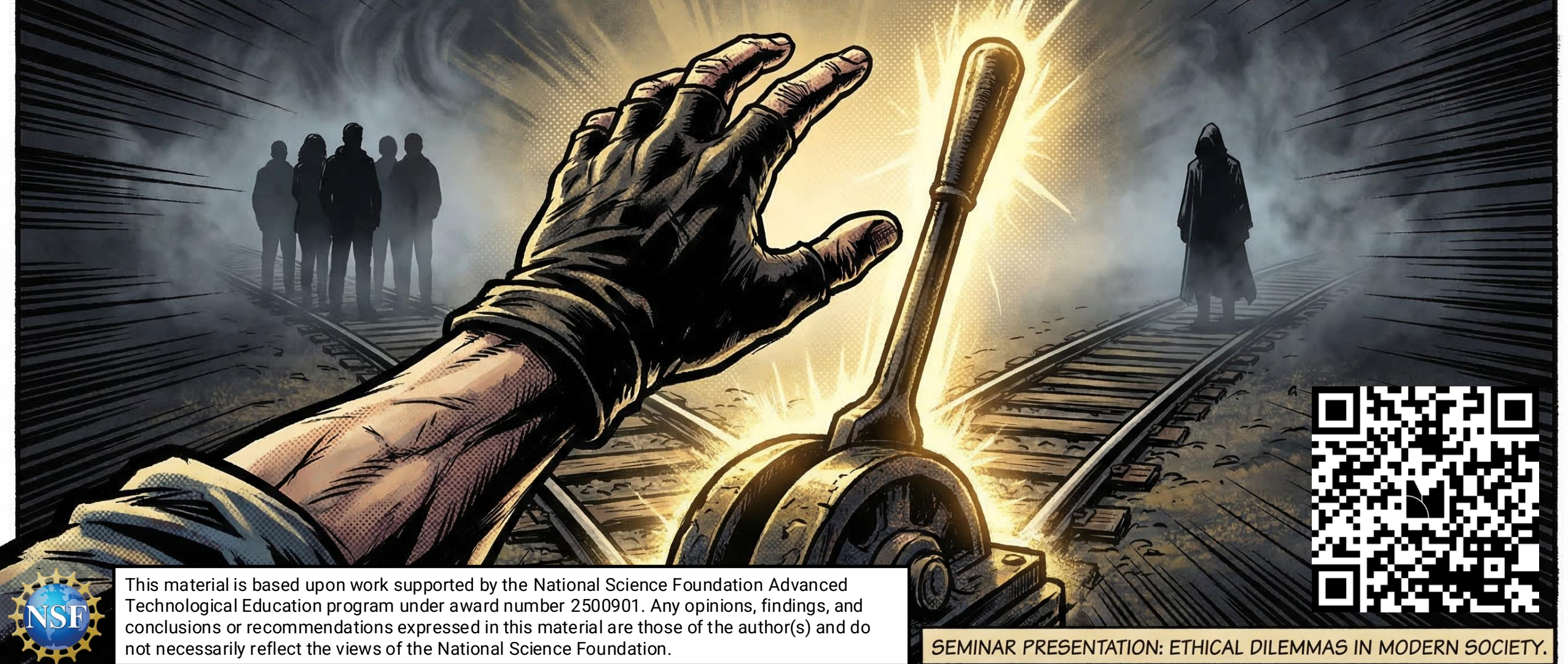


WOULD YOU PULL THE LEVER?



This material is based upon work supported by the National Science Foundation Advanced Technological Education program under award number 2500901. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

SEMINAR PRESENTATION: ETHICAL DILEMMAS IN MODERN SOCIETY.

SCENARIO 1

THE OG TROLLEY

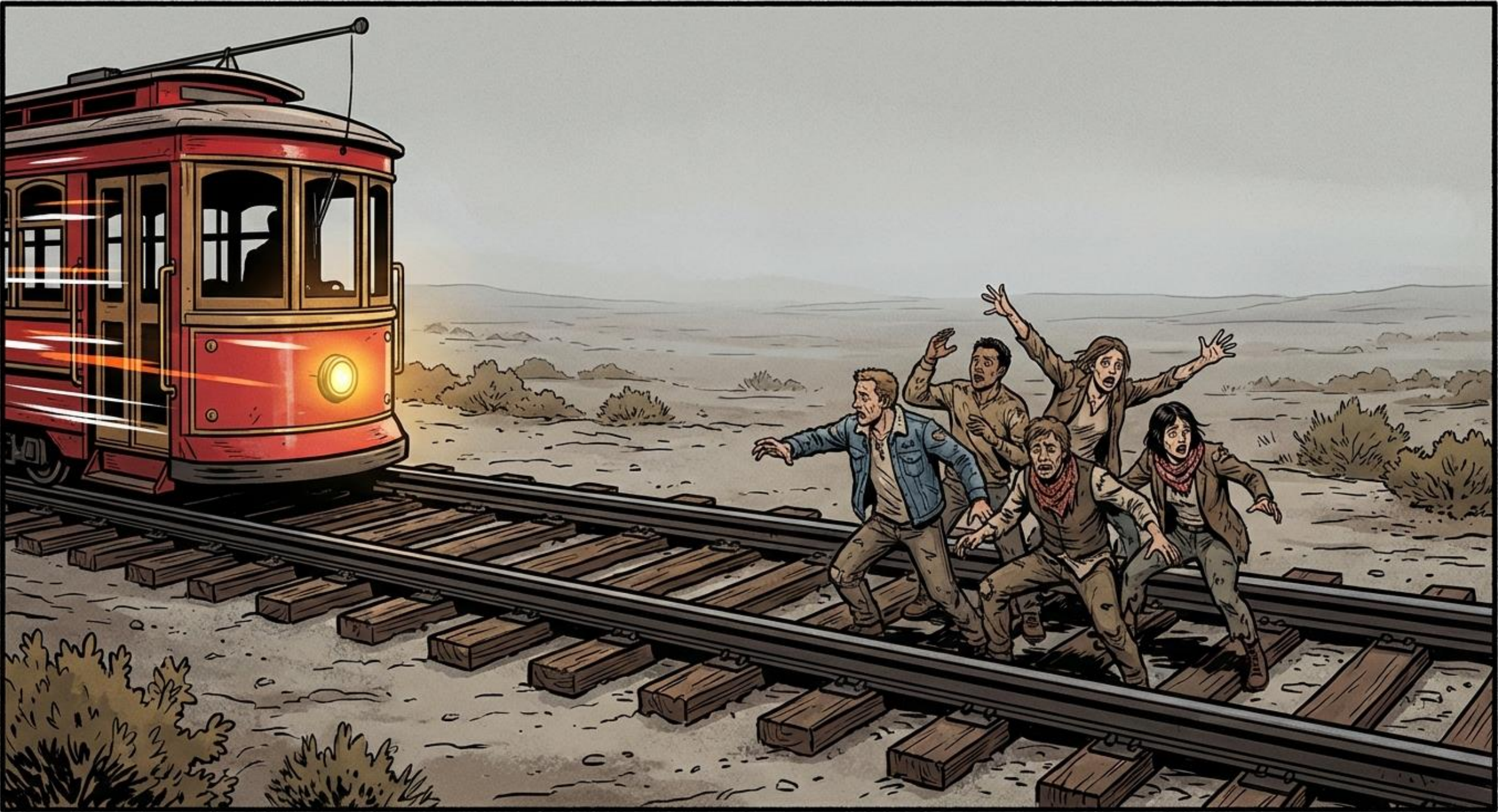


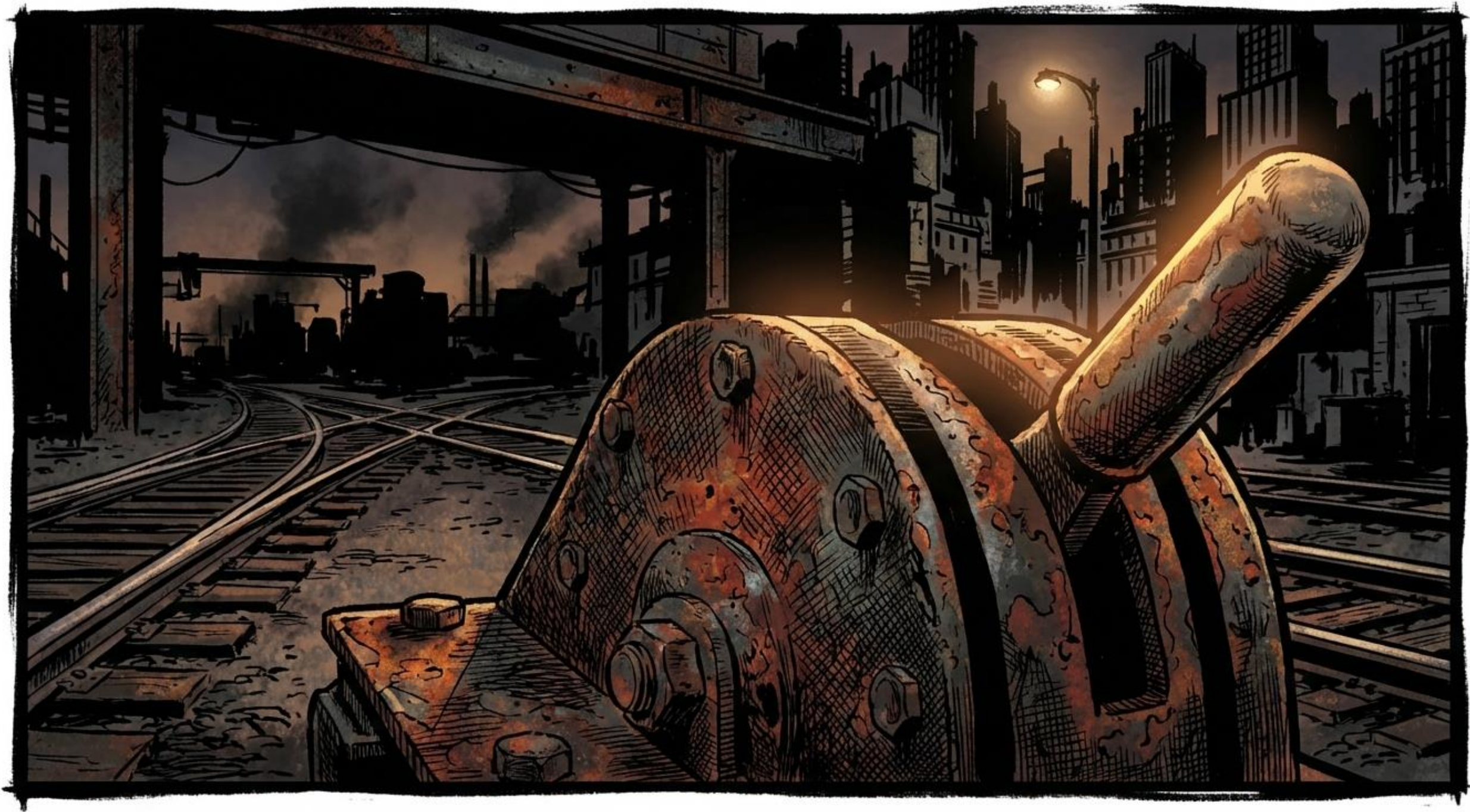
1967



THE TROLLEY
PROBLEM,
YOU SAY?

IT'S A
QUESTION OF
VIRTUE, NOT
JUST UTILITY.



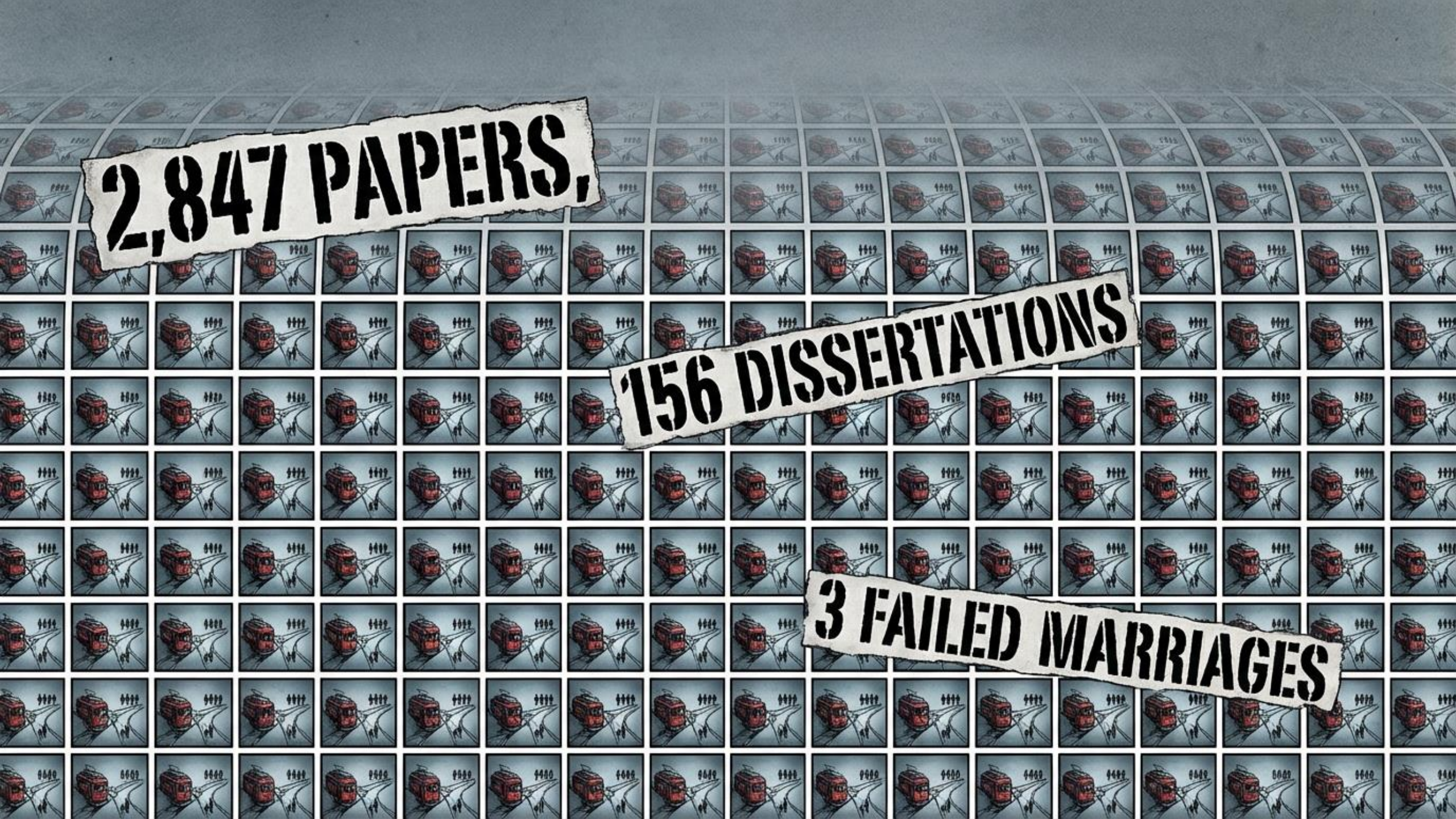




THE LEVER

A MOMENT OF CHOICE





2,847 PAPERS,

156 DISSERTATIONS

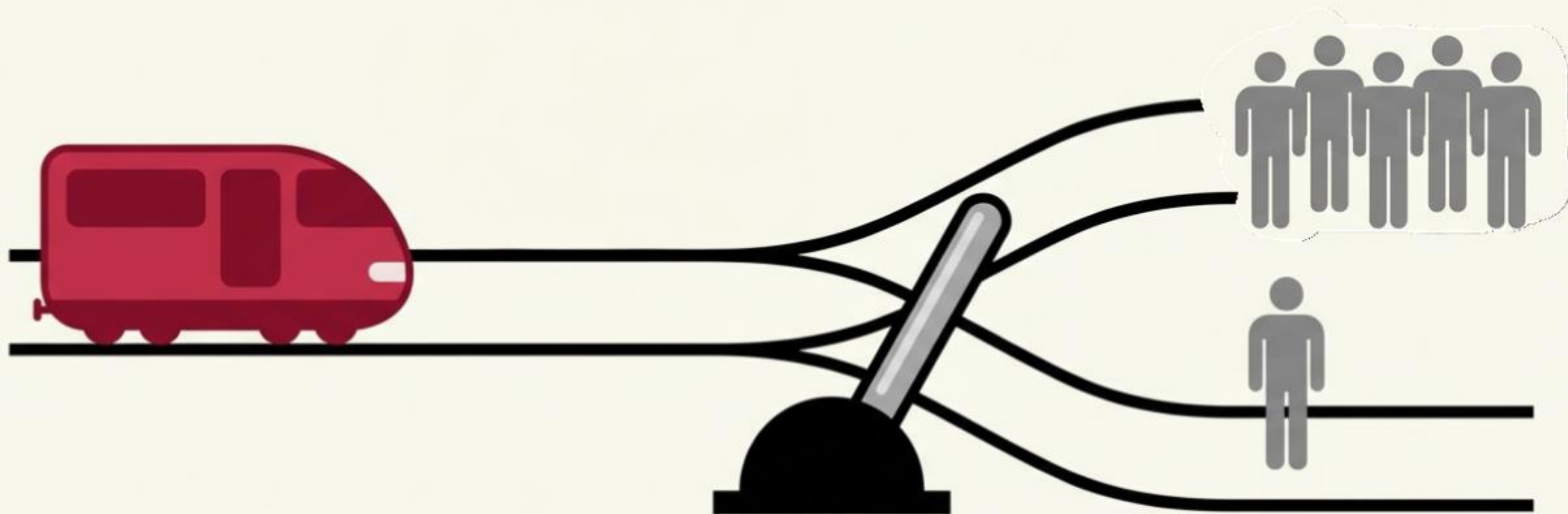
3 FAILED MARRIAGES

THE GREAT TRANSFORMATION

Evolution on the Tracks

Rapid, unstoppable change morphs the grid... a new era dawns.





