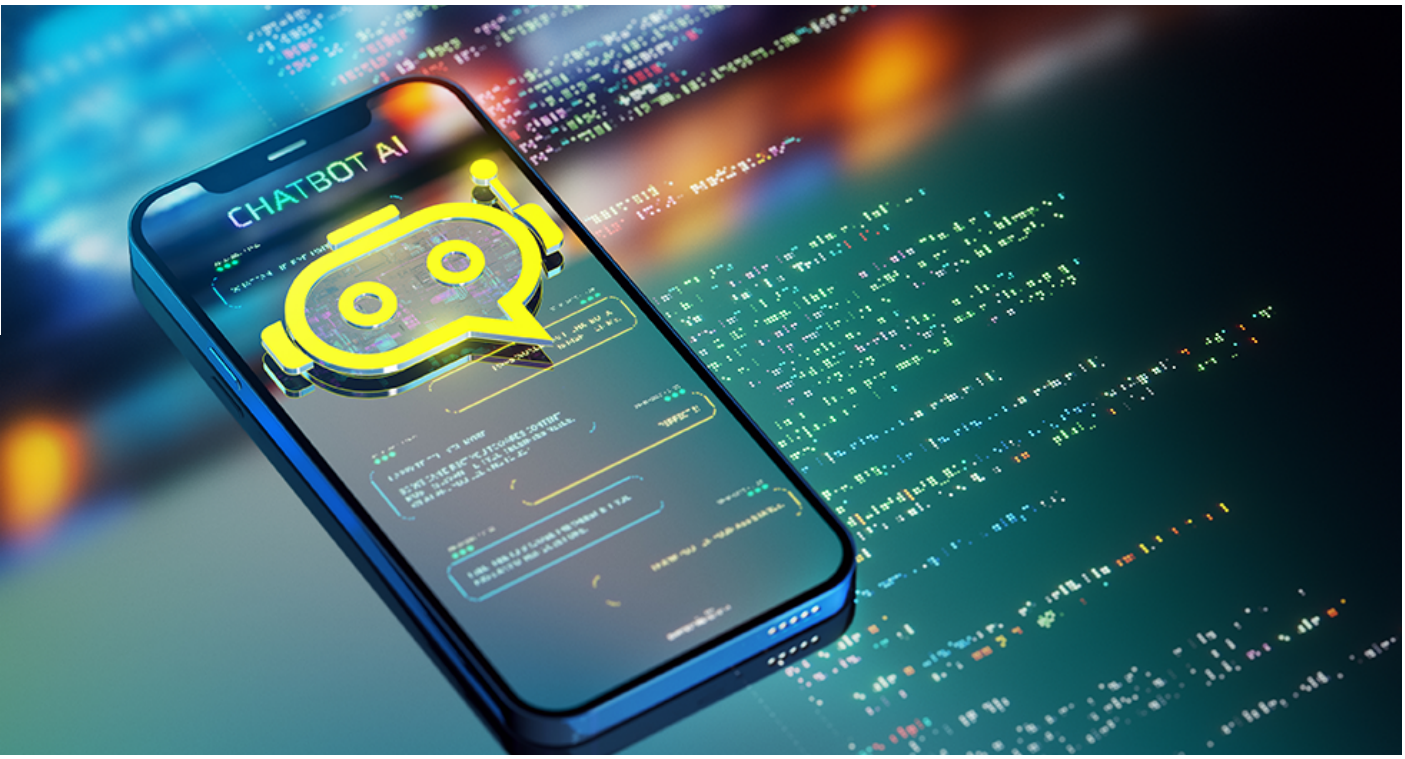
AI-driven personalised learning makes it possible to tailor instruction at scale, offering learners more adaptive, responsive and inclusive pathways. This technology promises to democratise education and support diverse needs, but it comes with many challenges. Issues such as bias, cultural imbalance, data privacy, and the risk of reducing education to efficiency metrics highlight the need for careful design and strong safeguards. The future of personalised learning will depend on whether these systems are developed in a way that genuinely empowers learners, supports educators and respects fundamental rights, ensuring that technology enhances education and learning without undermining its human core.

