



RIIKKA FORMULAHTI

The image and integration of artificial intelligence in society

DEGREE PROGRAMME IN DATA ENGINEERING
2025

ABSTRACT

Formulahti, Riikka: The image and integration of artificial intelligence in society
Bachelor's thesis
Data Engineering
April 2025
Number of pages: 50

The objective of the thesis was to examine the image of artificial intelligence in society, the fears associated with it and the extent to which these perceptions are based on reality. The research was conducted through a review of recent literature, industry reports and academic studies on the applications, biases and societal impacts of AI technologies. The analysis focused on several fields, including healthcare, creative industries, cybersecurity, sustainability and ethics.

The results indicated that while artificial intelligence offers significant advancements across various sectors, societal fears are often a blend of realistic concerns such as data privacy risks, workforce changes and in misconceptions fuelled by media portrayal. The thesis concluded that critical awareness, responsible development and transparent governance are essential in fostering public trust and mitigating unfounded fears related to artificial intelligence.

Keywords: Artificial intelligence, AI, coding, society, ethics, bias

CONTENTS

1	ARTIFICIAL INTELLIGENCE IN SOCIETY.....	6
1.1	Types of AI	7
1.1.1	Automation, machine learning and machine vision.....	8
1.1.2	Natural language processing	9
1.1.3	Robotics and manufacturing	9
1.1.4	Self-driving vehicles and transportation	9
1.1.5	AI in business and finance.....	10
1.1.6	AI in education	10
1.2	Augmented intelligence	10
1.3	AI and the workforce.....	11
1.4	AI and bias	11
2	AI IN THE MEDIA AND CREATIVE ARTS.....	14
2.1	AI in apps and social media.....	14
2.2	AI in entertainment and content creation	15
2.3	AI in creativity and the arts	16
2.4	ChatGPT and conversational AI	16
3	AI IN HEALTHCARE AND HEALTHIER LIVING.....	18
3.1	Predictive healthcare and early diagnosis	18
3.2	Remote patient monitoring and AI-driven care	20
3.3	AI in medical imaging and diagnostics.....	20
3.4	Chatbots and AI in mental health	21
3.5	Smart living and AI-powered devices	23
4	AI WITH SECURITY AND CYBERSECURITY	24
4.1	AI in threat detection and prevention.....	24
4.2	Deepfakes and the spread of misinformation	25
4.3	AI-driven cybersecurity risks.....	26
5	AI AND SUSTAINABILITY	28
5.1	AI in climate change mitigation.....	28
5.2	AI for optimizing energy use	30
5.3	AI in agriculture and resource management.....	30
6	ETHICAL AND PHILOSOPHICAL IMPLICATIONS OF AI.....	33
6.1	Ethical dilemmas in AI development.....	33
6.2	Philosophical questions about AI consciousness	34
6.3	Human-AI interaction and trust.....	35
6.4	Ethical considerations in AI for healthcare.....	35

7 REGULATORY AND POLICY PERSPECTIVES ON AI	37
7.1 Governmental approaches to AI regulation	37
7.2 AI accountability frameworks	38
7.3 AI in military and defence	39
8 CONCLUSION	41
REFERENCES	42
FIGURES	49

LIST OF SYMBOLS AND TERMS

AGI - Artificial general intelligence

AI - Artificial intelligence

AWS - Autonomous weapons systems

CEA - Controlled-environment agriculture

CRM - Customer relationship management

GAN - Generative adversarial network

IoT - Internet of things

LMM - Large multi-modal model

NLP - Natural language processing

OECD - The Organization for Economic Co-operation and Development

RPA - Robotic process automation

SIEM - Security information and event management

WHO - World Health Organization

XAI - Explainable AI

1 ARTIFICIAL INTELLIGENCE IN SOCIETY

Artificial intelligence is a term since 1950's meaning a simulation of the intelligence possessed by humans as it is processed by computer systems and machines (Burns 2023). AI can be applied to things such as natural language processing, machine vision and speech recognition. Artificial intelligence is not one single programming language but the most popular ones currently with AI developers are languages called Python, Java, C++, R and Julia. Many companies use machine learning, which requires fitting software and hardware for machine learning algorithm training. AI system takes in big volumes of labelled training data fed to it, which it then analyses for patterns and correlations and makes predictions based on, learning to generate lifelike exchanges in the cases of chatbots or learning to recognize and identify objects from images. Generative AI can also create by itself things like music or images after it has been trained to do so.

There are at least four categories for the cognitive skills artificial intelligence programming focuses on, which are learning, self-correction, reasoning and creativity (Burns 2023). Learning aspect consists of getting data and going through it with algorithms to make it usable information. Self-correction is for fine-tuning algorithms so that they stay accurate. Reasoning aspect is for picking the algorithm needed for getting to the desired results and goals. Creativity aspect is using statistical methods, neural networks and other rule-based systems to create new contents such as text generation.

Deep learning, machine learning and artificial intelligence can often be used interchangeably but they do differ, AI being the so-called umbrella holding the other two terms under it (Burns 2023). Machine learning uses historical data for its algorithms to learn upon and benefits on there being a large volume of data for its use while as deep learning is a subset of machine learning based

on human brain structure, including things such as neural networks and gets used for things such as self-driving cars.

The importance of artificial intelligence is its potential in changing the way humans work, live and play (Burns 2023). In business environments AI has been used to automate tasks such as customer service, quality control and fraud detection and it has been shown to outperform humans in certain tasks. For tasks which are repetitive and detailed, having to do with analysing and numbers an AI tool can make quick work without many errors. Generative AI tools especially are becoming important in fields such as education, marketing and product design.

Drawbacks of using artificial intelligence currently include it being expensive to process the amounts of data it would require to properly train an algorithm (Burns 2023). AI can also be biased and discriminatory either by accident or on purpose so it requires checking and tuning. Advantages of AI include things such as increased productivity and reduced time spent on tasks, it can be consistent and personalized. While AI worker can work basically around the clock, it also can be expensive to run and create. Another concern has been that it eliminates human jobs, therefore increasing unemployment. Lots of big corporations such as Apple, Microsoft and Google use AI as key parts of their operations and for outpacing their competitors.

1.1 Types of AI

Artificial intelligence can be categorized into four different types, which are reactive machines, limited memory, theory of mind and self-awareness (Burns 2023). Task-specific reactive machines are things like chess playing system with no memory. Limited memory systems use their past experiences for their future decisions and include applications such as self-driving cars. Theory of

mind system can understand emotions and has social intelligence, it can predict human intentions and behaviour. Self-awareness gives consciousness but this type has not been reached yet.

AI Model Types

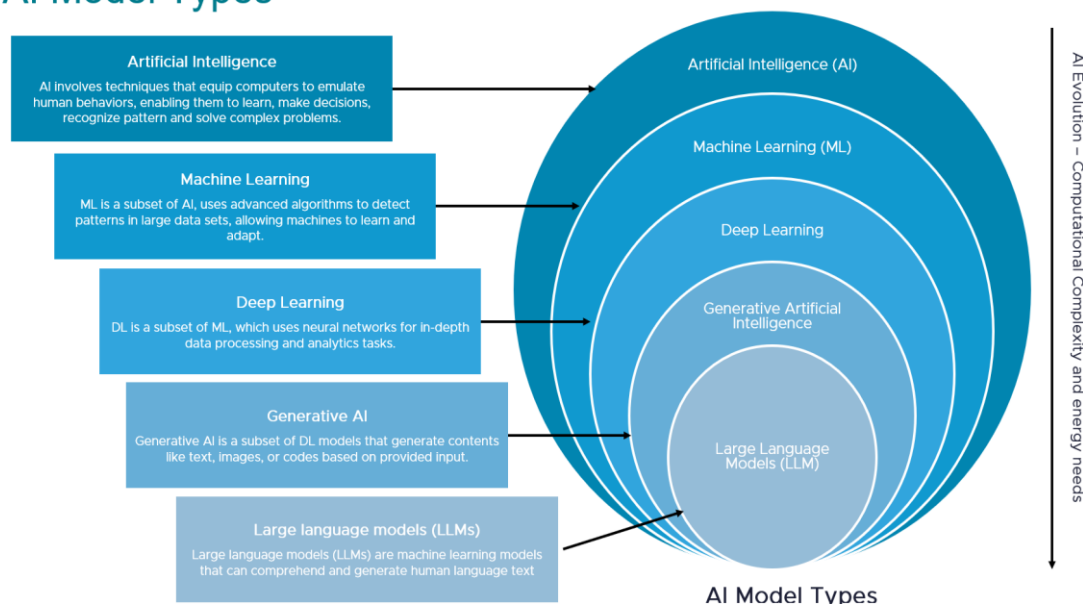


Figure 1. AI Model Types (Virtualizationteam.com 2024).

AI can be divided into five different model types as shown in Figure 1 (Virtualizationteam.com 2024). The types are artificial intelligence, machine learning, deep learning, generative artificial intelligence and large language models.

1.1.1 Automation, machine learning and machine vision

Robotic process automation, also called RPA, can be used to automate rule-based and repetitive tasks (Burns 2023). Machine learning algorithms consists of three types which are: unsupervised, supervised and reinforcement learning. Machine learning also has deep learning which can be described as automation of predictive analytics. Machine vision has to do with granting sight to a machine, letting it analyse and capture visual data through aids like cameras.

