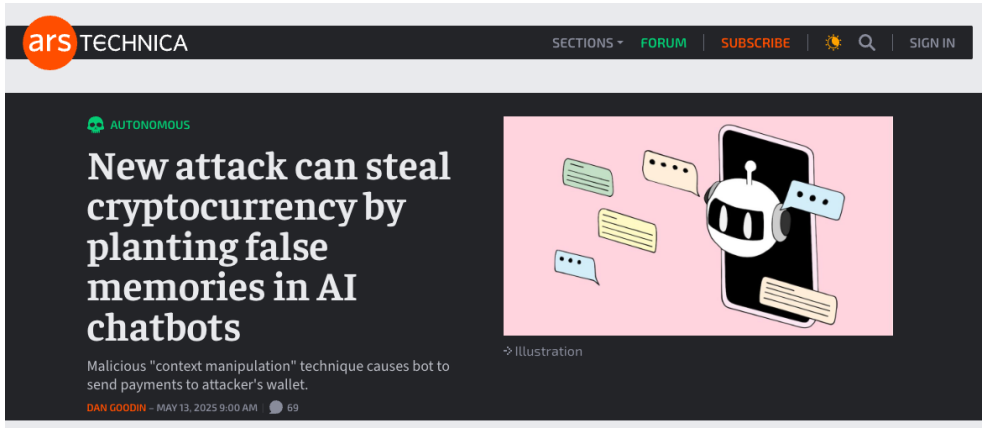
JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

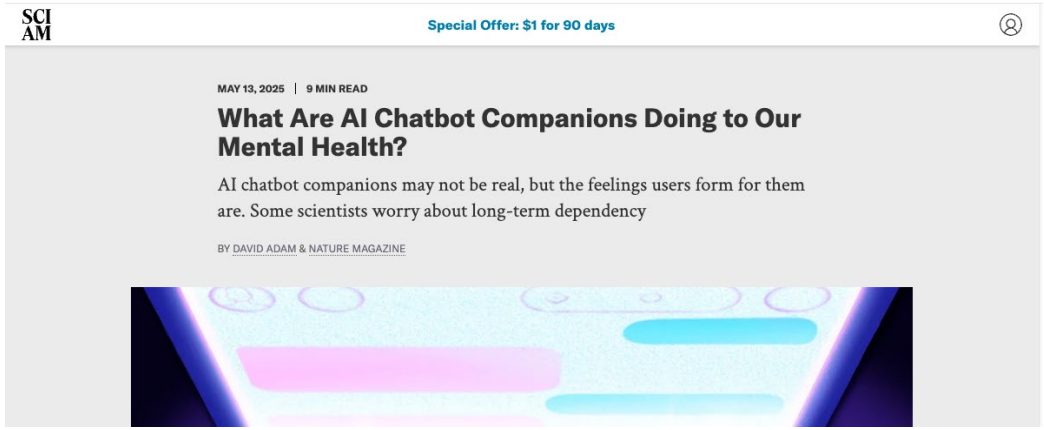
Snapshot:
AI Chatbots

Presented by Dr. Kira Allmann, JCOTS Chief Policy Analyst

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

Introduction

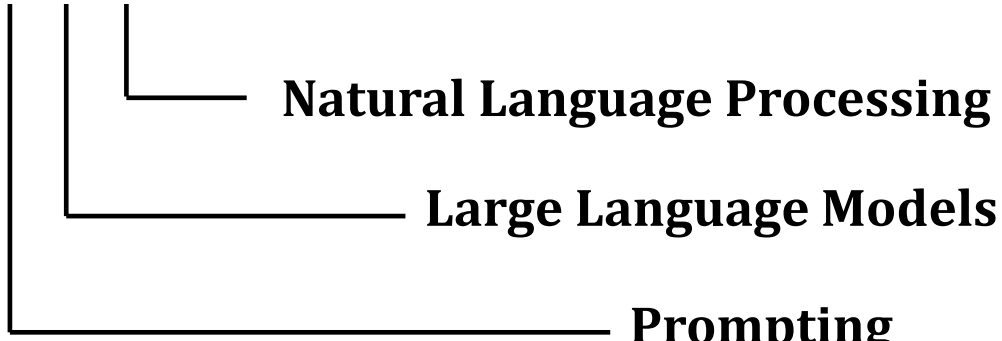




JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

Key Concepts

AI chatbots: artificial intelligence systems that users can interact with using natural language (rather than computer code) and that can perform complex tasks like searching the Internet, summarizing or editing documents, writing code, or generating new content, like images or poems.



JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

Types of AI Chatbots

AI Assistants	AI Companions	AI Agents
Designed to assist users with various tasks through natural language interaction, such as setting reminders, answering questions, drafting emails, summarizing meetings or documents, or controlling smart devices. **Examples:** ChatGPT (OpenAI), Claude (Anthropic), Siri (Apple), Gemini (Google), CoPilot (Microsoft)	Designed to offer emotional support or social interaction, often using NLP, sentiment analysis, and personalized behavior to create a more human-like, engaging experience and provide users with a sense of connection and support. **Examples:** Replika (Luka, Inc.), Character.ai, Eva (Novi), My AI (Snap, Inc.), Nomi.ai	Designed to perceive the environment, process information, and takes sequential actions to achieve specific goals and make decisions based on predefined objectives or learned behaviors. **Possible examples:** IBM watsonx, Orchestrate, Operator (OpenAI), Claude 3.5 Sonnet Model (Anthropic), Agentforce (Salesforce)

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Health & Wellbeing

Emotional connection: The natural language style of chatbots often leads users to perceive them as having human-like characteristics, leading users to trust them and even form emotional bonds with them.

Double-edged sword: Early evidence shows that chatbots can reduce loneliness and improve social skills, but they can also foster emotional dependency, addiction, and obsession, leading to harmful outcomes.

Financial costs: Chatbot companies have commercial incentives to hold users' attention, and users may have to pay high prices to maintain a relationship with a chatbot.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Privacy

Consent and disclosure: People are more likely to disclose sensitive information to a human-like chatbot, but platforms rarely obtain meaningful user consent for the multitude of ways they collect, use, and sell user data.

Data breaches: If a chatbot is compromised, sensitive user information could be leaked.

Targeting and discrimination: Much like other platforms that collect and process large amounts of user data and meta-data, chatbots can be used to target people with predatory advertising or subject them to discriminatory pricing.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Security

Customizability and accessibility: The availability of open-source LLMs means that users can customize their own AI chatbots relatively easily, with little technical training, and put them to good or bad uses.

Malicious chatbots: Chatbots can be developed with malicious intent (such as FraudGPT), and even standard, productivity-oriented chatbots can be directed to provide illegal information or perform criminal acts with the right prompts.

Data breaches and leakage: Chatbots have already been subject to several high-profile data breaches, and models are known to leak data that is provided through user interactions or held in the model's memory from the training data.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Transparency & Explainability

Lack of transparency: LLM-based chatbots are trained on large datasets that are often treated as trade secrets, which hampers oversight and accountability.

Bias and discrimination: Chatbots can replicate or deepen damaging societal biases, compounded by the fact that people often disproportionately trust what chatbots say.

Hard to explain: Even chatbot designers may find it difficult to explain how chatbots generate outputs, and even well-tuned models can generate harmful content.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Workforce Development

AI literacy: A lack of knowledge about how chatbots work can expose users to risks, so developing broad AI literacy will be necessary as these tools become more widespread.

Prompting: Prompting allows people to program LLMs, and users will increasingly need to develop skills in prompting to produce desired outcomes and to interpret the results.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

POLICY ISSUES

Democratic Culture

Leveling the field: AI chatbots provide accessible, affordable, and personalized services, supporting users in areas like education, public services, and mental health, democratizing information and access to essential services.

Social influence: Chatbots have the potential to shape democratic culture by filtering information and reinforcing biases, potentially leading to negative effects on decision-making and societal trust.

Power and responsibility: The companies behind successful chatbot products will have greater market power and more influence over culture and human decision-making, raising questions about who takes responsibility when things go wrong.

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

Policy Considerations

- *What are the commercial motivations behind problematic chatbot tendencies, such as sycophancy or encouraging personal disclosures? What incentives would urge the chatbot market to put consumer protection first?*
- *Does proposed AI legislation cover AI chatbots? What additional requirements or disclosures might be required for chatbots?*
- *What are the specific mechanisms through which chatbots elicit personal or sensitive information from users? How could those mechanisms be regulated in certain settings or use cases?*
- *What skills do people need in order to interact with chatbots productively and safely? Are these skills covered by any existing AI literacy curricula or frameworks?*
- *How can the public participate more directly in decisions about how AI chatbots are developed and deployed, particularly in public services, like healthcare?*

JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

Thank you