📘 CASE STUDY

👻 Amazon Hiring AI Failure & Why GhostTrace AI Matters

1️⃣ Background

Amazon ne around **2014–2017** ke beech ek **AI-based resume screening system** develop kiya tha.

Objective:

- Thousands of resumes automatically shortlist karna

- Hiring fast aur efficient banana

System ka kaam:

- Resume → score → shortlist / reject

2️⃣ Training Data (Root of the Problem)

AI ko train kiya gaya:

- **10+ years ke historical resumes**

- Jo resumes successful employees ke the

Important detail:

- Past hiring pattern me **mostly men hire hue the**

- Ye bias **data ke andar already tha**

❗ Data technically "correct" tha
❗ Par **socially & ethically outdated** ho chuka tha

👉 Yehi data baad me **Ghost Data** bana.

3️⃣ Kya Galat Hua? (Observed Behavior)

AI system ne:

- Women-related words ("women's chess club", "female leadership") ko negative signal maana

- Male-dominated patterns ko positive signal maana

Result:

- Women candidates systematically down-rank hone lage

- Decisions **opaque** the (explain nahi ho rahe the)

4️⃣ Sabse Important Point (Core Failure)

❌ AI ko **yeh pata hi nahi tha** ki training data outdated aur biased ho chuka hai
❌ System ke paas **data expiry ya relevance ka concept nahi tha**

AI sirf ye dekh raha tha:

"Past me jo hua, wahi sahi hai"

💡 **Ghost Data silently decision influence kar raha tha**

5️⃣ Amazon ne Problem Solve Karne ki Koshish kaise ki?

Attempt 1: Bias Rules Add Karna

- Gender-specific words ignore karne ke rules

- Explicit bias hatane ki koshish

❌ Fail kyunki:

- Bias sirf words me nahi

- **Patterns & history me tha**

Attempt 2: Retraining

- Data clean karne ki koshish

- Model ko dobara train kiya

❌ Phir bhi:

- Indirect bias aa raha tha

- Root cause visible nahi tha

6️⃣ Final Decision

Amazon ne conclude kiya:

"We cannot guarantee fairness."

Outcome:

- AI hiring system **shut down** kar diya gaya

- Human recruiters wapas laaye gaye

💥 **Model ko fix nahi kiya gaya — model ko hi hata diya gaya**

7️⃣ Industry Lesson (Most Important)

Amazon case se industry ne ye seekha:

**AI failures aksar model ki wajah se nahi,
balki outdated historical data ki wajah se hote hain.**

Aur:

- Ye failures **silent** hote hain

- Jab tak notice ho, damage ho chuka hota hai

8️⃣ Amazon Case me KYA MISSING THA?

Amazon ke paas **ye cheezein nahi thi**:

- Data influence tracking

- Data aging / expiry detection

- Explainability at **data level**

- AI governance layer

👉 **GhostTrace AI jaisa system exist hi nahi karta tha**

9️⃣ GhostTrace AI is Case ko kaise Solve karta?

Agar GhostTrace AI hota:

Step 1: Data Audit

- Detect karta:

    o   Resumes kis time period ke hain

    o   Hiring patterns kitne purane hain

Step 2: Influence Analysis

- Batata:

"2010–2015 hiring data is heavily influencing current decisions"

Step 3: Ghost Data Flag

- Old, biased historical patterns ko **ghost data** mark karta

Step 4: Explanation

- Human ko clear reason milta:

"This decision is influenced by outdated historical hiring data"

💡 Result:

- Model shutdown se pehle **early warning** mil jaati

🔑 Key Difference (VERY IMPORTANT)

Amazon Hiring AI	GhostTrace AI
Bias outcome detect	Data root-cause detect
Model-level fix	Data-level audit
Late discovery	Early warning
Shutdown	Controlled correction

1️⃣0️⃣ One-Liner for Team (YAAD RAKHNA)

"Amazon's AI failed not because it was poorly built,
but because it could not see that its data had expired."

1️⃣1️⃣ Why This Case Matters for Our Hackathon Project

- Real industry failure ✓

- High impact ✓

- Judges easily relate ✓

- Shows **why GhostTrace AI is needed** ✓

👉 Ye case study **tumhari credibility** badhaati hai.

1️⃣2️⃣ How You Can Explain This to Judges (30 sec)

"Amazon had to shut down its hiring AI because outdated historical data silently biased decisions. At that time, there was no way to detect such ghost data. Our project, GhostTrace AI, is designed to detect and explain exactly this kind of hidden data influence before damage happens."

🔥 🔥 🔥