1.1.2 Natural language processing

Natural language processing, also known as NLP, is having a computer program process human language (Burns 2023). Spam detection in emails is a longer existing example of NLP, where the program can decide if the texts in the incoming email correspond with what is considered spam. NLP type programs are able to translate text, recognize speech and analyse sentiments.

1.1.3 Robotics and manufacturing

Robotics is a field of engineering which has to do with designing and manufacturing robots (Burns 2023). Robots are used generally for things where humans cannot produce consistent results and tasks that are otherwise difficult for humans to perform, such as NASA space robots moving large things in space. Robots are also getting their place in social interaction settings.

1.1.4 Self-driving vehicles and transportation

Artificial intelligence can make for smarter, more efficient and safer vehicles and it is used on self-driving cars (Rafalski 2025). AI can personalize in-car experiences, predict maintenance and offer assistance systems such as voice controlled infotainment and navigation. In self-driving systems AI helps with seeing where the road is and with decision making, also handling things such as lane-keeping and emergency braking, reacting faster than human is capable of. Self-driving uses the help of cameras, sensors and computers. It can park and follow traffic rules automatically and even work as taxis. Lidar and radar additions make the vehicle see in 3D and work even if cameras are not operational. AI cameras inside the vehicle can also track the behaviour or the driver, noticing distractions, drowsiness, impairment and the pull to the side of the road accordingly.

1.1.5 AI in business and finance

Customer relationship management, also known as CRM, uses machine learning algorithms to analyse and find out information on customers and their needs (Burns 2023). Chatbots can provide instantaneous customer service when added to websites. Generative AI technology is however expected to end up eliminating jobs and disrupting business models.

Artificial intelligence applications are used in the field of finance for giving financial advice and collecting personal data, and can also help with buying properties (Burns 2023). Wall Street trading is currently largely performed by AI software. Banks use AI for their decision making when giving out loans and setting credit limits, as well as identifying potential investments.

1.1.6 AI in education

Artificial intelligence can be used for automating grading tasks, freeing up time for educators (Burns 2023). AI tutors are used for guiding students on tasks and providing support, they can assess needs and adapt to them. AI also has the possibility of replacing human teachers depending on the subject and to create teaching materials and coursework. AI tools have also had an impact on homework and plagiarism policies and testing.

1.2 Augmented intelligence

Artificial intelligence as a term can be seen as too tied to popular culture, creating a certain raised image and expectations of its capabilities and use (Burns 2023). The term augmented intelligence has been suggested to be used for describing AI tools that are used to support humans while as artificial intelligence would be used for the more science fiction shaded examples in the field, especially when having to do with pop culture.

Augmented intelligence is supposed to showcase a neutral connotation, driving home the image that AI implementations are weak and improve services and products. True artificial intelligence, known as artificial general intelligence, also called AGI, is more concept and future based. AGI is considered mostly to be in science fiction as an artificial superintelligence does not yet exist to shape our reality. Quantum computing is expected to hasten the arrival of AGI.

1.3 AI and the workforce

AI is expected to replace many jobs in the year 2025, as much as 41% of companies having plans to reduce their workforce due to AI expansion (Robinson 2025). The emphasis is that AI is to be used as a tool for efficiency but nevertheless workers believe they will be replaced fully instead. Most concerns seemed to have arisen for the professions of programming, software development, accounting, lawyers, data analysts and doctors.

There's calculations that AI will be displacing 92 million people in the workforce but that it will also be creating 170 million new roles around years 2025 to 2030 (World Economic Forum 2025). The most transformative expected trend is the broadening of digital access overall and across technology related avenues. For technology the most growth is seen to be happening in artificial intelligence, information processing, automation and robotics and within these fields the effect will be both fastest growing roles and fastest declining roles for workers. Big data with AI, cybersecurity and networks as well as technological literacy will be the skills that grow the fastest. The importance of these skills is to be complimented with lifelong learning, flexibility, creative thinking, resilience, agility and curiosity.

1.4 AI and bias

Machine learning and algorithms have bias, which means the results they give can be skewed (IBM 2023a). The skewedness depends on the data the AI has

been fed in training and it might lead to harmful or disoriented output. Bias reduces accuracy and potential, hindering people's ability to be participating in the society and the economy. Biased AI breeds mistrust between people and can cause scandals. AI models absorb the biases already existing in our society, especially when they take historical data where certain types of people have been harmed. This then leads to further harm when AI is used in things like policing, hiring and credit scoring.

In healthcare using AI that has bias can lead to people who have underrepresented data – such as women – having much less accurate predictions and results (IBM 2023a). In law enforcement bias can lead to disproportionate targeting, racial profiling and patterns that work almost like a self-fulfilling prophecy. In generating images AI has been seen to make certain genders, ages and professions in a way that it thinks is typical according to the data it was fed, even though they didn't actually end up reflecting the real life due to the smaller portion size of the data it was given, so reality was in the end differing.

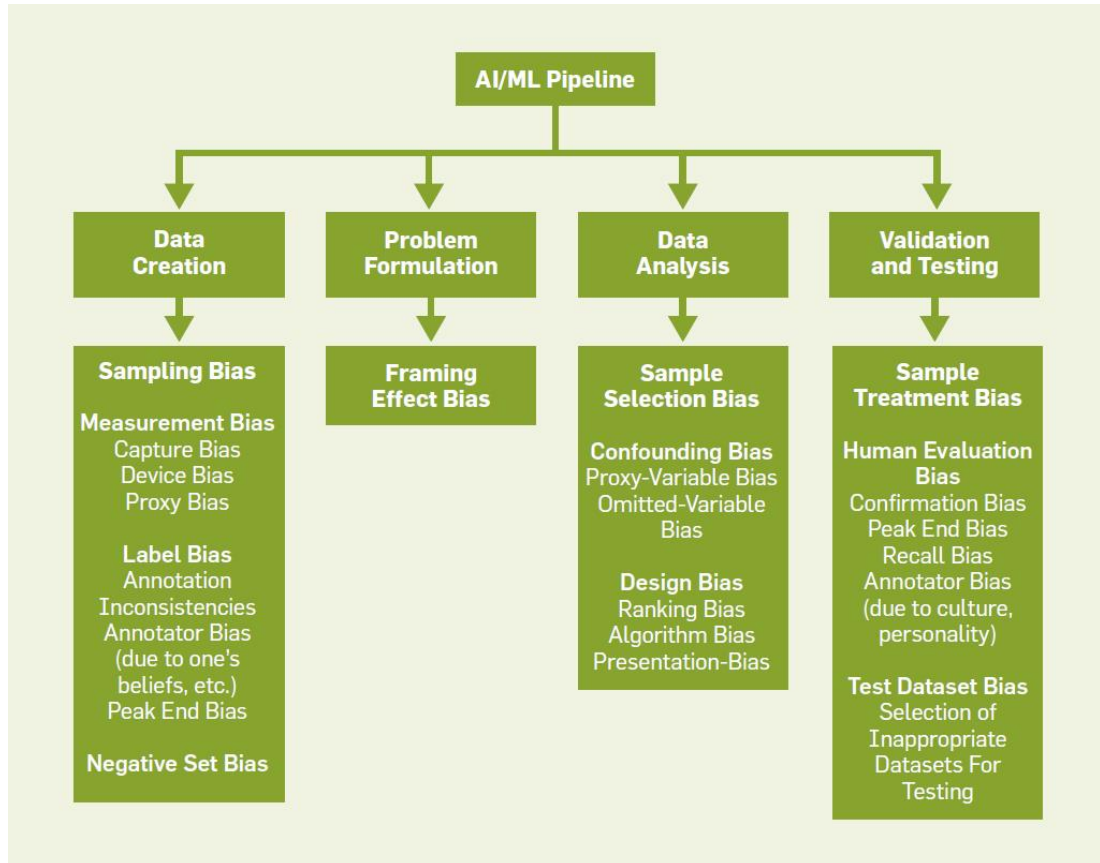


Figure 2. Biases in AI Systems (Srinivasan and Chander 2021).

Figure 2 showcases the AI/ML pipeline, where data creation leads to sampling bias, problem formulation to framing effect bias, data analysis to sample selection bias and validation and testing leading to sample treatment bias (Srinivasan and Chander 2021).

Algorithm bias occurs when the asked question is not correct or specific enough (IBM 2023a). Cognitive bias is when the input comes from humans and humans are naturally fallible and equipped with personal bias without even cognitively understanding they are. Confirmation bias has AI relying too strongly on data trends and pre-existing beliefs. Exclusion bias has important information missing and left out and measurement bias is because of incomplete data due to something like oversight.

Out-group homogeneity bias is not knowing what you don't know, people understanding within their limitations of their surroundings and growth (IBM 2023a). Prejudice bias has harmful societal stereotypes and recall bias is when labels are not correct in early build phase of the AI. Sample or selection bias has too small dataset to build on and leads to lacking representation. Stereotyping bias reinforces stereotypes in a harmful manner.

When it comes to mitigating bias in building AI, it is often more efficient to consider things beforehand than to fix them alter on (IBM 2023a). Sometimes financial decisions cause these models to come out in a rush and with errors. Direction, monitoring and management are key in finding out the biases in the product. Frameworks, policies and practises are supposed to guideline use and development of responsible AI. Fairness, inclusion and equity are to be assessed.

2 AI IN THE MEDIA AND CREATIVE ARTS

Artificial intelligence has been applied across industries, helping with reducing human errors, amount of work needed and going through large amounts of data and analysing it (Marsner n.d). Voice-based assistants are one of the main focuses in the industry adaptation process, especially in retail and automotive, while chatbots are a similar option in smaller scale. Apps and mobile devices are convenient mediums for AI deployment, offering things like personalized shopping, warehouse management and smart monitoring. The computing power of artificial intelligence has been doubling every 3.4 months.

Targeted advertising is one of the uses of artificial intelligence in entertainment business side (Burns 2023). AI can also help with distribution, recommending content, creating scripts and making movies. Automated journalism streamlines workflows, reducing complexity, time and money spent. AI can proofread news articles and make data entries, research topics and create headlines.

2.1 AI in apps and social media

Apple's voice assistant Siri uses AI in understanding habits and commands and Google Translate analyses language patterns for more accurate translations (Marsner n.d). Gmail has use for machine learning when filtering spam and unwanted messages. Amazon makes use of AI to give customers suggestions on their products and predict the types of things the customer would be interested in purchasing. Netflix uses AI for similarly suggesting its viewers content they might be interested in watching, basing it on metrics such as the current time of the day and viewer's previous interests and ratings, understanding the viewing habits. Google Maps also benefits having AI change route suggestions.

AI can be used by brands to perform analysis and do competitor profiles (Tiwari 2025). It can also work in real time, boosting social media strategies. Tasks

such as monitoring mentions can be automated, freeing workers from repetitive monotonous tasks, alerts can be arranged when trending goes negative to be quick with corrective brand image measures. AI as pattern recogniser can see trends that are only just arriving. Generating scripts, content ideas and even the actual publishable content can be done fast with AI, all the way down to generated images. Chatbots can reply to customers all hours of the day and content in general can be personalized, making it more engaging and effective.

2.2 AI in entertainment and content creation

Generative AI can create things such as scripts, stories, reviews and copies, marketing campaigns, images and videos, garnering a lot of attention and concern especially in Hollywood (Davenport and Bean 2023). The entertainment industry as a whole is facing financial pressures as well as the pressure of producing content in a timely manner. As it is an artistic field, many professionals in it see AI as something incapable of original creativity.

AI can even match actor's lip movements to a different language dubbed version (Crabree-Ireland 2025). Ethical and collaborative ways of integrating AI into the industry are being favoured by the people who and are forced to pick it up as a productivity tool. The Screen Actors Guild – American Federation of Television and Radio Artists (also known as SAG-AFTRA) after a long and very public process agreed about performers having their digital replicas be used but in a manner that grants the people their due compensation and rights. The person whose likeness is used will retain ownership of their AI voice model, at least on a specific agreed upon platform.

SAG-AFTRA comprises of around 116 000 talents in the entertainment fields such as television, games, broadcasting, music and commercials and managed in 2023 to negotiate guidelines and rules for the use of AI in entertainment (Crabree-Ireland 2025). Talents are to be given fair compensation and they are to give explicit informed consent for the AI use and creation of their

likeness. One of the points was to make sure that the people will not be losing their next employments to their digital versions.

2.3 AI in creativity and the arts

Ai-image generators such as DALL-E 2 and Midjourney sparked debate on whether art created by these algorithms can be considered real art (Mineo 2023). For writing there seems to be lack of real life experience that reflects from the generated products, also word choices and lack of linguistic originality make it distinguishable from something written by a human. When it comes to music, emphasis is put on emotion. For collaborative live music performance there needs to be curiosity. Film music done with AI seems to lack silences, emotions and surprise.

When it comes to animation, AI could learn an animator's style and be used like an assistant (Mineo 2023). That might threaten animator jobs and gloss over the creative process of making art. Some artists have gone to producing with more analogue techniques to differ clearly and people seem to think it would be best to be the guiding hand in control of how AI is to be used as a tool as it being there either way seems unavoidable. For things like architecture AI would need to consider a large variety of factors but it is also capable of being a useful tool for improving on designs but still lack a bit of magic.

Intellectual property infringement is still as issue, especially with AI creations (Appel et al. 2023). There's questions whether patent, copyright and trademark infringement apply. It might not be clear who owns what generative AI makes. There's also the big question of the training data that has been used to make AI models having been acquired without license.

2.4 ChatGPT and conversational AI

Conversational artificial intelligence comprises of things like chatbots and virtual agents and voice assistants that can be talked with (IBM 2025). They are

created from big amounts of data, using machine learning and natural language processing and are able to imitate human interaction. They recognize text and speech and can translate meanings in multiple languages.

Conversational AI is considered weak AI as its tasks are narrow and more focused, while as strong AI would have a human-like consciousness and be able to solve a wider variety of problems (IBM 2025). Even if considered weak, these AIs can perform money saving tasks for companies, such as online customer support around the clock, reducing barriers of entry by being an assistive technology, being part of internet of things devices such as Alexa speakers, Siri and Google Home. Conversational AI can be used in HR for orientation processes and employee training and even something as commonplace these days as autocomplete function can be an example of AI used in computer software.

Since late 2022 ChatGPT has been available to the public, being capable of lifelike human conversation (De Witte 2023). It can help writing cover letter, emails, analyse dating profiles and give advice although the results may vary. It has become increasingly difficult to distinguish what has been made with ChatGPT and what by a human. AI-mediated communication is still seen as a powerful tool or in assistant role. ChatGPT being open to public also means that it has users who do not have training to use it effectively or people who are not being ethical about the way they are using it.

3 AI IN HEALTHCARE AND HEALTHIER LIVING

Artificial intelligence in healthcare is mostly expected to reduce costs and to improve patient outcomes (Burns 2023). Machine learning can help with getting medical diagnoses faster and more accurately than if being diagnosed by a human. IBM Watson is a healthcare technology that can understand natural language and mine patient data for forming a hypothesis which it then confidence scores – aka evaluates the accuracy of. Chatbots and other virtual health assistants can help with scheduling appointments, finding medical information. AI can also be used for understanding and predicting pandemics.

3.1 Predictive healthcare and early diagnosis

Artificial intelligence was found to boost diagnostics, personalized care, efficiency, treatment planning, and patient outcomes significantly (Khalifa and Albadawy 2024b). Most significant helpful results were found in the fields of prognosis of disease course and outcomes, diagnosis and early detection of disease, risk assessment of future disease, disease progression, treatment response for personalized medicine, readmission risks, mortality prediction and complication risk. Radiology and oncology benefitted the most from the skills of an AI.

For AI to be better integrated into healthcare, there would need to be more interdisciplinary collaboration, accessibility of data of good quality, as well as investments in AI education and ethical AI practises (Khalifa and Albadawy 2024b). There would also need to be regulatory oversight, expanded clinical trials and involving patients in the AI integration process.

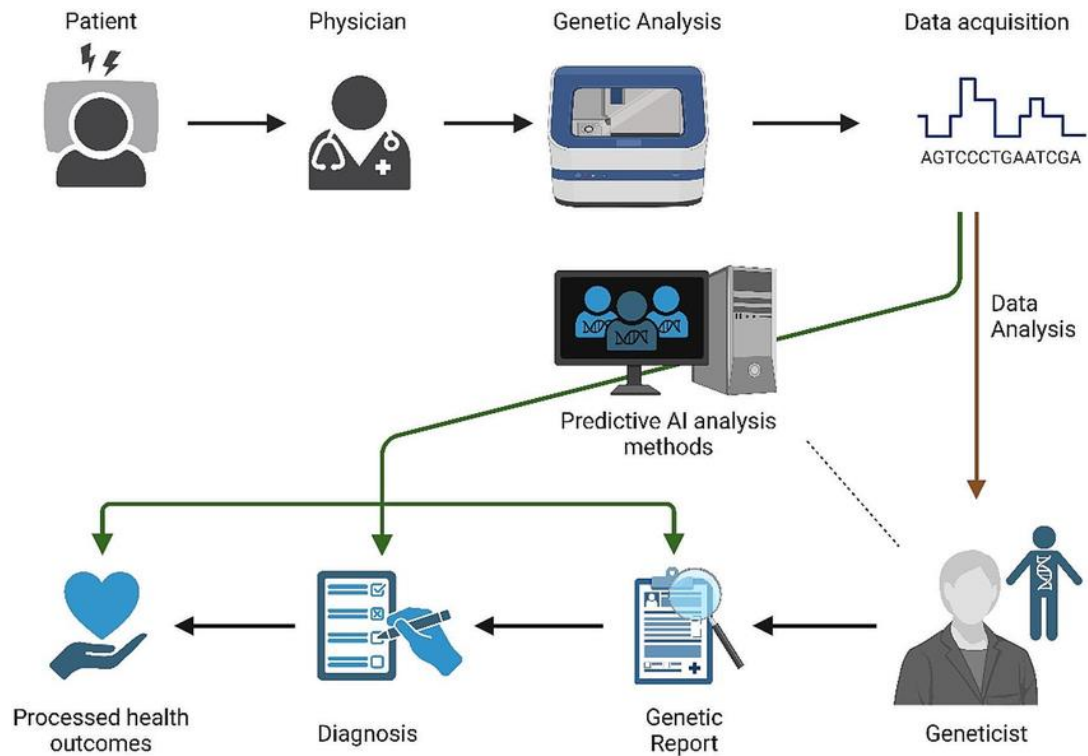


Figure 3. Diagnostic genetic testing workflow (Research Gate 2024).