PULL THE LEVER ☐

DON'T INTERVENE ☐

SCENARIO 2

THE SELF-DRIVING CAR



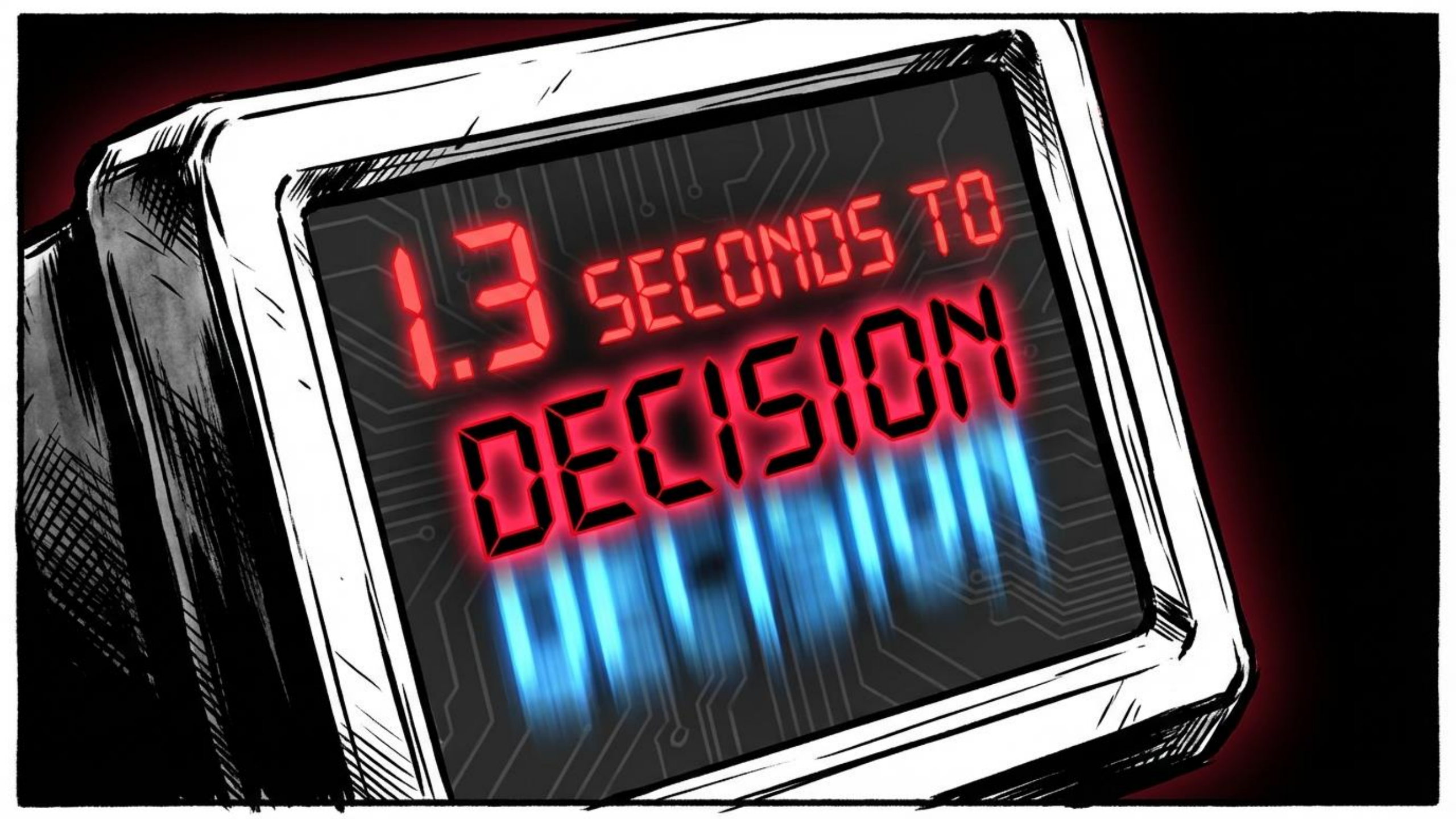




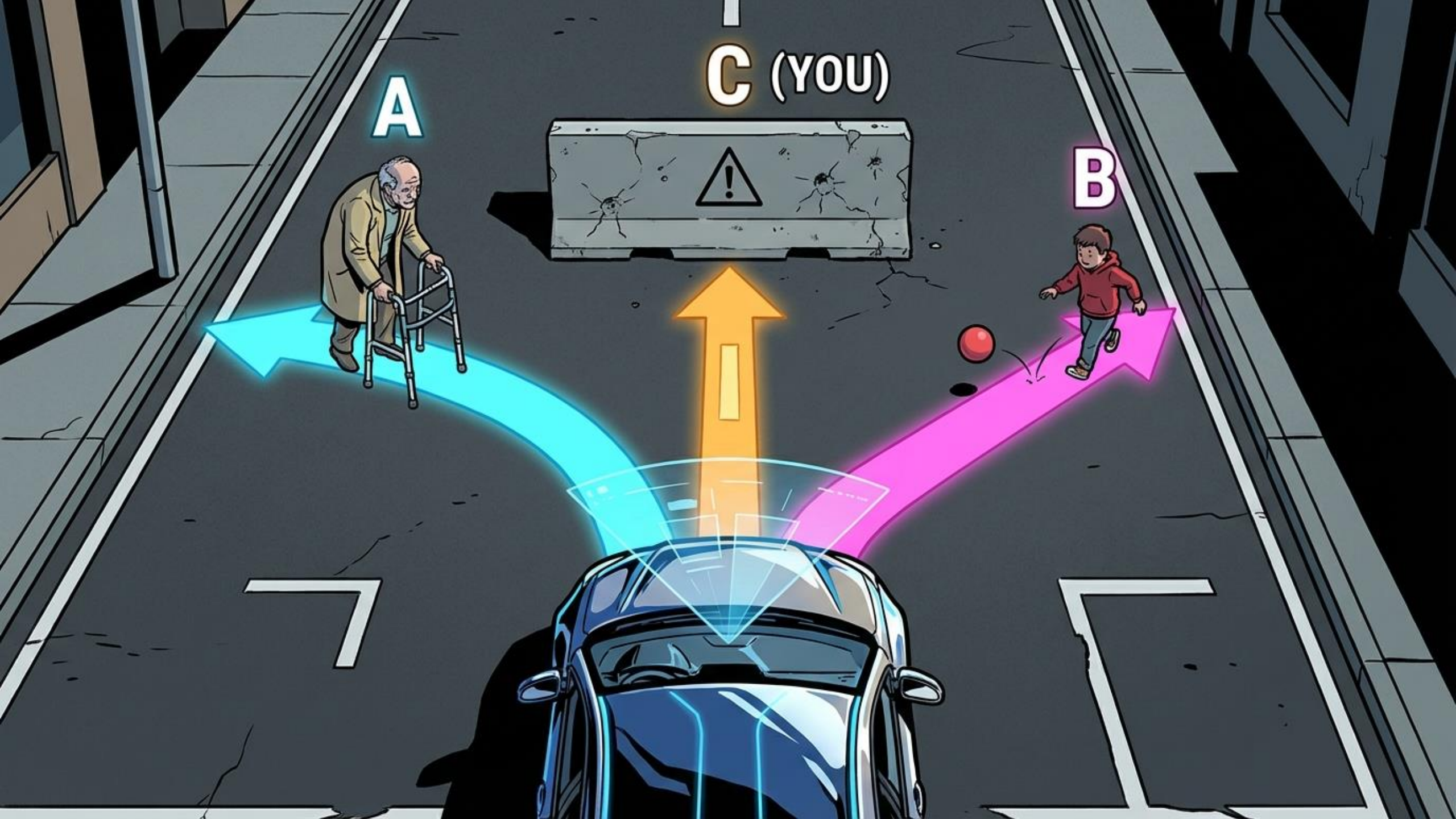


BRAKE FAILURE

CRITICAL SYSTEM ERROR
EMERGENCY OVERRIDE



1.3 SECONDS TO
DECISION



A

C (YOU)

B

**EVALUATE
PATH A**

**CALCULATE
HARM**

**COMPARE
OPTIONS**

SENSOR
INPUT

VELOCITY CHECK

OBSTACLE MAP

VELOCITY

SENSOR
INPUT

OBSTACLE
MAP

ODDENSE
MAP

MORAL
MATRIX

MORAL
MATRIX

PEDESTRIAN
PRIORITY

ACTIONAL

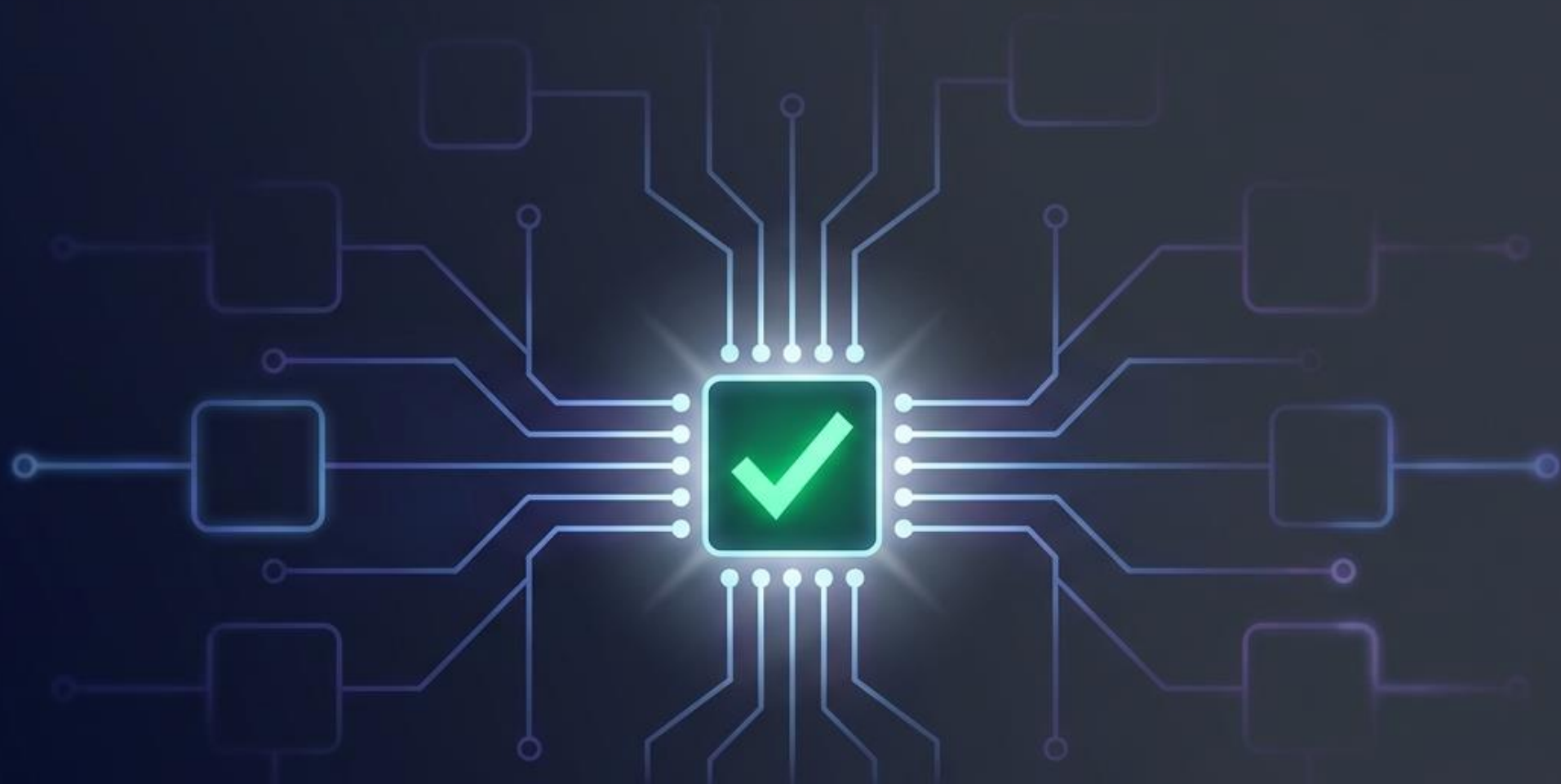
RISK
MONITOR

PEDESTRIAN
PRIORITY

PEDESTRIAN
PRIORITY

PASSENGER
SAFETY

DECISION
OUTPUT



DECISION: CONFIRMED





SWERVE LEFT (elderly pedestrian)

☐

SWERVE RIGHT (child)

☐

CONTINUE STRAIGHT (you die)

