CHECK POINT

CYBER WARFARE 2023: AI, CHATGPT, AND BEYOND

25th VIPSS Conference 22-24 FEB 2023

Michele Boland Evangelist and Architect
Check Point Software Technologies Office of the CTO
linkedin.com/in/micheleboland

YOU DESERVE THE BEST SECURITY

AGENDA

* Cyber Warfare 2022 year in review: deepfakes, automated bots, Generative AI
* 2023: ChatGPT (OpenAI) risk, reward, tradeoff
* Call to action for cybersecurity, GRC, and data protection and privacy professionals

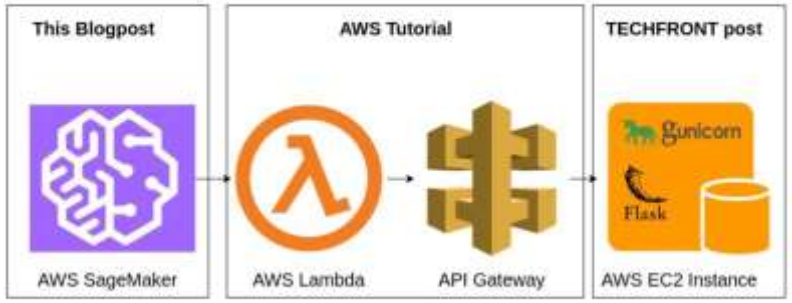
DEEPFAKES, VOICEFAKES, NEWSFAKES

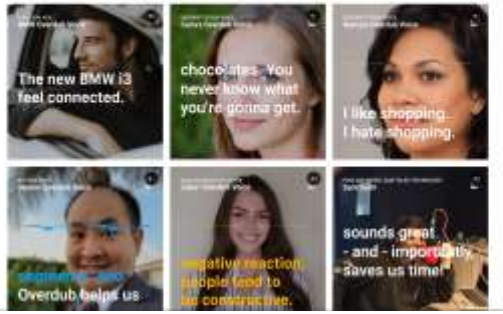
DEEPFAKES

#deepfake #morganfreeman
This is not Morgan Freeman - A Deepfake Singularity

A nonspeaking valedictorian with autism gives her college's commencement speech

In this blogpost, we will cover the first task in detail. Two others are covered in AWS Tutorial, TECHFRONT post. The final architecture will look like this:

