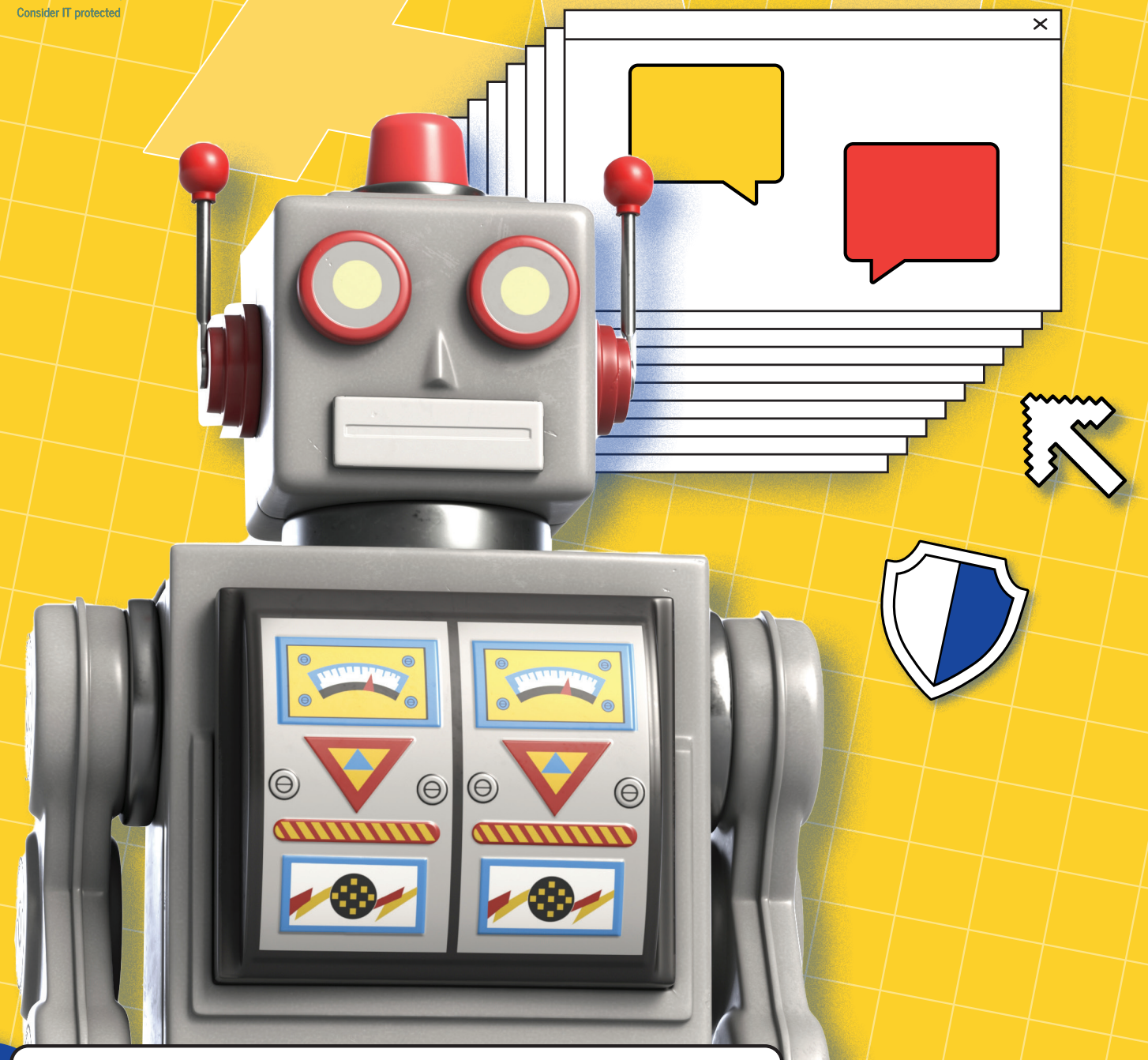
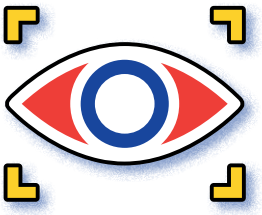
BHTC

Consider IT protected



PROTECTING YOUR BUSINESS

NAVIGATING AI SAFETY



CONTENTS

INTRODUCTION

Technology is advancing at a rapid pace, and at the forefront of this revolution is artificial intelligence (AI), changing the landscape for businesses and individuals alike. AI offers several exciting promises, such as streamlined processes, personalized virtual assistance and much more. However, as AI's influence grows, so do the challenges it poses.