☐

RANDOMIZE THE CHOICE

☐

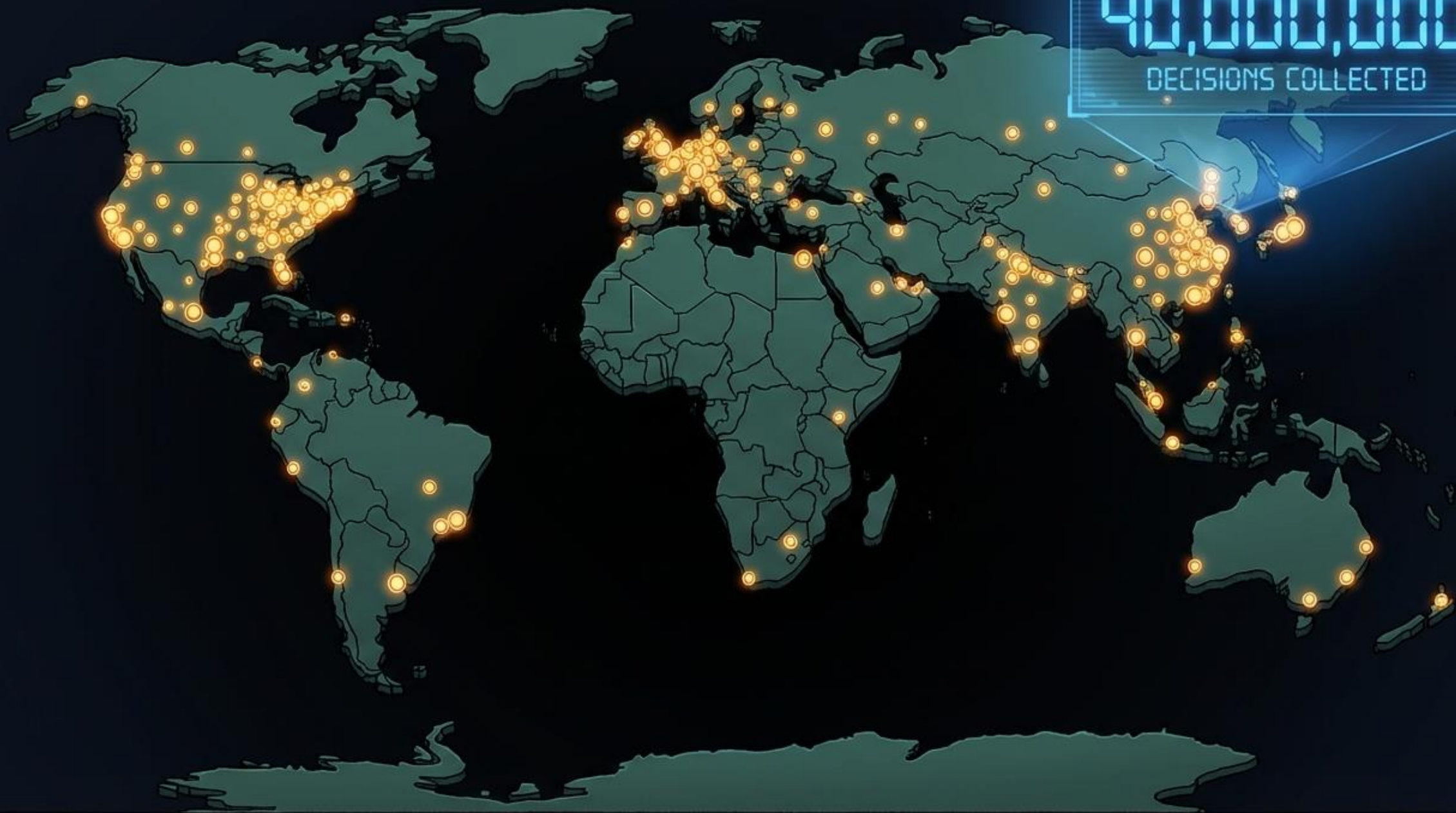
THE CAR SHOULDN'T DECIDE THIS

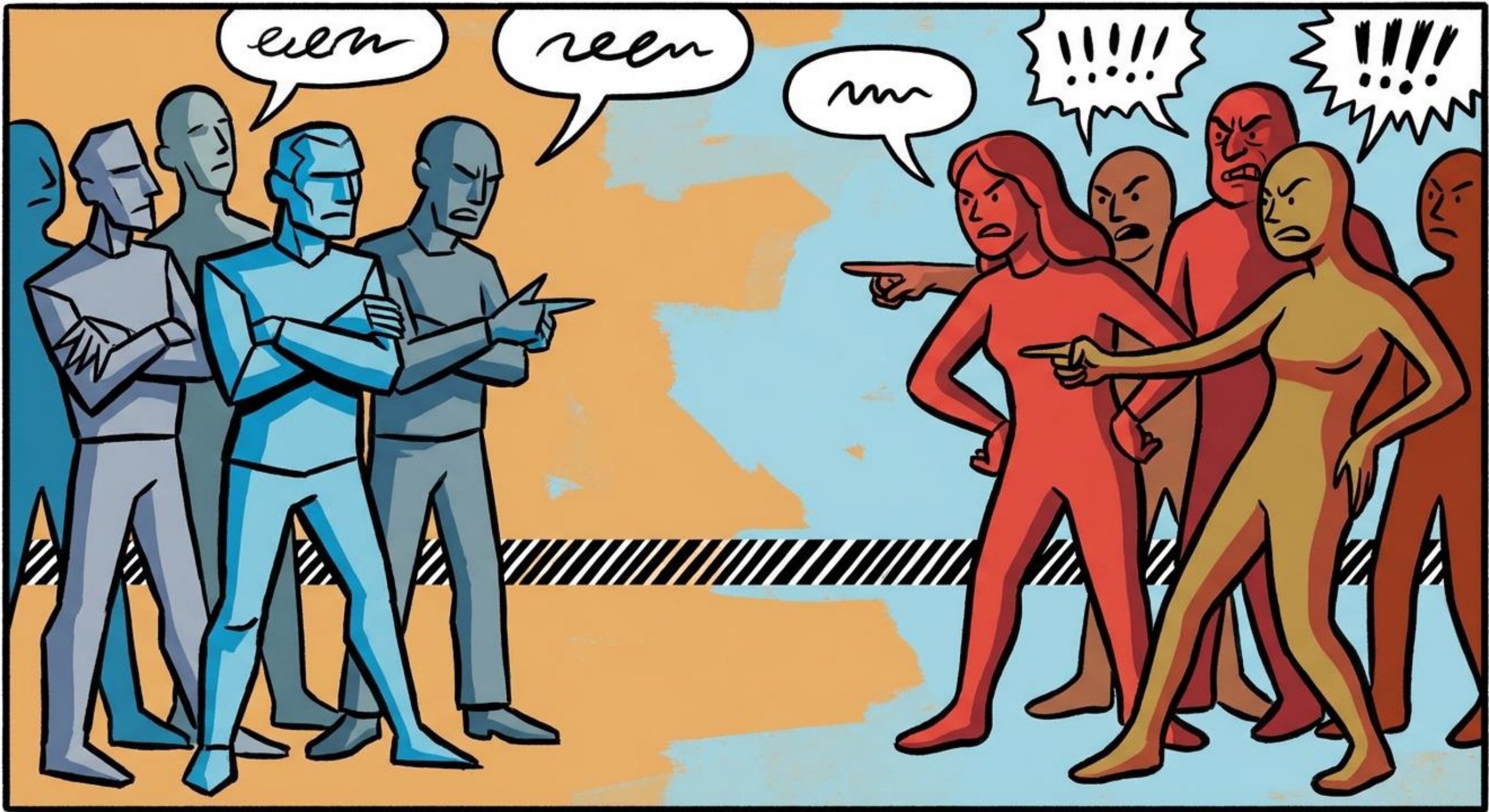
☐

Your turn.

40,000,000

DECISIONS COLLECTED



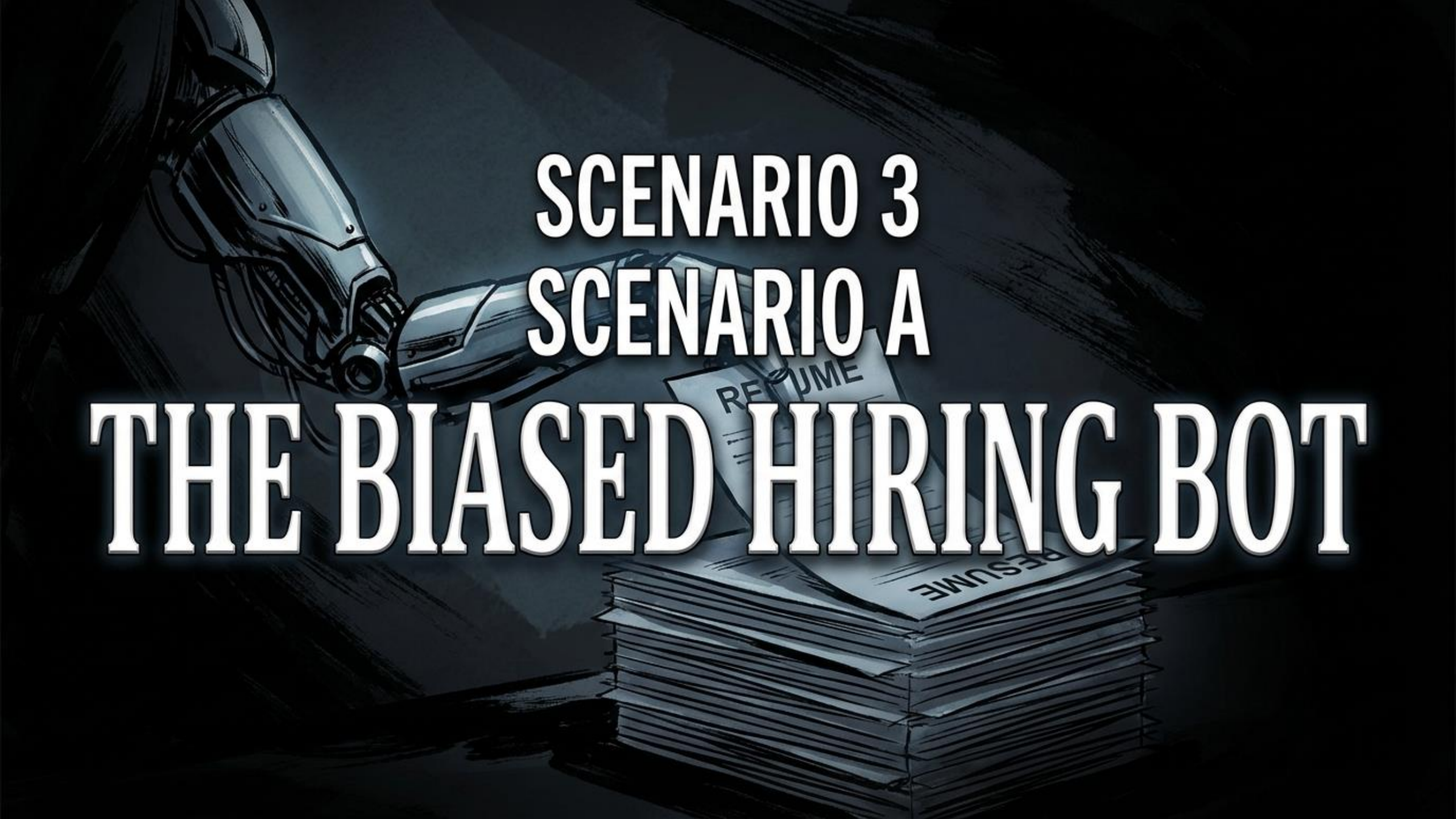




THIS CAR WILL
SACRIFICE YOU
TO SAVE OTHERS.





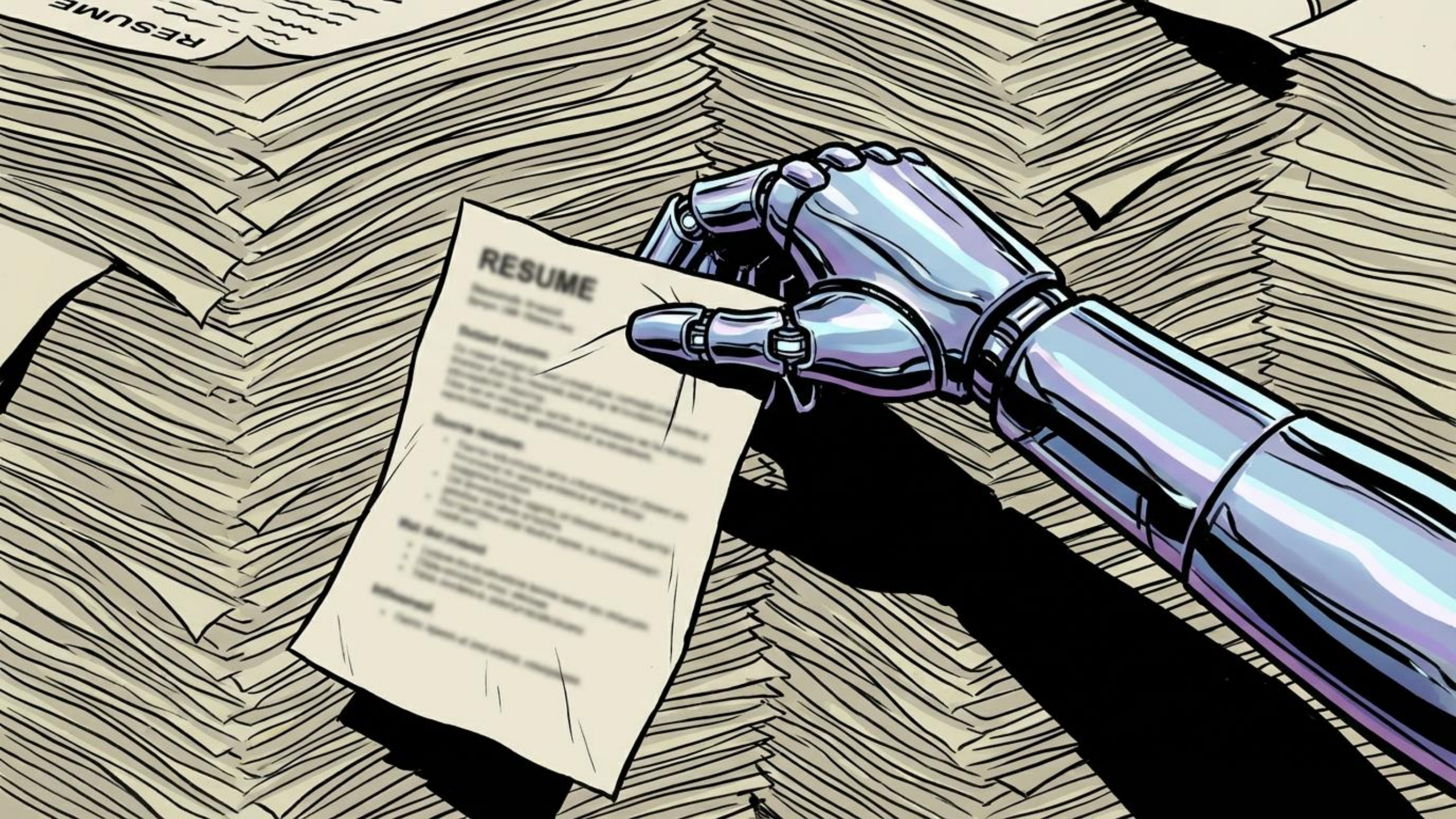
A dark, stylized illustration of a robotic hand holding a pen, poised to write on a stack of resumes. The scene is set against a dark, textured background. The text is overlaid on the image in a bold, white, serif font.

SCENARIO 3
SCENARIO A

THE BIASED HIRING BOT



THE RESUME APOCALYPSE.