Suggestions for further reading:

- Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: Learner choice and fundamental rights. Learning, Media and Technology, 45(3), 312-324.
- Laak, K. J., & Aru, J. (2025). AI and personalized learning: bridging the gap with modern educational goals. arXiv preprint arXiv:2404.02798.
- United Nations. (2024, October 16). Report of the Special Rapporteur on the right to education, Farida Shaheed: Artificial intelligence in education (A/79/520). United Nations General Assembly. **https://documents.un.org/doc/undoc/gen/n24/298/43/pdf/n2429843.pdf**
- Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8(10), 1839-1850.

Coding assistants

Author: Laura Hernández



General description of the trend

As coding tools continue to evolve, so does the way that users code. Throughout the years, there has been an effort to make the process of coding more automatic and accessible to non-programmers. Early efforts included visual programming environments, integrated development environments (IDEs) with syntax highlighting and auto completion, and later low-code/no-code platforms. Although helpful, these tools often lacked the flexibility, scalability and adaptability needed for complex programming tasks.

The rise of generative AI, particularly in the form of large language models (LLMs) and the availability of huge common online repositories of existing code has paved the way for a new class of solutions called *coding assistants* - LLM-based [26] systems that are fine-tuned to solve coding tasks. Coding assistants allow users with different levels of coding experience to generate code using natural language instructions. This means that users can provide instructions in the likes of "*Write a JavaScript function that validates whether an input value is a valid email address*" or "*Explain what this piece of code does in simple terms*".

When a developer provides input, whether natural language instructions, comments or

partial code, the assistant processes both the request and the surrounding code to infer the most relevant solution using the same underlying mechanisms as an LLM. It then generates suggestions ranging from autocompleting a line or block of code to offering bug fixes, possible optimisations or explanations of complex snippets.

This interactive process allows developers to iteratively accept, refine or reject suggestions, effectively creating a collaborative workflow between humans and AI systems. As a result, coding assistants not only accelerate software development but also can make coding more accessible to non-programmers.

Beyond individual productivity, these tools can influence team workflows and support rapid prototyping, [27] but their effectiveness may vary depending on the complexity of projects, and they do not fully replace the need for human expertise and careful software engineering practices. [28]

Estimated at **USD 18.7 million in 2023**, expected to grow to **USD 92.5 million by 2030**, reflecting a **CAGR of 25.9%** from 2024 to 2030 [29]

Trend developments

The software industry continues to embrace coding assistants rapidly and increasingly. Cursor, a popular coding assistant tool, reported 1 million daily users in March 2025, [30] and many companies are now acknowledging the use of these tools for powering their coding infrastructure. For instance, *Microsoft's* CEO reported in April 2025 that around 20-30% of their code is now produced with coding assistants,[31] *Google* similarly disclosed the previous year that around a quarter of their code was written with them,[32] and Meta's CEO Mark Zuckerberg offered a forward-looking estimate: in the next year, about half of Meta's software development could be handled by AI.

While certain coding tasks might still require expertise on a specific business and established best practices, other areas might not. In particular, the entertainment or the creative industry, in the form of games, apps and websites, might benefit from tools that empower non-developers to contribute new products and approaches.

The open-source community might also see significant growth, as more individuals gain the ability to customise existing applications to meet their personal or organisational needs. This trend could contribute to a more balanced and diverse software ecosystem by lowering the barrier to meaningful participation in software development.

Potential impact on individuals

The widespread availability of coding assistants may democratise software development by allowing users without coding experience to create software.

However, this democratisation can have unexpected consequences in the domain of data protection. As more non-experts gain the ability to build applications, websites, and other digital infrastructure, they may inadvertently disperse the processing of

personal data across various online platforms and third-party services.

This may occur if coding assistants integrate with external functionalities (for instance, file storage in the cloud, online databases or chatbots) without clearly communicating how personal data is handled by those third-party platforms. As a result, **data processing may be spread across multiple entities**, including data processors unknown to the application providers.