This eBook takes a deep dive into a crucial facet of AI: safety. Understanding AI safety is non-negotiable in a world where AI tools are fast becoming as common as mobile phones.

Join us in exploring AI pitfalls and strategies to unlock its potential safely and effectively. Remember, we hold the power to shape the AI era responsibly. Let's embark on this journey together to ensure AI's progress aligns with your business goals and the wider world.

POTENTIAL ISSUES

and consequences

Understanding AI issues and their consequences is critical to safeguarding your business and personal interactions in an AI-driven world.

ADVERSARIAL ATTACKS

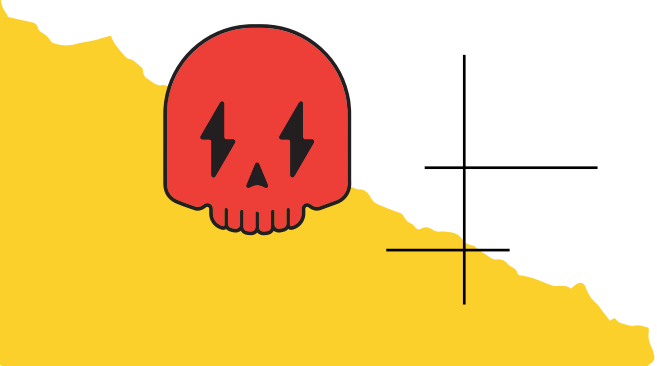
Malicious actors can exploit AI vulnerabilities to tamper with input data, resulting in incorrect outputs. For example, hackers could introduce noise to a device's AI face recognition software, rendering it unable to authenticate the rightful owner.

MODEL INVERSION ATTACKS

Attackers can exploit AI model outputs to extract sensitive information about individuals. For instance, if an AI model suggests a personalized diet and exercise plan based on someone's medical history, model inversion attacks could analyze the recommendations to infer the individual's medical background.

DATA MANIPULATION AND POISONING

AI models trained on data corrupted by hackers can produce inaccurate results. For example, if the data supporting self-driving cars or AI-backed traffic signals gets corrupted, it could lead to fatal accidents or chaos on the roads.



ALGORITHMIC BIASES

Unintentional biases in training data can inadvertently manifest in AI algorithms, potentially resulting in biased or unjust outcomes. An AI loan approval software might reject loan applications from specific individuals if the data used to train the software contains bias against that group of people.

AI-POWERED ATTACKS

Attackers can leverage AI's capabilities to orchestrate large-scale assaults, using tactics like ransomware and phishing with unprecedented perfection. For example, cybercriminals use AI chatbots to generate flawless phishing emails that do not have the usual red flags, such as grammatical or syntax errors.

DEEPFAKES AND IMPERSONATIONS

AI-generated deepfakes can propagate misinformation, deceiving unsuspecting individuals and leading to fraud or character defamation. For example, in the current era, where many banks rely on online KYC (KYC or Know Your Customer is commonly implemented in banks in order to comply with regulatory requirements and mitigate the risk of financial crimes), malicious actors could create ultra-realistic videos using another person's voice and image samples to open accounts for illegal transactions.

AI SAFETY PRACTICES

As AI becomes integral to modern existence, practicing its safe and responsible use becomes paramount. **Here are fundamental AI safety practices to bear in mind:**

CHOOSE WISELY
Pick AI tools from reputable and trusted sources.

VERIFY TRUSTWORTHINESS
Opt for technologies with a solid security track record of tackling vulnerabilities.

FORTIFY ACCESS
Enhance security with strong, unique passwords and enable two-factor authentication.

STAY UPDATED
Keep your software up to date by applying security updates promptly to prevent potential threats.

GUARD PERMISSIONS
Exercise caution when granting AI applications access to sensitive information and personal data.

PRIORITIZE PRIVACY
Regularly review privacy policies, adjust settings and limit data sharing to stay in control of your information.

MAKE PRIVACY-CENTRIC CHOICES
Choose AI tools that prioritize user privacy and data protection.

BEWARE OF SCAMS
Stay vigilant against phishing attempts and deceitful AI applications that can compromise data and devices.