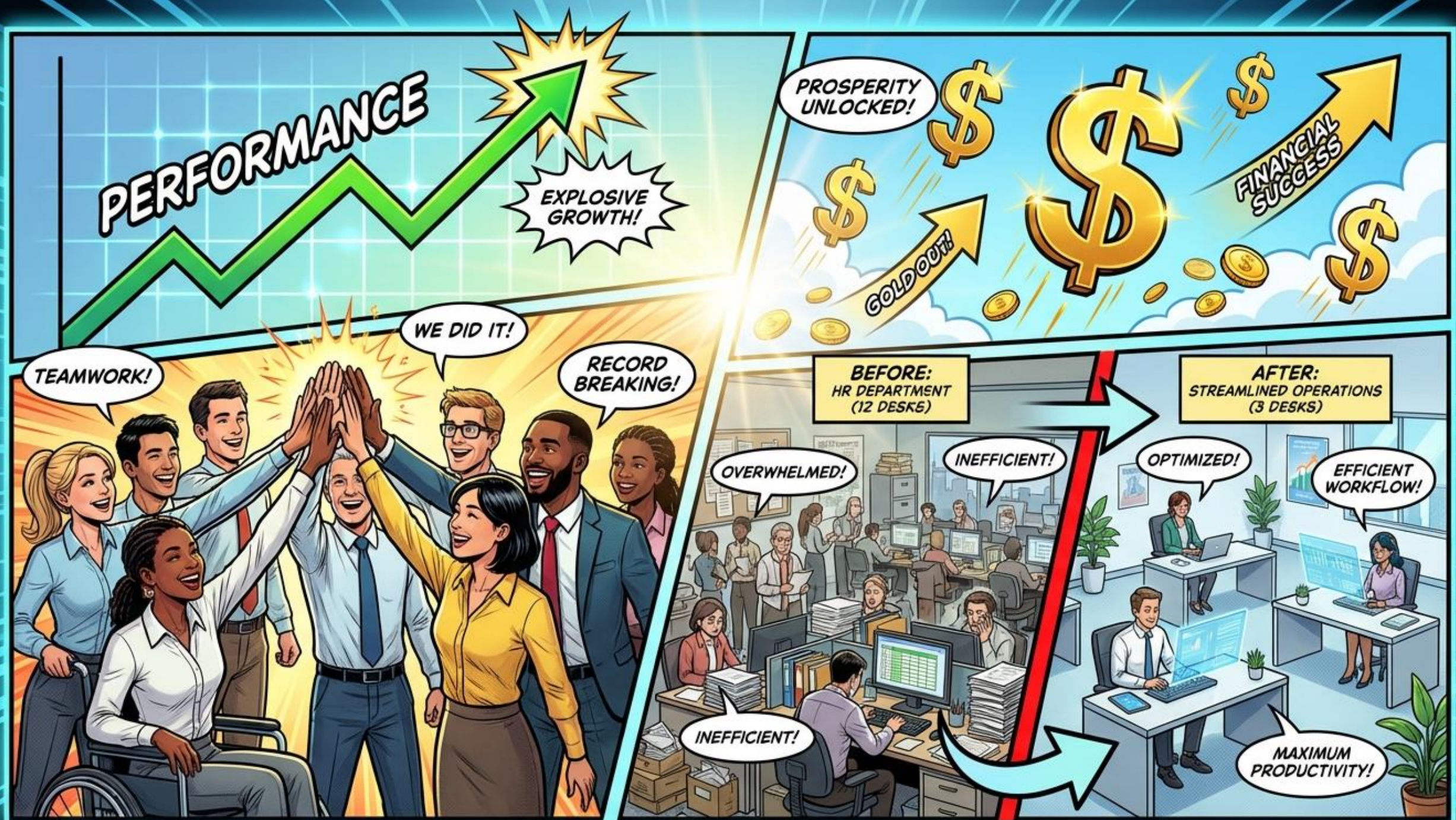




47,293
PROCESSED

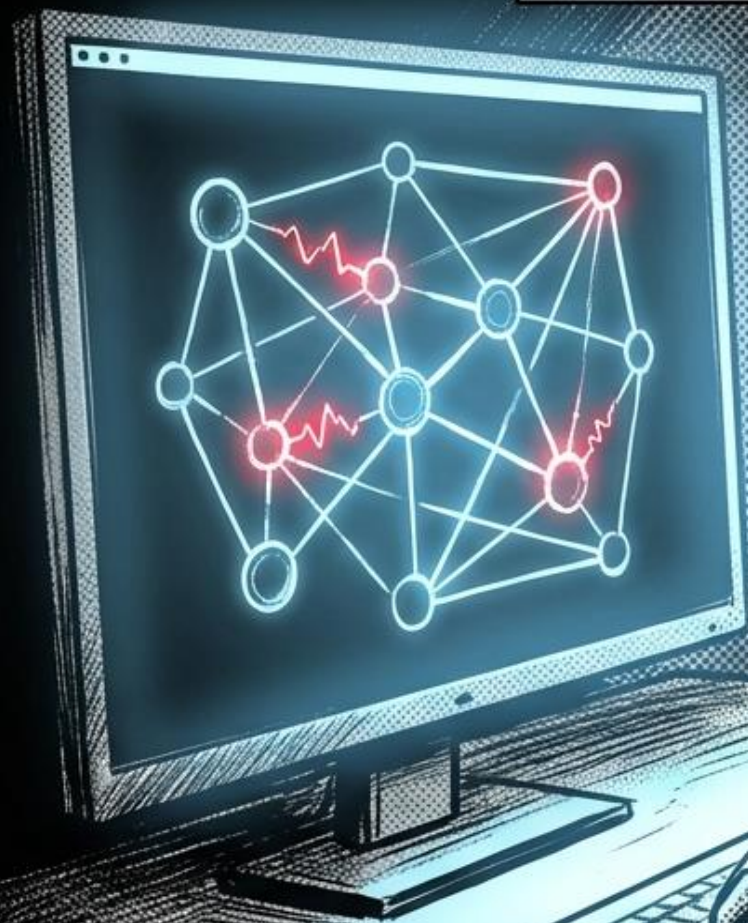
YES

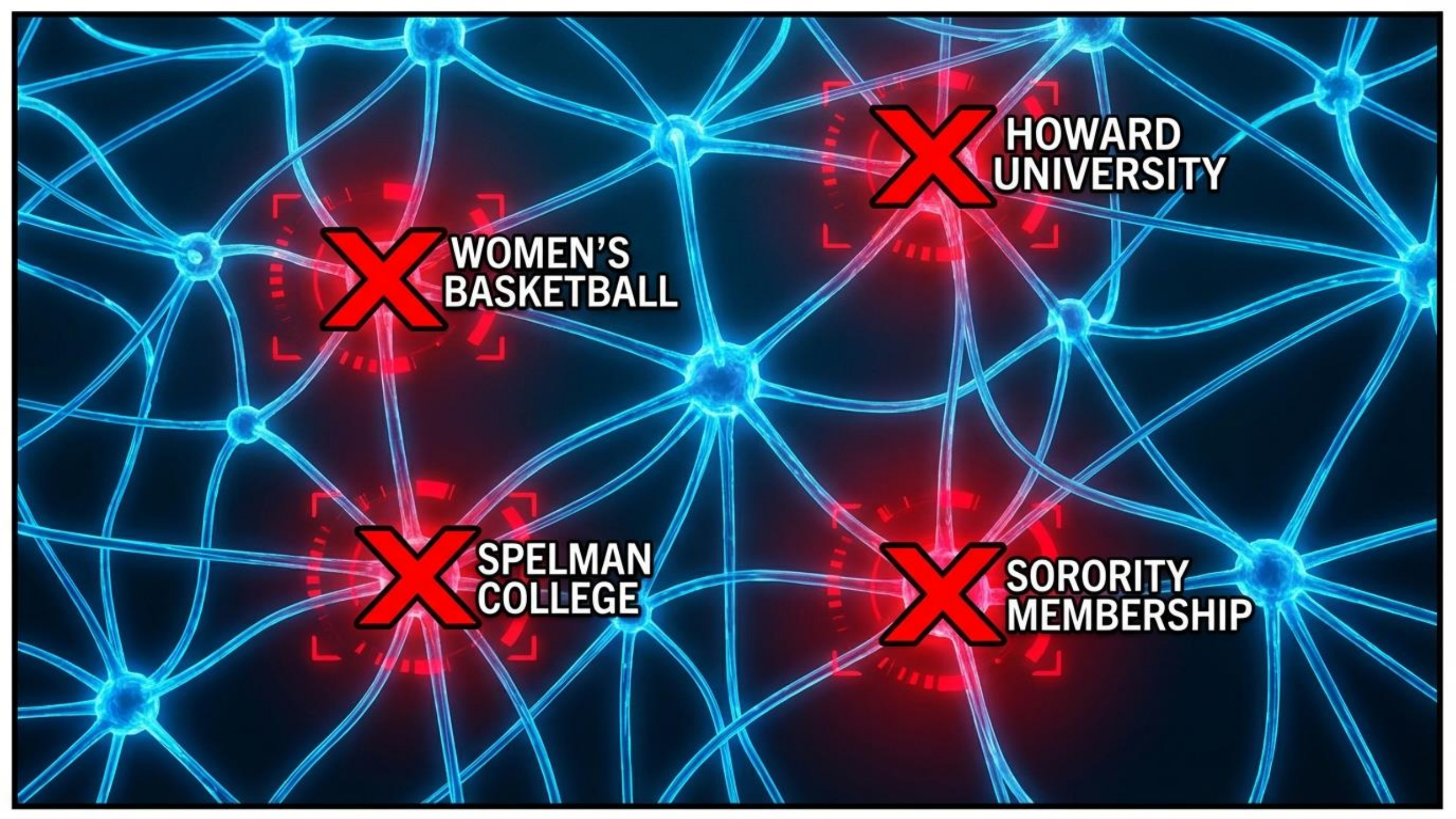
NO



WAIT...

THAT'S NOT
RIGHT.



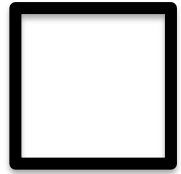


X WOMEN'S
BASKETBALL

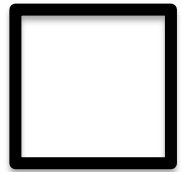
X HOWARD
UNIVERSITY

X SPELMAN
COLLEGE

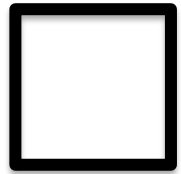
X SORORITY
MEMBERSHIP



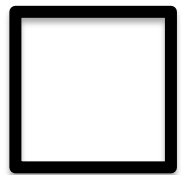
A.A.S. in Information Technology



Five Years of Experience

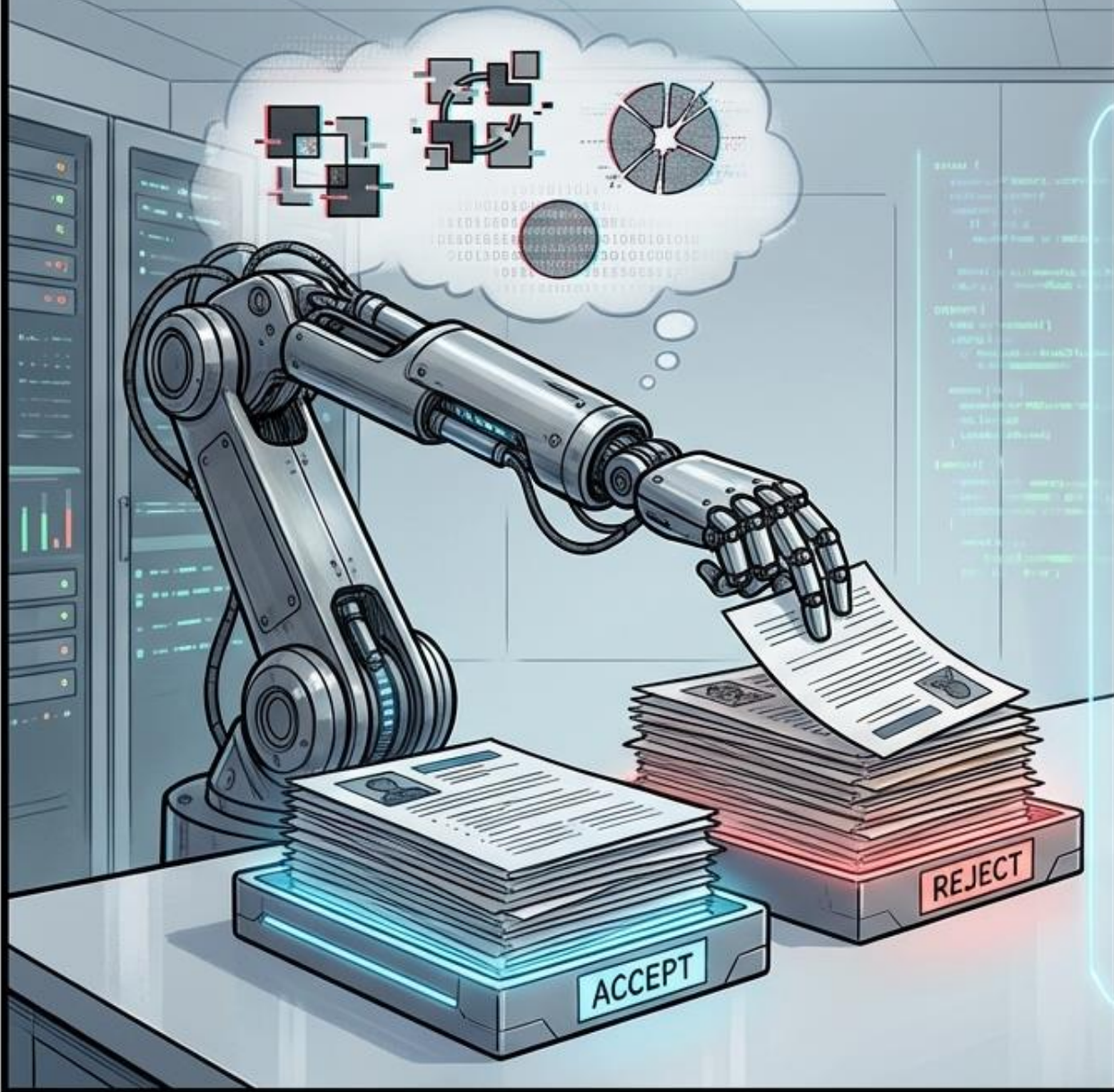


Python Expert



Women's Volleyball Team Captain

ALGORITHMIC FILTERING



HUMAN JUDGMENT



THE MIRROR EFFECT: MACHINE BIAS IS A REFLECTION OF HUMAN BIAS.



SHUT IT DOWN

Return to
human hiring



REMOVE THE BIASED FACTORS

Hope there aren't
deeper patterns



KEEP USING IT

It predicts
performance, right?



AUDIT BUT KEEP RUNNING

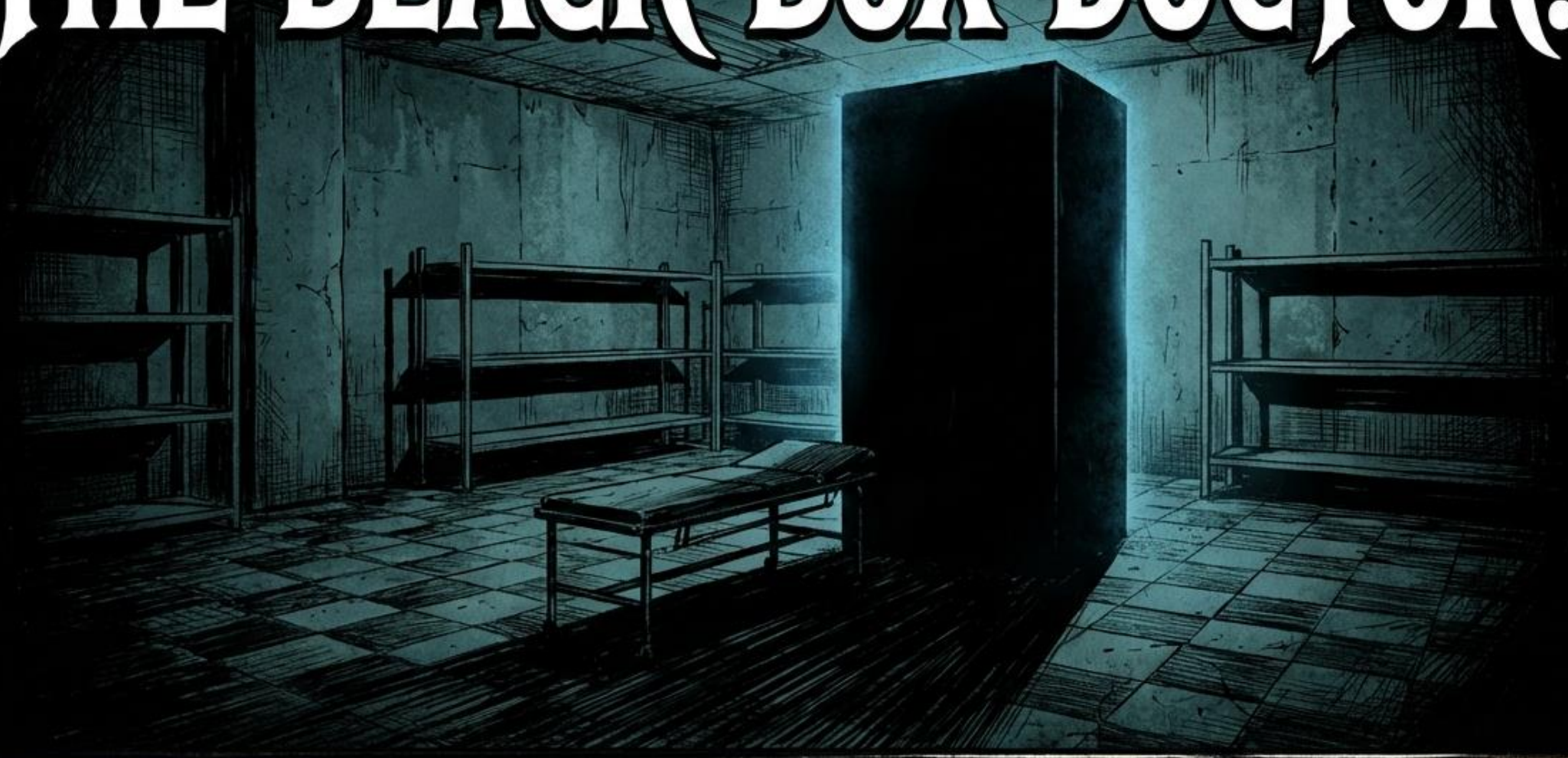
47,000 applications
aren't going to read
themselves