Figure 3 showcases genetic testing flow starting from patient, moving to physician, genetic analysis, data acquisition where it can move through data analysis to predictive AI analysis methods and diagnosis or to geneticist, genetic report, again diagnosis step and then the ending which is processed health outcomes (Research Gate 2024).

Artificial intelligence is going to affect all aspects of biomedical research and medicine (Duong and Solomon 2025). The speed of these changes and the affect it will have on patients, clinicians, families and researchers is still unknown but expected to be rather dramatic. Factors such as logistic, ethical, financial, regulatory and legal will be playing a role. The amount of people working in genetics is limited, as is their knowledge of all the often rare and numerous genetic conditions which are multisystem and severe.

3.2 Remote patient monitoring and AI-driven care

The global market for wearable devices and technology is estimated to reach 265 billion dollars in 2026 (Deljkic 2024). Wearables are equipped with things like health monitoring, fitness features and virtual assistants. As a market the wearables are still immature and could benefit from the integration of AI in them, giving more performance and functionality. Each vendor collects their own versions of user data, having their own metrics and ways to process data. Health data especially can be sensitive information so there are concerns about privacy when it comes to data collection.

Advanced telemedicine technologies and AI are changing remote healthcare when it comes to diagnostics, patient care, monitoring and medical treatment (Chaturvedi et al. 2024). Changes have already been seen especially in real-time monitoring of health status, patient engagement and connectivity and accuracy of diagnostics. AI has been helpful in overcoming limitations of traditional models of healthcare, with things such as predictive analysis, diagnostic systems and teleconsultation platforms.

3.3 AI in medical imaging and diagnostics

Artificial intelligence can enhance the efficiency and accuracy of interpretation of medical imaging such as MRIs, CT scans and X-rays (Khalifa and Albadawy 2024a). For image analysis and interpretation AI can detect even minor anomalies and discrepancies, reducing human error, oversight and fatigue. Operational efficiency also gets enhanced, making for a speedier diagnostic process and reducing cost. Predictive and personalized healthcare can be benefitted via predictive analysis and historical data used for early diagnosis and personalized medicine. For clinical decision support AI can help by integrating with other technologies, assisting in complex procedures by giving precise imaging support.

Diagnostic errors are a big challenge in the healthcare system, causing negative outcomes, costing a lot of resources and undermining trust in healthcare professionals (Jiwa and Cameron 2024). Around 10% of patient deaths and 17% negative events in hospitals are calculated to be due to diagnostic error. As a tool AI can go through significant amounts of data and identify patterns, supporting decision making. For diagnostic errors there are three distinct categories which are failure to detect, delayed diagnosis and misdiagnosis. Common reasons for these three are human factors such as workload and time pressure, cognitive bias and communication failure, equipment failures due to outdated technologies and ageing and logistical failures with waiting lists, paper systems and staffing limitations.

AI is good at data processing and pattern recognition, being able to detect more and faster than a human would (Jiwa and Cameron 2024). It also can reduce cognitive bias which human decision making in healthcare is often heavily influenced by. Humans tend to rely too much on initial information and have availability bias, their recent experiences shaping their decision making too strongly. AI can cross reference patient data with databases, being especially valuable when the case is complex. When it comes to training AI has the potential to be revolutionary, optimizing healthcare workers and upping patient safety. Surgeons can even train by having AI go through videos of their work, pointing out areas of improvement and tailoring it on a personal level.

3.4 Chatbots and AI in mental health

Artificial intelligence has been used more in psychiatry as the need for access to mental health services has been on the rise, especially following the COVID-19 pandemic (Pham et al. 2022). AI applications can assist in symptom tracking, psychoeducation, psychiatric diagnoses and disease course prediction. People have been reached through smartphone applications, internet and digital gaming. Conversational chat and therapy bots are able to teach their user

emotional coping mechanisms, avatar therapy by generating images of people's faces and support with communication difficulties. Intelligent animal-like robots have also been used in digital psychiatry.

Chatbot users enjoy having personalized humanlike interactions but are often losing interest in chatting due to the bots making incorrect assumptions and having improper responses (Haque and Rubya 2023). Due to the constant availability of chatbots their users can become attached to them in excess and ignore interactions with the people around them. With their availability they can be used for crisis care although they might not be very good at spotting what would count as a crisis themselves. As the bots are able to foster a judgement free environment they also can get users to more comfortably share information that would be considered private.

Chatbots that are based on generative AI such as ChatGPT are emerging being considered solutions to global mental health crisis (Siddas et al. 2024). Users have seen positive impact and their engagement with the chatbots have been high, helping them with healing from trauma and loss and improving relationships. Experiences have been that generative chatbots have been able to offer a safe and validating space readily available, give insightful guidance and new perspectives. There's also been joy of connection as these chatbots are human-like companions.

As a therapist chatbots have been augmenting therapy sessions, enhancing and replacing but however not matching therapy entirely (Siddas et al. 2024). Some find human empathy and connection to be lacking in their conversations and some have been able to use chatbots as a stepping stone to getting themselves a human therapist. Some find chatbots a good alternative when traditional therapy is not an option due to cost and availability and also when therapy has not been helpful in the way it had been expected to be. Chatbots can also be used to create roleplay scenarios and symbolic imagery to help process through trauma and situations.

3.5 Smart living and AI-powered devices

People tend to spend most of their time at home and it is the single most prominent investment they are going to make (Flynn 2024). Surroundings have an effect of person's health and happiness, smart technology being able to be both convenient and to improve quality of life along with health. Illness prevention and cleanliness have become increasingly important after COVID-19 pandemic, things like UV disinfectant lights, infrared sensors and indoor air quality monitoring being incorporated into homes. For better sleep the lights can be set to follow a circadian rhythm automatically.

Virtual assistants learn the typical behaviour of their people and can create a calming environment for them, such as keeping the home safe and thus reducing anxiety over it (Flynn 2024). People who have difficulty living alone – such as the elderly – can also brave living alone with the help of telehealth services and remote monitoring. Voice controls that are particularly helpful with some of the disabilities that people have can be used to play music, set house temperature and get in touch with people. Smart home gyms offer the chance to participate in virtual workout classes and smart kitchen appliances can direct to using items that are about to perish next or they can help with following dietary restrictions.

Delivery robots have risen to the need for contactless deliveries (NACCU 2022). Students have benefitted from them, being able to spend more of their time studying, some claiming improved mental health and improved feeling of safety. Delivery robots have also helped with meal delivery, reducing the amount of skipped meals students might otherwise have. Delivery robots were perceived as helpful, convenient and friendly due to their design.

4 AI WITH SECURITY AND CYBERSECURITY

Artificial intelligence can be used for overwhelming and time-consuming tasks like going through documents (Burns 2023). Machine learning can predict outcomes and describe data while computer vision is used by law firms to extract information from the documents as well as for classifying that information. Natural language processing is used for interpreting information requests.

Artificial intelligence and machine learning are used as trendy words for marketing in the security field (Burns 2023). AI techniques are used in cybersecurity for anomaly detection, behavioural threat analytics and solving false-positive problems where threats or vulnerabilities are incorrectly flagged and marked. Security information and event management, also known as SIEM, uses machine learning for anomaly detection and finding out threats, since it can quickly identify malicious code.

4.1 AI in threat detection and prevention

Phishing and cyber assaults are becoming complex and more common (Naseer 2024). Artificial intelligence can be used to detect and block threats real-time. AI can detect anomalies in network traffic, recognize malicious URLs and phishing emails while also being able to learn from them. Combined with human oversight AI can be used as a tool for fixing weaknesses and bolstering defence systems.

Advanced algorithms can shift through massive amounts of data sets, unearth hidden risks and automate critical processes such as pre-emptively mitigating cybersecurity threats (Palo Alto Networks 2025). Systems can accurately assess and respond to a broad spectrum of threats, being especially good when it comes thwarting evolving threat tactics that are usually difficult to pinpoint. There's an ongoing competition between threat actors and security measures, leading to ongoing evolution as each time there are advances in technology there are also more sophisticated ways of bypassing defences.

4.2 Deepfakes and the spread of misinformation

Scammers are able to use artificial intelligence to generate images that they then use to impersonate people such as celebrities and other public figures (Puzder 2024). The images and videos look real enough to fool people into thinking they are real, being able to also mimic a person's voice and movement convincingly. Financial information can be stolen when people purchase things from deepfake endorsers, which can also be used to lead people to websites and fake news articles. One of the more common scams is making fake dating profiles which are then used to get money and information from targets. Creating realistic appearing imagery of tragedies and using it to gather donations has also become a thing.

Malicious use of deepfakes has the potential to spread misinformation, erode trust in elections and undermine national security (U.S. Government Accountability Office 2024). Authentication technology such as watermarking can be used to help with identifying deepfakes from real content. AI can be trained for detection, looking at colour abnormalities, evidence of deepfake generation process and vocal and facial inconsistencies. Digital watermarks add audio or pixel patterns that are imperceptible to human eye, the pattern being disturbed when modifications are made. Some watermarking technologies can also make deepfake look or sound unrealistic. Metadata that describes characteristics of a media being missing or incomplete can also showcase that the media has been tampered with. Uploading media to public blockchain also creates a version that is relatively secure and it could be used as comparison when checking for fakes.

