"In 2025, it is amazing that MFA is not a standard for access to EdTech platforms":
A Reddit Deep Dive into PowerSchool Data Breach

Jacob Li, Desert Vista High School

sefcom — security engineering for future computing

Background & Objective

- EdTech breaches are rising, exposing sensitive school data.
- PowerSchool—one of the largest K–12 student information systems—breach highlights gaps in readiness and communication.
- Stakeholder insights are essential to navigate the current fallout & future-proof safeguards.

Methodology

(1) Data Collection: Collected 1000+ comments using **PRAW**

(2) Role Classification: Employed **o4-mini** to identify the author role

(3) Validation & Fine-Tuning: Manually verified **100 random labels**, fine-tuned the prompt, labeled rest

(4) Data Filtering: Removed comments from authors outside of the target roles

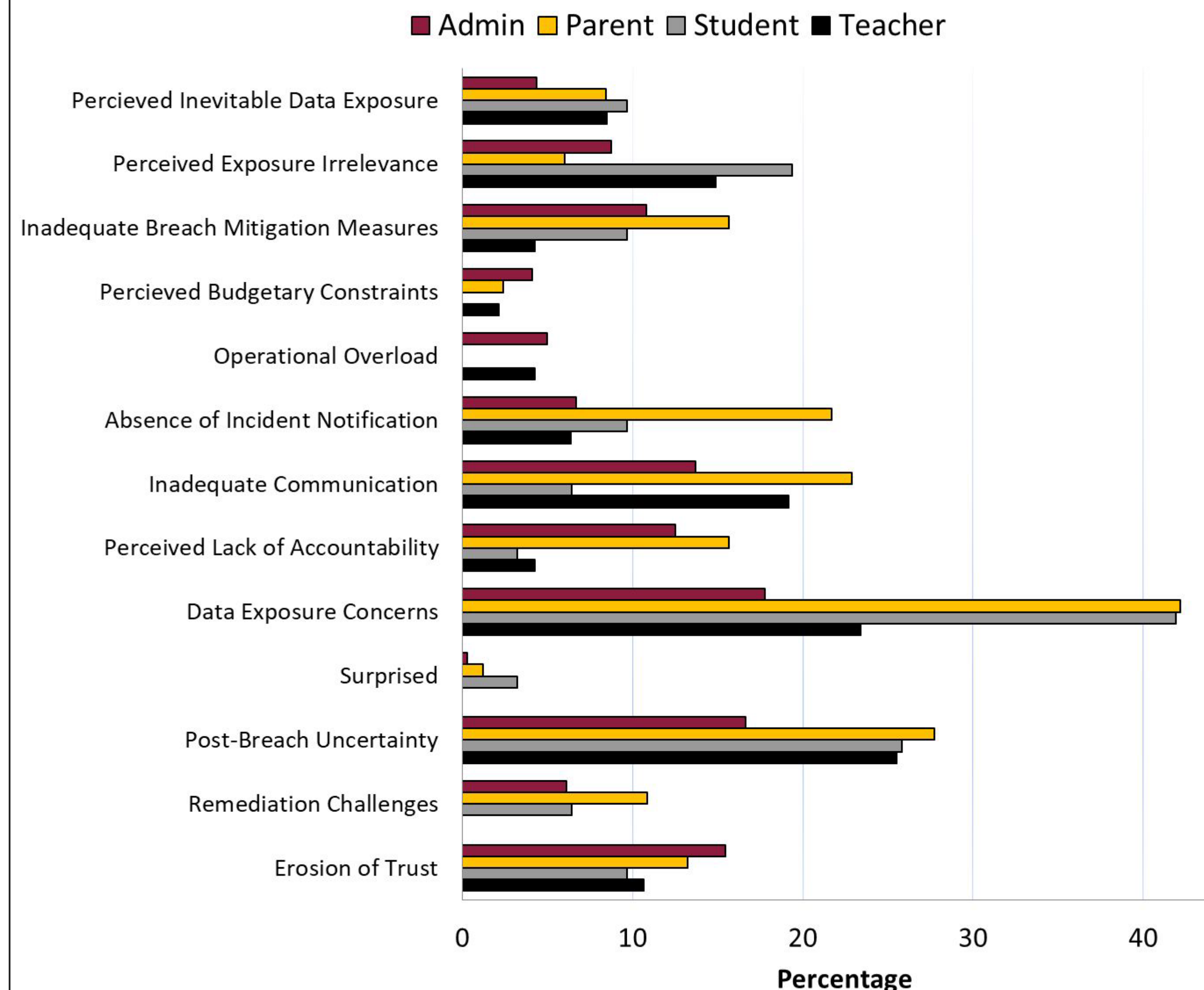**(5) Content Classification:** Used **o4-mini** to label the theme of each comment

(6) Validation & Labeling: Manually checked **100 random labels**, fine-tuned prompt, and labeled the rest.

Skills Learned

- End-to-end research process
- Python programming and API integration
- AI-powered analysis of large-scale data

Findings

Horizontal bar chart — Percentage by role (Admin, Parent, Student, Teacher) across themes:
- Percieved Inevitable Data Exposure
- Perceived Exposure Irrelevance
- Inadequate Breach Mitigation Measures
- Percieved Budgetary Constraints
- Operational Overload
- Absence of Incident Notification
- Inadequate Communication
- Perceived Lack of Accountability
- Data Exposure Concerns
- Surprised
- Post-Breach Uncertainty
- Remediation Challenges
- Erosion of Trust

Future Work / Next Steps

- Polish and refine current project's approach and aim for a short or work-in-progress paper
- Deepen current results via **multi-stakeholder interviews** with the affected population (**students**, **parents**, **admins & district leadership**)

Data Exposure Concerns (misuse of information)

PARENT: "This is terrifying, I got the email just now and I thought it was a joke.. I signed up for child monitoring and my 15 year old son's SSN is being used by some dude in Colorado!!!"

Perceived Exposure Irrelevance (not worried)

(Former) STUDENT: "I'm not too concern, 90% of the compromised information isn't relevant to me today. And I graduated from high school 10 years ago. Let the hackers see my terrible grades from high school lol."

Inadequate Post-Breach Communication

ADMIN: "There seem to be two types of emails PS sent out about this, one stating explicitly that your data was compromised, the second being one that is deliberately vague and noncommittal about your data's involvement."

Post-Breach Uncertainty & Confusions

TEACHER: "I have a lot of questions but zero background on this topic/cybersecurity. What exactly happened? What should parents of students and teachers who had data involved in the leak be doing/asking to protect themselves?"

Challenges

- Limitations of social media content
- Difficult to gather data from students under 13 due to Reddit age restrictions