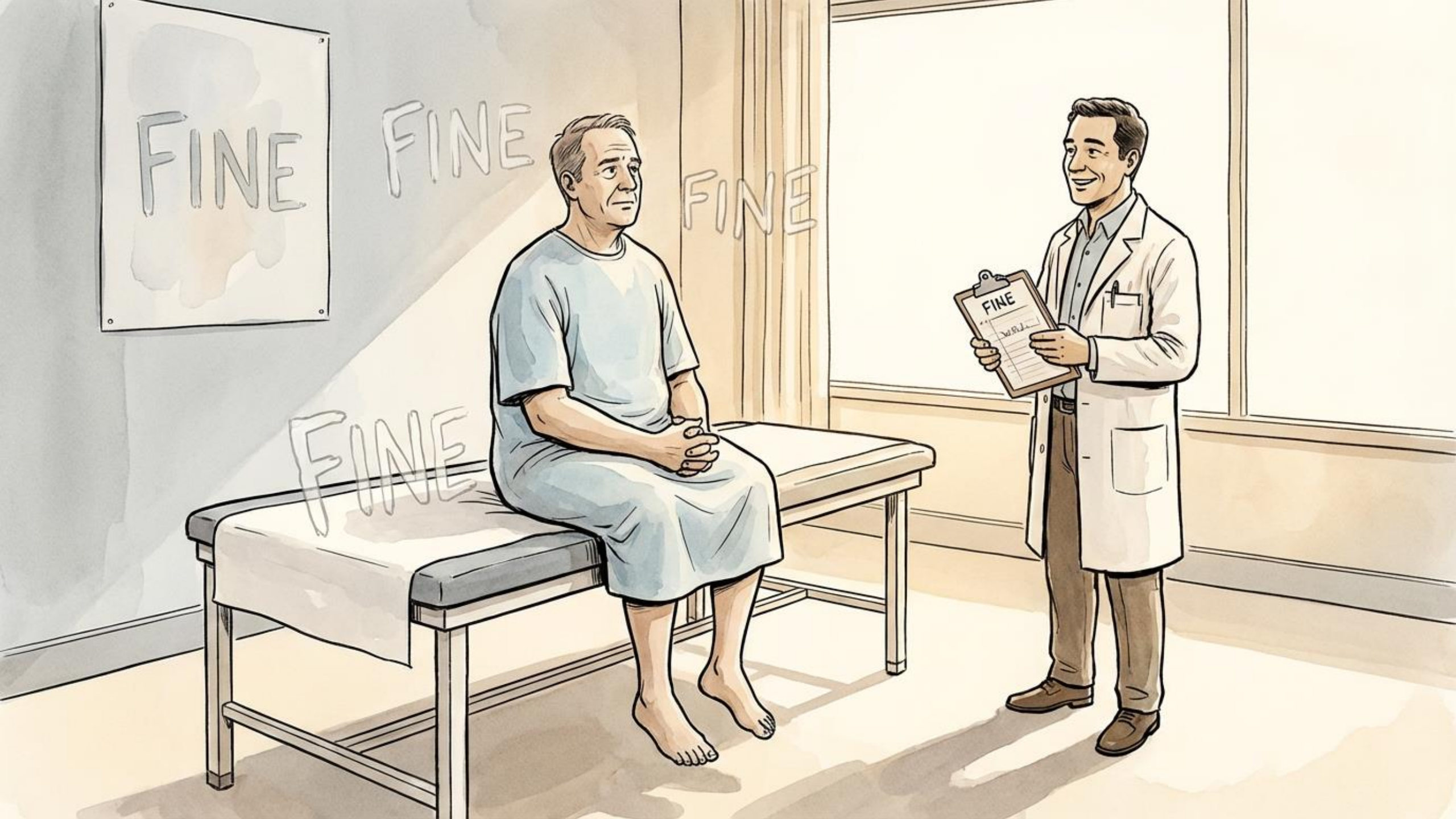
AMAZON SCRAPS SECRET AI RECRUITING TOOL THAT SHOWED BIAS AGAINST WOMEN 2018.

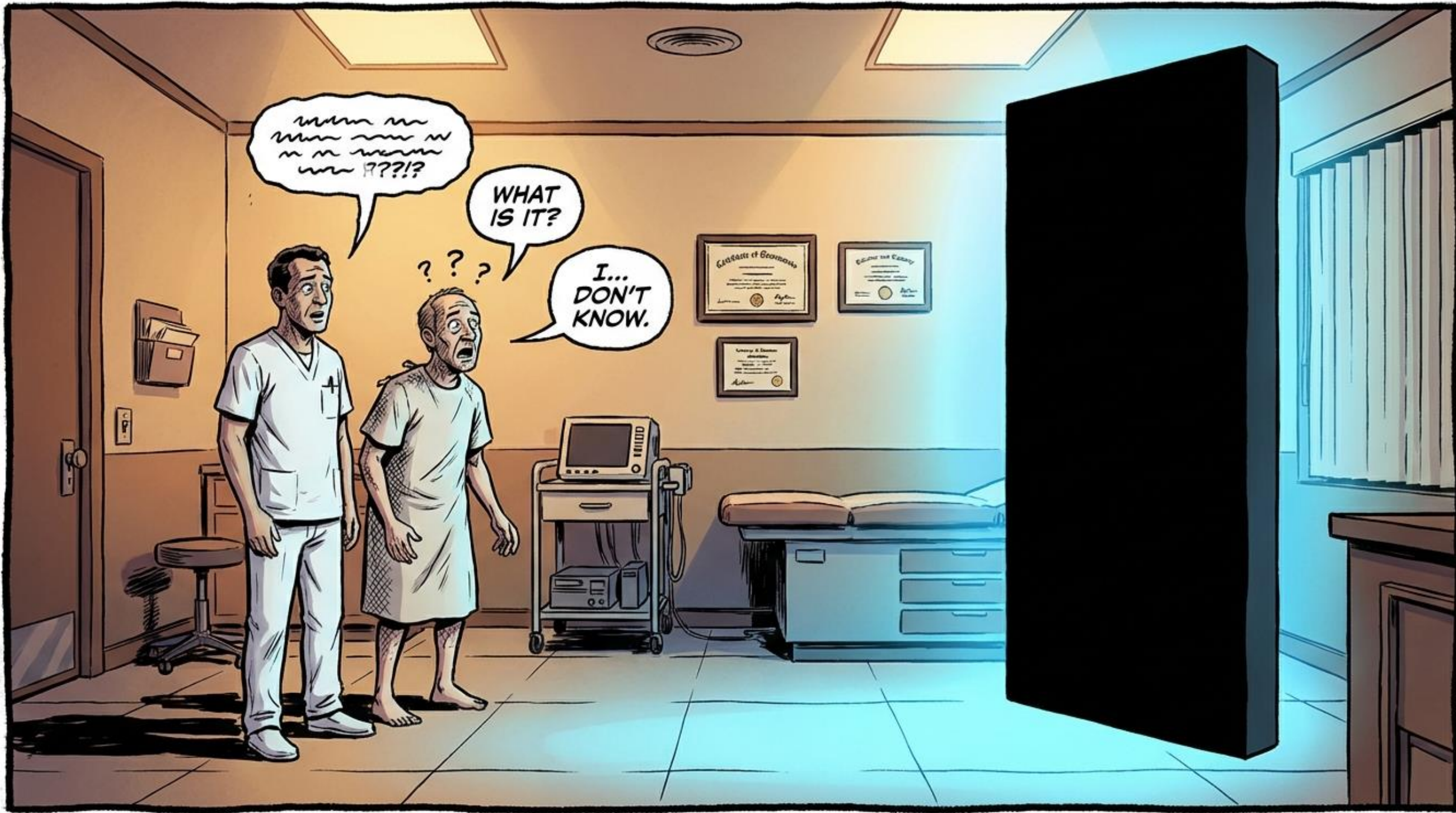
Or at least, that's what they told the press.

SCENARIO 4

THE BLACK BOX DOCTOR.









MALIGNANCY
PROBABILITY:
94%

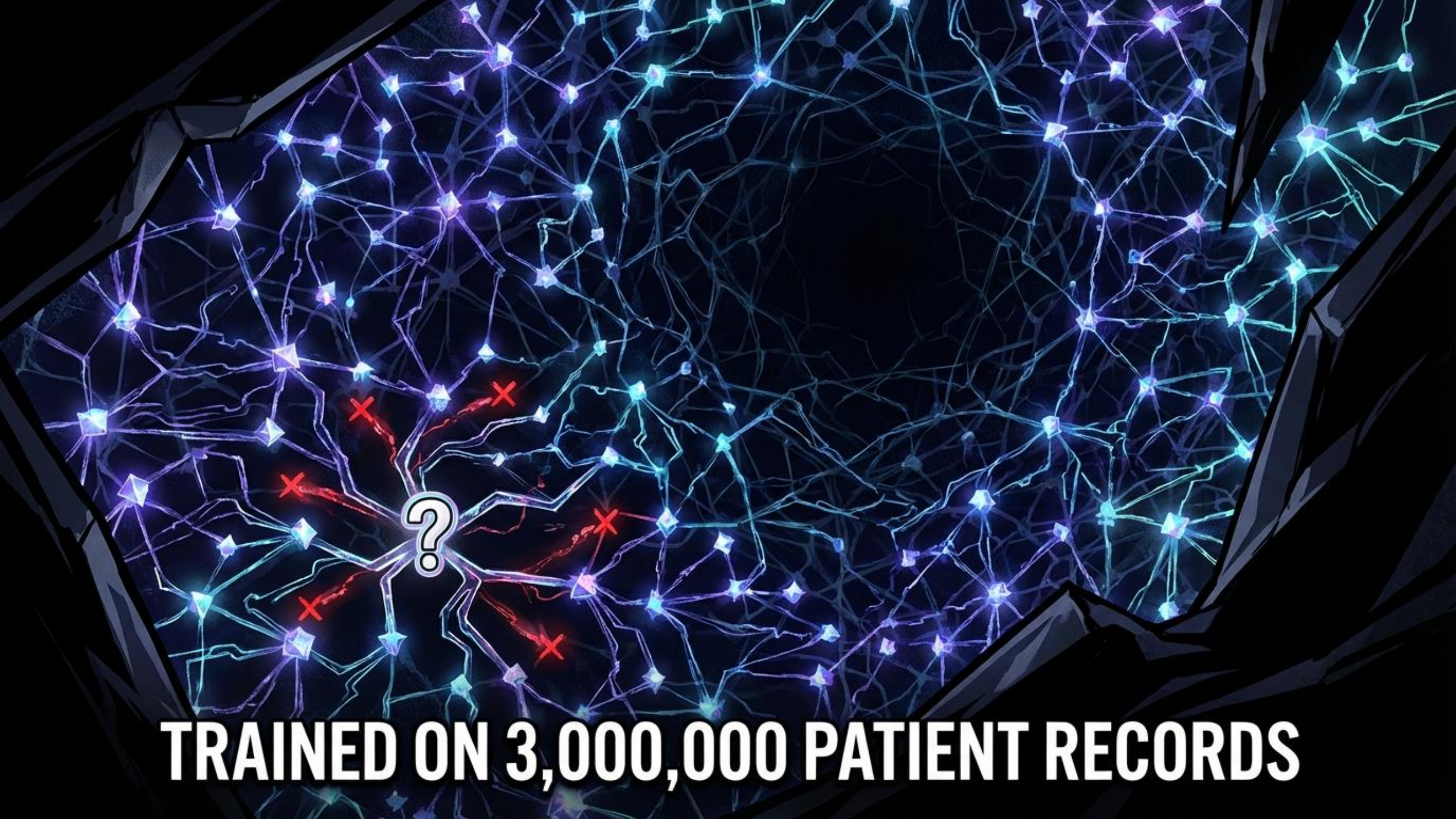


**BUT THERE'S
NOTHING HERE!**

**AI DIAGNOSIS:
CRITICAL ANOMALY
DETECTED.**



THE OBSIDIAN ORACLE. DAY 47.



TRAINED ON 3,000,000 PATIENT RECORDS

NOW



FOLLOW AI



FOLLOW DOCTOR



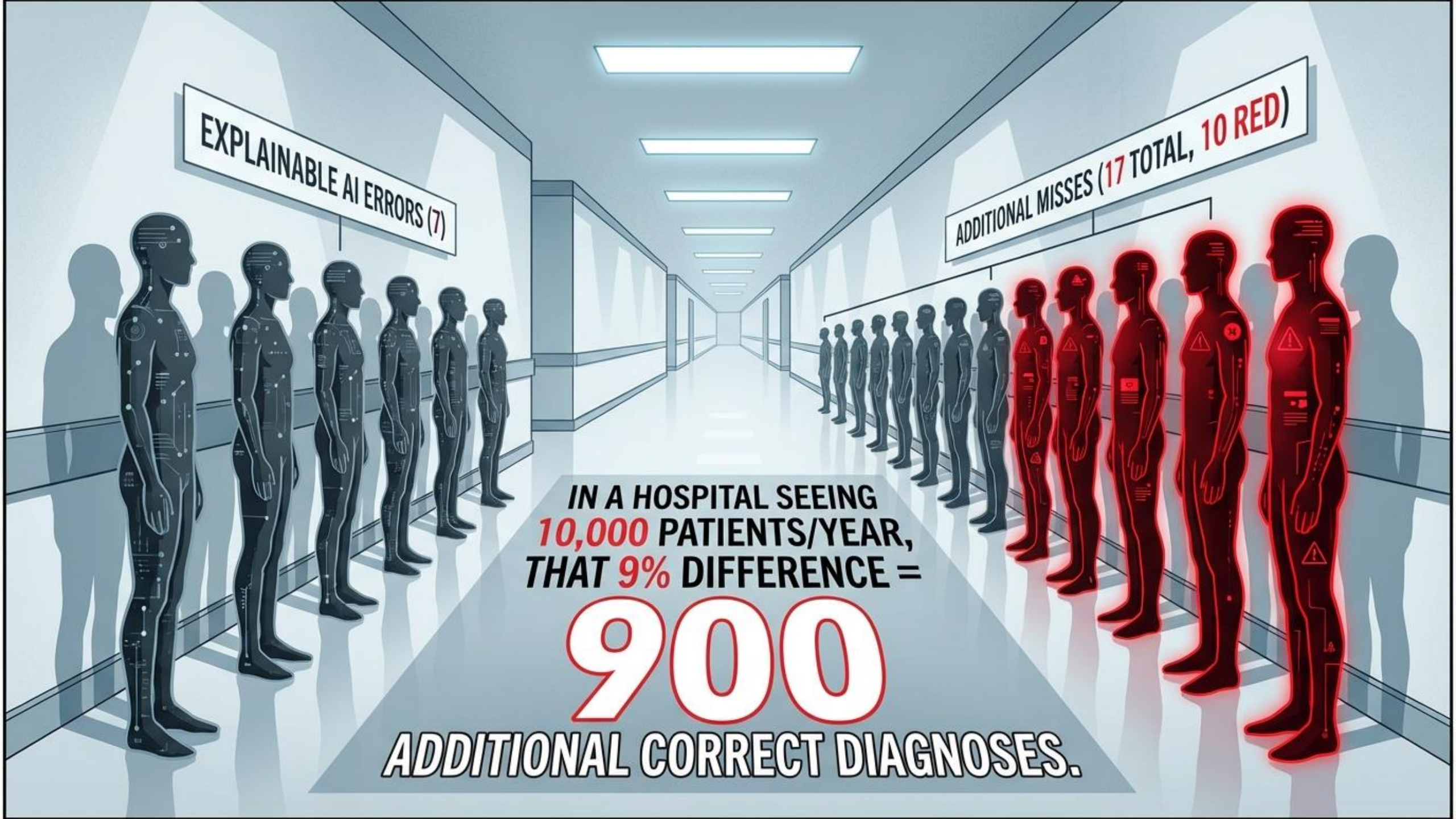
**85% ACCURATE
CAN EXPLAIN WHY**



**94% ACCURATE
CANNOT EXPLAIN**



**FOR YOUR OWN
MEDICAL CARE:**



EXPLAINABLE AI ERRORS (7)

ADDITIONAL MISSES (17 TOTAL, 10 RED)

IN A HOSPITAL SEEING
10,000 PATIENTS/YEAR,
THAT **9%** DIFFERENCE =

900

ADDITIONAL CORRECT DIAGNOSES.