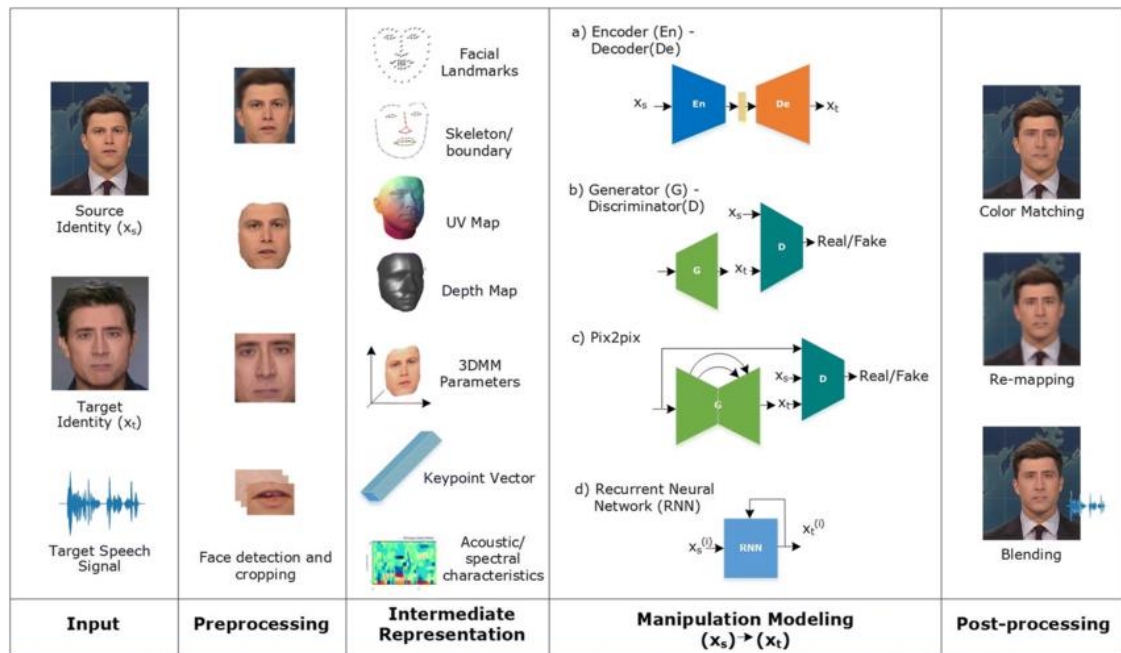


Figure 4. Processing pipeline of audio-visual deepfakes generation approaches (ResearchGate 2022).

Figure 4 shows the pipelines of making deepfakes, from input to pre-processing, intermediate representation, manipulation modelling and post-processing (ResearchGate 2022).

4.3 AI-driven cybersecurity risks

Cybercriminals have also gotten their hands on artificial intelligence, enabling AI-powered cyberattacks (Stanham 2025). This technology enhances, accelerates and automates cyberattack phases. AI can identify vulnerabilities and deploy campaigns along already identified attack vectors and paths, establish backdoors in systems, tamper with and exfiltrate data and interfere with system operations. These algorithms learn and evolve over time, adapting to avoid detection and being able to create new patterns of attack to bypass defences.

AI-powered cyberattacks have five main characteristics which are customization, reinforcement learning, employee targeting, attack automation and efficient data gathering (Stanham 2025). Customization is when the AI is able to

gather data from public pages such as social media platforms, gather and analyse, creating personalized and timely messages for phishing attacks and leveraging social engineering techniques. Employee targeting means the AI identifies high value individuals in an organization, people who are privy to sensitive data or have broad system access and also those who would have a lower technological aptitude and therefore fall easier for scams. Also people who are close to the identified targets could become targets themselves.

Ethical hacking is when there is an authorized attempt to get unauthorized access to data, systems or applications while using techniques that are considered same as what malicious attackers would use (Black Duck 2025). This type of hacking can be used to train for malicious events and to pinpoint vulnerabilities before they are breached. Ethical hackers are experts in security who hack to improve the security of an organization. They need to respect data sensitivity, sometimes having to agree to nondisclosure agreements. They also need to disclose their findings and provide advice on fixing the weak points they managed to find.

5 AI AND SUSTAINABILITY

Innovations and improvements have switched by a large margin to being created by industry instead of academia, mostly due to money and resources needed to make advancements in the field (Stanford University 2023). When it comes to environmental load produced by AI, carbon emissions have been calculated to be significant while it is argued that AI can also optimize energy usage. When it comes to science, AI models have been used to rapidly accelerate progress, leading to generation of new antibodies and even helping with hydrogen fusion.

5.1 AI in climate change mitigation

Artificial intelligence has been able to improve Earth system sciences, improving weather forecasting, parameter estimation, model emulation and prediction of extreme events (Camps-Valls et al. 2025). The data to base predictions on is unfortunately heterogeneous, noisy, has a small sample size and limited annotations. AI can help with predicting wildfires, droughts, floods and heat-waves which are all considered extreme climate events, also improving disaster response and communication.

Artificial intelligence deals with large volumes of data which requires large volumes of data stored and exchanged worldwide, requiring computing power that creates its own environmental impact (Climate Seed 2024). However, AI is also used as a tool in achieving sustainable development goals. It can help with rethinking business models and resource use, optimizing processes and automating tasks. When it comes to logistics AI can be used to optimize routes and load rates, especially helpful when most trucks drive with less than 15% full loads.

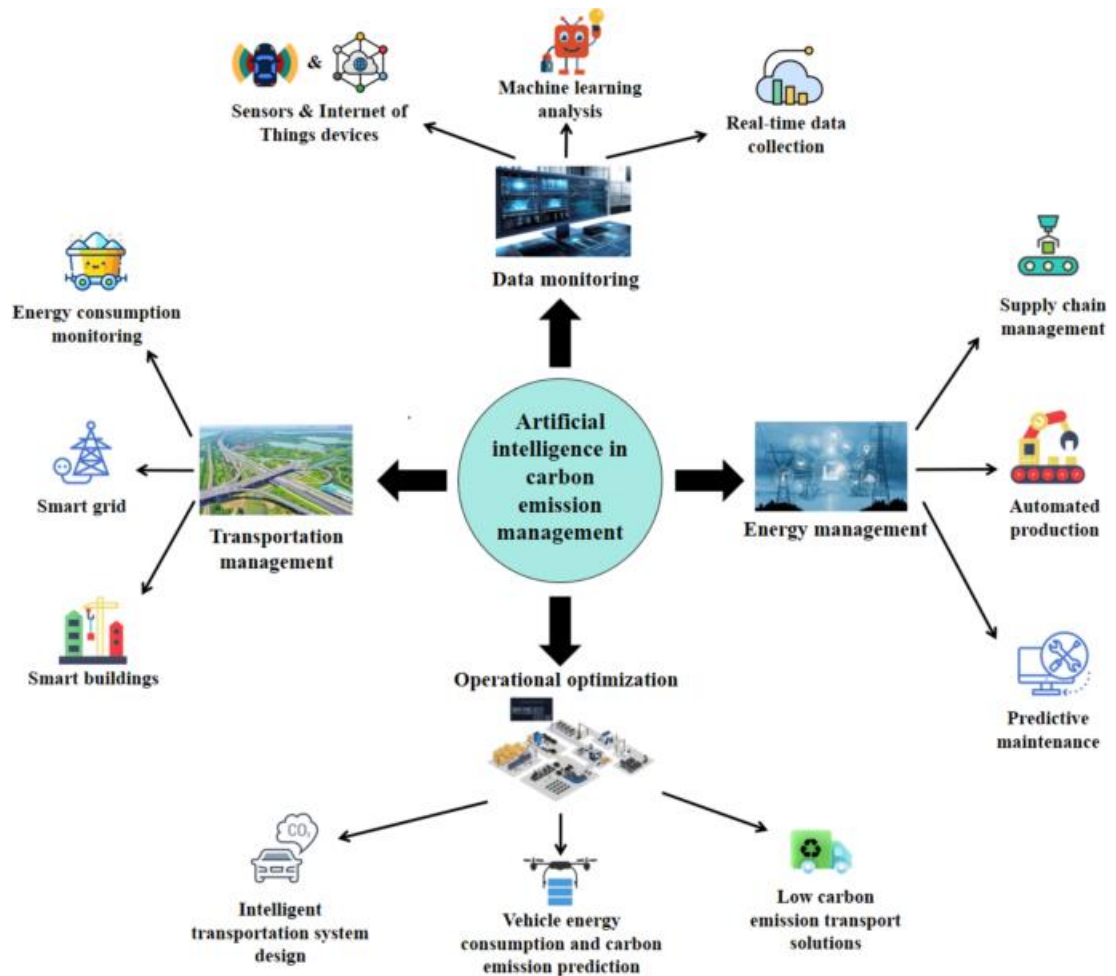


Figure 5. Utilization of artificial intelligence in carbon emissions management (Springer 2025).

Figure 5 showcases artificial intelligence in carbon emission management, operational optimization having to do with intelligent transportation system design, vehicle energy consumption and carbon emission prediction and low carbon emission transport solutions (Springer 2025). Energy management has to do with supply chain management, automated production and predictive maintenance. Data monitoring splits to sensors and internet of things devices, machine learning analysis and real-time data collection. Transportation management has to do with energy consumption monitoring, smart grid and smart buildings.

5.2 AI for optimizing energy use

In the energy sector renewable energy being rather intermittent, it is difficult to balance in terms of production and consumption (Climate Seed 2024). AI can help analyse the demand or production in real time, predicting and noticing patterns. It can improve efficiency of energy system and therefore also reduce carbon intensity.