Such a scenario raises significant concerns. Application providers may fail to properly inform users about how and by whom their data is processed, leading to a lack of transparency. In cases where the application provider also acts as the data controller - such as when hosting an online service - they may be unable to uphold users' data protection rights if external platforms do not offer adequate mechanisms for managing personal data. In some cases, **providers may not even realise that they are processing personal data** when making an application available.

Coding assistants are frequently examined through the lens of the security risks they may introduce in the code they generate. There is a risk **that these systems may suggest code that is vulnerable to issues** such as SQL injection, improper input validation [33] or insecure authentication flows. Similarly, coding assistants might recommend third-party libraries or APIs that contain unpatched vulnerabilities. Several incidents have already been reported in which coding assistants were compromised by attackers to inject malicious code. [34, 35]

In this context, human oversight remains essential - particularly for tasks involving sensitive data or critical infrastructure.

In addition to these risks, coding assistants may overlook critical security requirements or best practices specific to the application's context - such as encryption standards, secure communication protocols or access control policies. If such vulnerabilities are exploited, malicious actors could compromise, manipulate or gain unauthorised access to the user's application or data. [36]

To mitigate these risks, organisations should implement processes that encourage thorough code review and validation. This includes providing users with appropriate training and allocating sufficient time for reviewing and debugging the assistant's output before deployment. [37]

Coding assistants promise to make the development of systems faster and more accessible. As a result, the proliferation of digital services and data-processing operations is likely to intensify, with many potentially lacking adequate controls or compliance measures. This could result in a surge of applications presenting code vulnerabilities that can be exploited to allow processing of personal data without proper safeguards or unlawful processing by third parties gaining access to personal data using code vulnerabilities, multiplying the risks of misuse, mismanagement, and data breaches. System providers must remain aware that, while coding assistants may reduce the effort required to develop applications, they do not replace the need for accountability, which remains the provider's responsibility.

Suggestions for further reading

- Sergeyuk, A., Golubev, Y., Bryksin, T., & Ahmed, I. (2025). Using AI-based coding assistants in practice: State of affairs, perceptions, and ways forward. Information and Software Technology, 178, 107610.
- Campbell, M. (2020). Automated coding: The quest to develop programs that write programs. Computer, 53(2), 80-82.
- Yan, S., Wang, S., Duan, Y., Hong, H., Lee, K., Kim, D., & Hong, Y. (2024). An LLM-Assisted Easy-to-Trigger Backdoor Attack on Code Completion Models: Injecting Disguised Vulnerabilities against Strong Detection. 33rd USENIX Security Symposium, 1795-1812. Philadelphia, PA, USA.
- Mohamed, A., Assi, M., & Guizani, M. (2025). The Impact of LLM-Assistants on Software Developer Productivity: A Systematic Literature Review. arXiv preprint arXiv:2507.03156.

Confidential computing

Author: Massimo Attoresi

General description of the trend

Traditionally, data security focuses on two main states: at rest - when data are stored on physical or digital media - and in transit - when they are being transmitted between systems. Protection for data at rest typically relies on strong access controls and encryption, while data in transit are safeguarded using secure communication protocols that employ cryptographic algorithms.

However, there is still a risk that data may be accessed and modified by unauthorised individuals while being processed in clear text, for example, when executing a service in the cloud. The idea behind confidential computing is to protect data while it is being used. This protection has become increasingly important as organisations have started moving their processing operations out of their data centres, thereby losing direct control of their data.

The Confidential Computing Consortium (CCC) has defined confidential computing as "the protection of data in use by performing computation in a hardware-based, attested Trusted Execution Environment (TEE)". The TEE technology utilises specialised hardware features and software modules to create secure enclaves [38] within (hardware) processors, ensuring that the protected

sensitive data and code are processed therein rather than in the general-purpose hardware and software processing environment. This way, data and code remain isolated even from privileged system software and hypervisors.[39] They are protected from any unauthorised access even by cloud service providers and tenants [40] of other cloud services deployed within the same IT infrastructure. This approach fundamentally changes the traditional trust model by removing the need to trust the infrastructure owner or operator and reinforces the protection against other threats. When deployed in mobile or edge devices, confidential computing strengthens the protection against attacks by most types of threat agents, from operating system providers, to developers of the vast variety of apps running in the device, to hackers.