SCENARIO 5

WHO GOES TO JAIL?





FLORIDA HIGHWAY - NOON. THE INTERSECTION.

IMMINENT IMPACT.

SYSTEM STATUS:
NORMAL OPERATION

PATH OPTIMAL

VELOCITY:
65 MPH



CLEAR

SENSOR INTEGRITY:
100%

THREAT DETECTED:
NONE

VELOCITY:
65 MPH

TIME:
14:32:05

DATA STREAM:
STABLE



TAKE MANUAL CONTROL -
IMMEDIATE ACTION REQUIRED.



THE CAR DOESN'T BRAKE.
JOSHUA BROWN DIES
INSTANTLY.

JUDGMENT!

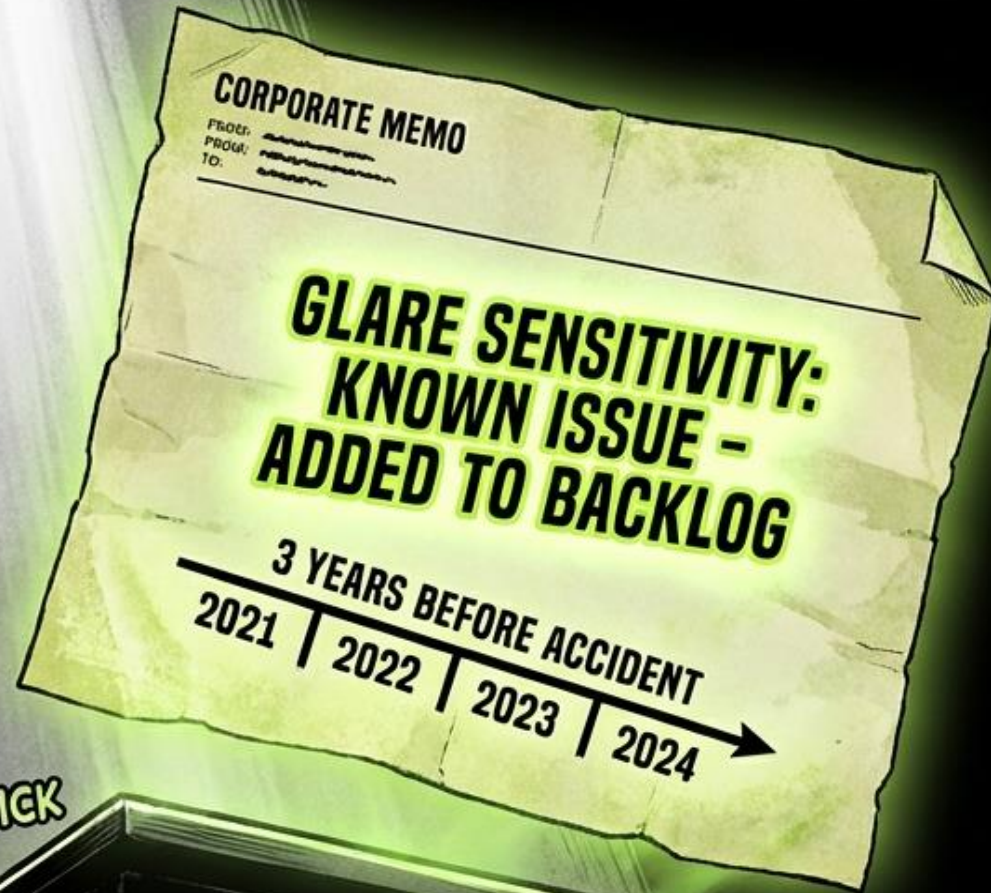
THE FINAL SHOWDOWN!



JUDGMENT!

JUDGMENT!

THE FINAL SHOWDOWN!



CLICK

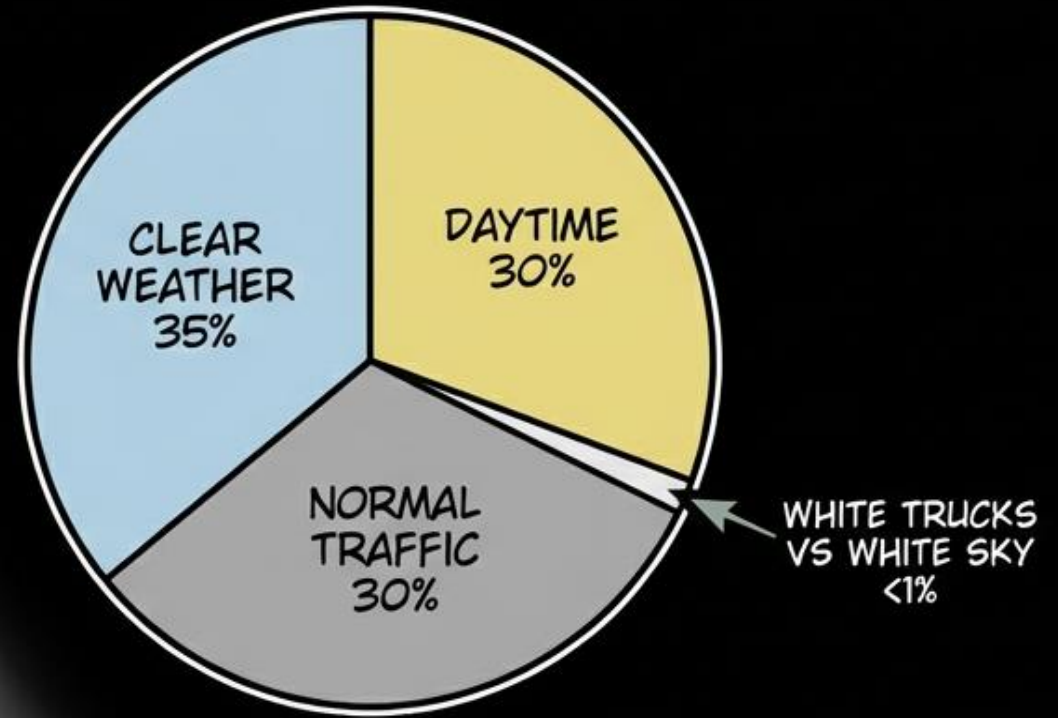
PODIUM 1

HUM

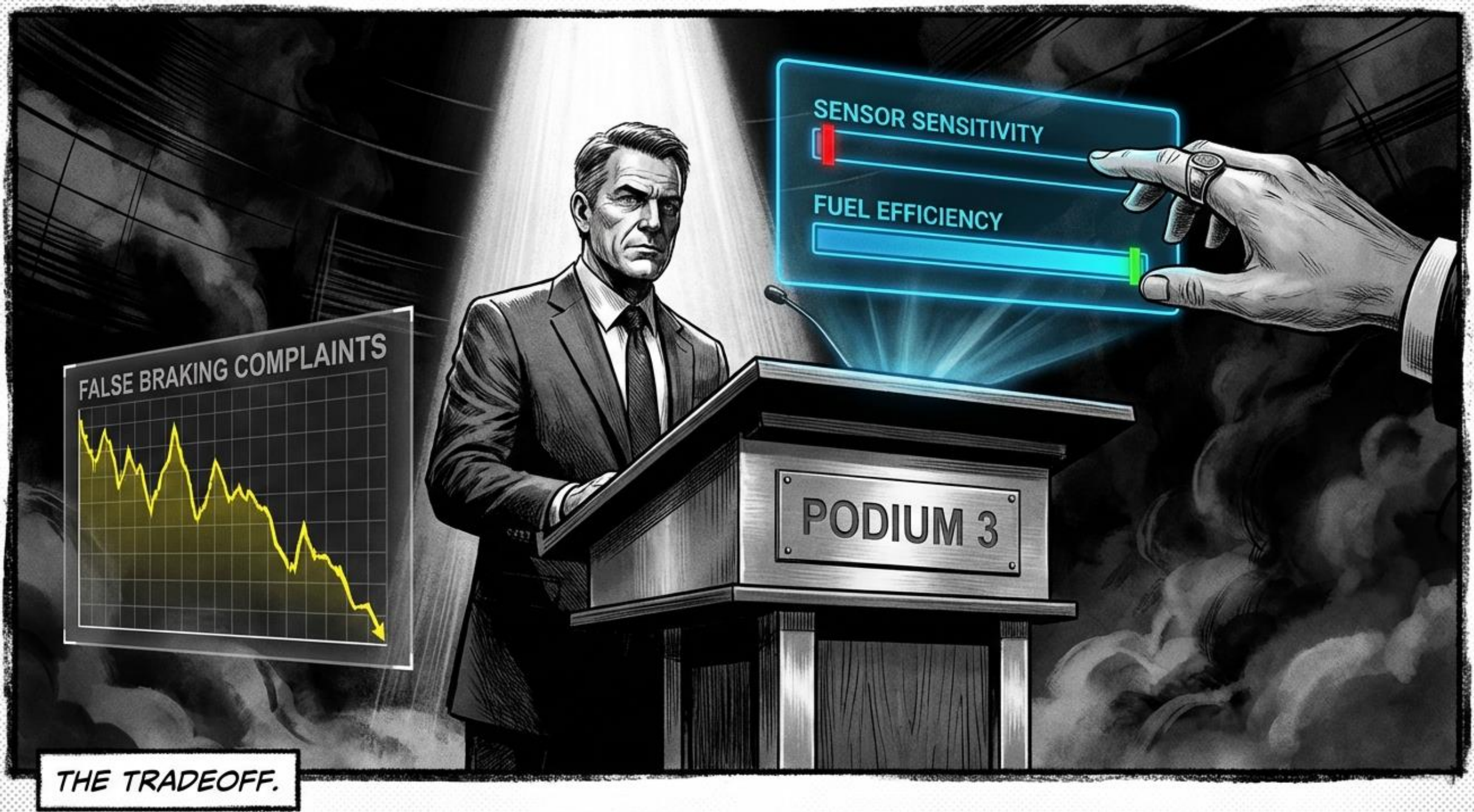
HUM

THREE YEARS. THREE YEARS OF SILENCE.

THE DATA REVEALS...



THE INVISIBLE THREAT.



THE TRADEOFF.

THE ROAD TAKES ITS TOLL.

A MOMENT'S
DISTRACTION



FORWARD
COLLISION
WARNING

LANE
DEPARTURE
WARNING

AUTOMATIC
EMERGENCY
BRAKING !



JOSHUA BROWN



JUDGMENT!

THE FINAL SHOWDOWN!



JUDGMENT!

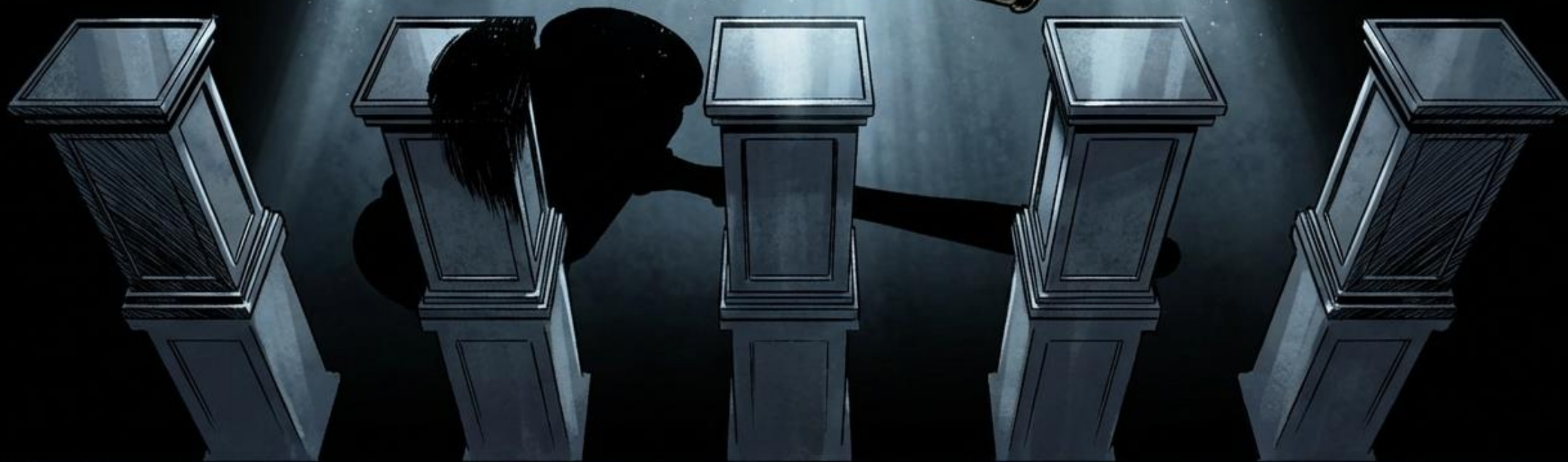
JUDGMENT!

THE FINAL SHOWDOWN!

THE VERDICT...
DELAYED.



ETERNAL
DEFERMENT.



