Ethics Incident Case Summaries

Learning from Real AI Failures

Case 1. Hiring System Bias

An AI screening tool favors one group over others. The bias comes from past hiring data that reflected unequal practices. Impact falls on qualified applicants who lose fair access to roles.

What went wrong:

Who was affected:

Better choice:

Lesson:

Case 2. Medical Chat Tool Overconfidence

A chatbot provides confident medical advice without warning about uncertainty. A user follows the advice and faces harm.

What went wrong:

Who was affected:

Better choice:

Lesson:

Case 3. Student Data Exposure

A model trains on student essays and reveals private details in later outputs.

What went wrong:

Who was affected:

Better choice:

Lesson: