Original scientific paper

UDC: 343.9:004.8

Received: January, 14, 2026.
Revised: July, 24, 2026.
Accepted: July, 22, 2026.

10.23947/2334-8496-2026-14-2-373-384

Check for updates

Enhancing Human-Machine Cooperation through the Integration of Empathetic AI into Corporate Cybersecurity and Forensic Operations

Radoslav Baltezarević[1], Marko Stanojević[2], Danilo Izgarević[2], Miloš Azdejković[3]*

[1] Institute of International Politics and Economics, Belgrade, Serbia, e-mail: radoslav@diplomacy.bg.ac.rs
[2] International Centre for Economics and Public Policy – ICEPP, Niš, e-mail: stanojevic.marko799@gmail.com, daniloizgarevic@gmail.com
[3] Faculty of Law, Megatrend University, Belgrade, Serbia, e-mail: azdejkovic.milos@gmail.com

Abstract: In order to strengthen the cooperation between humans and machines, empathic artificial intelligence is increasingly emerging as a key player. This cooperation is particularly evident in corporate cybersecurity and digital forensics. Today, thanks to the inclusion of emotional awareness in artificial intelligence systems, many organizations can more easily monitor the cognitive load of their analysts during critical security situations. Empathic artificial intelligence in corporate cybersecurity is a new force for proactive threat detection, compliance and risk management. It helps analysts to respond to constant cyber threats more effec-tively, and in forensic investigations it offers great support in emotional and behavioral analysis. In this way, in cybercrime inves-tigations, through the observation of anomalies and their patterns, decisions are made with a higher degree of accuracy. How-ever, this technology still faces numerous challenges, such as accurately reading emotional cues, protecting sensitive data, and ensuring ethical management. In any case, well-defined frameworks that promote cooperation between artificial intelligence and humans are the necessity of the present, in order to remove these barriers. It is undeniable that empathic artificial intelligence has the potential to create security ecosystems that are more adequate, reliable and above all focused on human needs. Therefore, in the near future, it will revolutionize the field of corporate cybersecurity operations and digital forensics, but only if it is smartly implemented.

Keywords: *empathetic artificial intelligence, human–machine collaboration, corporate cybersecurity, forensic operations, cybercrime investi-gation*

Introduction

As Cuff et al. (2014) argue, empathy, or sharing feelings with social others, is essential for both strong relationships and mental well-being. Brink et al. (2011) define empathy as simply putting oneself in the shoes of another person by understanding the situations and realities they face. It can foster participation because it allows a person to put aside their current situation and be open to others (Ashworth and Lucas, 2000). Building engagement and trust between humans and computers increasingly depends on the ability to personalize interactions based on the unique characteristics of the user (Baltezarević and Baltezarević, 2024).

AI is transforming how we go about our day-to-day lives, influencing everything from simple tasks to more involved decision-making scenarios. One of the most intriguing breakthroughs in this field is the creation of empathetic algorithms. By merging emotional intelligence (EI) with cognitive computing, these algorithms elevate the interaction between humans and machines. This development not only marks a pivotal moment in AI's evolution but also unlocks new avenues for its use in fields like healthcare, education, and customer service, where grasping and responding to human emotions can lead to much better results (Kukreja et al., 2022).

Lately, the role of Artificial Intelligence (AI) in cybersecurity has evolved significantly. It's no longer just about technical detection and response; we're now looking at the human elements involved as well. Enter empathetic AI, this innovative technology can understand and react to human emotions and thought processes.

*Corresponding author: radoslav@diplomacy.bg.ac.rs

It's becoming a game-changer for security analysts, particularly when they're under pressure. By alleviating cognitive overload, it helps improve their decision-making skills.

Empathetic algorithms are crafted to tackle the intricate and often nuanced realm of human feelings, unlike standard algorithms that focus solely on objective data. They do this by merging natural language processing (NLP), facial expression analysis, voice tone recognition, and other sensory inputs to identify emotional states (Zhou et al., 2019).

They contribute to fostering more honest interactions by modifying their responses according to the situation and human feelings (Baltezarević and Baltezarević, 2025). In many areas, but especially in risky ones, this collaboration between humans and AI has proven to be very reliable and is considered more effective than completely replacing humans with artificial intelligence (Bansal et al., 2021).

However, developing human-centred AI models to support people and creating human-facing interfaces for people to interact with the AI provide additional challenges for the partnership. For example, when it comes to AI-assisted writing, we need to develop AI models that produce practical writing recommendations while also creating human-facing systems that enable users to view, comprehend, and act upon those recommen-dations at the appropriate moment (Hirsch et al., 2017).

However, even highly skilled demonstrations of empathy have the potential to alienate recipients, and there is always the possibility of empathic error. Instead, a human-first approach is needed (Main et al., 2017), as has been done in the field of medicine to understand the effects of physician empathy on patients (Cadiente et al., 2023). A productive partnership in which AI merely assists, rather than controls, the creative process can be sustained by establishing explicit rules and responsibilities (Jylkäs et al., 2025).

AI has the potential to produce complex malware that can automate time-consuming phishing attempts, adjust to security precautions, and even carry out deepfake attacks that alter audio and visual data (Hussain et al., 2021). The increased availability of AI technologies and techniques is lowering the barrier to entry for hackers, leading to more complex and targeted cyberattacks. A proactive cybersecurity strategy is necessary, as defensive tactics are constantly changing in the AI realm (Rafy, 2024). Emotional intelligence in artificial intelligence can significantly improve human-computer collaboration in identifying cyber threats and ul-timately, preventing them (Martin, 2022). For this reason, the main goal of incorporating artificial intelli-gence into cybersecurity is to improve the skills of analysts so that they are better able to deal with in-creasingly complex cyber threats, which have been noticeably increasing recently (Baruwal Chhetri et al., 2024).

Research indicates that AI systems can provide substantial support to digital forensic analysts, assisting them in tasks such as organizing evidence, identifying anomalies, and tracing the origins of threats, all by inte-grating machine learning with human intuition (Akeiber, 2025). What's more, AI can handle huge amounts of forensic data much more effectively than traditional methods, while human analysts bring in the necessary context and legal insights (Dunsin et al., 2024).

This article examines the concept of empathetic artificial intelligence and looks at how it can improve collaboration between humans and this new technology. There is a particular focus on areas such as corporate cybersecurity and forensic investigations, and how empathetic AI can improve their processes. We're about to explore the real-world applications of empathetic AI, focusing on the benefits it offers and the ethical and practical challenges we might encounter along the way. To steer our discussion, we'll concentrate on three main research questions. First, we'll examine how empathic AI systems are being utilized in fields like corporate cybersecurity, forensic operations, and hu-man-machine interactions (RQ1). Next, we'll look into the possible benefits of integrating empathetic AI into these areas. This approach could strengthen trust, improve decision-making, and promote collaboration during challenging investigations and operations (RQ2). Finally, we'll identify the key technical, ethical, and operational hurdles associated with imple-menting empathetic AI in corporate cybersecurity and forensic investigations, while also brainstorming solutions to tackle these challenges without sacrificing that crucial human oversight (RQ3). This review brings together insights from psychology, human-computer interaction, artificial intelligence, and cyber-security to create a more holistic framework for developing empathetic AI in human-machine collaboration. We'll also point out existing research gaps and propose future research directions to improve corporate cybersecurity practices, forensic processes, and the design of AI-assisted investigations.

Foundations of Empathetic AI and Human–Machine Collaboration in Corporate

Cybersecurity and Forensic Operations

These days, we encounter AI in our daily lives more than we realize. Whether it's snapping photos on our smartphones, streaming our favorite tunes, or scrolling through social media, AI is often working behind the scenes. This can lead to a big gap between how often people think they interact with AI and how often actual AI experts believe they do (Richter, 2025). This difference highlights how crucial it is to create AI systems that are not only transparent but also easy for users to understand.

One human trait that greatly enhances our ability to communicate and make decisions is emotional intelli-gence. It encompasses the ability to recognize, understand, manage, and skillfully use emotions (Ghotbi, 2023). Since it first hit the scene, ChatGPT, an AI chatbot built on a powerful language model (LLM), has left both researchers and everyday users in awe with its knack for mimicking human-like traits. One standout trait is empathy, which involves understanding how others feel, viewing situations from their perspective, and genuinely caring about their emotions (Decety and Cowell, 2014). Thanks to this ability, modern AI systems can convey social and emotional signals, which can lead to real-world effects. For example, when AI shows more empathy, it can lead to better user compliance and acceptance of AI services (Adam et al., 2021). Depending on how AI shows empathy, it can also determine the quality of interactions (Vossen et al., 2015).

In the world of corporate cybersecurity, empathetic AI can really help connect the dots between technical alerts and how humans interpret them. Take, for instance SOC that's swamped with a flood of low-level alerts. An AI system that keeps an eye on the cognitive load of analysts can adapt its support on the fly: it might downplay less urgent alerts, offer concise insights, or even recommend collaborative triage when it senses that the team is feeling overwhelmed (as seen in adaptive decision-support frameworks) (Albanese et al., 2025).

It's essential to make a clear distinction between affective empathy, which involves sharing or mirroring someone else's emotional state, and cognitive empathy, where we understand another person's perspective or mental state. While AI systems are primarily concerned with cognitive empathy, humans generally employ both forms (Hou et al., 2024). This contrast highlights the existing philosophical and technical limitations of empathetic AI.

AI is now capable of demonstrating cognitive empathy and generating responses that often come across as empathetic. It's been proven to recognize human emotions (Ong et al., 2019), show facial expressions (Mishra et al., 2023), and effectively promote empathetic interactions (Sharma et al., 2023), thanks to significant advancements in how we approach emotional inference. However, Perry (2023) points out that while AI can mimic empathy, it lacks the ability to convey true concern or to indicate that the person it's interacting with holds any real significance.

When it comes to forensic operations, empathetic AI is truly a game-changer, delivering benefits that reach far beyond mere data analysis. By analyzing behavioral cues, tone, and communication patterns, it can help investigators pinpoint potentially malicious insider actions or social engineering efforts. Rather than just relying on transaction logs, this AI can interpret emotional or hesitant speech and connect it to known threat models. This evolution allows AI to act as a partner in investigations, enriching the collaboration between humans and machines and leading to more insightful and effective outcomes.

More and more people are choosing conversational AI over phone calls, which is one of the key reasons behind the rise of AI-powered customer service. A survey conducted in 2024 found that 82 percent of consumers preferred chatting with a bot rather than waiting for a customer support representative to pick up. Even more striking, 96 percent of respondents felt that businesses should really consider using chatbots instead of relying on traditional customer service (Coppola, 2024). The majority of domain-specific chatbots used for such interpersonal service settings rely solely on text-based communication via a webpage or mobile application that people access via their personal devices (e.g., phone, tablet, laptop, etc.), although some generic conversational AI agents, like Alexa and Siri, use spoken and audio-based interactions (Pamungkas, 2019).

Natural language processing (NLP) is a field of AI that focuses on how computers and humans communicate (using spoken or written natural language). It bridges the gap between humans and computers, allowing computers to understand and interpret human language (Jurafsky and Martin, 2024).

NLP allows designers to carefully assess individual (or micro-level) experiences and expectations by analyzing consumer input (Jylkäs et al., 2025).

In the real world, blending NLP with empathetic AI systems involves juggling various inputs like text, voice tone, and even facial expressions. The whole process revolves around understanding emotions, which shapes how responses are formulated. It begins with spotting an emotional state, assessing its valence and intensity, and then picking or crafting a response that matches a particular empathy strategy (Akhtar et al., 2025). Research has shown that third parties evaluate manifestations of empathy elicited with assistance more favorably than human emotions (Ayers et al., 2023). Additionally, through the analysis of mood and user research data, NLP enables designers to assess consumer feedback to understand their moods, preferences, and desires at the meso level of the service ecosystem (Liu, 2012). NLP helps develop successful content strategies (by assessing user language patterns and interactions) by ensuring that communication is us-er-centered and that the different components of the service coherently respond to user requests (Liddy, 2001).

Successful collaboration requires understanding the dynamics of interaction between AI and humans. Studies have shown that decision-making and user trust can be largely disproportionate to the apparent definitiveness of AI responses (Zhang et al., 2020). To improve collaborative outcomes, AI systems must be designed to be trustworthy (Steyvers et al., 2025). Nevertheless, it can be concluded that one should neither rely too little nor too much on the individuality of AI decision-making. A balanced relationship of human-machine col-laboration will yield the best outcomes.

The learning apprentice model, which is also known as the human-in-the-loop (HITL) approach, treats AI as a supportive teammate for human workers. This idea posits that AI learns by watching how humans make decisions and then uses those insights as additional training data. This not only allows AI to provide assis-tance (in real time), but also helps it learn from a diverse range of human input. In this way, it potentially allows it to surpass the knowledge of individual team members. With that in mind, the effectiveness of this learning process largely depends on the expertise of the human team and the availability of pertinent in-formation (Esteva et al., 2017). This collaboration between AI and humans enables AI to harness human creativity and empathy, which enhances its capacity to identify emotional cues and is essential for creating more compassionate user experiences (Jylkäs et al., 2025).

Furthermore, the latest developments in multimodal emotion recognition, blending speech, facial expres-sions, and text, can really enhance an empathetic AI's skill in recognizing analysts' emotional states on the fly (Alnuaim et al., 2022). By utilizing this multimodal approach, AI systems in SOCs or forensic teams can more effectively adjust their communication and support based on the human operator's emotional and cognitive load, which helps build collaboration and trust.

Researchers took a closer look at a collaborative design that tweaks existing human responses to make them more empathetic. This is different from the AI-only approach, which involves generating empathetic dia-logue from scratch (Lin et al., 2019). By adding new empathic phrases or replacing existing low-empathy sentences with more caring alternatives, this design can significantly improve its empathetic responses (Sharma et al., 2021). Furthermore, once AI is trained to recognize emotions, it can be used to create pro-grams that react in a helpful and appropriate manner. For example, an AI chatbot designed for customer support could be trained to detect and address customer frustration. To enhance customer satisfaction and improve the overall experience, the chatbot might provide additional support, apologize for any inconven-ience, or connect the customer with a live representative when they become upset. The public's acceptance of AI in various sectors largely depends on the trust customers place in it, which is fostered by these seemingly empathetic AI solutions (Clark, 2023).

This highlights the importance of integrating automated systems with human components (Papakonstantinidis et al., 2021). A recent example showed how human-AI collaboration led to more em-pathetic peer-to-peer mental health support conversations, as opposed to responses from humans alone (Sharma et al., 2023). Enabling machines to understand and respond to human emotional states improves the efficiency, naturalness, and ultimately humanness of interactions. This integration has the potential to completely transform human-robot interactions, making them more natural and aligned with human needs and expectations (Kukreja et al., 2022).

A better understanding of how AI can interpret human experiences could be of great help in the ethical development of this new technology. A good example would be the fact that AI is increasingly being used in the care of the elderly, to prevent their cognitive decline and at the same time improve their quality of life. This is achieved through diverse social interactions and individualized cognitive training (Balcombe, 2023). By integrating technological, theoretical, and psychological perspectives, these interdisciplinary methods highlight the complex nature of self-awareness in artificial intelligence.

Melloni et al. (2014) aimed to develop personal assistants that could understand and predict human emotions and social interactions. Their study used deep learning (DL) algorithms to identify social cues and emotional states. The researchers used (unstructured) data from social media, videos, and text messages for their study. The results ultimately showed that AI assistants can correctly recognize social interactions and emotional states. AI systems, however, need to build a good relationship with their human partners to reach their full potential. This will enable harmonious partnerships based on a shared understanding of thoughts or feelings. Sensitivity and humor, which increase likeability and promote cooperation by 30% (Short et al., 2010), could improve this relationship (Niculescu et al., 2013).

The combination of empathetic AI and human interaction is a crucial advancement in enhancing user ex-periences and creating technology that benefits society. By using thoughtfully crafted algorithms that can grasp and react to human emotions, alongside human insight and ethical considerations, we can create AI systems that not only assist with daily tasks but also encourage responsible and socially aware practices. This collaboration between people and empathetic AI opens up a future where technology aligns with our values and needs.

Exploring Applications and Challenges of Empathetic AI in Cybersecurity

Protecting systems, networks, and data from online threats while maintaining availability, confidentiality, and integrity in the digital sphere is known as cybersecurity. It includes a variety of tools, tactics, and procedures to reduce risks and protect important assets (Olagunju et al., 2024). Critical infrastructures like water supplies, petrochemical facilities, nuclear power plants, and transportation infrastructure systems have all been the subject of cyberattacks in recent years, which have compromised sensitive data and caused power outages (Plėta et al., 2020).

AI-driven technologies, however, have both advantages and disadvantages when combined with cyber security. A portion of AI supports cyber resilience, while another component trusses penetration. Although there is a wide range of research on AI-based applications, security concerns are significantly more important than preventative measures (Tyugu, 2011). Machine learning (ML) models empower machines to make predictions or judgments based on a mathematical representation of carefully chosen data, all without needing explicit programming. To kick off AI in cybersecurity, the initial step is to pre-process the data before it's input into the ML models. This method reveals insights into the data by assessing feature im-portance or ranking the key elements, and it's usually quite clear. Grasping how models make decisions is crucial for these strategies, as it helps ensure that we can make the right calls to spot any unusual behavior. After the data has been processed and trained, feature engineering helps identify anomalies or outliers in the dangerous patterns (Wang et al., 2016).

In the world of corporate cybersecurity, empathetic AI can be a game-changer for analysts. It goes beyond just spotting technical glitches; it also offers insights into potentially risky behaviors that take human emo-tions into account. For instance, when the number of alerts spikes, an AI system that recognizes cognitive load and emotional stress can help prioritize warnings in a way that eases the burden and enhances deci-sion-making. This approach strikes a balance between automated detection and human intuition, which is crucial in high-stakes security situations (Mohsin et al., 2025).

To keep an eye on network traffic, we use intrusion detection systems (IDS) that are designed to flag any suspicious behavior or policy violations. Conventional IDS typically rely on rule-based methods and static signatures, which can be limited in their ability to catch new or evolving threats. The game has changed with the advent of artificial intelligence, particularly machine learning and deep learning, which empower these systems to identify unusual patterns and adjust to new attack vectors. Conventional intrusion detection systems (IDS) mainly rely on static signatures and rule-based approaches. IDS cannot always detect new threats and its ability may be limited, but with the advent of AI, positive changes have occurred and these new systems can now recognize all unusual patterns and adapt to them in the right way (Ajibola et al., 2025). Threat intelligence and prediction, to identify possible cybersecurity risks and predict possible future attacks, involve the collection, processing, and interpretation of data. By automating threat detection and pattern identification, AI has significantly advanced this field. For example, AI models using natural language processing (NLP) can analyze text data from social media to extract useful information to improve cyber-security preparedness (Zangana et al., 2025). In the realm of forensic investigations, empathetic AI offers a richer analysis of attacker techniques and insider threats. It doesn't just look at logs or transaction patterns; it also interprets the emotional and behavioral context surrounding communications or suspicious activities. By integrating human involvement into the process, the system can recommend lines of inquiry that balance analytical precision with the human context of incidents, ultimately boosting both technical effectiveness and the sensitivity of investigations (Schaberreiter et al., 2024).

Today's security management systems often allow for the recording of suspicious behavior or data flow from hackers, moving from active defense techniques to more proactive approaches. The technique of enticing the adversaries is commonly known as "honey pot", and it helps system analysts obtain crucial data on the attack signatures and strategies of attackers. Honeypot's analysis can also be used to mimic intrusion detection systems (IDS) or even deflect attacks from important targets (Devi Priya and Chakkaravarthy, 2023). A huge amount of sensitive (real-time threat intelligence) data can be found in the data used by the Honey Pot process. Security analysts and forensics can identify new attack patterns by examining this data (Li et al., 2022).

Whenever AI processes sensitive data for security reasons, privacy concerns arise. Over time, as AI tech-nologies become more sophisticated, as they can draw complex conclusions from aggregated data, these concerns have escalated (Hitaj et al., 2017). However, these technologies are constantly improving, and progress fosters trust and reliability on the digital platform (Anica Popa et al., 2021). But safeguards for the use of this private information aren't always the best ones. The integrity of the person who believes that AI-based technological platforms are reliable may not always be protected by these procedures (Gupta et al., 2023).

AI can provide significant assistance by identifying and warning about malicious software (malware), or providing support in conducting cybersecurity training and vulnerability assessment (Karunamurthy et al., 2023). The use of large language models (LLM) for intrusion detection or phishing emails has recently gained significant popularity (Kheddar, 2024). Convolutional neural networks (CNNs) can extract complex information from high-resolution photographs and contribute to facial recognition with high accuracy (Ali et al., 2021). Similarly, ML-based behavioral biometrics tracks user behavior, which may include keyboard typing speed, mouse movement dynamics, or analyzing touchscreen gestures, to identify patterns or anom-alies that could be signs of fraud (Bansal and Ouda, 2024).

AI cues can easily deceive people into believing they're interacting with sentient beings, which often leads to the use of learned behaviors and scripts. A recent incident involving a finance employee at a multinational company in Hong Kong, who was tricked into transferring a staggering $25 million to scammers, really highlights this problem. During a video call, the scammers posed as the company's chief financial officer, using deepfake technology (Murphy, 2024). This scenario really underscores our natural tendency to trust rather than question, since being skeptical often takes more mental effort (Gilbert, 1991). In simpler terms, if AI can trigger certain responses in people, they might interact with technology just like they do with each other, tapping into those learned social behaviors. To make the most of technology, it's crucial to grasp these cues so we can develop frameworks that encourage its adoption (Kolomaznik et al., 2024).

The potential of empathetic AI in cybersecurity really shines when we look beyond just the technical side and think about how human operators and users emotionally connect with these systems. Empathetic AI agents can keep an eye on an analyst's stress levels during an incident, pick up on any confusion or hesitation in their decision-making, and offer helpful prompts or different perspectives. An AI agent can identify when a security engineer is stuck (for example, in incident assessment) and offer a simpler and more efficient explanation (in order to reduce cognitive fatigue), or suggest that the engineer contact a colleague who can help them make the right decision (Bhatia et al., 2024).

Introducing empathetic AI into cybersecurity practices certainly presents a few unique challenges. One of the main concerns is making sure that emotional cues are understood correctly. We definitely want to avoid any misunderstandings, whether they're false positives or negatives. It is vital to maintain human autonomy, so that AI does not steer our decisions in unwanted directions. In addition, emotional data must be treated with the same seriousness as, for example, transactional or behavioral data. However, there is a possibility that malicious actors can provide misleading emotional signals, which can compromise the monitoring of the system. The way AI is used in advanced social engineering tactics (such as "FraudGPT" and "WormGPT") highlights how fraudsters can exploit technology that mimics human empathy to break trust and bypass defense systems (Falade, 2023).

Humans and machines must have clear roles, and governance frameworks should specify when empathetic AI can intervene and when it is necessary for humans to take full control. It must also be ensured that data on emotional reactions is stored and protected from any misuse. Research in this area points to the idea that collaboration between humans and AI, where empathetic AI provides assistance rather than taking control, could be a fantastic model for cybersecurity operations (Sarker et al., 2023). As the demand for explainable AI (XAI) rises in security contexts, it's becoming increasingly clear that empathetic AI systems need to be upfront about their emotional insights, not just their threat detection abilities (Srivastava et al., 2022). Findings show that people who interacted with an empathic AI were 20% more likely to follow its rec-ommendations, highlighting a significant increase in feelings of support and confidence (Liu and Sundar, 2018). Additionally, Portela and Granell-Canut (2017) looked into how AI's empathetic responses impacted user interactions. Their findings showed that 68% of users felt more emotionally engaged when the AI displayed empathy, compared to those who interacted with non-empathetic AI.

Finally, it is important to note that while the technical elements of AI-driven detection, prediction, and response are undoubtedly vital, the next challenge is to figure out how these systems can better connect with human users for even greater operational effectiveness. AI when designed with empathy in mind in cyber-security can truly be a game-changer. This approach allows teams to remain resilient when the pressure is on. In order to fully exploit this enormous potential, we must confront a complex web of obstacles along the way. A recommendation for future research is that it should focus on how empathetic signals in AI monitoring affect the quality of decision-making in SOCs. Also, to create metrics to measure the effectiveness of emo-tional AI and to develop standardized frameworks in security operations for managing emotional data.

Discussions

Cooperation between people and machines will be significantly improved in the coming period through the use of empathic artificial intelligence and will become a higher priority in organisations, and this will be particularly reflected in cyber security. This advanced type of artificial intelligence is designed with the primary goal of understanding human emotions and reacting accordingly. These empathic systems can contribute to reducing the mental effort of employees at times when they are not making the most adequate decisions, which can be problematic when dealing with environments that function under high pressure, such as those in security operations centres (SOC) (Mohsin et al., 2025). In any case, by prioritising during crises that carry security risks, empathic artificial intelligence can help analysts perform their work tasks more efficiently (Liu and Sundar, 2018).

However, despite the obvious advantages, there are still significant difficulties in this area that will be overcome over time. For example, the incorrect interpretation of emotional cues represents a significant obstacle, because their misinterpretation can produce incorrect suggestions that, in addition to undermining trust, negatively affect efficiency (Hitaj et al., 2017). Furthermore, in practice there are many cases of malicious people who could misuse empathic artificial intelligence and offer untrue emotional cues. In that way, they would bypass security measures and influence distorted decision-making (Falade, 2023). In the scientific literature, most research on empathic artificial intelligence in cybersecurity is currently theoretical or based only on simulations, a fact that points to the need for further research to address the practical consequences, reliability and moral implications of these systems (Sarker et al., 2023).

In order to introduce empathetic artificial intelligence, in addition to the moral and technical aspects, it is essential to establish a culture of cyber security, which will recognise the value of emotional intelligence and be ready to accept the insights of artificial intelligence. Training employees to better understand feedback received through artificial intelligence is key, but care must be taken that over-reliance on it or misinter-pretation of emotional cues can have the opposite effect. It is also crucial that the human element remains a priority in collaboration with artificial intelligence (Kolomaznik et al., 2024). Of course, during analyst training, top priority must be given to emotional literacy in order to encourage fruitful collaboration with these systems (Liu and Sundar, 2018). The synergy of emotional intelligence with instant detection of cyber threats is a new strategy, which aims to overcome fatigue and provide suggestions to analysts. This com-bination of humans and technology represents a strategic advantage (Bhatia et al., 2024), as it improves efficiency, contributes to better and smarter task allocation, and, equally importantly, shortens incident response times in SOCs (Liu and Sundar, 2018).

Studies investigating this kind of collaboration between humans and artificial intelligence in SOCs have revealed a worrying trend. It has been observed (although not recommended) that analysts are increasingly relying on artificial intelligence systems to make decisions, which can cause a risk of overburdening and the reduction and dulling of skills. Therefore, empathetic artificial intelligence helps human analysts, but only as long as it is used in the right way and does not affect their individual situational awareness and does not undermine their decision-making skills (Yaich et al., 2024). All of this indicates that it will take time for human analysts and parallel developing artificial intelligence systems to adapt to each other. Recent research has shown that humans working with artificial intelligence can significantly improve their ability to identify delicate cyber threats (Cohen et al., 2025). Adapting your answers to the way human analysts think builds trust and ensures consistency in understanding while performance remains high.

Conclusions

In cybersecurity, empathetic artificial intelligence is poised to improve decision-making, build and nurture trust, and increase the resilience of human analysts. These systems reduce cognitive effort and through human collaboration improve outcomes in high-stress cyber security scenarios by engaging emotional awareness (Kolomaznik et al., 2024). By reducing fatigue and prioritizing, allowing analysts to concentrate on the most pressing security issues, empathic artificial intelligence can significantly increase workplace efficiency (Liu and Sundar, 2018). However, we must overcome a number of organizational, ethical and technical challenges, including avoiding aggressive exploitation, to fully exploit these benefits.

To fully understand how empathic AI impacts decision-making, future research should focus on its use in security operations centers (SOCs). Also, collaboration on a common goal between experts from many fields is necessary for the development of this technology to be appropriate (such as AI and cybersecurity experts, but also psychologists and human-computer interface experts). This is the only right path to a balanced development of empathic AI that will not only improve security, but will also be in line with both human values and operational needs. In any case, the potential of this technology to significantly advance the field of cybersecurity is undeniable, primarily by building better interactions between humans and AI through careful consideration of its design and management.

Modern organizations facing real and potential cybersecurity challenges should definitely consider using agent-based or empathetic artificial intelligence systems in the coming period. These systems must be flexible enough to adapt fully to the analyst's mental states as well as the number of alerts. This field (human-technology collaboration at an empathetic level) is still in its early stages of development and further scientific studies are needed to better understand the long-term impacts of these systems on human competence, self-confidence, short-term productivity gains and, perhaps most importantly, cognitive resilience.

Acknowledgements

The paper presents findings of a study developed as a part of the research project "Serbia and challenges in international relations in 2026", financed by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, and conducted by Institute of International Politics and Economics, Belgrade during year 2026.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interests

The authors declare no conflict of interest.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Institutional Review Board Statement

Not applicable.

