Examining Artificial Intelligence and its Role in Malevolent Creativity and Innovation: How AI Can Enhance Novel Threats and Attacks

Alexis L. d'Amato
Madison Scott
Emma Theobald
Averie E. Linnell
Joel Elson

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unomaha.edu/ncitereportsresearch
Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE
Authors
Alexis L. d'Amato, Madison Scott, Emma Theobald, Averie E. Linnell, Joel Elson, Sam T. Hunter, and Jack Rygg
Examining Artificial Intelligence and its Role in Malevolent Creativity and Innovation: How AI Can Enhance Novel Threats and Attacks

Systematic Review Plain Language Summary
For references, see longer report.

Alexis d’Amato, Madison Scott, Emma Theobald, Averie Linnell, Jack Rygg, Joel Elson, Ph.D., & Sam Hunter, Ph.D.

February 2024
OVERVIEW

- This review attempts to prepare the Homeland Security Enterprise for the impending use of AI in generation of novel and malign threats.
- AI broadly refers to technology that can infer patterns from external data sources with numerous subfields including machine learning, logical agents, generative AI, and robotics.
- The benefits of AI may be overshadowed by potential harms such as AI-based autonomous weaponry and malign AI interferences such as disruption of political information integrity.
- The intersection of artificial intelligence and malevolent creativity involves using technology to generate and implement novel ideas to cause harm.

IMPLICATIONS

1. By itself, AI is not creative but is used as a tool to aid and enhance creativity, hence AI-facilitated creativity.
2. AI should not be viewed as standalone technology, but can express multiple capabilities (e.g. metaverse and gaming).
3. As malign use of AI increases, tools to combat these malign uses will also emerge, meaning that security and counterterrorism professionals can think and act creatively to predict and prepare for novel threats.
4. Future research should focus on AI-generated deepfakes specifically within radicalization, recruitment, fundraising, and pornography made for malicious use.

WHAT IS ARTIFICIAL INTELLIGENCE?

AI broadly refers to technology that can infer patterns from external data sources, make predictions, and draw conclusions. Various definitions of AI exist depending on where it is being used. For the purpose of this review, AI will be referred to as a software system designed by humans that, given a complex goal, is able to make a decision based on a process of perception, interpretation and reasoning drawn from data collected about the environment and that meets the properties of autonomy, adaptability, and interactivity.
Part I: AI Overview

Machine Learning (ML)
A subfield of AI, employs algorithms and techniques that enable machines to identify patterns, make predictions, and adapt to novel environments. Common sub-classifications include: supervised learning, unsupervised learning, and reinforcement learning.

Knowledge Representation, Reasoning, and Planning
A subfield of AI aims to equip artificial intelligence systems with the capacity to effectively manage and manipulate information, draw meaningful inferences, and formulate strategic plans.

- Logical Agents
  A type of AI designed to perform reasoning and make decisions based on formal logic rather than internal knowledge.

- Fuzzy Logic
  A type of logic that deals with the gray in-between rather than other systems that use classical logic.

- Expert Systems
  Emulate human experts’ decision-making abilities by solving complex problems using knowledge to formulate predictions and solutions.

Communicating, Perceiving, and Acting
A subfield of AI seeks to imbue machines with the ability to interact seamlessly with the world around them.

- Generative AI (GAI)
  A subset of AI that can be learned and trained on data including text, images, and audio to reproduce and create new content.

- Robotics
  Can enable machines to not only match human abilities in the physical world but surpass human performance.
Creativity: the generation of ideas that are novel and useful in a given context in response to an ill-defined or ambiguous problem.

Malevolent Creativity is the generation of ideas that are novel and useful with the prime goal of causing physical or psychological harm.

Problem Finding and Construction
To find, define, and develop their own understanding of a problem before they can attempt to solve it requires finite resources, such as attention and perception.

Information Gathering
Information gathering allows individuals to obtain preexisting information using searches, usually dependent on the internet, the expertise of their team, or their own knowledge.

Idea Generation
The core phase of the creative process where multiple alternate ideas or solutions are created that are both novel and useful.

Idea Evaluation
The intermediate process between idea generation and idea selection, where ideas are determined to viable options that are worth pursuing and allocating necessary resources.

Idea Selection
The outcome of idea evaluation, in which the idea that most adequately addresses the goals of the creative effort to ensure the selection of the most viable idea for effective implementation.

Implementation Planning
Implementation of a creative idea to maximize resources, ensure implemented idea aligns with organization’s mission or goal, and ensures the implementation over a set time.
AI Enhancing the Malevolently Creative Process

Problem Finding and Construction
AI, being algorithmically based, can understand complex environments and sort and filter information instantly, expediting the creative process. Extremists could leverage AI to improve problem construction through identification of individuals to target for radicalization and recruitment.

Information Gathering
AI can eliminate the time consumption of information gathering and be used to generate and select ideas. Also, human dependence on AI to conduct the information gathering phase could result in the spread of mis-dis-mal (MDM) information.

Idea Generation
GAI platforms (e.g. ChatGPT) generate and summarize data so users can understand comprehensible information quickly. GAI could also enhance idea generation that would allow for development of ideas the user did not think of beforehand.

Idea Evaluation
AI provides users with more available data to evaluate ideas based on their novelty and quality, allowing users to visualize the outcomes, in turn increasing the computing power and available data that contributes to evaluating and selecting ideas.

Idea Selection
AI implications fall more so on the preceding stages, but can assist in selecting which idea is most likely to succeed and by expediting the preceding stages, allow idea selection to happen quickly.

Implementation Planning
This phase of the creativity process is where most ideas fail, causing an individual or group to begin the creative process over. AI could help avoid idea failure by gathering information, generating ideas, and evaluating ideas that would effectively match the desired outcome.

The material in this report is based on work supported by the U.S. Department of Homeland Security under Grant Award Number 20STTPC00001-04-00. The views and conclusions included here are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.