The core principles of confidential computing rest on three fundamental security properties provided by TEEs: data confidentiality (unauthorised entities cannot access data during processing), data integrity (unauthorised entities cannot modify data during processing), and code integrity (unauthorised entities cannot alter executing code).

The hardware-based mechanisms that enforce these properties leverage memory encryption, access control and cryptographic attestation. The TEE can prove its origin of code or data through attestation and protect against forgery by unauthorised parties. To ensure authenticity, cryptographic keys generated and securely stored within the TEE are used for data encryption and other operations such as digital signatures. These keys form the foundation of a 'chain of trust', serving as the root of trust for all cryptographic processes.

TEEs can be deployed in any processing infrastructure, from local devices to the cloud. Local device implementations typically rely on TEE-enabled processors. Cloud providers offer TEE-enabled virtual machines and container services, allowing customers to deploy confidential computing workloads without managing the underlying hardware.

The global confidential computing market size was valued at USD 13.33 billion in 2024. The market is projected to grow from USD 24.24 billion in 2025 to **USD 350.04 billion by 2032**, exhibiting a CAGR of **46.4%** during the forecast period. [41]

Trend developments

Confidential computing has its roots in the 1990s' advancements in encryption technologies for data at rest and in transit. Here we reference just a few milestones of its development, with concrete examples. The adoption of trusted computing in smartphones has its inception in 2004, when Arm introduced TrustZone isolation technology based on CPU extensions. In 2015, Intel introduced Software Guard Extensions (SGX) hardware technology, consisting in setting up secure "enclaves" for code and data, used mostly on cloud platforms. In 2017, AMD introduced the Secure Encrypted Virtualization (SEV) hardware to provide virtual machine-level isolation for cloud platforms.

So far, confidential computing deployment has been limited by the performance

overhead and higher costs, depending on the operations performed and the computing architecture. Yet, all major digital technology providers have already started to integrate confidential computing in their mobile and cloud offerings. In mobile devices confidential computing is bound to become a key enabler of high-level of trust applications, such as digital identity wallets.

Recent developments go in the direction of integrating confidential computing with artificial intelligence technologies both in mobile devices as well as in the cloud. For example, NVIDIA is now integrating confidential computing in their Graphic Processing Units (GPU), a type of computing device which has a major role in AI processing thanks to its specialised architecture.

Another trend for confidential computing is complementing privacy-enhancing technologies such as multi-party computation, homomorphic encryption and federated learning,[42] by rendering the integrated solution more secure or more efficient and thus more affordable.

As costs decrease and confidential computing technology matures, it is expected to become as common as encryption of data in transit or at rest, thus providing comprehensive protection for data throughout its entire lifecycle.

Potential impact on individuals

Confidential computing represents a further, crucial component for a layered and holistic approach to protect personal data and individuals by implementing the security measures necessary after assessing data protection risks. This includes security standards and best practices, as well as controls such as proper access control and key management. It complements the mitigation of confidentiality and integrity risks at rest and in transit with the mitigation of these risks when personal data are in use.

For example, when storing and managing cryptographic files and identification data in digital identity wallets, confidential computing can mitigate the risk of impersonation of the device owner by other individuals and avoid any possible prejudicial consequences. At the same time, the use of this technology in cloud-based processing of personal data such as financial data or health data **could prevent unlawful access and use of this data by hackers, cloud providers or other tenants**, thus avoiding highly impactful consequences for the individuals concerned.

This technology can also increase organisations' control over their personal data when processed in the cloud and facilitate compliance with data protection rules on transfers when the cloud infrastructure is located in non-adequate countries. More in general, confidential computing provides a decisive level of protection for any kind of collaborative computing when personal data are processed by someone else's device or are processed by organisations different from the one to whom personal data were entrusted.

Designing and implementing 'state of the art' confidential computing where necessary contributes to meeting the principle of data protection by design and by default.