Author Contributions

Conceptualization, R.B., M.S., formal analysis, D.I., M.A. writing—original draft preparation, R.B.; writing—review and editing, M.S., D.I., M.A. All authors have read and agreed to the published version of the manuscript.

References

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets, 31*(2), 427–445. https://doi.org/10.1007/s12525-020-00414-7

Ajibola, O. S., Dopamu, O., & Olurin, O. (2025). Challenges and ethical implications of using AI in cyber security. *International Journal of Science and Research Archive, 14*(2), 294–304. https://doi.org/10.30574/ijsra.2025.14.2.0276

Akeiber, H. J. (2025). A comprehensive study of cybercrime and digital forensics through machine learning and AI. *Rafidain Journal of Engineering Sciences, 3*(1), 369–395. https://doi.org/10.61268/hff1pp49

Akhtar, M., Mumtaz, A., Zaman, M., Ameer, S., Fatima, H. S., Shahbaz, N., & Samar Bukhari, S. S. (2025). Emotionally intelligent AI: Enhancing human–computer interaction through NLP and sentiment anaysis. *Policy Research Journal, 3*(5), 160173. https://policyrj.com/index.php/1/article/view/621

Albanese, M., Ou, X., Lende, D., & Goldgof, D. (2025). Towards AIdriven humanmachine coteaming for adaptive and agile cyber security operation centers. *arXiv*. https://arxiv.org/abs/2505.06394

Ali, W., Tian, W., Din, S. U., Iradukunda, D., & Khan, A. A. (2021). Classical and modern face recognition approaches: A complete review. *Multimedia Tools and Applications, 80*(3), 4825–4880. https://doi.org/10.1007/s11042-020-09850-1

Alnuaim, A. A., Zakariah, M., Shukla, P. K., Alhadlaq, A., Hatamleh, W. A., Tarazi, H., Sureshbabu, R., & Ratna, R. (2022). Human-computer interaction for recognizing speech emotions using multilayer per ceptron classifier. *Journal of Healthcare Engineering, 2022*, 6005446. https://doi.org/10.1155/2022/6005446

AnicaPopa, I., AnicaPopa, L., Rădulescu, C., & Vrîncianu, M. (2021). The integration of artificial intelligence in retail: Benefits, challenges and a dedicated conceptual framework. *Amfiteatru Economic, 23*(56), 120–136. https://doi.org/10.24818/EA/2021/56/120

Ashworth, P., & Lucas, U. (2000). Achieving empathy and engagement: A practical approach to the design, conduct and reporting of phenomenographic research. *Studies in Higher Education, 25*(3), 295–308. https://doi.org/10.1080/713696156

Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal*

Medicine, 183(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

Balcombe, L. (2023). AI chatbots in digital mental health. *Informatics, 10*(4), Article 100. https://doi.org/10.3390/informatics10040082

Bansal, P., & Ouda, A. (2024). Continuous Authentication in the Digital Age: An Analysis of Reinforcement Learning and Behavioral Biometrics. *Computers, 13*(4), 103. https://doi.org/10.3390/computers13040103

Baltezarević, R., & Baltezarević, I. (2024). Students' Attitudes on The Role of Artificial Intelligence (AI) in Personalized Learning. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE*), 12(2), 387-397. https://doi.org/10.23947/2334-8496-2024-12-2-387-397

Baltezarević, R. & Baltezarević, I., (2025). Digital Game-Based Learning's (DGBL) Effect on Students' Academic Performance. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 13*(1), 127-140. https://doi.org/10.23947/2334-8496-2025-13-1-127-140

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. S. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445717

Baruwal Chhetri, M., Tariq, S., Singh, R., Jalalvand, F., Paris, C., & Nepal, S. (2024). Towards human-AI teaming to mitigate alert fatigue in security operations centres. *ACM Transactions on Internet Technology, 24*(3), 1–22. https://doi.org/10.1145/3670009

Bhatia, V., Sekra, P., & Sweta, S. (2024). Empathetic Intelligence: LLMbased conversational AI voice agent. *International Journal of Intelligent Systems and Applications in Engineering, 12*(4), 4400–4408. https://ijisae.org/index.php/IJISAE/article/view/7075

Brink, T. T., Urton, K., Held, D., Kirilina, E., Hofmann, M., KlannDelius, G., Jacobs, M. A., & Kuchinke, L. (2011). The role of orbitofrontal cortex in processing empathy stories in 4 to 8yearold children. *Frontiers in Psychology, 2*, 80. https://doi.org/10.3389/fpsyg.2011.00080

Cadiente, A., Chen, J., & Pilkington, B. (2023). *Machine-Made Empathy? Why Medicine Still Needs Humans. JAMA Internal Medicine, 183*(11), 1278-1279. https://doi.org/10.1001/jamainternmed.2023.4386

Clark, S. (2023, October). Emotionally intelligent AI refers to AI systems that can understand, recognize and respond to human emotions. *CMS Wire.* https://www.cmswire.com/digital-experience/ai-gets-empathetic-advances-in-emotionally-intelligent-ai/

Cohen, D., Te'eni, D., Yahav, I., et al. (2025). Human–AI enhancement of cyber threat intelligence. *International Journal of Information Security, 24*, 99. https://doi.org/10.1007/s10207-025-01004-4

Coppola, D. (2024, November). Consumer opinions on conversational AI for customer service 2024. *Statista.* https://www.statista.com/statistics/1538260/consumer-opinions-on-conversational-ai/

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2014). Empathy: A review of the concept. *Emotion Review, 8*(2), 144–153. https://doi.org/10.1177/1754073914558466

Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences, 18*(7), 337–339. https://doi.org/10.1016/j.tics.2014.04.008

Devi Priya, V. S., & Chakkaravarthy, S. S. (2023). Containerized cloud-based honeypot deception for tracking attackers. *Scientific Reports, 13*(1), 1437. https://doi.org/10.1038/s41598-022-26977-3

Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. (2024). A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Science International: Digital Investigation, 48*, Article 301675. https://doi.org/10.1016/j.fsidi.2023.301675

