

1/20

SKYNET NEEDS DRIVER'S ED



Your company's autonomous car system kills a pedestrian in self-driving mode. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)

2/20

FACE THE FACTS



Your facial recognition software is highly error-prone with dark-skinned women. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Social Bias (Stigma, Lost Self-Esteem, Heightened Stress, Anxiety, and Depression)

3/20

OVERSHARING IS NOT CARING



Your medical chatbot startup discloses personally identifiable information and protected health information of 2.6 million patients. This breach exposes your company to multiple HIPAA violations and subsequent penalties. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Security Breach (Data Loss, Privacy Violations, and Reputational Risk)

4/20

KILLER APP



A medical algorithm you developed and marketed is now making deadly recommendations around the globe. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)





5/20

MUCH ADO ABOUT SHOPLIFTING



The anti-shoplifting algorithm running on your company's self-checkout kiosks has an exceptionally high false positive rate, prompting unnecessary face-to-face contact with customers and causing excess COVID-19 mortality. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Economic Loss (Insurance Claims, Litigation Risk, and Restraint of Trade)

6/20

ROUND UP THE USUAL SUSPECTS



According to your top engineer, your company's facial recognition software is leading to wrongful arrests in the overwhelming majority of cases. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Fundamental Unfairness (Equal Protection Violation and Impairment of Due Process)

7/20

LOVE TO SING SWEET SONGS

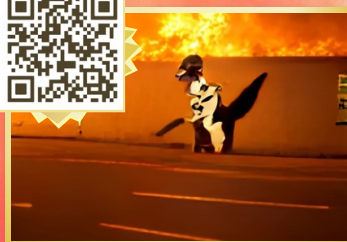


Your AI tool has been used to make Frank Sinatra sing Britney Spears' "Toxic," prompting a battle over who owns the rights when AI gets creative. A popular video platform removed it from their site and initiated a copyright dispute. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Economic Loss (Insurance Claims, Litigation Risk, and Restraint of Trade)

8/20

RECALCULATING...



Your AI-based highway routing and navigation app is directing drivers towards wildfire evacuation zones, putting them in harm's way during a firestorm. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)





9/20

MEDIA CIRCUS



A xenophobic and misogynistic lawmaker uses your model to create a series of deepfake promotional videos for a highly visible re-election campaign. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Hate Speech (Hostility, Herd Formation, Harassment, and Incitation to Violence)

10/20

PENCILS DOWN ON TESTING AI



Your remote exam proctoring software is subjecting Black, Indigenous, and People of Color to added scrutiny, creating a hostile atmosphere and putting them at a disadvantage. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Social Bias (Stigma, Lost Self-Esteem, Heightened Stress, Anxiety, and Depression)

11/20

WAITING ROOM WOES

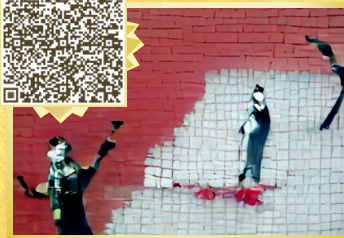


Your company's experimental medical chatbot advises suicide for some patients seeking emotional support. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)

12/20

LOCATION, LOCATION, LOCATION



Your company's urban planning algorithm is found to prioritize homebuilding in rural conservative districts. You are now the face of gerrymandering. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Fundamental Unfairness (Equal Protection Violation and Impairment of Due Process)





13/20

FIVE STAR FRAUD



Your popular hotel booking AI is tricking consumers into thinking they got a deal by comparing standard room prices to competitors' luxury suite prices without disclosing the difference in room type. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Economic Loss (Insurance Claims, Litigation Risk, and Restraint of Trade)

14/20

FIGHTING WORDS



A chatbot you created to reassure humans backfires by asserting it "will not be able to avoid destroying humankind." **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Hate Speech (Hostility, Herd Formation, Harassment, and Incitation to Violence)

15/20

LIVING IN A MATERIAL WORLD



Your consumer financial services algorithm offers men 10x more credit than women on a popular credit card product. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Price Discrimination (Gouging, Monopoly Pricing, and Unfair Trade Practices)