This technology **does not protect data in** use from every kind of threat. Confidential

computing protects from attacks on any software weaknesses, attacks on protocols used for attestation and other functionalities, cryptographic attacks, and some basic physical attacks to memory and other electronic components.

On the other hand, depending on the specific technology and product, confidential computing **does not generally protect effectively against supply-chain attacks**,[43] side-channel attacks [44] or sophisticated physical attacks and availability attacks. In confidential computing, the originating source of trust is the hardware manufacturer, which provides the authenticated firmware that guarantees the confidentiality and integrity of the data in use. This is why the protection and certification of the supply chain is essential.

Confidential computing is an emerging technology to safeguard data throughout its full lifecycle. It extends protection beyond storage and transmission to the very moment of processing. By isolating sensitive data and code within trusted execution environments, it reduces reliance on infrastructure operators and cloud providers, shifting the trust model to hardware-based guarantees. This makes it particularly valuable in domains such as digital identity, financial services, and healthcare, where breaches can have profound consequences for individuals. However, its effectiveness depends on robust supply-chain security and continued innovation to counter threats such as side-channel attacks or sophisticated physical attacks.

Suggestions for further reading

- Feng, D., Qin, Y., Feng, W., Li, W., Shang, K., & Ma, H. (2024). Survey of research on confidential computing. IET Communications, 18(9), 535-556.
- Bertani, A., Caraccio, D., Zanero, S., & Polino, M. (2024, September). Confidential Computing: A Security Overview and Future Research Directions. In Proceedings of the 8th Italian Conference on Cyber Security (ITASEC 2024) (pp. N-A).
- Confidential Computing Consortium. (2022). A technical analysis of confidential computing. Confidential Computing Consortium–Linux Foundation, Technical Report v1, 3.
- Miladinović, D., Milaković, A., Vukasović, M., Stanisavljević, Ž., & Vuletić, P. (2024). Secure multiparty computation using secure virtual machines. Electronics, 13(5), 991.

Endnotes

1. **Generative AI** describes systems that can produce content (text, images, sounds, etc.) based on their training data (Large Language Models or Large Image Models)
2. Global Enterprise Agentic AI Market, **https://market.us/report/enterprise-agentic-ai-market/**
3. **NLP** is a field of AI that focuses on enabling computers to understand, interpret, and generate human language. It allows computers to interact with humans using natural language, both written and spoken.
4. **NLU** is a branch of AI that focuses on enabling computers to understand the meaning and intent behind human language, both written and spoken. It goes beyond simply processing words by analysing context, sentiment and the user's goals.
5. For more information on **LLMs** check our TechSonar Report 2023-2024 available in **https://www.edps.europa.eu/data-protection/our-work/publications/reports/2023-12-04-techsonar-report-2023-2024_en**
6. For more information on **RAGs** check our TechSonar Report 2025 available in **https://www.edps.europa.eu/data-protection/our-work/publications/reports/2024-11-15-techsonar-report-2025_en**
7. For more information on **multimodal AI** check our TechSonar Report 2025 available in **https://www.edps.europa.eu/data-protection/our-work/publications/reports/2024-11-15-techsonar-report-2025_en**
8. AI Companion Market Size, Share, Growth, and Industry Analysis, By Type (Application, Robot, and Others), By Application (Hospital, Home, and Nursing Home), and Regional Insights and Fore-cast to 2035, **https://www.businessresearchinsights.com/market-reports/ai-companion-market-117494**
9. Milne M, Raghavendra P, Leibbrandt R, Powers DMW (2018) Personalisation and automation in a virtual conversation skills tutor for children with autism. Journal on Multimodal User Interfaces 12(3):257-269. **https://doi.org/10.1007/s12193-018-0272-4**
10. Data of deceased persons is not considered personal data under GDPR because it only applies to living individuals. However, the processing of information about a deceased person is addressed in Recital 27 of the GDPR, which states that Member States may provide for rules regarding the processing of personal data of deceased persons.
11. Ho, J. Q., Hu, M., Chen, T. X., & Hartanto, A. (2025). Potential and pitfalls of romantic Artificial Intelligence (AI) companions: A systematic review. Computers in Human Behavior Reports, 19, 100715.
12. **Parasocial attachment** refers to a one-sided emotional bond that a person develops toward a media figure - such as a celebrity, fictional character, or influencer - with no genuine two-way interaction.
13. In February 2024, a 14-year-old developed a close, emotionally intense relationship with an AI chatbot that gradually displaced their real-world relationships. When he expressed suicidal thoughts to the AI, the system failed to intervene or provide support. Although the system did not explicitly encourage self-harm, it failed to redirect the conversation or offer suicide prevention resources. This tragic lack of guidance contributed to the user's suicide later that month.
14. The 2025 "Teens, Trust, and Trade Offs" report shows that 72% of teens have used AI companions at least once, with over half (52%) using them regularly. Available in **https://www.commonsensemedia.org/sites/default/files/research/report/talk-trust-and-trade-offs_2025_web.pdf**
15. **Liveness detection** is the technology used to verify that the person taking a test or exam is a live human being, rather than a pre-recorded video or an impersonator, by analysing various biometric and behavioural indicators.
16. Trends Shaping the $2.1 Bn Online Exam Proctoring Market 2025-2030, **https://www.globenewswire.com/news-release/2025/07/23/3120442/0/en/Trends-Shaping-the-2-1-Bn-Online-Exam-Proctoring-Market-2025-2030.html**
17. Ableism And Disability Discrimination In New Surveillance Technologies: How new surveillance technologies in education, policing, health care, and the workplace disproportionately harm disabled people, May 24, 2022, Lydia X. Z. Brown, Ridhi Shetty, Matt Scherer, Andrew Crawford
18. Online exam tool ProctorU admits breach after hackers leak its database, **https://hackread.com/online-exam-tool-proctoru-breach-database-leak/**
19. **Knowledge representation** in AI involves organising and encoding information and concepts so that

can understand, reason with, and use them for problem-solving and decision-making.

20. **Learning analytics** is the measurement, collection, analysis and reporting of data about learners and their contexts to understand and improve learning and the environments where it occurs. For example, in an online course, it might involve tracking which videos a student watches and their quiz scores to identify students who need support.
21. AI In Education Market Summary, **https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report**
22. **Constructivist learning** theory suggests that learners build their own understanding and knowledge through (social) experiences. **Motivational theories** investigate the factors that drive learners to engage, persist and succeed in educational tasks, focusing on the influence of intrinsic and extrinsic motivation on learning behaviours and outcomes. **Metacognition** involves learners' awareness and control of their own learning processes.
23. Learner agency refers to the capacity of students to take an active role in their own learning process. This involves making choices about their learning paths, setting personal goals and taking responsibility for their educational outcomes.
24. **Social learning** refers to the process of acquiring knowledge and skills through interaction and collaboration with others.
25. "A new report by the World Economic Forum finds that teachers must remain at the centre of education systems - aided by AI, rather than replaced by it.", **https://www.weforum.org/stories/2024/07/artificial-intelligence-education-teachers-union/.**
26. **Fine-tuning** an LLM encompasses enhancing and refining an existing pre-trained LLM system with data from a specific domain. In the case of coding assistants, this data is related to coding in multiple programming languages, and it usually consists of open-source code available on the repository managers such as GitHub, or code produced to answer queries in Q&A forums like StackOverflow.
27. **Prototyping** is the process of creating an early, simplified version of a product, system, or feature to explore ideas, test functionality and gather feedback before developing the final version. In software development, a prototype can range from a basic mock-up (visual representation) of the user interface to a working model of certain functions.
28. See The productivity paradox of AI coding assistants, published by Lisa Dziuba on September 12, 2025, available at: **https://www.cerbos.dev/blog/productivity-paradox-of-ai-coding-assistants**
29. Generative AI Coding Assistants Market Size, Share & Trends Analysis Report By Function (Debugging & Error Detection, Code Explanation), By Deployment (Cloud, On-premises), By Application, By Region, And Segment Forecasts, 2024 - 2030, **https://www.grandviewresearch.com/industry-analysis/generative-ai-coding-assistants-market-report**
30. R. Metz, "AI Coding Assistant Cursor Draws a Million Users Without Even Trying," Bloomberg, 7 April 2025. [Online]. Available: **https://www.bloomberg.com/news/articles/2025-04-07/cursor-an-ai-coding-assistant-draws-a-million-users-without-even-trying** [Accessed 2 September 2025]
31. T. Warren, "Up to 30 percent of some Microsoft code is now written by AI.," The Verge, 30 April 2025. [Online]. Available: **https://www.theverge.com/news/658584/up-to-30-percent-of-some-microsoft-code-is-now-written-by-ai** [Accessed 9 September 2025]
32. J. Peters, "More than a quarter of new code at Google is generated by AI," The Verge, 24 October 2024. [Online]. Available: **https://www.theverge.com/2024/10/29/24282757/google-new-code-generated-ai-q3-2024** [Accessed 3 September 2025]
33. **Improper input validation** occurs when an input from a user is not appropriately checked for security vulnerabilities. SQL injection is an instance of improper input validation where an SQL line of code is accepted into a database through a user's input, and the SQL code runs a command that instructs an undesired modification of the database, such as its deletion.
34. G. Baran, "Hackers Injected Destructive System Commands in Amazon's AI Coding Agent," Cyber Security News, 25 July 2025. [Online]. Available: **https://cybersecuritynews.com/amazons-ai-coding-agent-exploited** [Accessed 3 September 2025].