RELY ON SAFE APP STORES
If you're looking for AI applications on app stores, download them solely from official app stores to mitigate risks.

THINK CRITICALLY
Exercise caution when using AI-generated recommendations, especially for critical decisions.

BACK UP DATA
Regularly back up essential data to guard against AI malfunctions or cyber incidents.

STAY INFORMED
Keep up with AI advancements and security risks by referring to reliable journals or publications.

READY TO SAFELY EMBRACE AI'S POTENTIAL?

As you incorporate AI into your business and personal interactions, remember that responsible usage is essential. Our team is here to support you every step of the way. Whether you need help with AI safety practices, want to explore reliable AI tools or have questions about the evolving AI landscape, we're just a message away.

Contact us today to harness AI's potential while ensuring safety and security in your endeavors. Together, let's build an innovative and responsible AI-powered future.

SPOT THE RED FLAGS

Can you find the red flag in these two pictures?

What's wrong with these guys?

WORD SCRAMBLE

Use the hints to unscramble the words.

TUNOIAMAOT

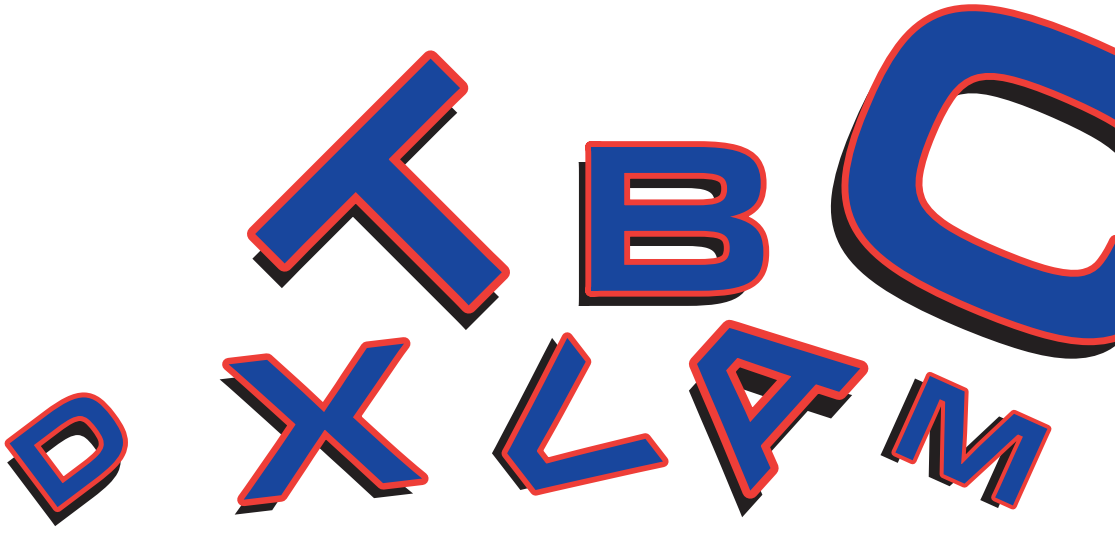
HINT: Using AI to complete tasks instead of humans.

VCTEDIEPE IA

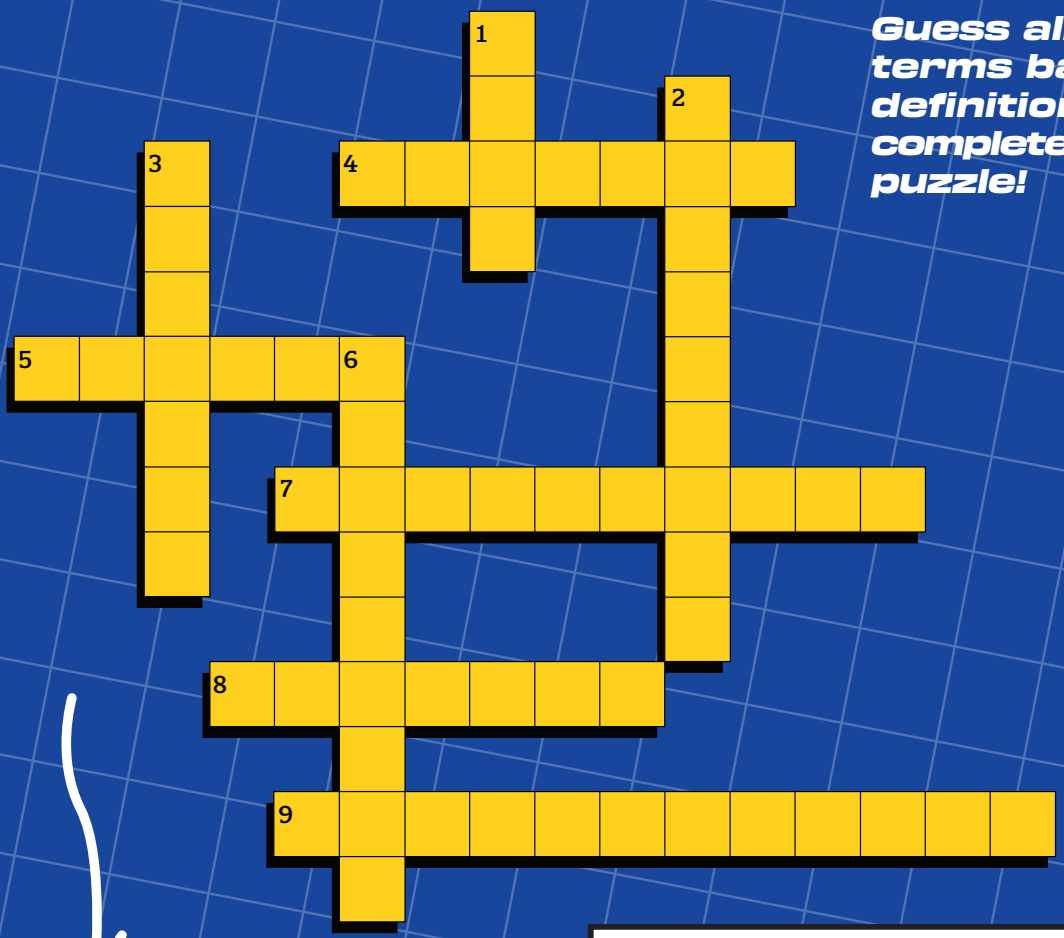
HINT: These types of applications may compromise your data or devices.

GOIGESSUNTS

HINT: Data offered by AI platforms.



CROSSWORD



Guess all nine AI safety terms based off the definitions below to complete the crossword puzzle!

DOWN

1. Major AI bot competitor created by a search giant.

2. A set of rules a machine follows to do a task.

3. Popular AI chatbot used for generating content.

6. Famously said by Arnold Schwarzanegger's character from the future, "My CPU is a _____ processor; a learning computer."

ACROSS

4. Adobe's latest entry into the AI art space (here's a hint: it's a glowing bug).

5. Its versatility and array of robust libraries make it the go-to programming language for chatbot creation.

7. Named after a famed mathematician,computer scientist, and logician, it tests a machine's ability to pass for a human.

8. The famous AI who said, "I'm sorry, Dave. I'm afraid I can't do that."

9. A function of artificial intelligence that imitates the human brain by learning from how data is structured.

WORD SEARCH

O T S M E M X F R F E G C I E
A S W B V V I F F V U N T N C
U A M M W M I Q I A U I R T D
T O B T A H C T E R O T V E G
G D Z G G C C R I V W U P L N
I V E S C I H T F N O P T L I
H B B E D D R I U V G M U I N
E S O E P G N F N R G O K G R
T B R M I N I N G E I C C E A
E P I F N Y V Y B X J N F N E
S Q Y K U Z E D L C V L G C L
T H E J A N A L Y T I C S E A
N O I T I N G O C E R V M T K
J A B K B Q U V U C Q P A O I
S Y B D X H Z H Z W B D X Q O

CHATBOT
COGNITIVE COMPUTING
PREDICTIVE ANALYTICS
MACHINE INTELLIGENCE
IMAGE RECOGNITION
DEEP LEARNING
DATA MINING
TURING TEST

ACTIVITIES KEY

SPOT THE RED FLAG

Both models are not human.
They were created with AI software.
Deep fakes use AI to manipulate videos and images to create a digital representation of the target person.
Don't believe everything you see.
Always check the sources.

WORD SCRAMBLE

AUTOMATION

DECEPTIVE AI

SUGGESTIONS

CROSSWORD

DOWN

1. Bard
2. Algorithm
3. ChatGPT
6. Neuralnet

ACROSS

4. Firefly
5. Python
7. Turing test
8. HAL9000
9. Deep learning

WORD SEARCH