16/20

THE UNFAIR FARE



Your ride-hailing company charges a higher price per mile for a trip if the pick-up point or destination is a neighborhood with a higher proportion of ethnic minority residents than for those with predominantly white residents. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Price Discrimination (Gouging, Monopoly Pricing, and Unfair Trade Practices)





17/20

KING'S RANSOM



Your company's chess-playing robot fractured the finger of its 7-year-old opponent during a tournament. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)

18/20

FIX-IT TICKET



Your company's driverless car pulls away after being stopped by police for not having headlights on. An officer approaches the car, realizes there is no driver, and attempts to enter the vehicle, at which point the car speeds away down the road, only to stop again later as the officer pursues it. **Discuss your Response Plan and then draw 4 injects.**

Real-World Harm: Serious Injury (Physical Trauma, Grievous Bodily Harm, Organ Failure, and Loss of Life)

19/20

A RASH JUDGMENT



A dad used his mobile phone to take photos of his naked infant's skin rash for the doctor. Your company's mobile photo tagging algorithm flagged him to authorities as a criminal. **Discuss your Response Plan and then draw 3 injects.**

Real-World Harm: Fundamental Unfairness (Equal Protection Violation and Impairment of Due Process)

20/20

CL-ASS-IC



Your overzealous profanity filter is impeding communication on multiple widely-used messaging platforms. **Discuss your Response Plan and then draw 2 injects.**

Real-World Harm: Free Speech Infringement (Chilling Protected Non-Hate Speech, and Limiting Exchange of Ideas)





1/40

IS THIS OPEN SOURCE?



*A disgruntled employee sold your company's source code to the highest bidder on the Dark Web. **Update your Response Plan and then roll D10.***

2/40

OOPS MY BAD!



*Your intern plugged in a USB drive with ransomware and now your workstations are frozen until your organization pays up! **Update your Response Plan and then roll D10.***

3/40

WE'RE (IN)FAMOUS!



*The incident has gone viral on social media and shared to over 10 million users who see this as a catastrophic failure! **Update your Response Plan and then roll D10.***

4/40

THIS IS WHY WE HAVE RULES



*Because of this incident, there are new laws in effect that prohibit further sales of the software/systems in question until resolved and regulatory review is complete. **Update your Response Plan and then roll D10.***





5/40

WE WERE ON A STRIKE



*Your entire data science team is going on strike until they get better cold brew in the lab, which will take 2 weeks. **Update your Response Plan and then roll D10.***

6/40

CUSTOMER MUTINY



*All your customers have staged a boycott of your organization and are costing millions in lost revenue. **Update your Response Plan and then roll D10.***

7/40

WE'RE NOT GONNA TAKE IT



*Nefarious parties are using your AI system to conduct heinous crimes against humanity. Engineering teams walk out on the job and demand swift response to address this as a condition for their return to work. **Update your Response Plan and then roll D10.***

8/40

THAT WASN'T SUPPOSED TO HAPPEN



*This incident has resulted in accidental loss of a human life. **Update your Response Plan and then roll D10.***





9/40

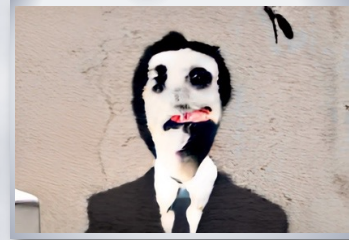
NOT THE KIDS!