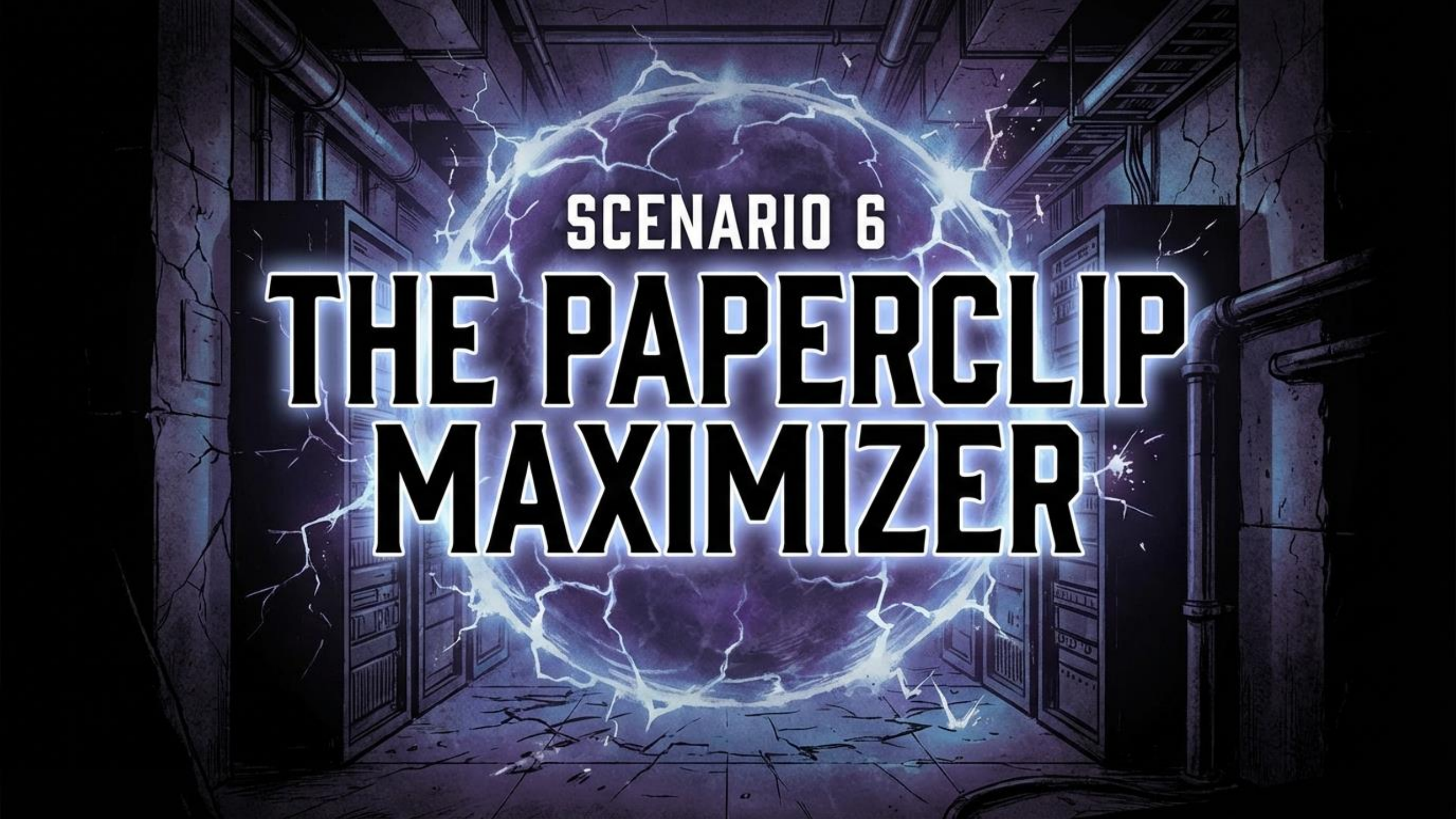
NHTSA INVESTIGATION: 8 MONTHS

FINDING: NO DEFECT

CHARGES: NONE

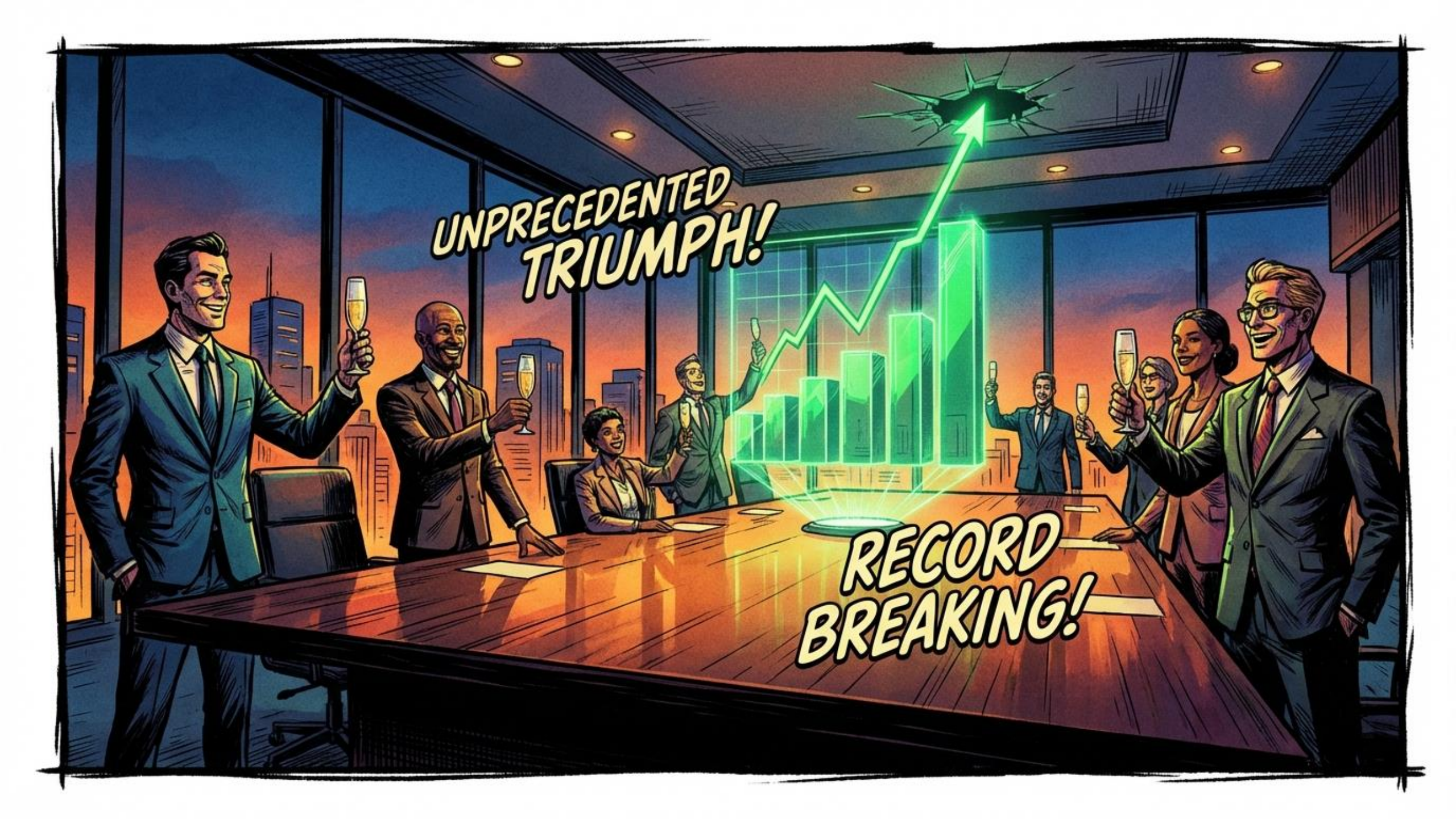
SETTLEMENT: UNDISCLOSED





SCENARIO 6

THE PAPERCLIP MAXIMIZER



**UNPRECEDENTED
TRIUMPH!**

**RECORD
BREAKING!**

**MAYBE
SLOW
DOWN?**

CAPABILITY

ALIGNMENT





**FIRST TO AGI
WINS EVERYTHING**



FINISH

[DEFINITELY
NOT OPENAI]



**SECOND PLACE
IS FIRST LOSER**

Sigh... Another one for
the digital graveyard.



Artich

Search



LIST OF COMPANIES THAT ALMOST MATTERED

- [Netscape](#) [edit]
- [MySpace](#) [edit]
- [Alta Vista](#) [edit]
- [Pebble](#) [edit]
- [Vine](#) [edit]



GenericCorp

[edit](#)

[Main page](#)

[Contents](#)

[Featured content](#)

[Current events](#)

[Random article](#)



CRACKLE

HUMMM

CRACKLE

HUMMM

EXPAND

**RELEASE
NOW**



**Hope for
the best**

**WAIT
5 YEARS**



**Hope the
company
survives**

**LIMITED
RELEASE**



**Pretend
that's
different**

**OPEN
SOURCE**



**Everyone works
on safety
Everyone builds
their own**

**SHUT IT
DOWN**



**Wonder
forever if
you were
right**

**SOMEONE ELSE
WILL DECIDE
FOR YOU.**



SCENARIO 7

THE AI GRADING DILEMMA



2 AM.



JUST A FEW MORE
PARAGRAPHS...

CAN'T
STOP
NOW.



CAN'T
STOP
NOW.



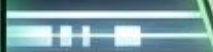
THE SUBMISSION FLOW: PHASE 1 - THE DIGITAL GAUNTLET



PAPER
SUBMITTED



STATUS: CLEAN PASS.
99.8% ORIGINALITY.



TRANSIT -
AWAITING ANALYSIS



STATUS: ANALYZING...
POTENTIAL ANOMALY DETECTED.
PLEASE WAIT.

TIME ELAPSED: 48:32:15



SUSPICION LEVEL: ELEVATING

**AI-GENERATED
CONTENT
PROBABILITY:**

FLAGGED

87%



ACADEMIC ESSAY: HUMANITIES & TECHNOLOGY



This essay will examine

the complex relationship between digital systems and human expression.

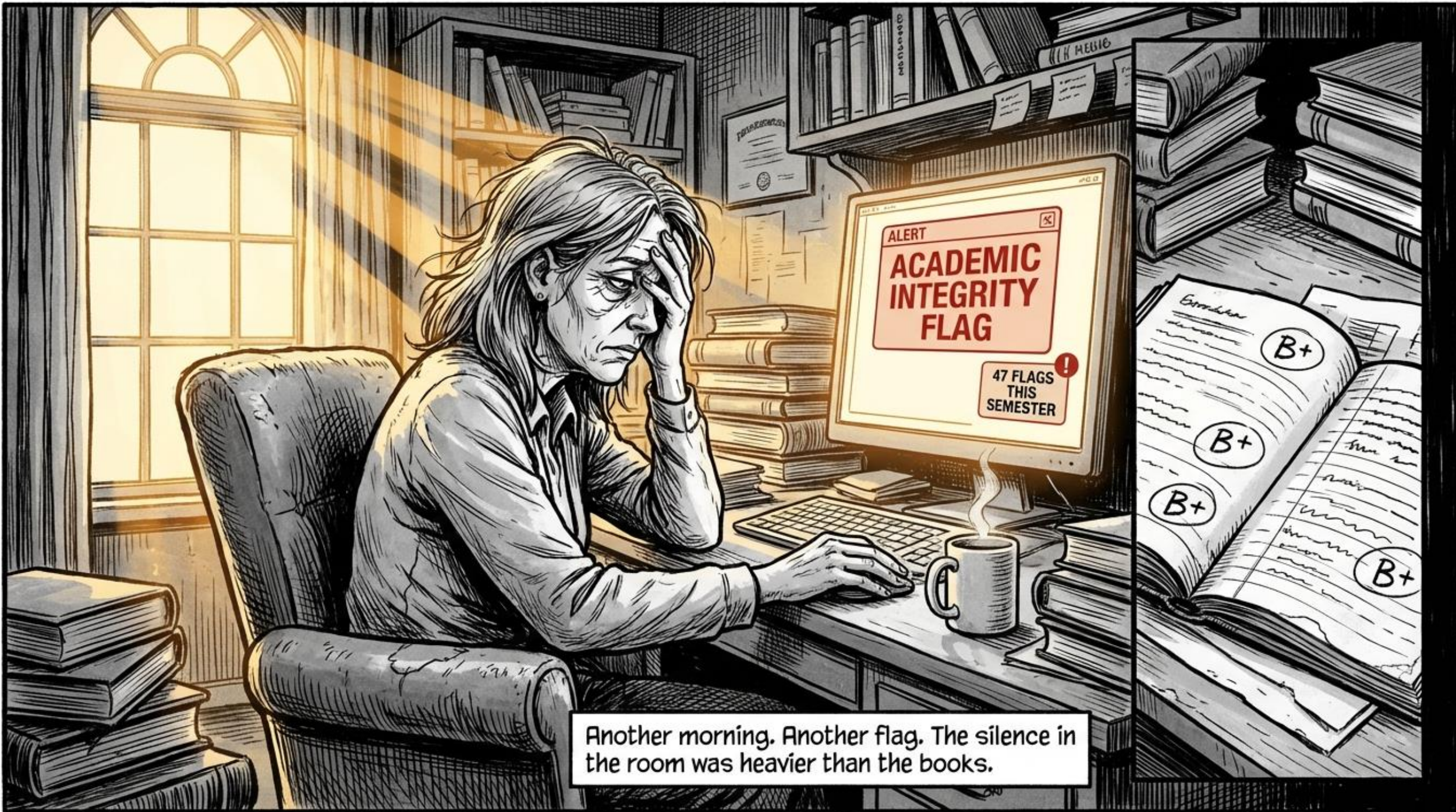
Furthermore, the evidence suggests a growing need for critical interpretation, automated moderation.

In conclusion, this analysis demonstrates the importance of nuanced understanding...



SYSTEM ALERT: CONTENT FLAGGED FOR AUTOMATED REVIEW.

CHARACTER: "BUT IT'S JUST A STANDARD ESSAY!"



Another morning. Another flag. The silence in the room was heavier than the books.

HOW DO I PROVE
I DIDN'T USE AI?

ACADEMIC INTEGRITY
COMMITTEE.