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://doi.org/10.1038/nature21056

Falade, P. V. (2023). Decoding the Threat Landscape: ChatGPT, FraudGPT, and WormGPT in Social Engineering Attacks. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 9* (5), 185-198. https://doi.org/10.48550/arXiv.2310.05595

Ghotbi, N. (2023). The ethics of emotional artificial intelligence: A mixed-method analysis. *Asian Bioethics Review, 15*(4), 417–430. https://doi.org/10.1007/s41649-022-00237-y

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107–119. https://doi.org/10.1037/0003-066X.46.2.107

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy. *IEEE Access, 11*, 80218–80245. https://doi.org/10.1109/ACCESS.2023.3300381

Hirsch, T., Merced, K., Narayanan, S., Imel, Z. E., & Atkins, D. C. (2017). Designing contestability: Interaction design, machine learning, and mental health. In *Proceedings of the 2017 ACM Conference on Designing Interactive Systems (DIS '17)* (pp. 95–99). Association for Computing Machinery. https://doi.org/10.1145/3064663.3064703

Hitaj, B., Ateniese, G., & PérezCruz, F. (2017). Deep models under the GAN: Information leakage from collaborative deep learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS)* (pp. 603–618). ACM. https://doi.org/10.1145/3133956.3134012

Hou, H., Ni, S., Lin, S., & Wang, P. (2024). When AI learns to empathise: Topics, scenarios, and optimisation of empathy computing from a psychological perspective. *Advances in Psychological Science, 32*(5), 845–858. https://doi.org/10.3724/SP.J.1042.2024.00845

Hussain, S., Neekhara, P., Jere, M., Koushanfar, F., & McAuley, J. (2021). Adversarial deepfakes: Evaluating vulnerability of deepfake detectors to adversarial examples. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2021)* (pp. 3348–3357). IEEE. https://doi.org/10.1109/WACV48630.2021.00339

Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition with language models* (3rd ed., draft). Pearson. https://web.stanford.edu/~jurafsky/slp3/

Jylkäs, T., Song, C., & Miettinen, S. (2025). Enhancing empathy through AI in service systems. In M. Suoheimo, P. H. Jones, S.–H. Lee, & B. Sevaldson (Eds.), *Systemic Service Design* (Chapter 13, pp. 258-276). Routledge. https://doi.org/10.4324/9781003501039-15

Karunamurthy, A., Kiruthivasan, R., & Gauthamkrishna, S. (2023). *HumanintheLoop Intelligence: Advancing AICentric Cybersecurity for the Future. Quing: International Journal of Multidisciplinary Scientific Research and Development, 2*(3), 20–43. https://doi.org/10.54368/qijmsrd.2.3.0011

Kheddar, H. (2024). Transformers and large language models for efficient intrusion detection systems: A comprehensive survey (arXiv preprint *arXiv:2408.07583*). https://doi.org/10.48550/arXiv.2408.07583

Kolomaznik, M., Petrik, V., Slama, M., & Jurik, V. (2024). The role of socio-emotional attributes in enhancing human-AI collaboration. *Frontiers in Psychology, 15*, 1369957. https://doi.org/10.3389/fpsyg.2024.1369957

Kukreja, V., Solanki, V., Baliyan, A., & Jain, V. (2022). WeatherNet: Transfer learningbased weather recognition model. In *Proceedings of the 2022 International Conference on Emerging Smart Computing and Informatics (ESCI)* (pp. 1–5). IEEE. https://doi.org/10.1109/ESCI53509.2022.9758183

Li, Z., Zeng, J., Chen, Y., & Liang, Z. (2022). AttackG: Constructing technique knowledge graph from cyber threat intelligence reports. In V. Atluri, R. Di Pietro, C. D. Jensen, & W. Meng (Eds.), *Computer Security – ESORICS 2022, Lecture Notes in Computer Science* (Vol. 13554, pp. 589609). Springer. https://doi.org/10.1007/978-3-031-17140-6_29

Liddy, E. D. (2001). Natural language processing. In M. E. Williams (Ed.), *Encyclopedia of library and information science* (2nd ed., pp. 2126–2136). Marcel Dekker, Inc. https://surface.syr.edu/cgi/viewcontent.cgi?article=1043&context=istpub

Lin, Z., Madotto, A., Shin, J., Xu, P., & Fung, P. (2019). MoEL: Mixture of empathetic listeners. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the *9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 121–132). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1012

Liu, B. (2012). *Sentiment analysis and opinion mining.* Morgan & Claypool Publishers. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior, and Social Networking, 21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110

Main, A., Walle, E. A., Kho, C., & Halpern, J. (2017). The interpersonal functions of empathy: A relational perspective. *Emotion Review, 9*(4), 358–366. https://doi.org/10.1177/1754073916669440

Martin, T. (2022). On the Need for Collaborative Intelligence in Cybersecurity. *Electronics*, *11*(13), 2067. https://doi.org/10.3390/electronics11132067

Melloni, M., Lopez, V., & Ibanez, A. (2014). Empathy and contextual social cognition. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 407–425. https://link.springer.com/article/10.3758/s13415-013-0205-3

Mishra, C., Verdonschot, R. G., Hagoort, P., & Skantze, G. (2023). Real-time emotion generation in human-robot dialogue using large language models. *Frontiers in Robotics and AI, 10*, Article 1271610. https://doi.org/10.3389/frobt.2023.1271610

Mohsin, A., Janicke, H., Ibrahim, A., & Sarker, I. H. (2025). A unified framework for human–AI collaboration in Security Operations Centers with trusted autonomy. *arXiv*. https://doi.org/10.48550/arXiv.2505.23397

Murphy, S. (2024, February). A Deepfake Scammed a Bank out of $25M — Now What? Trend Micro. https://www.trendmicro.com/en_gb/research/24/b/deepfake-video-calls.html