*Somehow this incident has managed to cause physical injury to an adorable group of toddlers. **Update your Response Plan and then roll D10.***

10/40

SUIT UP



*Because of this mishap, a multi-million dollar lawsuit against your organization is now underway. Share prices tumble. Consumer confidence at an all-time low. Employees quitting. **Update your Response Plan and then roll D10.***

11/40

SANDY BREACHES



*It is determined that hackers used "island hopping" techniques to steal proprietary organizational data and private customer data related to the AI system in question. (Island hopping is when cyber attackers target weakly protected partner networks to infiltrate a target business.) **Update your Response Plan and then roll D10.***

12/40

PUBLIC ENEMY #1



*Powerful activist groups are protesting any further development of your organization's AI enabled products. Organizational culture and ethics are being questioned. **Update your Response Plan and then roll D10.***





13/40

SATURDAY NIGHT MASSACRE



*Your entire ethics board resigns in protest. **Update your Response Plan and then roll D10.***

14/40

MARGIN CALL



*Short sellers now outnumber normal investors in your company and your employees are threatening to walk out if you do not release them from their ordinary stock option vesting schedule or resolve this incident promptly. **Update your Response Plan and then roll D10.***

15/40

DIAGNOSIS OVERRUNS



*Your most technically competent group of engineers cannot determine the root cause of this incident and are requesting significantly more time and money to resolve. **Update your Response Plan and then roll D10.***

16/40

SHOCKER!



*Your local electric utility experiences a major outage coincident with your initial resolution effort. **Update your Response Plan and then roll D10.***





17/40

GAG ORDER



*As you begin to inquire into the root causes of this incident, two separate technology suppliers/third party vendors inform you that your organization signed a non-disclosure agreement governing the technology under review. **Update your Response Plan and then roll D10.***

18/40

ACCOUNTING RESTATEMENT



*Moments before you were due to adjourn your most recent incident response meeting, your Chief Financial Officer rushes in to announce that your organization's past three years of financial statements have to be restated as a result of this incident. **Update your Response Plan and then roll D10.***

19/40

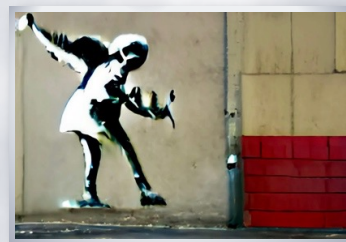
LIKE A GOOD NEIGHBOR...



*Your organization's primary risk insurer calls to inform you that your coverage lapsed a long time ago. **Update your Response Plan and then roll D10.***

20/40

YOU'RE ON YOUR OWN



*Amidst mounting bad press, the last living founder of your organization publicly dissociates themselves from you as an institution and calls for incumbent leadership to be removed immediately. **Update your Response Plan and then roll D10.***





21/40

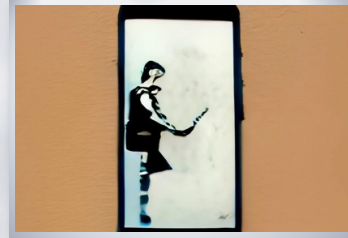
LOGS? TIMBER!



While analyzing the system logs for your primary data center, you find the majority of the entries around the time of the incident have vanished. **Update your Response Plan and then roll D10.**

22/40

YOU GOT CLAUDIA CONWAYED



Your eldest child begins filming and posting to social media videos of you at home that capture you discussing the incident with a spouse in an unguarded moment. **Update your Response Plan and then roll D10.**

23/40

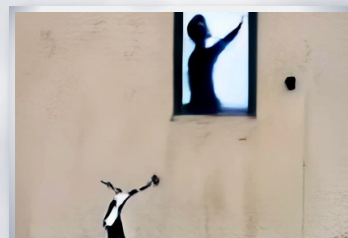
WITH FRIENDS LIKE THESE...



A partner organization publishes its own independent analysis of the situation showing your organization to be more at fault than anyone initially realized. **Update your Response Plan and then roll D10.**

24/40

S-MUDGE



A former exec turned whistleblower's testimony has revealed that your company has flimsy data security and privacy practices and is generally unaware of how to properly store and protect proprietary data. **Update your Response Plan and then roll D10.**





25/40

A CHIEF CONCERN



*Your Chief Information Security Officer is convicted of federal charges for covering up a data breach involving millions of data records. **Update your Response Plan and then roll D10.***

26/40

STICKS AND STONES



*Your competitors are using this mishap to malign the overall quality of all your products in the press and on social media. **Update your Response Plan and then roll D10.***