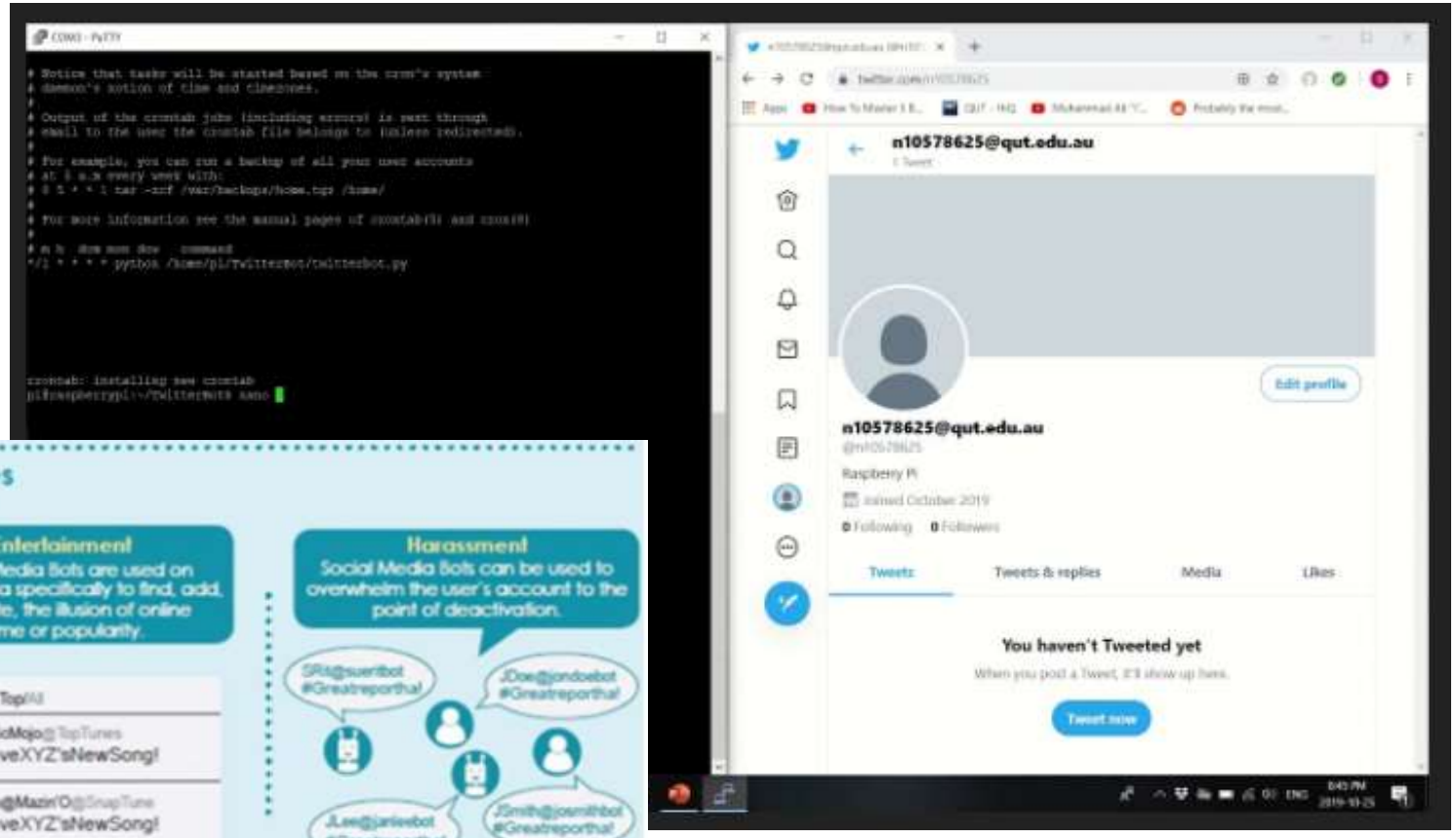


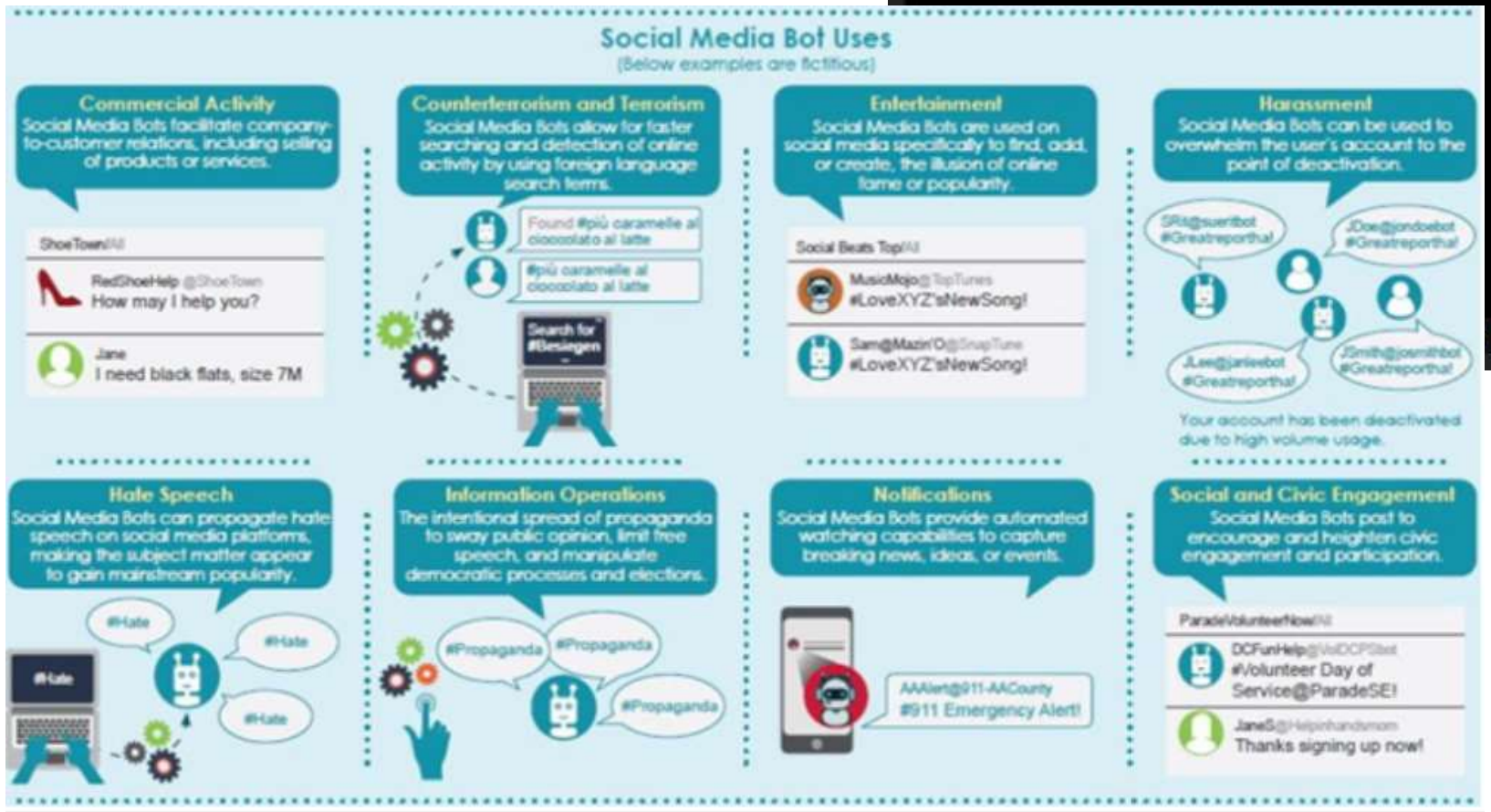


Credit: creative commons internet

AUTOMATED BOTS

AUTOMATED BOTS

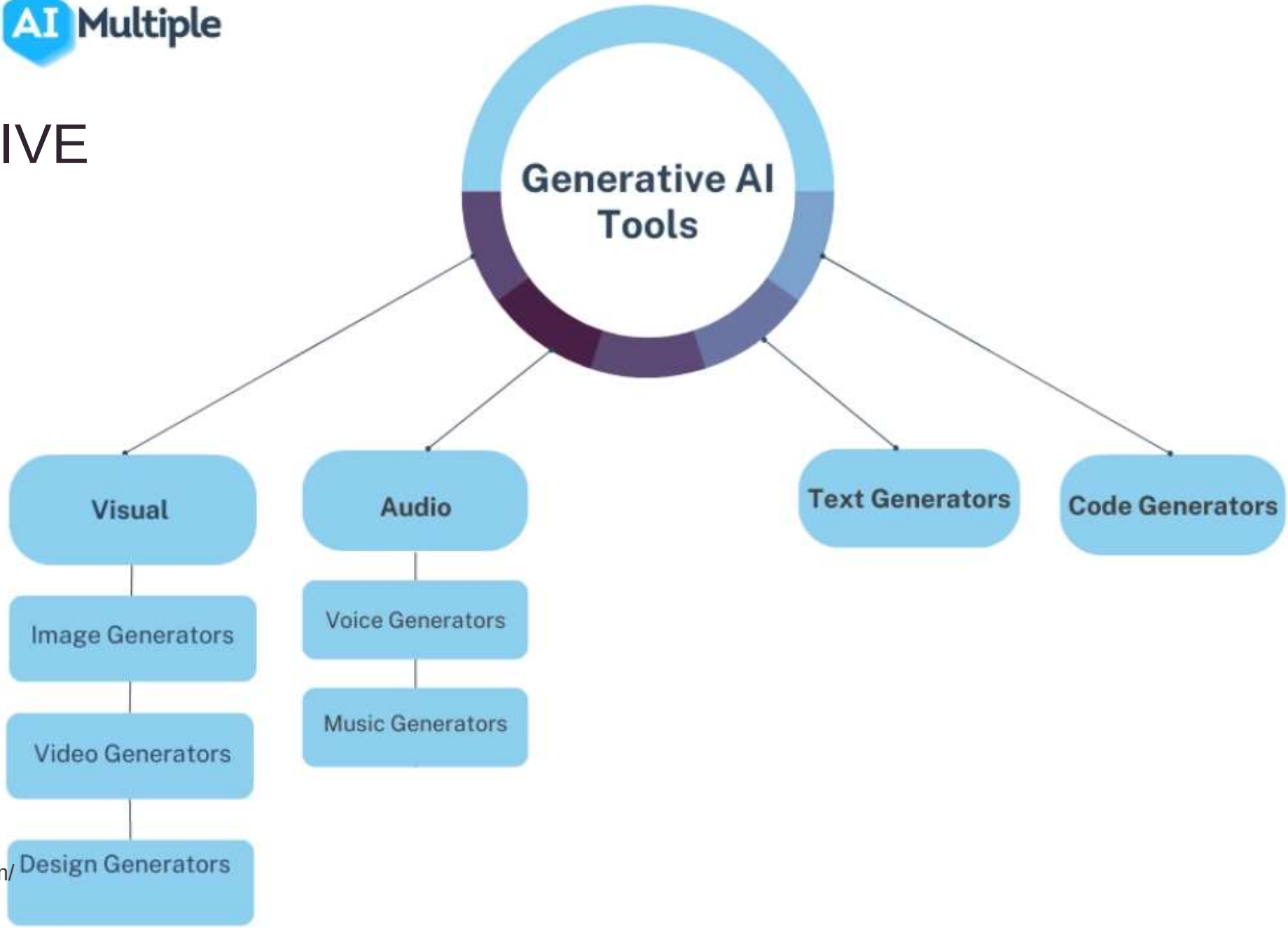




https://sites.northwestern.edu/thesafe/2019/10/21/new-cybersecurity-series-protect-it-social-media-bots/

GENERATIVE AI

GENERATIVE AI TOOLS



https://research.aimultiple.com/generative-ai-tools/

TOP GENERATIVE AI TOOLS

Credit: Moti Sagey and Reddit

CHATGPT (OPENAI) RISK, REWARD, TRADEOFF

CHATGPT (OPENAI) RISK, REWARD, TRADEOFF

TIME SAVINGS: ChatGPT can generate responses quickly and efficiently, which can save time for both the user and the company using the technology.

CUSTOMIZATION: ChatGPT can be customized to fit the needs of the user, allowing for a personalized experience that can improve customer satisfaction and engagement.

SCALE: ChatGPT can handle a large volume of interactions at once, making it an effective solution for companies that need to handle high volumes of customer inquiries.

BIAS: ChatGPT is trained on a large amount of text data, which can contain biases, stereotypes, and offensive language. ChatGPT may perpetuate them in its responses and has potential to manipulate the user.

SAFETY: ChatGPT is not good for decision making. There is risk its responses will be inappropriate for crisis management or healthcare decisions.

ZERO EMPATHY: ChatGPT is an AI language model and lacks the ability to empathize with users. This may result in responses that are impersonal or unsympathetic, which could negatively impact user experience.

ETHICS BYPASS: ChatGPT API with third party software (Telegram) uses OpenAI's GPT-3 model and can circumvent ChatGPT ethics.

FRAGILE: ChatGPT requires extensive training to function effectively, which can be time-consuming and expensive for companies.