When it comes to energy use AI has the potential to reduce it, having already reduced building energy consumption by 20% in a residential area in Sweden and 50% for low traffic mobile networks in Japan (Vattenfall Media Relations 2023). Within EU area buildings take up 40% of total energy consumption and also generate 36% of carbon dioxide emissions. Rising energy prices are one motivator to wasting energy while its use could have been optimized better. AI can be applied in controlling household temperatures and ventilation around the clock as required energy amounts fluctuate with the fluctuating day and night-time temperatures and weather conditions.

5.3 AI in agriculture and resource management

AI-powered precision agriculture is a technology that increases crop yield and utilizes resources efficiently (Олена et al. 2025). AI can help with estimating the best timings for sowing seeds and for harvesting and it can also be used to minimize fertilizer, water and pesticide use. Better informed farming decisions can be made with the combination of machine learning, internet of things sensors and autonomous precision farming systems. Implementation costs for AI related technologies in farming are still high, however they can create new opportunities for production process optimization, food security and the reduction of environmental impact.

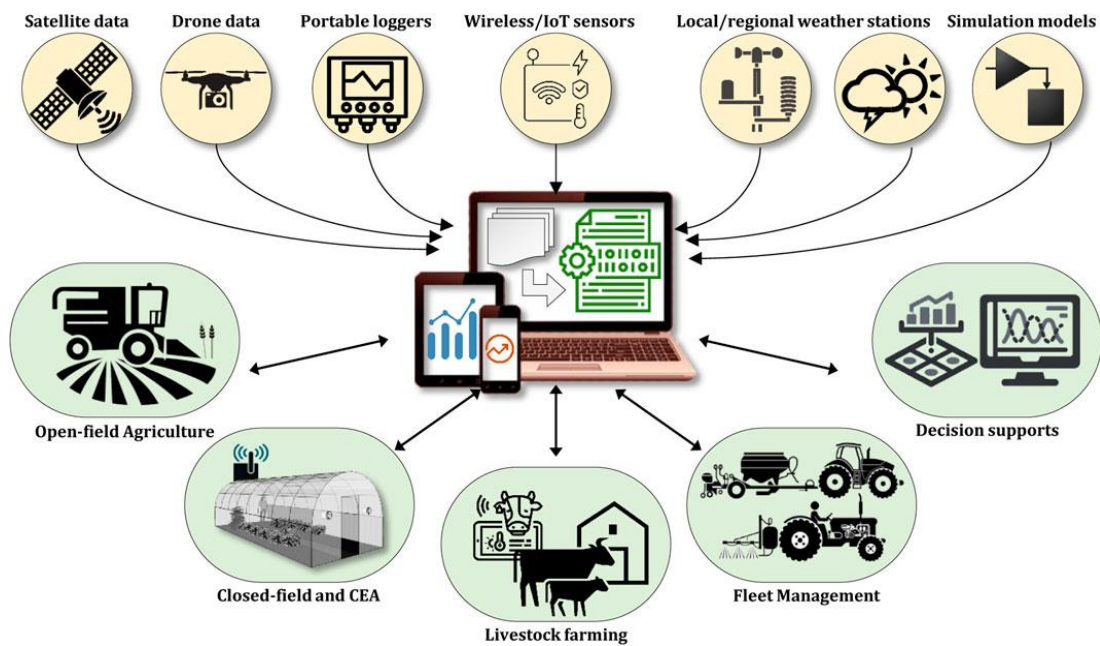


Figure 6. A conceptual view of data-sharing in connected farms for the implementation of digital agriculture (Frontiers 2024).

Figure 6 showcases satellite data, drone data, portable loggers, wireless and IoT sensors, local and regional weather stations and simulations models intermingling through AI with open-field agriculture, closed-field and controlled-environment agriculture (CEA), livestock farming, fleet management and decision supports (Frontiers 2024).

Agriculture is under pressure to perform more production and use less resources as the estimated global population is estimated to reach over 9 billion people by the year 2050, placing demand on food systems (Hassan 2024). Sustainability in food production, reduced environmental impact and boosted productivity are looked at to be affected with smart technologies. AI can be applied to predictive analytics, precision agriculture, livestock monitoring, automated machinery and supply chain optimization.

Precision agriculture uses things like drones, satellite imagery and sensors (Hassan 2024). For predictive analysis there's pest outbreaks, weather patterns and timing recommendations for tasks. AI can drive automated harvesters, drones and tractors, reducing human error and reducing needed human labour. Animal monitoring helps with ensuring animal welfare, monitoring real

time around the clock. Improved logistics and traceability of food can be optimized with AI across supply chains.

6 ETHICAL AND PHILOSOPHICAL IMPLICATIONS OF AI

Artificial system reinforces the things it learns, raising ethical questions as the tools become only as good and just as the data it has been fed to grow upon (Burns 2023). Human bias enters the process already at the training stage as humans select the type of data that is used for training. It is very important to monitor potential machine learning bias for deep learning and generative adversarial network applications, also known as GAN.

Ethical misuse of AI has increased 26 times since 2012 (Stanford University 2023). Deepfake videos are one facet of this, allowing realistic looking image of events that are not true. U.S. prisons have also been using call-monitoring on their prisoners. As the use of AI technologies increases, so seems to increase the misuse possibility awareness of it.

6.1 Ethical dilemmas in AI development

The concept of the artificial intelligence we use these days has existed since 1940s, becoming famous in the 1950s via Turing test (Office of the Victorian Information Commissioner n.d). Improved algorithms, increased ability to store and capture data in large quantities and increased networked computing power have been some of the core reasons for the fast development in recent years. The way people think about intelligent machines has also been developing since 1960s.

AI technology generally stops being referred to as AI once it works and just becomes mainstream computing (Office of the Victorian Information Commissioner n.d). Automated voice greetings and being recommended things based on previous preferences are some of these type of technologies. Voice recognition, predictive analysis and natural language processing have become so widespread that they have been starting to blend to the background.

In the future less people who work with AI might necessarily need access to raw data, helping with reducing privacy and breach safety concerns when it comes to the contents of the data (Office of the Victorian Information Commissioner n.d). Considering ethics of technology alongside with solving privacy challenges will shape the future of AI. Information privacy already provides framework for ethical choices.

World already is a playground of big data enhanced with artificial intelligence, everything from the internet of things to smart cities technology, expected higher standard of living, increased efficiency and dynamic use of resources (Office of the Victorian Information Commissioner n.d). AI technology will be making advancements to healthcare and justice systems as well as to other government services. AI however still has technological, social and legal challenges with information privacy.

6.2 Philosophical questions about AI consciousness

In the future AI is hypothesised to surpass human consciousness, leading to fast innovation or destruction (Orf 2025). This event, also known as singularity, is expected to occur within the current century, perhaps soon. A neuroscientist claims AI will not reach the complexity of a brain since it lacks consciousness, which is not the same as intelligence. Neuroscience could say that consciousness is a series of predictions that brains make to aid in their organism's survival, a type of controlled hallucination. Biology, such as having a metabolism and being able to reproduce also currently differentiate humans and machines.

AI will also be changing the way humans think and behave, some people afraid that relying too much on AI will be having negative effects (Duffy 2025). Skills such as deep thinking and empathy and other core skills of being human are feared being on the chopping block. People would also be delegating things like creativity, agency and decision making to AI's that are not advanced

enough to handle them properly. Becoming too reliant on AI seems like a legitimate issue. AI will change how people gather information, how they communicate and how they work.

6.3 Human-AI interaction and trust

Explainable AI, also known as XAI, is processes and methods that allow a human to trust and comprehend the output of machine learning algorithms (IBM 2023b). Explainable AI helps characterize the fairness, accuracy, outcomes and transparency of AI-powered decision making. Explainability is important in building trust and having a responsible approach when developing artificial intelligence. With more advanced AI the process the AI does to arrive to a conclusion is more difficult to understand. The main difference between XAI and regular AI is in its process traceability and backtracking ability.

AI is already a big part of people's daily lives through services and applications and distrust in AI has the potential to slow down the rate of AI adoption (Afroogh 2024). Trust as a consideration seems to be key when training users, AI and robot designers and theorizing about AI. Different cultures have different thresholds, laws and principles so there will be differences country to country.

6.4 Ethical considerations in AI for healthcare

Many of the artificial intelligence technologies tend to be controlled and owned by private entities (Murdoch 2021). Clinics, public bodies and corporations will have a larger role in utilizing, obtaining and protecting people's health information. This could lead to privacy issues with data security and implementation.

Potential risks of AI misdiagnosis is also a concern (News Medical 2024). AI technology has been adopted quicker than anything in the history of consumer applications. World Health Organization (WHO) has given guidance on the governance and ethics of large multi-modal models (LMMs). AI's use in

healthcare is already widespread, covering fields like scientific research, diagnosis, medical training and drug development. In some areas doctor shortages have been patched up with the use of AI.

WHO also warned about the risks associated, such as misdiagnosis when overestimating LMM capabilities especially when a LMM is not properly updated or maintained (News Medical 2024). Reliance on LMMs can also lead to firing healthcare workers and require extensive retraining in using AI. Equal access to AI models is also difficult to guarantee, things such as subscription fees further dividing those who can afford it from those who can't. LMMs trained on biased data could also do more harm than good and compensation strategies for these harms have already been planned.