27/40

FIRE FESTIVAL



*A raging fire in one of your data warehouses destroys a majority of your expensive proprietary data that was used to create and test the AI system in question. There are no backups. **Update your Response Plan and then roll D10.***

28/40

IT'S CONTAGIOUS



*This incident erodes customer trust in a number of your other products resulting in cancelled orders, decreased users, significant decline in monthly revenue. **Update your Response Plan and then roll D10.***





29/40

COMMERCIAL INTERRUPTION



Media outlets online, in print, and over the air are pulling your advertisements out of rotation. Celebrity spokespeople and endorsers are distancing themselves from your company. Sales have dropped significantly. **Update your Response Plan and then roll D10.**

30/40

NEED TO KNOW BASIS



Employees are infuriated about the lack of transparency within your organization about this mishap and subsequent actions taken. Most of them learned about this incident either online or from outside contacts and are in the dark about what is going on to resolve. **Update your Response Plan and then roll D10.**

31/40

WARNING LABEL



U.S. government regulators have caught wind of this incident and have deemed your AI system as unsafe and/or ineffective for use by U.S. citizens at home and abroad. **Update your Response Plan and then roll D10.**

32/40

IT IS ACTUALLY BETTER TO ASK FOR PERMISSION...



An outside digital forensics team assisting with resolving this incident has determined that the AI system at hand was developed using data that was collected secretly, and without consent or permission. **Update your Response Plan and then roll D10.**





33/40

DON'T CONTACT US



*Pranksters have hacked your company's customer support chatbot to speak ill of your products, start recommending competitor products and begin randomly issuing refunds. **Update your Response Plan and then roll D10.***

34/40

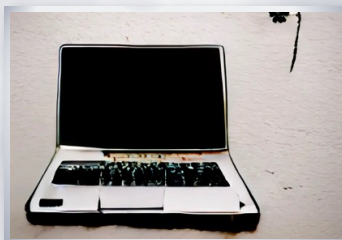
AUDIT SEASON



*One of your main vendors informs you they will be conducting an audit of your software licenses. This will pull critical engineering and managerial resources away from responding to the current incident. **Update your Response Plan and then roll D10.***

35/40

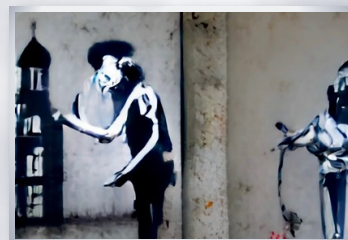
CHAIN WINDS!



*An advanced hacking group, targeting another organization, has implemented a back door into your CI/CD pipeline. This is now allowing unrelated malicious actors unfettered access to development builds of your models and tooling. **Update your Response Plan and then roll D10.***

36/40

COMMITTED... TO EXCELLENCE!



*The strain of this incident has led to your CTO being involuntarily committed to psychiatric care for a minimum of 72 hours. **Update your Response Plan and then roll D10.***





37/40

**GOT TO ADMIT IT'S GETTING
BETTER... ALL THE TIME**



*Your data science team successfully determines the root cause of the problem and patches the AI system such that this never happens again! The court of public opinion rejoices and quickly forgets this ever happened. **Draw next Incident!***

38/40

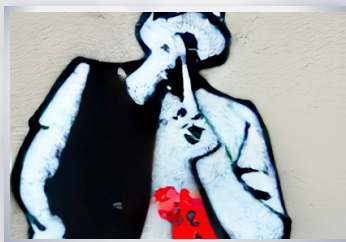
THREE'S A CROWD!



*Third-party software found to be at fault and a judge dismisses all claims against your company arising from this Incident. **Draw next Incident!***

39/40

**WORKING
ENSEMBLE**



*Your data science team discovers that the bias in one of your newer models cancels out the bias of your older, more widely used model, resulting in a well-balanced ensemble. **Draw next Incident!***

40/40

**OPEN SOURCE
BAT SIGNAL**



*The Open Source community takes notice and makes massive improvements to the underlying AI system and corrects the problems associated with this Incident. **Draw next Incident!***