CONTEXT PROBLEMS AND BAD FOR DECISION MAKING: ChatGPT can generate responses based only on the information it has been given, it cannot understand the full context of a situation. This can lead to inaccurate or inappropriate responses.

Micki created with DALL-E and ChatGPT

CALL TO ACTION FOR CYBERSECURITY, GRC, AND DATA PROTECTION AND PRIVACY PROFESSIONALS

CORPORATIONS WORRY ABOUT...

DATA PRIVACY: The GPT-3 API endpoint exposed by OpenAI should not retain or save any part of training data provided to it as part of the model fine-tuning/training process. No third party has access to the data shown to the model as a part of the training prompt by providing any kind of input to the exposed API endpoint.

DATA RETENTION AND LEAKAGE: GPT-3 has a default “data retention” period which requires the model to keep the training data for detect/prevent misuse of the API capabilities. Corporate custom data privacy agreements this retention period can be adjusted. Requires Mutual MOU then data will be scrubbed from the OpenAI systems. Protect data leakage by creating OpenAI data and model silos. OpenAI will simply silo off the requests/asks data independent of retention period, third parties will never have access or be able to extract your data by providing any input to the GPT-3 API.

Micki created with DALL-E

https://chat.openai.com/chat

+ New chat

ChatGPT's Data Protecti

Generative AI Vendors Categor

Classical Mechanics Action Fun

Clear conversations

CHATGPT SAYS ALL IS WELL!

As an AI language model, ChatGPT takes data protection and privacy very seriously, and we understand that these concerns are crucial for corporations. ChatGPT is committed to protecting corporate data and maintaining privacy in accordance with governance, risk, and compliance. Here are some of the ways we do this:

1. Compliance with Applicable Laws and Regulations: ChatGPT adheres to all applicable data protection and privacy laws and regulations, such as the General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), and other global privacy laws.
2. Robust Security Measures: We have implemented robust security measures to ensure the confidentiality, integrity, and availability of corporate data. This includes encryption of data at rest and in transit, network segmentation, firewalls, access controls, and regular security audits.
3. Access Control: ChatGPT limits access to corporate data to only authorized personnel who require access to perform their jobs. We also have strict password policies, multi-factor authentication, and monitoring of access to corporate data.
4. Data Protection by Design and Default: We implement data protection by design and default by incorporating privacy and data protection into the development and implementation of our AI models and services.
5. Data Minimization: ChatGPT minimizes the collection, storage, and processing of corporate data to only what is necessary for the provision of our services.
6. Transparency and Accountability: We are transparent about our data protection and privacy practices and provide our clients with regular reports on how their data is being used. We are also accountable for any data breaches and take appropriate action to mitigate any potential harm.

Overall, ChatGPT is committed to protecting corporate data and maintaining privacy in

BATTLE OF THE AI (SHORT LIST)

DEFENDER

Check Point has 42 different AI threat engines in prevention first architecture with ThreatCloud (IOC/TTP data lake)

Effective and efficient for Malware DNA genotyping and analysis

Helps SOC analysts see attack vectors and landscape

Help write good software code

Fix bugs in source code

Auto-write cybersecurity policies and controls based on GRC framework(s)

Gamify cybersecurity training

Mickl created with DALL-E

ADVERSARY

Check Point Research saw instances of 5 major attacks created by ChatGPT since its inception

Create Deepfakes and bots

Malware writing for dummies (joke)

Phishing email creation for dummies (again a joke)

Easy to identify attack landscape based on vulnerabilities and find exploits to match (multiphased attacks)