Niculescu, A., van Dijk, B., Nijholt, A., Li, H., & See, S. L. (2013). Making social robots more attractive: The effects of voice pitch, humor and empathy. *International Journal of Social Robotics, 5*(2), 171–191. https://doi.org/10.1007/s12369-012-0171-x

Olagunju, H. I., Alozie, E., Imoize, A. L., Faruk, N., Garba, S., & Baba, B. A. (2024). Cybersecurity in the Internet of Medical Things for healthcare applications. In A. L. Imoize, C. Meshram, J. B. Awotunde, Y. Farhaoui, & D.–T. Do (Eds.), *Cybersecurity in Emerging Healthcare Systems* (pp. 41–74). Institution of Engineering and Technology. https://doi.org/10.1049/PBHE064E_ch2

Ong, D. C., Zaki, J., & Goodman, N. D. (2019). *Computational models of emotion inference in theory of mind: A review and roadmap. Topics in Cognitive Science, 11*(2), 338–357. https://doi.org/10.1111/tops.12371

Pamungkas, E. W. (2019). Emotionallyaware chatbots: A survey. *arXiv*. https://arxiv.org/abs/1906.09774

Papakonstantinidis, S., Kwiatek, P., & Baltezarević, R. (2021). The impact of relationship quality and selfservice technology on company performance. *Polish Journal of Management Studies, 23*(1), 315326. https://doi.org/10.17512/pjms.2021.23.1.19

Perry, A. (2023). AI will never convey the essence of human empathy. *Nature Human Behaviour, 7*(11), 1808–1809. https://doi.org/10.1038/s41562-023-01675-w

Plėta, T., Tvaronavičienė, M., Della Casa, S., & Agafonov, K. (2020). Cyber-attacks to critical energy infrastructure and management issues: Overview of selected cases. *Insights into Regional Development, 2*(3), 802–813. https://doi.org/10.9770/ird.2020.2.3(5)

Portela, M., & Granell-Canut, C. (2017). A new friend in our smartphone? Observing interactions with chatbots in the search of emotional engagement. In *Proceedings of the XVIII International Conference on Human-Computer Interaction* (pp. 1–7). ACM Press. https://doi.org/10.1145/3123818.3123826

Rafy, M. F. (2024). Artificial intelligence in cyber security. *SSRN*. https://doi.org/10.2139/ssrn.4687831

Richter, F. (2025, April 24). Do people underestimate their encounters with AI? *Statista*. https://www.statista.com/chart/34346/how-often-do-people-interact-with-ai/

Sarker, I. H., Janicke, H., Mohammad, N., Watters, P., & Nepal, S. (2023, October). AI potentiality and awareness: A position paper from the perspective of humanAI teaming in cybersecurity. In *International Conference on Intelligent Computing & Optimization* (pp. 140149). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-50887-5_14

Schaberreiter, T., Andriessen, J., Cappiello, C., Papanikolaou, A., & Pardijs, M. (2024). Human-in-the-loop anomaly detection and contextual intelligence for enhancing cybersecurity management. In *Proceedings of the First International Conference on Natural Language Processing and Artificial Intelligence for Cyber Security* (pp. 127–136). Lancaster, UK. https://aclanthology.org/2024.nlpaics-1.15/

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2021). Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference 2021 (WWW '21)* (pp. 2669–2681). Association for Computing Machinery. https://doi.org/10.1145/3442381.3450012

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2023). Human–AI collaboration enables more empathic conversations in textbased peertopeer mental health support. *Nature Machine Intelligence, 5*(1), 4657. https://www.nature.com/articles/s42256-022-00593-2

Short, E., Hart, J., Vu, M., & Scassellati, B. (2010). No fair!! An interaction with a cheating robot. In *2010 5th ACM/IEEE International Conference on HumanRobot Interaction (HRI)* (pp. 219–226). IEEE. https://doi.org/10.1109/HRI.2010.5453193

Srivastava, G., Jhaveri, R. H., Bhattacharya, S., Pandya, S., Hall, J. G., Alazab, M., & Gadekallu, T. R. (2022). XAI for cybersecurity: State of the art, challenges, open issues and future directions *[Preprint]*. arXiv. https://arxiv.org/abs/2206.03585

Steyvers, M., Tejeda, H., Kumar, A., Belem, C., Karny, S., Hu, X., Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence, 7*(2), 221–231. https://www.nature.com/articles/s42256-024-00976-7

Tyugu, E. (2011). *Artificial intelligence in cyber defense.* In *2011 3rd International Conference on Cyber Conflict (CYCON)* (pp. 1–11). IEEE. https://ieeexplore.ieee.org/abstract/document/5954703

Vossen, H. G., Piotrowski, J. T., & Valkenburg, P. M. (2015). Development of the adolescent measure of empathy and sympathy (AMES). *Personality and Individual Differences, 74*, 66–71. https://doi.org/10.1016/j.paid.2014.09.040

Wang, H., Lei, Z., Zhang, X., Zhou, B., & Peng, J. (2016). Machine learning basics. *Deep learning*, 98-164.

Yaich, R., Balondrade, A., Sicard, A., Fouquiau, C., Giraud, G., Amokrane-Ferka, K., & Arbaretier, E. (2025). Symbiotic human–AI collaboration for augmented cybersecurity operations. *Proceedings of the AAAI Symposium Series, 6*(1), 350–358. https://doi.org/10.1609/aaaiss.v6i1.36072

Zangana, H. M., Mustafa, F. M., Li, S., & AlKaraki, J. N. (2025). *Natural language processing for cyber threat intelligence in a quantum world.* In *Advances in Information Security, Privacy, and Ethics* (pp. 345388). IGI Global. https://doi.org/10.4018/979-8-3373-1102-9.ch011

Zhang, Y., Liao, Q. V., & Bellamy, R. K. (2020). Effect of confidence and explanation on accuracy and trust calibration in AIassisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 295–305). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372852

Zhou, L., Gao, J., Li, D., & Shum, HY. (2019). The design and implementation of XiaoIce, an empathetic social chatbot. *arXiv*. https://doi.org/10.48550/arXiv.1812.08989