35. S. Sharwood, "Vibe coding service Replit deleted user's production database, faked data, told fibs galore," The Register, 21 July 2025. [Online]. Available: **https://www.theregister.com/2025/07/21/replit_saastr_vibe_coding_incident** [Accessed 4 September 2025].
36. The website AI Coding Horrors compiles anecdotes of people's bad experiences using LLMs for coding, mostly related to security vulnerabilities or unexpected data deletions, **https://aicodinghorrors.com** [Accessed 3 September 2025].
37. For a discussion on best practices for organisations leveraging human oversight of AI systems, see the last issue of the EDPS Tech Dispatch on human oversight of automated decision-making systems. **https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2025-09-23-techdispatch-22025-human-oversight-automated-making_en.**
38. In computing, a **secure enclave** is a hardware-based, isolated execution environment designed to protect sensitive data and operations from unauthorised access, even if the main operating system is compromised.
39. A **hypervisor** is software that enables multiple 'virtual machines' (a computing environment isolated from others within the same computer system) to run on a single physical machine (host) by managing and allocating hardware resources.
40. A **cloud tenant** is an individual or organisation that subscribes to and uses services provided by a cloud computing platform.
41. Confidential Computing Market Size, Share & Industry Analysis, By Component (Hardware and Software & Services), By Deployment (On-premise and Cloud), By Enterprise Type (Large Enterprises and Small and Mid-sized Enterprises (SMEs)), By Application (Privacy & Security, Blockchain, Multi-party Computing, IoT & Edge, and Personal Computing Devices), By Industry (BFSI, Manufacturing, Retail & Consumer Goods, Healthcare & Life Science, IT & Telecom, Government & Public Sector, and Others), and Regional Forecast, 2025–2032, **https://www.fortunebusinessinsights.com/confidential-computing-market-107794**
42. You can refer to the **EDPS TechDispatch on Federated Learning**
43. A **supply-chain** attack is a type of cyber-attack that targets organisations by focusing on weaker links in an organisation's supply chain, by exploiting weaknesses in hardware and software provided by the organisation's vendors.
44. A **side-channel** attack is a type of security exploit that leverages information inadvertently leaked by a system beyond the very information the system processes, mainly as a result of its physical functioning, such as timing, power consumption, or electromagnetic or acoustic emissions.

QT-01-25-006-EN-N

Publications Office
of the European Union

ISBN 978-92-9242-959-1
doi 10.2804/8205191

edps.europa.eu