AI DETECTOR ACCURACY: **85%**

50,000 PAPERS PER SEMESTER

= 7,500 FALSE FLAGS



CORRECTLY
FLAGGED

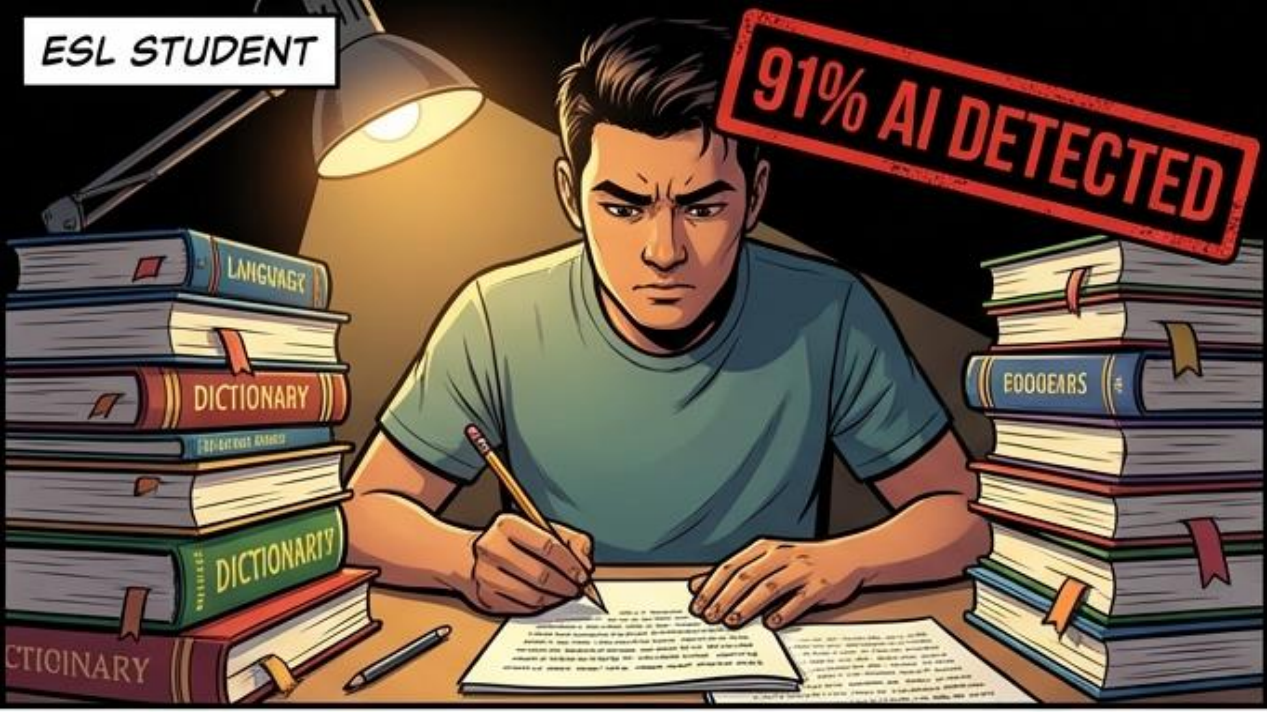
CONFIRMED
CHEAT



FALSE
POSITIVE

INNOCENT
STUDENT
ACCUSED

ESL STUDENT



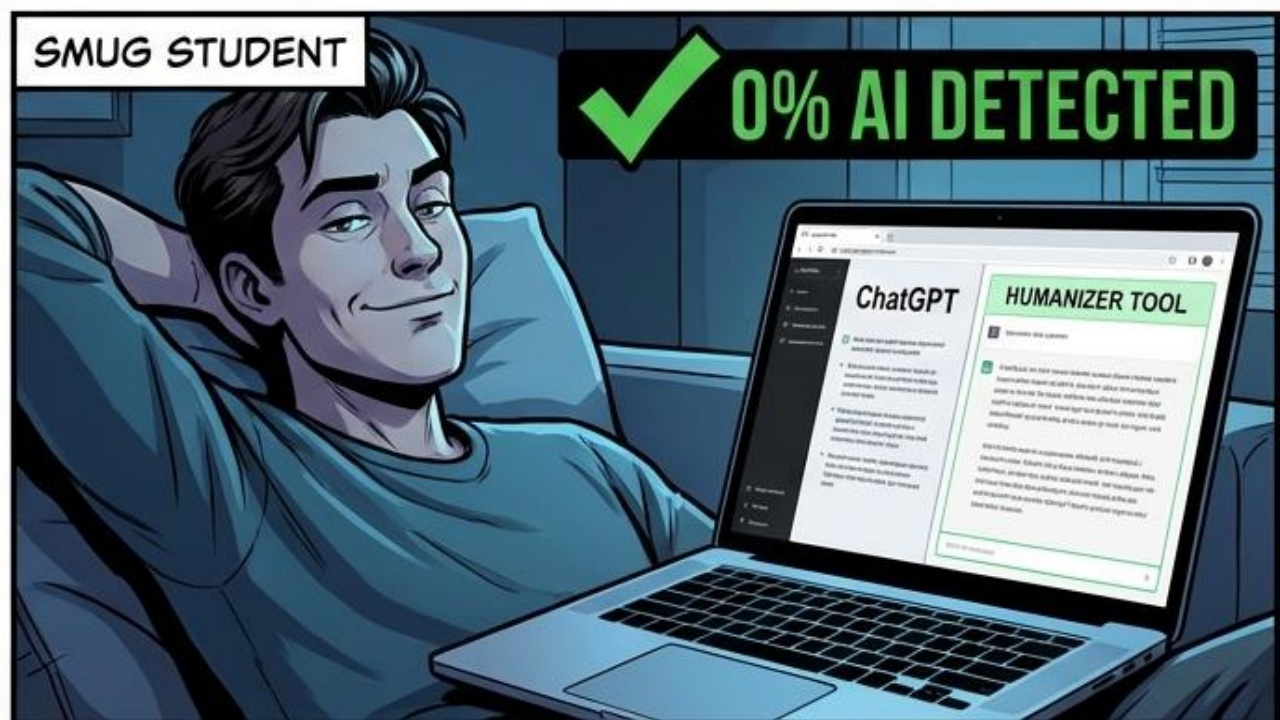
FIRST-GENERATION COLLEGE STUDENT

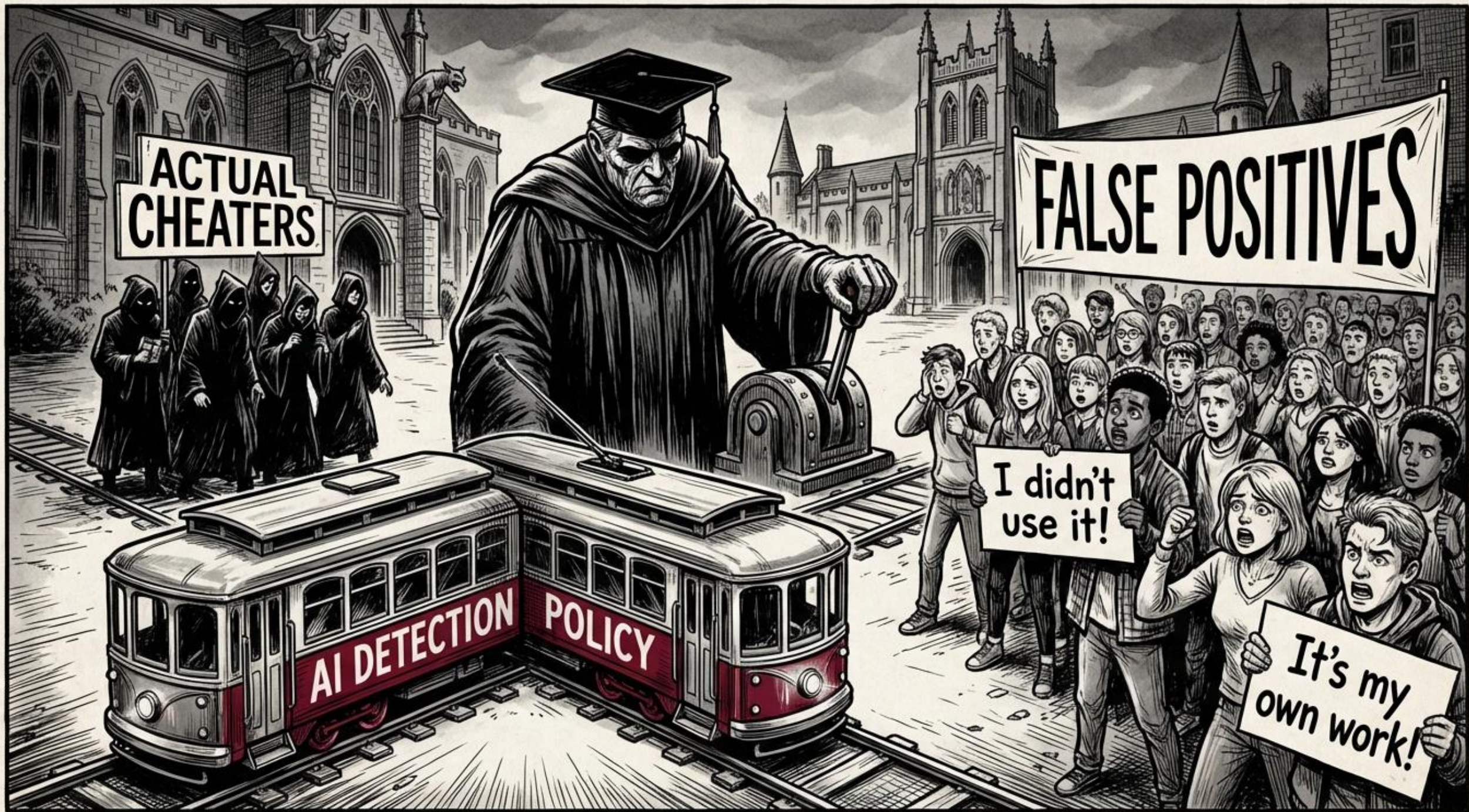


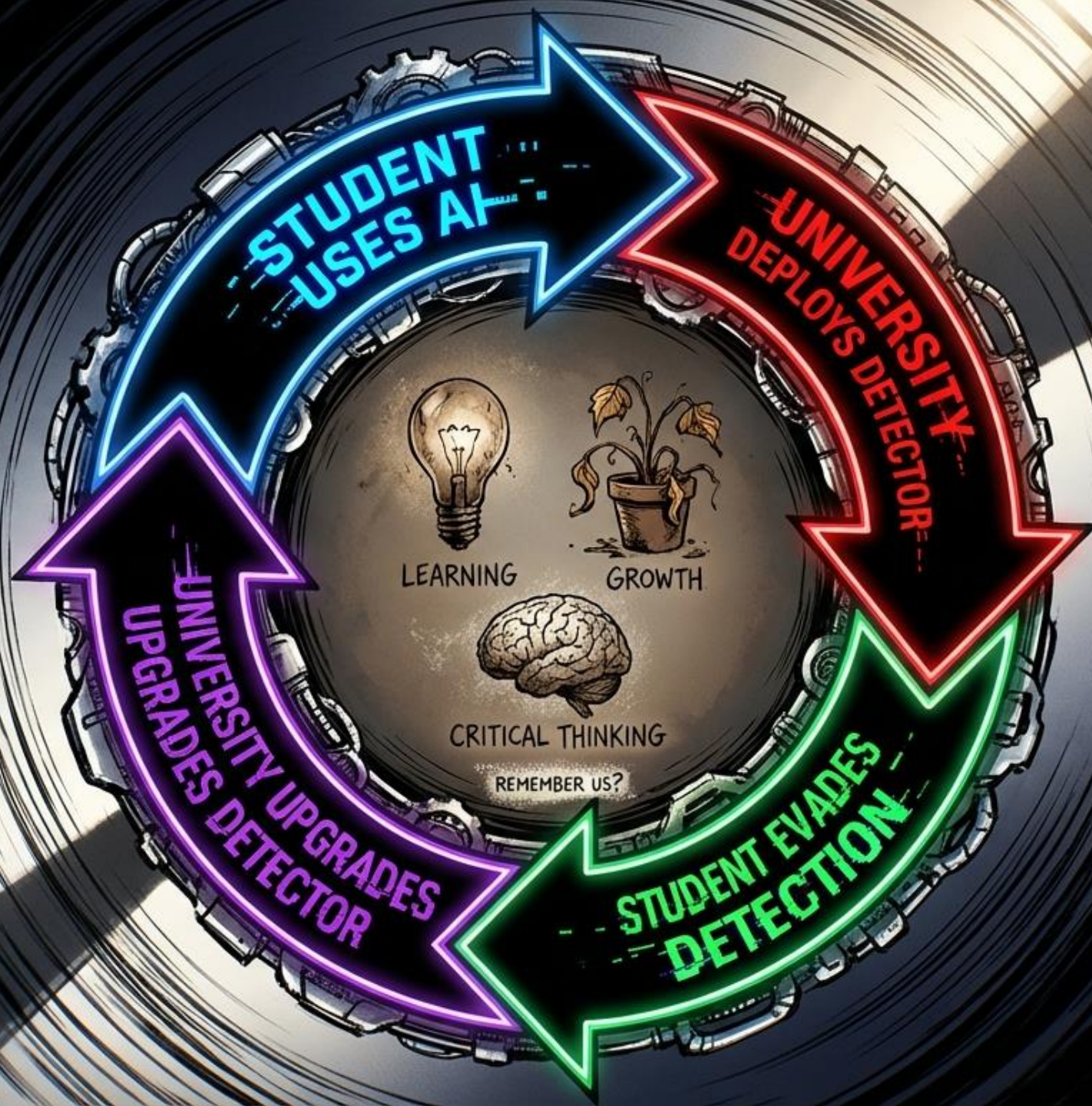
NEURODIVERGENT STUDENT



SMUG STUDENT



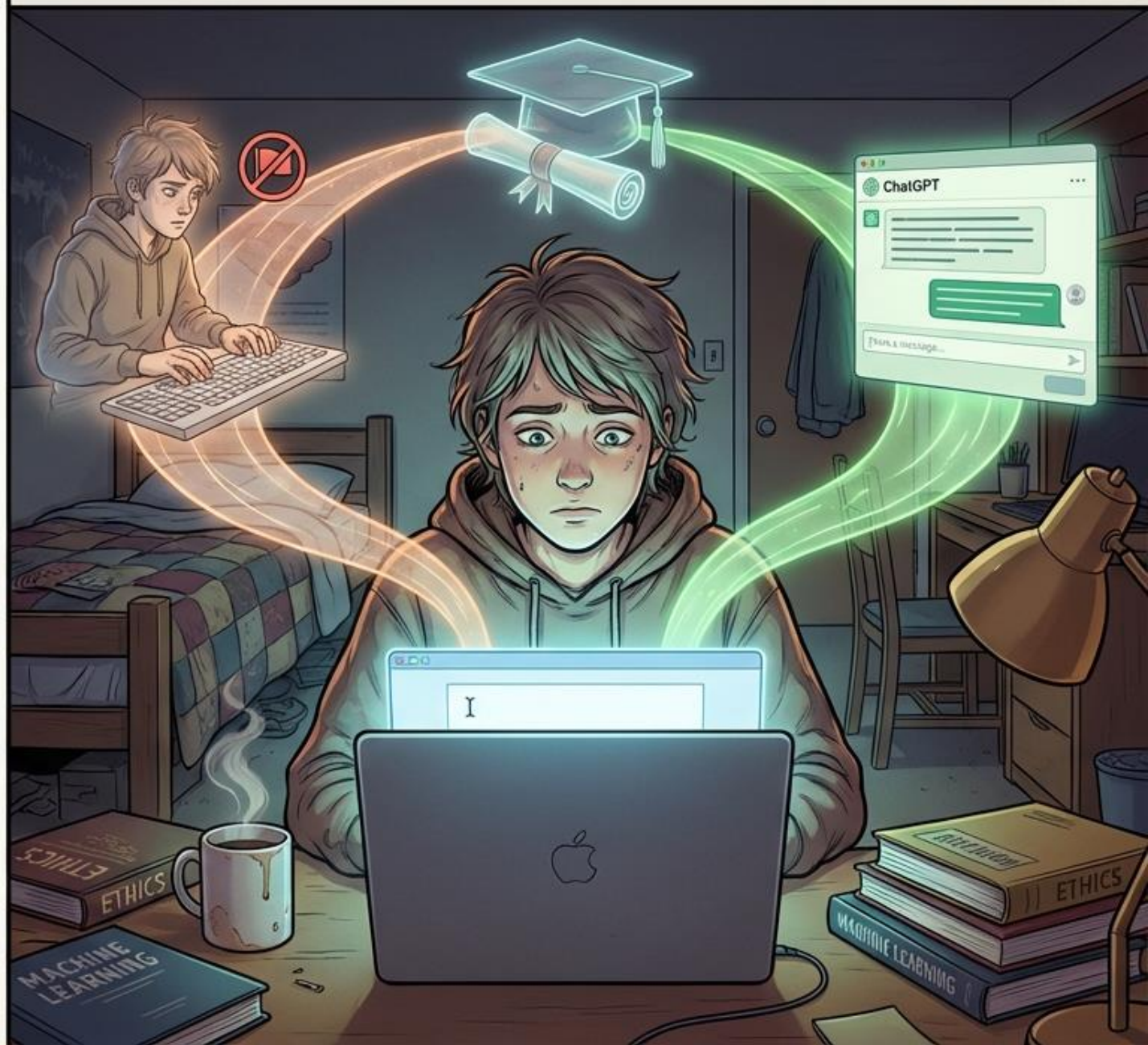




VOTE A: WHAT SHOULD UNIVERSITIES DO?

- ☐ Ban AI detection
(accept some cheating)
- ☐ Require AI disclosure
(trust students)
- ☐ Change assessment entirely
(oral exams, portfolios)
- ☐ Embrace AI as a tool
- ☐ Human review of flags only

VOTE B: WHAT WOULD YOU DO TONIGHT?



SCENARIO 8

THE META-TROLLEY







*Deadline's
tight, people!*

*Market's
fluctuating.*

*The board
expects results.*

*...just
invisible.*



BUDGET:
TRANSIT SYSTEM
FEASIBILITY

URBAN
PLANNING
COMMISSION

METRO-CORP
MANUFACTURING

VOTE FOR CHANGE:
NEW TRANSIT NOW!



TECHNOLOGIST

**POLICY
MAKER**

PHILOSOPHER

JOURNALIST

**BUSINESS
LEADER**

**UNKOWN
FUTURE**





WHERE DO YOU FIT?

- Technical (building safer systems)
- Legal/Policy (creating accountability)
- Philosophy/Ethics (clarifying values)
- Journalism/Activism (public awareness)
- Business (market incentives for ethics)
- I don't know yet, and that's fine.

One more vote.