7 REGULATORY AND POLICY PERSPECTIVES ON AI

Artificial intelligence technologies are impacting lives at a fast rate, changing how people conduct their work and personal lives (Jacob 2024). As artificial intelligence gain popularity and becomes more widespread, it raises legal, moral and social concerns. Therefore countries have begun to adopt legislation having to do with AI deployment, design and use.

7.1 Governmental approaches to AI regulation

Around the world countries have their own legal frameworks and ways of regulating AI (Jacob 2024). In The United States of America decentralized model of regulation is used, reflecting the general approach existing in terms of governance. Most practises are on sectoral level, several sector-specific organizations and agencies having been established to asses AI related challenges. Consumer protection, autonomous cars and processing of consumer data are some of the most looked after aspects.

The European Union has been taking a proactive approach when it comes to AI legislation (Jacob 2024). There are initiatives aiming for strict guidelines when it comes to preserving personal information, as well as its gathering and use. The European Union aims to restrict uncontrolled functioning of AI, complementing General Data Protection Regulation. Artificial Intelligence Act is supposed to extent control over development, use and regulation of artificial intelligence, accountability, transparency and ethics being the corner stones.

China has established itself as a leading global power when it comes to AI, their goal being to become leader in AI innovation by 2030 (Jacob 2024). There are patterns of control, ethics and security being some of the main areas, compliance and timely risk management ensured also with The Chinese Cybersecurity Law and New Generation AI Development Plan. International organizations such as the United Nations and the Organization for Economic Co-oper-

ation (OECD) are also setting international guidelines. OECD wants responsibility, transparency and inclusion in implementation and development of AI while United Nations focuses on sustainability and global benefits. Key principles in general seem to be data privacy, ethical principles, algorithm bias, international collaboration, transparency and explainability.

7.2 AI accountability frameworks

Accountability affects customer trust, legal liability, ethical considerations and brand reputation (Emergence Digital 2025). AI systems make decisions with real outcomes, handling customer interactions and strategic plans. The lack of clear accountability threatens damage to business reputation, operational risks and legal issues.

There are different layers when it comes to accountability and who holds it (Emergence Digital 2025). AI users are at the initial layer, their responsibility being to understand limitations and functionality of AI tools in their use. If AI users have managers, their job is to make sure they are properly trained to use AI tools responsibly. Companies and employees should establish clear guidelines, being accountable for AI use consequences for their organizations. AI developers have the responsibility to ensure AI gets trained and designed with safety measures and mitigated bias. AI vendors should provide secure, reliable and ethical AI and not fail to disclose potential limitations and risks. Data providers are important as AI training relies on the data it is fed, so attention should be paid to quality and accuracy of data, making sure it is adhering to privacy regulations and ethically sourced. Regulatory bodies have overarching accountability when it comes to enforcing and establishing regulations for AI. They should be protecting the public interests as well as businesses, define legal landscapes and ensure ethics are being upheld.

Using AI in legal decisions could noticeable decrease the time it takes to make judgements (Sreelatha and Choudhary 2023). AI can automate mundane tasks, facilitate legal research and augment the examination of legal aspects

as the volume of data to go through usually is large, therefore creating higher accuracy. Predictive capabilities of AI can also be worked with in analysing correlations and patterns but they can also predict case outcomes.

Biggest challenges when considering AI for legal systems stems from the lack of openness and comprehensibility of AI systems (Sreelatha and Choudhary 2023). Even if AI can predict results correctly, its thought process cannot be clearly defined and explained, undermining trust. Bias is also one of the problems, potentially increasing the effect of existing societal biases in cases. Regular audits would be needed to check for this, and the teams would need to be diverse themselves to consider multiple aspects. Legal data contains also sensitive data and requires data protection mechanics.

7.3 AI in military and defence

AI is changing modern warfare landscape, adding precision, complexity and speed to change (Nordic Defence Review 2024). Countries with resources and data have an advantage, as AI systems can be costly to make. AI-guided aerial and naval drones are already used in real life conflicts, AI being capable of advanced object recognition and problem solving. Geopolitical rivalries also advance the rate at which these technologies get developed. AI can be used in fields such as intelligent weapons systems, command and control, decision-support systems, simulations and training and predictive maintenance and logistics.

Autonomous weapons systems (AWS) pose ethical challenges, having become a policy debate. (Blanchard 2024). There's a lack of historical precedent in terms of ethics, regulatory debate being mostly legal arguments of international humanitarian law. Ethics in this case has also become something that has been made to compete with law. AWS has the ability to pick targets it engages without human intervention, arising questions over the role of a human in use of force. Using algorithms to decide targets could have widespread

consequences. Delegating decisions about life and death to machines poses moral issues when it comes to our relationship to technology and killing.

8 CONCLUSION

This thesis explored the image of artificial intelligence in society and examined the fears and hopes people associate with its increasing presence in daily life. Through an extensive review of current literature, reports and studies it was found that while AI offers considerable benefits across multiple sectors such as healthcare, creative industries, cybersecurity and business the societal concerns are often a combination of legitimate issues and misconceptions.

The fears regarding AI such as loss of privacy, job displacement, bias in decision-making and ethical misuse have concrete foundations, especially when development and deployment are rushed without appropriate oversight. The reality however often falls short of the extreme scenarios portrayed by media or popular culture. Most current AI systems are narrow in scope, require human oversight and depend heavily on the quality of their input data and ethical frameworks.

The findings highlight the importance of promoting transparency, explainability and ethical considerations in AI development to maintain public trust. It is equally important to foster critical media literacy, allowing individuals to assess AI-related information more objectively.

Further research could investigate the long-term societal impact of artificial intelligence, especially as large multimodal models and autonomous systems continue to evolve. There is also a need for continuous interdisciplinary collaboration between technologists, ethicists and policymakers to ensure that AI technologies serve the broader interest of humanity.

REFERENCES

Afroogh, S., Akbari, A., Malone, E., Kargar, M. and Alambeigi, H. 2024. *Trust in AI: progress, challenges, and future directions*. Nature. Retrieved April 26, 2025, from <https://www.nature.com/articles/s41599-024-04044-8>

Appel, G., Neelbauer, J. and Schweidel, D. A. 2023. *Generative AI has an intellectual property problem*. (2023). Harvard Business Review. Retrieved April 20, 2025, from <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>

Black Duck. 2025. *What is ethical hacking?* Retrieved April 26, 2025, from <https://www.blackduck.com/glossary/what-is-ethical-hacking.html>

Blanchard, A. 2024. *The road less travelled: ethics in the international regulatory debate on autonomous weapon systems*. International Committee of The Red Cross. Retrieved April 27, 2025, from <https://blogs.icrc.org/law-and-policy/2024/04/25/the-road-less-travelled-ethics-in-the-international-regulatory-debate-on-autonomous-weapon-systems/>

Burns, Ed. 2023. *Definition artificial intelligence*. TechTarget. Retrieved July 9, 2023, from <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>

Camps-Valls, G., Fernández-Torres, M.Á., Cohrs, K.H. and et al. (2025). *Artificial intelligence for modeling and understanding extreme weather and climate events*. Nature communications. Retrieved April 26, 2025, from <https://www.nature.com/articles/s41467-025-56573-8>

Chaturvedi, U., Chauhan, S. B. and Singh, I. 2024. *The impact of artificial intelligence on remote healthcare: Enhancing patient engagement, connectivity, and overcoming challenges*. ScienceDirect. Retrieved April 23, 2025, from <https://www.sciencedirect.com/science/article/pii/S2949866X24001230>

Climate Seed. (2024). *AI, an ally in reducing carbon footprint*. Retrieved April 26, 2025, from <https://climateseed.com/blog/reducing-carbon-footprint-with-artificial-intelligence>

Crabree-Ireland, D. 2025. *Labour leaders: Artificial intelligence must serve human creativity, not replace it*. World Economic Forum. Retrieved April 19, 2025, from <https://www.weforum.org/stories/2025/01/artificial-intelligence-must-serve-human-creativity-not-replace-it/>

Davenport, T. H and Bean, R. 2023. *The impact of generative AI on Hollywood and Entertainment*. MIT Sloan Management Review. Retrieved April 19, 2025, from <https://sloanreview.mit.edu/article/the-impact-of-generative-ai-on-hollywood-and-entertainment/>

De Witte, M. 2023. *How will ChatGPT change the way we think and work? Stanford scholar examines*. StanfordReport. Retrieved April 21, 2025, from <https://news.stanford.edu/stories/2023/02/will-chatgpt-change-way-think-work>

Deljkic, E. 2024. *How AI will transform wearables*. Forbes. Retrieved April 23, 2025, from <https://www.forbes.com/councils/forbestechcouncil/2024/02/08/how-ai-will-transform-wearables/>

Duffy, C. 2025. *The 'father of the internet' and hundreds of tech experts worry we'll rely on AI too much*. CNN. Retrieved April 26, 2025, from <https://edition.cnn.com/2025/04/02/tech/ai-future-of-humanity-2035-report/index.html>

Duong, D. and Solomon, B. D. 2025. *Artificial intelligence in clinical genetics*. European Journal of Human Genetics. Retrieved April 23, 2025, from <https://www.nature.com/articles/s41431-024-01782-w>

Emergence Digital. (2025). *AI accountability: Who's responsible when AI goes wrong?* Retrieved April 27, 2025, from <https://emerge.digital/resources/ai-accountability-whos-responsible-when-ai-goes-wrong/>

