

<b>A3 No. and Name</b>
899 and Group 1
<b>Team Leader (name &amp; 'phone ext)</b>
Srinu Babu Rai

<b>Team members (name &amp; role)</b>
1. Shiranth Stephen Sahaya Anbu Anitha
2. Bhupender Sejwal
3. Preetpal Singh
4.

<b>Stakeholders (role &amp; department)</b>
1. AI & ML Coordinator, Conestoga College
2. Potential Client name(s)
3. Other Conestoga College stakeholder(s)
4.

<b>Company objective</b>
Start date & planned duration

<b>SafeplayAI</b>
-------------------

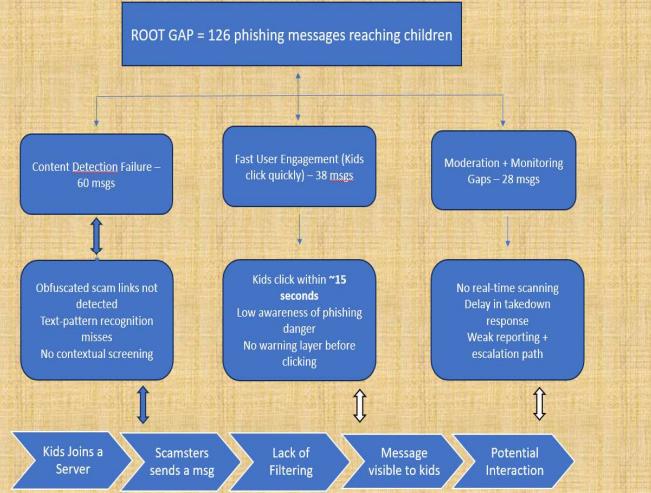
**1. Clarify the problem**

**Current Situation**  
Dataset contains 2000 total messages  
14% are phishing messages  
Current moderation captures only 50% of phishing content  
→ ~140 blocked / 140 total messages exposed to users (children)

**Ideal Situation**  
SafePlay AI will automatically detect and reduce phishing exposure from 140 → 14 messages, achieving 90% protection for child users.  
Final phishing rate should be ≤0.7%

**4. Analyse the Root Cause**

- Why do phishing messages bypass filters?  
Attackers obfuscate URLs + alter keywords
- Why do filters fail when obfuscation happens?  
Detection is rule-based and pattern dependent
- Why is the system only rule-driven?  
No adaptive or contextual scanning in place
- Why is context not analyzed?  
Filter architecture does not interpret intent
- Why is filtering system limited to static rules & cannot interpret evolving phishing patterns

**2. Breakdown the problem****3. Set the Target**

- 1 Improve phishing detection accuracy from 50% to 90%.
- 2 Reduce false positives from 20% to below 5%.
- 3 Cut scam exposure from 14% to under 5%.
- 4

**5. Develop Countermeasures****6. Implement Countermeasure****7. Monitor Results & Process****8. Standardise & Share Success**