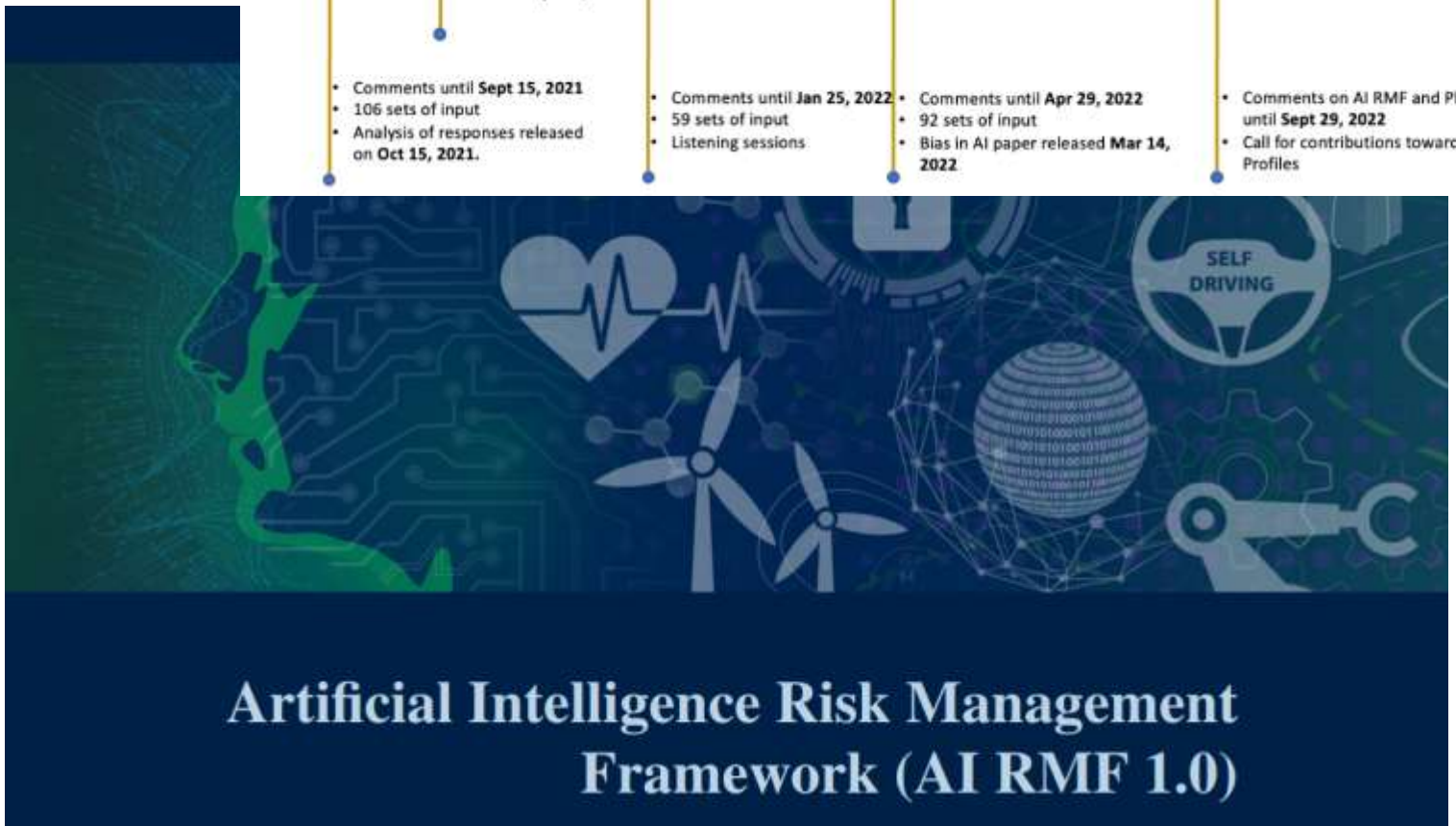
Circumvent AI ethics of GPT-3 by using API with Telegram or other software integration

SOME HELP...

NIST released on 26 JAN 2023 its 42 page Artificial Intelligence Risk Management Framework (AI RMF 1.0)

https://www.nist.gov/itl/ai-risk-management-framework

https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf



Mega thanks!

Michele Boland Evangelist and Architect
Check Point Office of the CTO
linkedin.com/in/micheleboland