Flynn, S. 2024. *The health benefits of smart home technology*. IoT for all. Retrieved April 26, 2025, from <https://www.iotforall.com/the-health-benefits-of-smart-home-technology>

Haque, M. D. and Rubya, S. (2023). *An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews*. National Library of Medicine. Retrieved April 26, 2025, from <https://pmc.ncbi.nlm.nih.gov/articles/PMC10242473/>

Hassan, T. 2024. *Revolutionizing farming with AI*. GriNext Conference. Retrieved April 26, 2025, from <https://agrinextcon.com/farming-with-aisustainable-agriculture-solutions/>

IBM. 2023a. *What is AI bias?* Retrieved April 19, 2025, from <https://www.ibm.com/think/topics/ai-bias>

IBM. 2023b. *What is explainable AI?* Retrieved April 26, 2025, from <https://www.ibm.com/think/topics/explainable-ai>

IBM. 2025. *What is conversational AI?* Retrieved April 20, 2025, from <https://www.ibm.com/think/topics/conversational-ai>

Jacob, S. (2024). *AI regulations around the world: A comprehensive guide to governing artificial intelligence*. Spiceworks. Retrieved April 27, 2025, from <https://www.spiceworks.com/tech/artificial-intelligence/articles/ai-regulations-around-the-world/>

Jiwa, A. and Cameron, M. 2024. *The potential of AI to help reduce diagnostic errors*. The Royal College of Surgeons of Edinburgh. Retrieved April 23, 2025, from <https://www.rcsed.ac.uk/news-resources/rcsed-blog/2024/september/the-potential-of-ai-to-help-reduce-diagnostic-errors>

Khalifa, M. and Albadawy, M. (2024a). *AI in diagnostic imaging: Revolutionising accuracy and efficiency*. ScienceDirect. Retrieved April 23, 2025, from <https://www.sciencedirect.com/science/article/pii/S2666990024000132>

Khalifa, M. and Albadawy, M. (2024b). *Artificial intelligence for clinical prediction: Exploring key domains and essential functions*. ScienceDirect. Retrieved April 22, 2025, from <https://www.sciencedirect.com/science/article/pii/S2666990024000156>

Marsner. n.d. *The current state of AI*. Retrieved October 10, 2023, from <https://marsner.com/blog/the-current-state-of-ai/>

Mineo, L. 2023. *If it wasn't created by a human artist, is it still art?* The Harvard Gazette. Retrieved April 20, 2025, from <https://news.harvard.edu/gazette/story/2023/08/is-art-generated-by-artificial-intelligence-real-art/>

Murdoch, B. 2021. *Privacy and artificial intelligence: challenges for protecting health information in a new era*. BMC. Retrieved April 27, 2025, from <https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-021-00687-3>

NACCU. 2022. *National study of college students shows delivery robots help them feel safer, improve mental health*. Retrieved April 26, 2025, from <https://www.naccu.org/news/national-study-of-college-students-shows-delivery-robots-help-them-feel-safer-improve-mental-health>

Naseer, I. 2024. *The role of artificial intelligence in detecting and preventing cyber and phishing attacks*. ResearchGate. Retrieved April 26, 2025, from <https://www.researchgate.net/publication/385077487> The role of artificial intelligence in detecting and preventing cyber and phishing attacks

News Medical. (2024). *AI risks in healthcare: Misdiagnosis, inequality, and ethical concerns*. Retrieved April 27, 2025, from <https://www.news-medical.net/news/20240123/AI-risks-in-healthcare-Misdiagnosis-inequality-and-ethical-concerns.aspx>

Nordic Defence Review. 2024. *The AI battlefield: How artificial intelligence is revolutionizing modern warfare*. Retrieved April 27, 2025, from <https://nordicdefencereview.com/the-ai-battlefield-how-artificial-intelligence-is-revolutionizing-modern-warfare/>

Office of the Victorian Information Commissioner. (n.d.). *Artificial intelligence and privacy – Issues and challenges*. Retrieved April 26, 2025, from <https://ovic.vic.gov.au/privacy/resources-for-organisations/artificial-intelligence-and-privacy-issues-and-challenges/>

Orf, D. 2025. *Human consciousness is a ‘controlled hallucination,’ scientist says—and AI can never achieve it*. Popular Mechanics. Retrieved April 26, 2025, from <https://www.popularmechanics.com/science/a64555175/conscious-ai-singularity/>

Palo Alto Networks. 2025. *What is the role of AI in threat detection?* Retrieved April 26, 2025, from <https://www.paloaltonetworks.com/cyberpedia/ai-in-threat-detection>

Pham, K. T., Nabizadeh, A. and Selek, S. (2022). *Artificial intelligence and chatbots in psychiatry*. National Library of Medicine. Retrieved April 25, 2025, from <https://pubmed.ncbi.nlm.nih.gov/35212940/>

Puzder, D. 2024. *The dangers of AI art and deepfakes*. Washington University in St. Louis. Retrieved April 26, 2025, from <https://informationsecurity.wustl.edu/the-dangers-of-ai-art-and-deepfakes/>

Rafalski, K. (2025). *AI in automotive: Driving the future of transportation*. Netguru. Retrieved April 27, 2025, from <https://www.netguru.com/blog/ai-in-automotive>

Robinson, B. 2025. *Fears about AI job loss: New study answers if they're justified*. Forbes. Retrieved April 18, 2025, from <https://www.forbes.com/sites/bryanrobinson/2025/02/09/fears-about-ai-job-loss-new-study-answers-if-theyre-justified/>

Siddas, S., Torous, J. and Coxon, A. (2024). *'It happened to be the perfect thing': Experiences of generative AI chatbots for mental health*. Springer Nature. Retrieved April 26, 2025, from <https://www.nature.com/articles/s44184-024-00097-4>

Sreelatha, A. and Choudhary, G. 2023. *Exploring the use of AI in legal decision making: Benefits and ethical implications*. Woxsen University. Retrieved April 27, 2025, from <https://woxsen.edu.in/research/white-papers/exploring-the-use-of-ai-in-legal-decision-making-benefits-and-ethical-implications/>

Stanford University. 2023. *The AI index report*. Retrieved October 10, 2023, from <https://aiindex.stanford.edu/report/>

Stanham, L. (2025). *AI-powered cyberattacks*. CrowdStrike. Retrieved April 26, 2025, from <https://www.crowdstrike.com/en-us/cybersecurity-101/cyberattacks/ai-powered-cyberattacks/>

Tiwari, R. 2025. *How to use AI in social media to drive business growth: 20 insights and tips*. Socialinsider. Retrieved April 19, 2025, from <https://www.socialinsider.io/blog/ai-in-social-media/>

U.S. Government Accountability Office. 2024. *Science & tech spotlight: Combating deepfakes*. Retrieved April 26, 2025, from <https://www.gao.gov/products/gao-24-107292>

Vattenfall Media Relations. 2023. *Yes, AI will improve energy efficiency – ‘Essentially all buildings can benefit from this technology’*. Vattenfall. Retrieved April 26, 2025, from <https://group.vattenfall.com/press-and-media/news-room/2023/yes-ai-will-improve-energy-efficiency--essentially-all-buildings-can-benefit-from-this-technology>

World Economic Forum. 2025. *Future of jobs report 2025*. Retrieved April 19, 2025, from https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf

Олена, А., Наталія, Б., Юлія, Д., Тетяна, І., Оксана, М., Adebayo, H. and Caleb, A. (2025). *AI-driven precision farming: Enhancing crop yield and resource efficiency*. ResearchGate, Retrieved April 26, 2025, from https://www.researchgate.net/publication/388528296_AI-Driven_Precision_Farming_Enhancing_Crop_Yield_and_Resource_Efficiency

FIGURES

Figure 1. AI Model Types (Virtualizationteam.com 2024).

Source: Virtualizationteam.com (2024) Retrieved April 27, 2025, from <https://virtualizationteam.com/artificial-intelligence-ai/artificial-intelligence-model-types.html>

Figure 2. Biases in AI systems (Srinivasan and Chander 2021).

Source: Srinivasan, R. and Chander, A. (2021) *Biases in AI systems*. Communications of the ACM. Retrieved April 19, 2025, from <https://cacm.acm.org/practice/biases-in-ai-systems/>

Figure 3. Diagnostic genetic testing workflow (Research Gate 2024).

Source: ResearchGate (2024) *Diagnostic genetic testing workflow*. Retrieved April 23, 2025, from https://www.researchgate.net/figure/An-illustration-of-a-diagnostic-genetic-testing-workflow-Using-AI-algorithms-to-compare_fig1_383152208

Figure 4. Processing pipeline of audio-visual deepfakes generation approaches (ResearchGate 2022).

Source: ResearchGate (2022) *Processing pipeline of audio-visual deepfakes generation approaches*. Retrieved April 26, 2025, from https://www.researchgate.net/figure/Processing-pipeline-of-audio-visual-deepfakes-generation-approaches_fig4_361086563

Figure 5. Utilization of artificial intelligence in carbon emissions management (Springer 2025).

Source: Springer (2025) *Utilization of artificial intelligence in carbon emissions management*. Retrieved April 26, 2025, from <https://link.springer.com/article/10.1007/s10311-024-01799-z>

Figure 6. A conceptual view of data-sharing in connected farms for the implementation of digital agriculture (Frontiers 2024).

Source: Frontiers (2024) *A conceptual view of data-sharing in connected farms for the implementation of digital agriculture*. Retrieved April 26, 2025, from <https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2024.1375193/full>